INTERNATIONAL STANDARD

ISO/IEC/IEEE 8802-3

Second edition
2017-03-01

AMENDMENT 3
2017-11

Information technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements —

Part 3: Standard for Ethernet

AMENDMENT 3: Physical layers and management parameters for 25 Gb/s and 40 Gb/s operation, types 25GBASE-T and 40GBASE-T

Technologies de l'information — Télécommunications et échange d'information entre systèmes — Réseaux locaux et métropolitains — Prescriptions spécifiques —

Partie 3: Norme pour Ethernet

AMENDEMENT 3: Couches physiques et paramètres de gestion pour l'exploitation des interfaces à 25 Go/s et 40 Go/s, types 25GBASE-T et 40GBASE-T

Reference number
ISO/IEC/IEEE 8802-3:2017/Amd.3:2017(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Institute of Electrical and Electronics Engineers, Inc
3 Park Avenue, New York
NY 10016-5997, USA

stds.ipr@ieee.org
www.ieee.org

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

IEEE Standards documents are developed within the IEEE Societies and the Standards Coordinating Committees of the IEEE Standards Association (IEEE-SA) Standards Board. The IEEE develops its standards through a consensus development process, approved by the American National Standards Institute, which brings together volunteers representing varied viewpoints and interests to achieve the final product. Volunteers are not necessarily members of the Institute and serve without compensation. While the IEEE administers the process and establishes rules to promote fairness in the consensus development process, the IEEE does not independently evaluate, test, or verify the accuracy of any of the information contained in its standards.

The main task of ISO/IEC JTC 1 is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is called to the possibility that implementation of this standard may require the use of subject matter covered by patent rights. By publication of this standard, no position is taken with respect to the existence or validity of any patent rights in connection therewith. ISO/IEEE is not responsible for identifying essential patents or patent claims for which a license may be required, for conducting inquiries into the legal validity or scope of patents or patent claims or determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance or a Patent Statement and Licensing Declaration Form, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from ISO or the IEEE Standards Association.

ISO/IEC/IEEE 8802-3:2017/Amd.3 was prepared by the LAN/MAN of the IEEE Computer Society (as IEEE STD 802.3bq-2016). It was adopted by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 6, *Telecommunications and information exchange between systems,* in parallel with its approval by the ISO/IEC national bodies, under the “fast-track procedure” defined in the Partner Standards Development Organization cooperation agreement between ISO and IEEE. IEEE is responsible for the maintenance of this document with participation and input from ISO/IEC national bodies.

IEEE Std 802.3bq™-2016
(Amendment to
IEEE Std 802.3™-2015
as amended by
IEEE Std 802.3bw™-2015, and
IEEE Std 802.3by™-2016)

IEEE Standard for Ethernet

Amendment 3: Physical Layer and Management Parameters for 25 Gb/s and 40 Gb/s Operation, Types 25GBASE-T and 40GBASE-T

Sponsor

LAN/MAN Standards Committee
of the
IEEE Computer Society

Approved 30 June 2016

IEEE-SA Standards Board

Abstract: This amendment to IEEE Std 802.3-2015 specifies new Physical Coding Sublayer (PCS) interfaces and new Physical Medium Attachment (PMA) sublayer interfaces for 25 Gb/s Ethernet and 40 Gb/s Ethernet. 25GBASE-T and 40GBASE-T specify LAN interconnects for up to 30 m of balanced twisted-pair structured cabling, for 25 Gb/s and 40 Gb/s, respectively.

Keywords: 25GBASE-T, 25 Gigabit Ethernet, 25GMII, 40GBASE-T, 40 Gigabit Ethernet, Auto-Negotiation, Ethernet, IEEE 802®, IEEE 802.3™, IEEE 802.3bq™, Physical Coding Sublayer (PCS), Physical Medium Attachment (PMA) sublayer, structured cabling, XLGMII

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

Print: ISBN 978-1-5044-2282-6 STD21086
PDF: ISBN 978-1-5044-2283-3 STDPD21086

Important Notices and Disclaimers Concerning IEEE Standards Documents

IEEE documents are made available for use subject to important notices and legal disclaimers. These notices and disclaimers, or a reference to this page, appear in all standards and may be found under the heading "Important Notice" or "Important Notices and Disclaimers Concerning IEEE Standards Documents."

Notice and Disclaimer of Liability Concerning the Use of IEEE Standards Documents

IEEE Standards documents (standards, recommended practices, and guides), both full-use and trial-use, are developed within IEEE Societies and the Standards Coordinating Committees of the IEEE Standards Association ("IEEE-SA") Standards Board. IEEE ("the Institute") develops its standards through a consensus development process, approved by the American National Standards Institute ("ANSI"), which brings together volunteers representing varied viewpoints and interests to achieve the final product. Volunteers are not necessarily members of the Institute and participate without compensation from IEEE. While IEEE administers the process and establishes rules to promote fairness in the consensus development process, IEEE does not independently evaluate, test, or verify the accuracy of any of the information or the soundness of any judgments contained in its standards.

IEEE does not warrant or represent the accuracy or content of the material contained in its standards, and expressly disclaims all warranties (express, implied and statutory) not included in this or any other document relating to the standard, including, but not limited to, the warranties of: merchantability; fitness for a particular purpose; non-infringement; and quality, accuracy, effectiveness, currency, or completeness of material. In addition, IEEE disclaims any and all conditions relating to: results; and workmanlike effort. IEEE standards documents are supplied "AS IS" and "WITH ALL FAULTS."

Use of an IEEE standard is wholly voluntary. The existence of an IEEE standard does not imply that there are no other ways to produce, test, measure, purchase, market, or provide other goods and services related to the scope of the IEEE standard. Furthermore, the viewpoint expressed at the time a standard is approved and issued is subject to change brought about through developments in the state of the art and comments received from users of the standard.

In publishing and making its standards available, IEEE is not suggesting or rendering professional or other services for, or on behalf of, any person or entity nor is IEEE undertaking to perform any duty owed by any other person or entity to another. Any person utilizing any IEEE Standards document, should rely upon his or her own independent judgment in the exercise of reasonable care in any given circumstances or, as appropriate, seek the advice of a competent professional in determining the appropriateness of a given IEEE standard.

IN NO EVENT SHALL IEEE BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE PUBLICATION, USE OF, OR RELIANCE UPON ANY STANDARD, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Translations

The IEEE consensus development process involves the review of documents in English only. In the event that an IEEE standard is translated, only the English version published by IEEE should be considered the approved IEEE standard.

Official statements

A statement, written or oral, that is not processed in accordance with the IEEE-SA Standards Board Operations Manual shall not be considered or inferred to be the official position of IEEE or any of its committees and shall not be considered to be, or be relied upon as, a formal position of IEEE. At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that his or her views should be considered the personal views of that individual rather than the formal position of IEEE.

Comments on standards

Comments for revision of IEEE Standards documents are welcome from any interested party, regardless of membership affiliation with IEEE. However, IEEE does not provide consulting information or advice pertaining to IEEE Standards documents. Suggestions for changes in documents should be in the form of a proposed change of text, together with appropriate supporting comments. Since IEEE standards represent a consensus of concerned interests, it is important that any responses to comments and questions also receive the concurrence of a balance of interests. For this reason, IEEE and the members of its societies and Standards Coordinating Committees are not able to provide an instant response to comments or questions except in those cases where the matter has previously been addressed. For the same reason, IEEE does not respond to interpretation requests. Any person who would like to participate in revisions to an IEEE standard is welcome to join the relevant IEEE working group.

Comments on standards should be submitted to the following address:

Secretary, IEEE-SA Standards Board
445 Hoes Lane
Piscataway, NJ 08854 USA

Laws and regulations

Users of IEEE Standards documents should consult all applicable laws and regulations. Compliance with the provisions of any IEEE Standards document does not imply compliance to any applicable regulatory requirements. Implementers of the standard are responsible for observing or referring to the applicable regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not in compliance with applicable laws, and these documents may not be construed as doing so.

Copyrights

IEEE draft and approved standards are copyrighted by IEEE under U.S. and international copyright laws. They are made available by IEEE and are adopted for a wide variety of both public and private uses. These include both use, by reference, in laws and regulations, and use in private self-regulation, standardization, and the promotion of engineering practices and methods. By making these documents available for use and adoption by public authorities and private users, IEEE does not waive any rights in copyright to the documents.

Photocopies

Subject to payment of the appropriate fee, IEEE will grant users a limited, non-exclusive license to photocopy portions of any individual standard for company or organizational internal use or individual, non-commercial use only. To arrange for payment of licensing fees, please contact Copyright Clearance Center, Customer Service, 222 Rosewood Drive, Danvers, MA 01923 USA; +1 978 750 8400. Permission to photocopy portions of any individual standard for educational classroom use can also be obtained through the Copyright Clearance Center.

Updating of IEEE Standards documents

Users of IEEE Standards documents should be aware that these documents may be superseded at any time by the issuance of new editions or may be amended from time to time through the issuance of amendments, corrigenda, or errata. An official IEEE document at any point in time consists of the current edition of the document together with any amendments, corrigenda, or errata then in effect.

Every IEEE standard is subjected to review at least every ten years. When a document is more than ten years old and has not undergone a revision process, it is reasonable to conclude that its contents, although still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to determine that they have the latest edition of any IEEE standard.

In order to determine whether a given document is the current edition and whether it has been amended through the issuance of amendments, corrigenda, or errata, visit the IEEE-SA Website at http://ieeexplore.ieee.org/Xplore/home.jsp or contact IEEE at the address listed previously. For more information about the IEEE-SA or IEEE's standards development process, visit the IEEE-SA Website at http://standards.ieee.org.

Errata

Errata, if any, for all IEEE standards can be accessed on the IEEE-SA Website at the following URL: http://standards.ieee.org/findstds/errata/index.html. Users are encouraged to check this URL for errata periodically.

Patents

Attention is called to the possibility that implementation of this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. If a patent holder or patent applicant has filed a statement of assurance via an Accepted Letter of Assurance, then the statement is listed on the IEEE-SA Website at http://standards.ieee.org/about/sasb/patcom/patents.html. Letters of Assurance may indicate whether the Submitter is willing or unwilling to grant licenses under patent rights without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination to applicants desiring to obtain such licenses.

Essential Patent Claims may exist for which a Letter of Assurance has not been received. The IEEE is not responsible for identifying Essential Patent Claims for which a license may be required, for conducting inquiries into the legal validity or scope of Patents Claims, or determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from the IEEE Standards Association.

Participants

The following individuals were officers and members of the IEEE 802.3 Working Group at the beginning of the IEEE P802.3by Working Group ballot. Individuals may have not voted, voted for approval, disapproval or abstained on this standard.

David J. Law, *IEEE 802.3 Working Group Chair*
Adam Healey, *IEEE 802.3 Working Group Vice-Chair*
Peter Anslow, *IEEE 802.3 Working Group Secretary*
Steven B. Carlson, *IEEE 802.3 Working Group Executive Secretary*
Valerie Maguire, *IEEE 802.3 Working Group Treasurer*

David Chalupsky, *IEEE P802.3bq 25G/40GBASE-T Task Force Chair*
George Zimmerman, *IEEE P802.3bq 25G/40GBASE-T Task Force Editor-in-Chief*

John Abbott
David Abramson
Shadi Abughazaleh
Faisal Ahmad
Dale Amason
J. Michael Andrewartha
Oleksandr Babenko
Kwang-Hyun Baek
Amrik Bains
Koussalya Balasubramanian
Thananya Baldwin
Denis Beaudoin
Christian Beia
Yakov Belopolsky
Michael Bennett
Vipul Bhatt
William Bliss
Brad Booth
Martin Bouda
David Brandt
Ralf-Peter Braun
Theodore Brillhart
Paul Brooks
David Brown
Matthew Brown
Thomas Brown
Phillip Brownlee
Mark Bugg
Juan-Carlos Calderon
J. Martin Carroll
Clark Carty
Mandeep Chadha
Jacky Chang
Xin Chang
David Chen
Wheling Cheng
Ahmad Chini
Golam Choudhury
Peter Cibula
Christopher R. Cole
Keith Conroy
Eugene Dai
Shaoan Dai
John D'Ambrosia
Mike Darling
Yair Darshan
Piers Dawe
Fred Dawson
Ian Dedic
Chris Diminico
Thuyen Dinh
Curtis Donahue
Dan Dove
Mike Dudek
David Dwelley
Hesham Elbakoury
David Estes
John Ewen
Josef Faller
Shahar Feldman
German Feyh
Alan Flatman
Howard Frazier
Richard Frosch
Michael Furlong
Andrew Gardner
Mike Gardner
Ali Ghiasi
Joel Goergen
Zhigang Gong
Steven Gorshe
James Graba
Robert Grow
Mark Gustlin
Marek Hajduczenia
Bernie Hammond
Jeffrey Heath
Carl Herman
David Hess
Yasuo Hidaka
Riu Hirai
Thomas Hogenmueller
Brian Holden
Rita Horner
Bernd Horrmeyer
Victor Hou
Rui Hua
Liang-wei Huang
Scott Irwin
Kazuhiko Ishibe
Hideki Isono
Tom Issenhuth
Kenneth Jackson
Jack Jewell
Wenbin Jiang
Andrew Jimenez
Chad Jones
Antony Joseph
Yasuaki Kawatsu
Michael Kelsen
Yongbum Kim
Jonathan King
Scott Kipp
Michael Klempa
Avi Kliger
Curtis Knittle
Shigeru Kobayashi
Keisuke Kojima
Paul Kolesar
Tom Kolze
Glen Kramer
Albert Kuo
Hans Lackner
Efstathios Larios
Wayne Larsen
Mark Laubach
Greg Le Cheminant
Arthur Lee
David Lewis
Lei Li
Mike Peng Li
Shaohua Li
Thomas Lichtenegger

Ru Jian Lin
Robert Lingle
James Liu
Zhenyu Liu
William Lo
Miklos Lukacs
Kent Lusted
Jeffery Maki
James Malkemus
Yonatan Malkiman
Edwin Mallette
Arthur Marris
Chris Mash
Kirsten Matheus
Erdem Matoglu
Laurence Matola
Brett Mcclellan
Thomas Mcdermott
John McDonough
Richard Mei
Richard Mellitz
Bryan Moffitt
Leo Montreuil
Paul Mooney
Charles Moore
Andy Moorwood
Thomas Mueller
Ron Muir
Dale Murray
Henry Muyshondt
Edward Nakamoto
Gary Nicholl
Paul Nikolich
Kevin Noll
Ronald Nordin
Mark Nowell
David Ofelt
Ichiro Ogura
Tom Palkert
Sujan Pandey
Sesha Panguluri
Carlos Pardo
Moon Park
Pravin Patel
Petar Pepeljugoski
Gerald Pepper
Ruben Perez De Aranda Alonso
Michael Peters
John Petrilla
Rick Pimpinella
Neven Pischl
Rainer Poehmerer
William Powell
Richard Prodan
Rick Rabinovich
Saifur Rahman
Adee Ran
Ram Rao
Alon Regev
Duane Remein
Victor Renteria
Michael Ressl
Poldi (Pavlick) Rimboim
Christopher Roth
Salvatore Rotolo
Hisaya Sakamoto
Vineet Salunke
Sam Sambasivan
Yasuo Sasaki
Fred Schindler
Stefan Schneele
Peter Scruton
Alexander Seiger
Naoshi Serizawa
Megha Shanbhag
Masood Shariff
Stephen Shellhammer
Bazhong Shen
Mizuki Shirao
Kapil Shrikhande
Jeff Slavick
Scott Sommers
Yoshiaki Sone
Xiaolu Song
Tom Souvignier
Bryan Sparrowhawk
Edward Sprague
Peter Stassar
Leonard Stencel
Robert Stone
Steve Swanson
Andre Szczepanek
William Szeto
Bharat Tailor
Akio Tajima
Takayuki Tajima
Tomoo Takahara
Satoshi Takahashi
Kiyoto Takahata
Alexander Tan
Toshiki Tanaka
Mehmet Tazebay
Brian Teipen
Geoffrey Thompson
Alan Tipper
Pirooz Tooyserkani
Nathan Tracy
David Tremblay
Albert Tretter
Stephen Trowbridge
Wen-Cheng Tseng
Yoshihiro Tsukamoto
Mike Tu
Alan Ugolini
John Ulm
Ed Ulrichs
Sterling A. Vaden
Stefano Valle
Paul Vanderlaan
Robert Wagner
Robert Wang
Tongtong Wang
Xiaofeng Wang
Xinyuan Wang
Zhong Feng Wang
Markus Weber
Brian Welch
Yang Wen
Matthias Wendt
Oded Wertheim
Natalie Wienckowski
Ludwig Winkel
Peter Wu
Yu Xu
Lennart Yseboodt
Liquan Yuan
Hayato Yuki
Garold Yurko
Andrew Zambell
Jin Zhang
Yan Zhuang
Helge Zinner
Pavel Zivny
Gaoling Zou

The following members of the individual balloting committee voted on this standard. Balloters may have voted for approval, disapproval, or abstention.

Shadi Abughazaleh
Thomas Alexander
Richard Alfvin
Carol Ansley
Peter Anslow
Butch Anton
Stefan Aust
Amrik Bains
Saman Behtash
Michael Bennett
Gennaro Boggia
Ralf-Peter Braun
Nancy Bravin
Theodore Brillhart
Matthew Brown
Jairo Bustos Heredia
William Byrd
Steven B. Carlson
Juan Carreon
Clark Carty
Mandeep Chadha
David Chalupsky
Minho Cheong
Keng Hua Chuang
Peter Cibula
Charles Cook
John D'Ambrosia
Christopher Diminico
Yezid Donoso
Sourav Dutta
German Feyh
Matthias Fritsche
Yukihiro Fujimoto
James Graba
Randall Groves
Marek Hajduczenia
Adam Healey
Marco Hernandez
David Hess
Yasuo Hidaka
Guido Hiertz
Werner Hoelzl
Rita Horner
Tetsushi Ikegami
Noriyuki Ikeuchi
Sergiu Iordanescu
Scott Irwin
Osamu Ishida
Atsushi Ito
Peter Jones
Adri Jovin
Shinkyo Kaku
Piotr Karocki
John Kay
Stuart Kerry
Yongbum Kim
Scott Kipp
Bruce Kraemer
Mark Laubach
David L. Law
David Lewis
Jon Lewis
Arthur H. Light
William Lo
Elvis Maculuba
Valerie Maguire
Jeffery Maki
Arthur Marris
Brett Mcclellan
Thomas Mcdermott
Richard Mellitz
Bryan Moffitt
Charles Moorwood
Michael Newman
Nick S. A. Nikjoo
Satoshi Obara
Arumugam Paventhan
Michael Peters
Adee Ran
R. K. Rannow
Alon Regev
Maximilian Riegel
Robert Robinson
Benjamin Rolfe
Martin Rossbach
Frank Schewe
Dieter Schicketanz
Takeshi Shimizu
Kapil Shrikhande
David Solomon
Ju-Hyung Son
Thomas Starai
Peter Stassar
Eugene Stoudenmire
Walter Struppler
Mitsutoshi Sugawara
Patricia Thaler
James Theodoras
David Thompson
Geoffrey Thompson
Michael Thompson
Sterling Vaden
Dmitri Varsanofiev
Prabodh Varshney
Balasubramanian Vittal
George Vlantis
Roy Wang
Hung-Yu Wei
Peter Wu
Oren Yuen
Andrew Zambell
Zhen Zhou
George Zimmerman

When the IEEE-SA Standards Board approved this standard on 30 June 2016, it had the following membership:

Jean-Philippe Faure, *Chair*
Ted Burse, *Vice Chair*
John D. Kulick, *Past Chair*
Konstantinos Karachalios, *Secretary*

Chuck Adams
Masayuki Ariyoshi
Stephen Dukes
Jianbin Fan
Ronald W. Hotchkiss
J. Travis Griffith
Gary Hoffman
Michael Janezic
Joseph L. Koepfinger*
Hung Ling
Kevin Lu
Gary Robinson
Annette D. Reilly
Mehmet Ulema
Yingli Wen
Howard Wolfman
Don Wright
Yu Yuan
Daidi Zhong

*Member Emeritus

Introduction

This introduction is not part of IEEE Std 802.3bq™-2016, IEEE Standard for Ethernet—Amendment 3: Physical Layer and Management Parameters for 25 Gb/s and 40 Gb/s Operation, Types 25GBASE-T and 40GBASE-T.

IEEE Std 802.3 was first published in 1985. Since the initial publication, many projects have added functionality or provided maintenance updates to the specifications and text included in the standard. Each IEEE 802.3 project/amendment is identified with a suffix (e.g., IEEE Std 802.3ba™-2010).

The half duplex Media Access Control (MAC) protocol specified in IEEE Std 802.3-1985 is Carrier Sense Multiple Access with Collision Detection (CSMA/CD). This MAC protocol was key to the experimental Ethernet developed at Xerox Palo Alto Research Center, which had a 2.94 Mb/s data rate. Ethernet at 10 Mb/s was jointly released as a public specification by Digital Equipment Corporation (DEC), Intel, and Xerox in 1980. Ethernet at 10 Mb/s was approved as an IEEE standard by the IEEE Standards Board in 1983 and subsequently published in 1985 as IEEE Std 802.3-1985. Since 1985, new media options, new speeds of operation, and new capabilities have been added to IEEE Std 802.3. A full duplex MAC protocol was added in 1997.

Some of the major additions to IEEE Std 802.3 are identified in the marketplace with their project number. This is most common for projects adding higher speeds of operation or new protocols. For example, IEEE Std 802.3u™ added 100 Mb/s operation (also called Fast Ethernet), IEEE Std 802.3z™ added 1000 Mb/s operation (also called Gigabit Ethernet), IEEE Std 802.3ae™ added 10 Gb/s operation (also called 10 Gigabit Ethernet), IEEE Std 802.3ah™ specified access network Ethernet (also called Ethernet in the First Mile), and IEEE Std 802.3ba added 40 Gb/s operation (also called 40 Gigabit Ethernet) and 100 Gb/s operation (also called 100 Gigabit Ethernet). These major additions are all now included in and are superseded by IEEE Std 802.3-2015 and are not maintained as separate documents.

At the date of IEEE Std 802.3bq-2016 publication, IEEE Std 802.3 is composed of the following documents:

IEEE Std 802.3-2015

Section One—Includes Clause 1 through Clause 20 and Annex A through Annex H and Annex 4A. Section One includes the specifications for 10 Mb/s operation and the MAC, frame formats, and service interfaces used for all speeds of operation.

Section Two—Includes Clause 21 through Clause 33 and Annex 22A through Annex 33E. Section Two includes management attributes for multiple protocols and speed of operation as well as specifications for providing power over twisted-pair cabling for multiple operational speeds. It also includes general information on 100 Mb/s operation as well as most of the 100 Mb/s Physical Layer specifications.

Section Three—Includes Clause 34 through Clause 43 and Annex 36A through Annex 43C. Section Three includes general information on 1000 Mb/s operation as well as most of the 1000 Mb/s Physical Layer specifications.

Section Four—Includes Clause 44 through Clause 55 and Annex 44A through Annex 55B. Section Four includes general information on 10 Gb/s operation as well as most of the 10 Gb/s Physical Layer specifications.

Section Five—Includes Clause 56 through Clause 77 and Annex 57A through Annex 76A. Clause 56 through Clause 67 and Clause 75 through Clause 77, as well as associated annexes, specify subscriber access and other Physical Layers and sublayers for operation from 512 kb/s to 10 Gb/s, and defines services and protocol elements that enable the exchange of IEEE 802.3 format frames between stations in a subscriber access network. Clause 68 specifies a 10 Gb/s Physical Layer specification. Clause 69 through Clause 74 and associated annexes specify Ethernet operation over electrical backplanes at speeds of 1000 Mb/s and 10 Gb/s.

Section Six—Includes Clause 78 through Clause 95 and Annex 83A through Annex 93C. Clause 78 specifies Energy-Efficient Ethernet. Clause 79 specifies IEEE 802.3 Organizationally Specific Link Layer Discovery Protocol (LLDP) type, length, and value (TLV) information elements. Clause 80 through Clause 95 and associated annexes include general information on 40 Gb/s and 100 Gb/s operation as well the 40 Gb/s and 100 Gb/s Physical Layer specifications. Clause 90 specifies Ethernet support for time synchronization protocols.

IEEE Std 802.3bw-2015

Amendment 1—This amendment includes changes to IEEE Std 802.3-2015 and adds Clause 96. This amendment adds 100 Mb/s Physical Layer (PHY) specifications and management parameters for operation on a single balanced twisted-pair copper cable.

IEEE Std 802.3by-2016

Amendment 2—This amendment includes changes to IEEE Std 802.3-2015 and adds Clause 105 through Clause 112, Annex 109A, Annex 109B, Annex 109C, Annex 110A, Annex 110B, and Annex 110C. This amendment adds MAC parameters, Physical Layers, and management parameters for the transfer of IEEE 802.3 format frames at 25 Gb/s.

IEEE Std 802.3bq-2016

Amendment 3—This amendment includes changes to IEEE Std 802.3-2015 and adds Clause 113 and Annex 113A. This amendment adds new Physical Layers for 25 Gb/s and 40 Gb/s operation over balanced twisted-pair structured cabling systems.

A companion document IEEE Std 802.3.1 describes Ethernet management information base (MIB) modules for use with the Simple Network Management Protocol (SNMP). IEEE Std 802.3.1 is updated to add management capability for enhancements to IEEE Std 802.3 after approval of the enhancements.

IEEE Std 802.3 will continue to evolve. New Ethernet capabilities are anticipated to be added within the next few years as amendments to this standard.

Contents

IEEE Standard for Ethernet

Amendment 3: Physical Layer and Management Parameters for 25 Gb/s and 40 Gb/s Operation, Types 25GBASE-T and 40GBASE-T

(This amendment is based on IEEE Std 802.3™-2015 as amended by IEEE Std 802.3bw™-2015 and IEEE Std 802.3by™-2016.)

NOTE—The editing instructions contained in this amendment define how to merge the material contained therein into the existing base standard and its amendments to form the comprehensive standard.

The editing instructions are shown in ***bold italic***. Four editing instructions are used: change, delete, insert, and replace. ***Change*** is used to make corrections in existing text or tables. The editing instruction specifies the location of the change and describes what is being changed by using ~~strikethrough~~ (to remove old material) and underscore (to add new material). ***Delete*** removes existing material. ***Insert*** adds new material without disturbing the existing material. Deletions and insertions may require renumbering. If so, renumbering instructions are given in the editing instruction. ***Replace*** is used to make changes in figures or equations by removing the existing figure or equation and replacing it with a new one. Editing instructions, change markings, and this NOTE will not be carried over into future editions because the changes will be incorporated into the base standard.[1]

Cross references that refer to clauses, tables, equations, or figures not covered by this amendment are highlighted in green.

[1]Notes in text, tables, and figures are given for information only, and do not contain requirements needed to implement the standard.

1. Introduction

1.3 Normative references

Insert the following references in alphanumeric order:

ANSI/TIA-568-C.2-1-2016, Balanced Twisted-Pair Telecommunications Cabling and Components Standard, Addendum 1: Specifications for 100 Ω Category 8 Cabling.[2]

ISO/IEC DIS 11801-1:2016, Information technology–Generic cabling for customer premises–Part 1: General Requirements.[3]

1.4 Definitions

Insert the 25GBASE-T definition after 1.4.64g 25GBASE-SR (added by IEEE Std 802.3by-2016) as follows:

1.4.64h 25GBASE-T: IEEE 802.3 Physical Layer specification for a 25 Gb/s LAN using four pairs of ANSI/TIA Category 8, ISO/IEC Class I, or ISO/IEC Class II balanced copper cabling. (See IEEE Std 802.3, Clause 113.)

Insert the 40GBASE-T definition after 1.4.72 40GBASE-SR4 as follows:

1.4.72a 40GBASE-T: IEEE 802.3 Physical Layer specification for a 40 Gb/s LAN using four pairs of ANSI/TIA Category 8, ISO/IEC Class I, or ISO/IEC Class II balanced copper cabling. (See IEEE Std 802.3, Clause 113.)

Insert the Category 8 balanced cabling definition after 1.4.131 Category 7A balanced cabling as follows:

1.4.131a Category 8 balanced cabling: Balanced 100 Ω cables and associated connecting hardware whose transmission characteristics are specified up to 2000 MHz (i.e., cabling components that meet the Category 8.1 or Category 8.2 requirements specified in ISO/IEC DIS 11801-1:2016 or Category 8 specified in ANSI/TIA-568-C.2-1-2016). In addition to the requirements outlined in ISO/IEC DIS 11801-1 and ANSI/TIA-568-C.2-1, IEEE Std 802.3 Clause 14, Clause 23, Clause 25, Clause 40, Clause 55, and Clause 113 specify additional requirements for this cabling when used with 10BASE-T, 100BASE-T, 10GBASE-T, 25GBASE-T, and 40GBASE-T.

Insert the Infofield definition into the list after 1.4.237 in-band signaling and insert the MultiGBASE-T definition into the list after 1.4.277 mixing segment as follows:

1.4.237a Infofield: A 16 octet frame transmitted at regular intervals containing messages for startup operation by certain PHYs. (See IEEE Std 802.3, Clause 55 and Clause 113).

1.4.277a MultiGBASE-T: PHYs that belong to the set of specific BASE-T PHYs at speeds in excess of 1000 Mb/s, including 10GBASE-T, 25GBASE-T, and 40GBASE-T. (See IEEE Std 802.3, Clause 55 and Clause 113).

[2]ANSI publications are available from the American National Standards Institute (http://www.ansi.org).
[3]ISO/IEC publications are available from the ISO Central Secretariat (http://www.iso.org/). ISO publications are also available in the United States from the American National Standards Institute (http://www.ansi.org/).

1.5 Abbreviations

Insert the following new abbreviation in alphanumeric order:

ACRF	attenuation to crosstalk ratio, far-end

28. Physical Layer link signaling for Auto-Negotiation on twisted pair

28.3 State diagrams and variable definitions

28.3.1 State diagram variables

Insert rows for 25GigT and 40GigT in the first list in 28.3.1 below the row for 10GigT as follows:

25GigT;represents that the 25GBASE-T PMA is the signal source.

40GigT;represents that the 40GBASE-T PMA is the signal source.

28.3.2 State diagram timers

Change last sentence in the first paragraph of definition of link_fail_inhibit_timer as follows:

link_fail_inhibit_timer
Timer for qualifying a link_status=FAIL indication or a link_status=READY indication when a specific technology link is first being established. A link will only be considered "failed" if the link_fail_inhibit_timer has expired and the link has still not gone into the link_status=OK state. The link_fail_inhibit_timer shall expire 750 ms to 1000 ms after entering the FLP LINK GOOD CHECK state for devices operating at 10/100/1000 Mb/s. The link_fail_inhibit_timer shall expire 2000 ms to 2250 ms after entering the FLP LINK GOOD CHECK state for devices ~~operating at 10 Gb/s~~ in the MultiGBASE-T PHY set.

Change description for the link_fail_inhibit_timer (10 Gb/s devices) Parameter in Table 28–9 as follows (unchanged rows not shown):

Table 28–9—Timer min./max. value summary

Parameter	Min.	Typ.	Max.	Units
link_fail_inhibit_timer (~~10 Gb/s~~ devices in the MultiGBASE-T set)	2000		2250	ms

28.5 Protocol implementation conformance statement (PICS) proforma for Clause 28, Physical Layer link signaling for Auto-Negotiation on twisted pair[4]

28.5.3 Major capabilities/options

Change row for *10G in 28.5.3 as follows (unchanged rows not shown):

Item	Feature	Subclause	Status	Support	Value/Comment
*~~10~~MG	Implementation supports a member of the MultiGBASE-T PHY set (see 1.4.277a) ~~10GBASE-T PHY~~	55 and	O		N/A

[4]*Copyright release for PICS proformas:* Users of this standard may freely reproduce the PICS proforma in this subclause so that it can be used for its intended purpose and may further publish the completed PICS.

28.5.4 PICS proforma tables for Physical Layer link signaling for Auto-Negotiation on twisted pair

28.5.4.8 State diagrams

Change rows for Items SD10 and SD11 in 28.5.4.8 as follows (unchanged rows not shown):

Item	Feature	Subclause	Status	Support	Value/Comment
SD10	link_fail_inhibit_timer (10/100/1000 Mb/s)	28.3.2	!~~10~~MG: M		Expires 750 ms to 1000 ms after entering the FLP LINK GOOD CHECK state
SD11	link_fail_inhibit_timer (MultiGBASE-T devices~~10 Gb/s~~)	28.3.2	~~10~~MG: M		Expires 2000 ms to 2250 ms after entering the FLP LINK GOOD CHECK state upon successful master/slave resolution

30. Management

30.2 Managed objects

30.2.5 Capabilities

Change column header from 10GBASE-T Operating Margin package (conditional) to MultiGBASE-T Operating Margin package (conditional) in Table 30-1e as follows (unchanged lines not shown):

Table 30–1e—Capabilities

			DTE								Repeater				MAU								
			Basic Package (mandatory)	Mandatory Package (mandatory)	Recommended Package (optional)	Optional Package (optional)	Array Package (optional)	Excessive Deferral Package (optional)	Multiple PHY Package (optional)	PHY Error Monitor Capability (optional)	Basic Control Capability (mandatory)	Performance Monitor Capability (optional)	Address Tracking Capability (optional)	100/1000 Mb/s Monitor Capability (optional)	1000 Mb/s Burst Monitor Capability (optional)	Basic Package (mandatory)	MAU Control Package (optional)	Media Loss Tracking Package (conditional)	Broadband DTE MAU Package (conditional)	MII Capability (conditional)	PHY Error Monitor Capability (optional)	40MultiGBASE-T Operating Margin package (conditional)	Forward Error Correction Package (conditional)	Energy-Efficient Ethernet (optional)	Auto-Negotiation Package (mandatory)

30.3 Layer management for DTEs

30.3.2 PHY device managed object class

30.3.2.1 PHY device attributes

30.3.2.1.2 aPhyType

Insert new entries in APPROPRIATE SYNTAX (as modified by IEEE Std 802.3bw-2015 and IEEE Std 802.3by-2016) after the entries for 25GBASE-R (25GBASE-T) and for 40GBASE-R (40GBASE-T) as follows:

25GBASE-T	Clause 113 25 Gb/s DSQ128
40GBASE-T	Clause 113 40 Gb/s DSQ128

30.3.2.1.3 aPhyTypeList

Insert new entries in APPROPRIATE SYNTAX (as modified by IEEE Std 802.3bw-2015 and IEEE Std 802.3by-2016) after the entries for 25GBASE-R (25GBASE-T) and for 40GBASE-R (40GBASE-T) as follows:

25GBASE-T	Clause 113 25 Gb/s DSQ128
40GBASE-T	Clause 113 40 Gb/s DSQ128

30.5 Layer management for medium attachment units (MAUs)

30.5.1 MAU managed object class

30.5.1.1 MAU attributes

30.5.1.1.2 aMAUType

Insert new entries in APPROPRIATE SYNTAX (as modified by IEEE Std 802.3bw-2015 and IEEE Std 802.3by-2016) after the entries for 25GBASE-SR (25GBASE-T) and for 40GBASE-FR (40GBASE-T) as follows:

25GBASE-T	Four-pair twisted-pair balanced copper cabling PHY as specified in Clause 113
40GBASE-T	Four-pair twisted-pair balanced copper cabling PHY as specified in Clause 113

30.5.1.1.4 aMediaAvailable

Change the sixth paragraph of BEHAVIOUR DEFINED AS (as modified by IEEE Std 802.3bw-2015 and IEEE Std 802.3by-2016) as follows:

> For 40 Gb/s and 100 Gb/s the enumerations map to value of the link_fault variable (see 81.3.4) within the Link Fault Signaling state diagram (see 81.3.4.1 and Figure 46-11) as follows: the values OK and Link Interruption maps to the enumeration “available”, the value Local Fault maps to the enumeration “not available” and the value Remote Fault maps to the enumeration “remote fault.”

Change the eighth paragraph of BEHAVIOUR DEFINED AS (as modified by IEEE Std 802.3by-2016) as follows:

> For 10 Gb/s and 25 Gb/s the enumerations map to value of the link_fault variable within the Link Fault Signaling state diagram (Figure 46–11) as follows: the values OK and Link Interruption maps to the enumeration “available”, the value Local Fault maps to the enumeration “not available” and the value Remote Fault maps to the enumeration “remote fault”.

30.5.1.1.19 aSNROpMarginChnlA

Change text to include both 25GBASE-T and 40GBASE-T (as part of the MultiGBASE-T operating package) as follows:

ATTRIBUTE

APPROPRIATE SYNTAX:
INTEGER

BEHAVIOUR DEFINED AS:

The current SNR operating margin measured at the slicer input for channel A for the ~~10GBASE-T~~ MultiGBASE-T PMA. It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –12.7 dB to 12.7 dB. If a Clause 45 MDIO Interface to the PMA/PMD is present, then this attribute maps to the SNR operating margin channel A register (see 45.2.1.66).;

30.5.1.1.20 aSNROpMarginChnlB

Change text to include 25GBASE-T and 40GBASE-T (as part of the MultiGBASE-T operating package) as follows:

ATTRIBUTE

APPROPRIATE SYNTAX:

INTEGER

BEHAVIOUR DEFINED AS:

The current SNR operating margin measured at the slicer input for channel B for the ~~10GBASE-T~~ MultiGBASE-T PMA. It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –12.7 dB to 12.7 dB. If a Clause 45 MDIO Interface to the PMA/PMD is present, then this attribute maps to the SNR operating margin channel B register (see 45.2.1.67).;

30.5.1.1.21 aSNROpMarginChnlC

Change text to include 25GBASE-T and 40GBASE-T (as part of the MultiGBASE-T operating package) as follows:

ATTRIBUTE

APPROPRIATE SYNTAX:

INTEGER

BEHAVIOUR DEFINED AS:

The current SNR operating margin measured at the slicer input for channel C for the ~~10GBASE-T~~ MultiGBASE-T PMA. It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –12.7 dB to 12.7 dB. If a Clause 45 MDIO Interface to the PMA/PMD is present, then this attribute maps to the SNR operating margin channel C register (see 45.2.1.68).;

30.5.1.1.22 aSNROpMarginChnlD

Change text to include 25GBASE-T and 40GBASE-T (as part of the MultiGBASE-T operating package) as follows:

ATTRIBUTE

APPROPRIATE SYNTAX:

INTEGER

BEHAVIOUR DEFINED AS:

The current SNR operating margin measured at the slicer input for channel D for the ~~10GBASE-T~~ MultiGBASE-T PMA. It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –12.7 dB to 12.7 dB. If a Clause 45 MDIO Interface to the PMA/PMD is present, then this attribute maps to the SNR operating margin channel D register (see 45.2.1.69).;

30.5.1.1.24 aLDFastRetrainCount

Change text of aLDFastRetrainCount to include 25GBASE-T and 40GBASE-T (as part of the Energy Efficient Ethernet package (optional)) as follows:

ATTRIBUTE

APPROPRIATE SYNTAX:

Generalized nonresettable counter. This counter has a maximum increment rate of 1000 counts per second

BEHAVIOUR DEFINED AS:

A count of the number of ~~10GBASE-T~~ fast retrains initiated by the local device. ~~The indication reflects the state of the PHY event~~ This counter can be derived from fr_tx_counter (see ~~45.2.1.79.2 and~~55.4.5.4~~1~~ and 113.4.5.4~~.~~). If a Clause 45 MDIO Interface to the PMA/PMD is present, then this attribute can be derived from the LD fast retrain count register (see 45.2.1.79.2).;

30.5.1.1.25 aLPFastRetrainCount

Change text of aLPFastRetrainCount to include 25GBASE-T and 40GBASE-T (as part of the Energy Efficient Ethernet package (optional)) as follows:

ATTRIBUTE

APPROPRIATE SYNTAX:

Generalized nonresettable counter. This counter has a maximum increment rate of 1000 counts per second

BEHAVIOUR DEFINED AS:

A count of the number of ~~10GBASE-T~~ fast retrains initiated by the link partner. ~~The indication reflects the state of the PHY event~~ This counter can be derived from fr_rx_counter (see ~~45.2.1.79.1 and~~55.4.5.4~~1~~ and 113.4.5.4~~.~~). If a Clause 45 MDIO Interface to the PMA/PMD is present, then this attribute can be derived from to the LP fast retrain count register (see 45.2.1.79.1).;

30.6 Management for link Auto-Negotiation

30.6.1 Auto-Negotiation managed object class

30.6.1.1 Auto-Negotiation attributes

30.6.1.1.5 aAutoNegLocalTechnologyAbility

Insert new entries in APPROPRIATE SYNTAX (as modified by IEEE Std 802.3by-2016) after the entries for 10GBASE-KRFD (25GBASE-T) and for 40GBASE-CR4 (40GBASE-T) as follows:

25GBASE-T	25GBASE-T as specified in Clause 113
40GBASE-T	40GBASE-T as specified in Clause 113

45. Management Data Input/Output (MDIO) Interface

45.2 MDIO Interface Registers

45.2.1 PMA/PMD registers

Change the names for register addresses 1.129 to 1.144, 1.145 through 1.146, and 1.147 in Table 45–3 as follows (unchanged rows not shown):

Table 45–3—PMA/PMD registers

Register address	Register name	Subclause
1.129	~~10~~MultiGBASE-T status	45.2.1.62
1.130	~~10~~MultiGBASE-T pair swap and polarity	45.2.1.63
1.131	~~10~~MultiGBASE-T TX power backoff and PHY short reach setting	45.2.1.64
1.132	~~10~~MultiGBASE-T test mode	45.2.1.65
1.133	~~10GBASE-T~~ SNR operating margin channel A	45.2.1.66
1.134	~~10GBASE-T~~ SNR operating margin channel B	45.2.1.67
1.135	~~10GBASE-T~~ SNR operating margin channel C	45.2.1.68
1.136	~~10GBASE-T~~ SNR operating margin channel D	45.2.1.69
1.137	~~10GBASE-T m~~Minimum margin channel A	45.2.1.70
1.138	~~10GBASE-T m~~Minimum margin channel B	45.2.1.71
1.139	~~10GBASE-T m~~Minimum margin channel C	45.2.1.72
1.140	~~10GBASE-T m~~Minimum margin channel D	45.2.1.73
1.141	~~10GBASE-T~~ RX signal power channel A	45.2.1.74
1.142	~~10GBASE-T~~ RX signal power channel B	45.2.1.75
1.143	~~10GBASE-T~~ RX signal power channel C	45.2.1.76
1.144	~~10GBASE-T~~ RX signal power channel D	45.2.1.77
1.145 through 1.146	~~10~~MultiGBASE-T skew delay	45.2.1.78
1.147	~~10~~MultiGBASE-T fast retrain status and control register	45.2.1.79

45.2.1.6 PMA/PMD control 2 register (Register 1.7)

Change the indicated two reserved rows and insert four new rows (two of which are reserved) in alphanumeric order in the 1.7.5:0 row of Table 45-7 (as modified by IEEE Std 802.3bw-2015 and IEEE Std 802.3by-2016), as follows (unchanged rows not shown):

Table 45–7—PMA/PMD control 2 register bit definitions

Bit(s)	Name	Description	R/W[a]
1.7.5:0	PMA/PMD type selection	1 1 0 1 1 1 = 25GBASE-T PMA 1 1 0 1 1 0 = reserved 1 1 0 1 0 x = reserved 1 1 0 ~~x~~0 x x = reserved~~ for future use~~ 1 0 0 1 1 1~~x~~ = reserved 1 0 0 1 1 0 = 40GBASE-T PMA	R/W

[a]R/W = Read/Write, RO = Read only

45.2.1.7 PMA/PMD status 2 register (Register 1.8)

45.2.1.7.4 Transmit fault (1.8.11)

Insert description locations for 25GBASE-T after 25GBASE-SR (see IEEE Std 802.3by-2016) and before 40GBASE-KR4 and description for 40GBASE-T after the row for 40GBASE-FR in Table 45–9 as follows (unchanged rows not shown):

Table 45–9—Transmit fault description location

PMA/PMD	Description location
25GBASE-T	113.4.2.2
40GBASE-T	113.4.2.2

45.2.1.7.5 Receive fault (1.8.10)

Insert description locations for 25GBASE-T after 25GBASE-SR (see IEEE Std 802.3by-2016) and before 40GBASE-KR4 and description for 40GBASE-T after the row for 40GBASE-FR in Table 45–10 as follows (unchanged rows not shown):

Table 45–10—Receive fault description location

PMA/PMD	Description location
25GBASE-T	113.4.2.4
40GBASE-T	113.4.2.4

45.2.1.8 PMD transmit disable register (Register 1.9)

Insert description location for 25GBASE-T after row for 25GBASE-SR (see IEEE Std 802.3by-2016) and before 40GBASE-KR and 40GBASE-T after the row for 40GBASE-FR in Table 45–12 as follows (unchanged rows not shown):

Table 45–12—Transmit disable description location

PMA/PMD	Description location
25GBASE-T	113.4.2.3
40GBASE-T	113.4.2.3

45.2.1.12 40G/100G PMA/PMD extended ability register (Register 1.13)

Change the name and description for bit 1.13.6 in Table 45–16 as follows (unchanged rows not shown):

Table 45–16—40G/100G PMA/PMD Extended Ability register bit definitions

Bit(s)	Name	Description	R/W[a]
1.13.6	~~Reserved~~40GBASE-T ability	~~Value always 0~~ 1 = PMA/PMD is able to perform 40GBASE-T 0 = PMA/PMD is not able to perform 40GBASE-T	RO

[a]RO = Read only

45.2.1.9a 40GBASE-T ability (1.13.6)

Insert 45.2.1.12.9a after 45.2.1.12.9 as follows:

When read as a one, bit 1.13.6 indicates that the PMA/PMD is able to operate as a 40GBASE-T PMA type. When read as a zero, bit 1.13.6 indicates that the PMA/PMD is not able to operate as a 40GBASE-T PMA type.

45.2.1.14b 25G PMA/PMD extended ability register (Register 1.19)

Change the reserved row, and insert a row below it with the description for bit 1.19.5 in Table 45–17bb (inserted by IEEE Std 802.3by-2016) as follows (unchanged rows not shown):

Table 45–17b—25G PMA/PMD extended ability register bit definitions

Bit(s)	Name	Description	R/W[a]
1.19.15:6 ~~5~~	Reserved	Value always 0	RO
1.19.5	25GBASE-T ability	1 = PMA/PMD is able to perform 25GBASE-T 0 = PMA/PMD is not able to perform 25GBASE-T	RO

[a]RO = Read only

45.2.1.14b.a 25GBASE-T ability (1.19.5)

Insert 45.2.1.14b.a before 45.2.1.14b.1 (inserted by IEEE Std 802.3by-2016) as follows:

When read as a one, bit 1.19.5 indicates that the PMA/PMD is able to operate as a 25GBASE-T PMA type. When read as a zero, bit 1.19.5 indicates that the PMA/PMD is not able to operate as a 25GBASE-T PMA type.

Change the title and text of 45.2.1.62 to include 25GBASE-T and 40GBASE-T as follows:

45.2.1.62 ~~10~~MultiGBASE-T status (Register 1.129)

The assignments of bits in the ~~10~~MultiGBASE-T status register is shown in Table 45–54.

45.2.1.62.1 LP information valid (1.129.0)

Change the text of 45.2.1.62.1 to include 25GBASE-T and 40GBASE-T as follows:

When read as a one, bit 1.129.0 indicates that the startup protocol defined in 55.4.2.5 (for 10GBASE-T) or 113.4.2.5 (for 25G/40GBASE-T) has been completed, and that the contents of bits 1.130.11:0, 1.131.15:10, 1.145.14:8, 1.146.14:8, and 1.146.6:0, which are established during the startup protocol, are valid. When read as a zero, bit 1.129.0 indicates that the startup process has not been completed, and that the contents of these bits that are established during the startup protocol are invalid. A ~~10GBASE-T~~ PMA in the MultiG-BASE-T set shall return a value of zero in bit 1.129.1 if PMA link_status=FAIL.

Change title of Table 45–54 to include 25GBASE-T and 40GBASE-T as follows:

Table 45–54—~~10~~MultiGBASE-T status register bit definitions

Change the title of 45.2.1.63 to include 25GBASE-T and 40GBASE-T as follows:

45.2.1.63 ~~10~~MultiGBASE-T pair swap and polarity register (Register 1.130)

Change the title in Table 45–55 to include 25GBASE-T and 40GBASE-T as follows:

Table 45–55—~~10~~MultiGBASE-T pair swap and polarity register bit definitions

Change the title and text of 45.2.1.64 to include 25GBASE-T and 40GBASE-T as follows:

45.2.1.64 ~~10~~MultiGBASE-T TX power backoff and PHY short reach setting (Register 1.131)

The complete assignment of bits in the ~~10~~MultiGBASE-T TX power backoff and short reach mode settings register is shown in Table 45–56.

Change the title in Table 45–56 to include 25GBASE-T and 40GBASE-T as follows:

Table 45–56—~~10~~MultiGBASE-T TX power backoff and PHY short reach setting register bit definitions

Change the title and text of 45.2.1.64.1 to include 25GBASE-T and 40GBASE-T as follows:

45.2.1.64.1 ~~10~~MultiGBASE-T TX power backoff settings (1.131.15:10)

The ~~10~~MultiGBASE-T TX power backoff settings reflects the TX power backoff selected during the startup negotiation process. The 10GBASE-T startup negotiation process and all TX power backoff settings are defined in 55.4.2.5 and 55.4.5.1. The 25GBASE-T and 40GBASE-T startup negotiation process and all TX power backoff settings are defined in 113.4.2.5 and 113.4.5.1. If LP information valid bit, 1.129.0, is set to one then bits 1.131.15:13 indicates the TX power backoff setting of the link partner and bits 1.131.12:10 indicates the TX power backoff setting of the local device.

45.2.1.64.2 PHY short reach mode (1.131.0)

Change the text of 45.2.1.64.2 to include 25GBASE-T and 40GBASE-T as follows:

The short reach mode of the 10GBASE-T PHY provides a means for operation on a cable plant that has parametric performance equivalent to 30 m of Class F and Class EA cabling as defined in 55.5.4.5. The short reach mode of the 25GBASE-T and 40GBASE-T PHYs provides for operation on a direct attach link segment that has parametric performance defined in 113.7.4. If bit 1.131.0 is a one, the PHY is in short reach mode. If bit 1.131.0 is a zero, the PHY is not in short reach mode~~the PHY is operating in normal mode~~. The default value for this bit is zero. For 25GBASE-T and 40GBASE-T, setting this bit to a one puts the PHY in short reach mode, and setting this bit to a zero puts the PHY into normal (non-short reach) mode.

Change the title and text of 45.2.1.65 to include 25GBASE-T and 40GBASE-T as follows:

45.2.1.65 ~~10~~MultiGBASE-T test mode register (Register 1.132)

The assignment of bits in the ~~10~~MultiGBASE-T test mode register is shown in Table 45–57. The default values for each bit should be chosen so that the initial state of the device upon power up or reset is a normal operational state without management intervention.

Change the title in Table 45–57 to include 25GBASE-T and 40GBASE-T as follows:

Table 45–57—~~10~~MultiGBASE-T test mode register bit definitions

Change text of 45.2.1.65.1 and 45.2.1.65.2 to include references to 25G/40GBASE-T and Clause 113 as follows:

45.2.1.65.1 Test mode control (1.132.15:13)

Transmitter test mode operations defined by bits 1.132.15:13~~,~~ are described for 10GBASE-T in 55.5.2 and Table 55–12, and for 25G/40GBASE-T in 113.5.2 and Table 113–17. The default value for bits 1.132.15:13 is zero.

45.2.1.65.2 Transmitter test frequencies (1.132.12:10)

When test mode 4 is selected by setting bits 1.132.15:13 to one, zero, zero respectively, bits 1.132.12:10 select the transmit test frequency as shown in Table 45–57. Detailed use and operation of these transmitter test frequencies is described in 55.5.2 and 113.5.2.

45.2.1.66 SNR operating margin channel A register (Register 1.133)

Change the text of 45.2.1.66 to include 25GBASE-T and 40GBASE-T as follows:

Register 1.133 contains the current SNR operating margin measured at the slicer input for channel A for the ~~10GBASE-T~~ PMAs in the MultiGBASE-T set. It is reported with 0.1 dB of resolution to an accuracy of 0.5 dB within the range of –12.7 dB to 12.7 dB. The number is in offset two's complement notation, with 0.0 dB represented by 0x8000. Implementation of this register is optional.

45.2.1.67 SNR operating margin channel B register (Register 1.134)

Change the text of 45.2.1.67 to include 25GBASE-T and 40GBASE-T as follows:

Register 1.134 contains the current SNR operating margin measured at the slicer input for channel B for the ~~10GBASE-T~~ PMAs in the MultiGBASE-T set. It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –12.7 dB to 12.7 dB. The number is in offset two's complement notation, with 0.0 dB represented by 0x8000. Implementation of this register is optional.

45.2.1.68 SNR operating margin channel C register (Register 1.135)

Change the text of 45.2.1.68 to include 25GBASE-T and 40GBASE-T as follows:

Register 1.135 contains the current SNR operating margin measured at the slicer input for channel C for the ~~10GBASE-T~~ PMAs in the MultiGBASE-T set. It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –12.7 dB to 12.7 dB. The number is in offset two's complement notation, with 0.0 dB represented by 0x8000. Implementation of this register is optional.

45.2.1.69 SNR operating margin channel D register (Register 1.136)

Change the text of 45.2.1.69 to include 25GBASE-T and 40GBASE-T as follows:

Register 1.136 contains the current SNR operating margin measured at the slicer input for channel D for the ~~10GBASE-T~~ PMAs in the MultiGBASE-T set. It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –12.7 dB to 12.7 dB. The number is in offset two's complement notation, with 0.0 dB represented by 0x8000. Implementation of this register is optional.

45.2.1.74 RX signal power channel A register (Register 1.141)

Change the text of 45.2.1.74 to include 25GBASE-T and 40GBASE-T as follows:

The RX signal power channel A register is read only and contains the receive signal power measured at the MDI during training as described in 55.4.3.1 and 113.4.3.1. The RX signal power should reflect the power measured when the device transitions out of the state PMA_Training_Init_M or MA_Training_Init_S (as appropriate, see 55.4.6.1 and 113.4.6.1), when the link partner is transmitting with PBO_tx = 4 (8 dB power backoff). It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –20 dBm to 5.5 dBm. The number is in offset two's complement notation, with 0.0 dB represented by 0x8000. Implementation of this register is optional.

45.2.1.75 RX signal power channel B register (Register 1.142)

Change the text of 45.2.1.75 to include 25GBASE-T and 40GBASE-T as follows:

The RX signal power channel B register is read only and contains the receive signal power measured at the MDI during training as described in 55.4.3.1 and 113.4.3.1. The RX signal power should reflect the power measured when the device transitions out of the state PMA_Training_Init_M or MA_Training_Init_S (as appropriate, see 55.4.6.1 and 113.4.6.1), when the link partner is transmitting with PBO_tx = 4 (8 dB power backoff). It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –20 dBm to 5.5 dBm. The number is in offset two's complement notation, with 0.0 dB represented by 0x8000. Implementation of this register is optional.

45.2.1.76 RX signal power channel C register (Register 1.143)

Change the text of 45.2.1.76 to include 25GBASE-T and 40GBASE-T as follows:

The RX signal power channel C register is read only and contains the receive signal power measured at the MDI during training as described in 55.4.3.1 and 113.4.3.1. The RX signal power should reflect the power measured when the device transitions out of the state PMA_Training_Init_M or MA_Training_Init_S (as appropriate, see 55.4.6.1 and 113.4.6.1), when the link partner is transmitting with PBO_tx = 4 (8 dB power backoff). It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –20 dBm to 5.5 dBm. The number is in offset two's complement notation, with 0.0 dB represented by 0x8000. Implementation of this register is optional.

45.2.1.77 RX signal power channel D register (Register 1.144)

Change the text of 45.2.1.77 to include 25GBASE-T and 40GBASE-T as follows:

The RX signal power channel D register is read only and contains the receive signal power measured at the MDI during training as described in 55.4.3.1 and 113.4.3.1. The RX signal power should reflect the power measured when the device transitions out of the state PMA_Training_Init_M or MA_Training_Init_S (as appropriate, see 55.4.6.1 and 113.4.6.1), when the link partner is transmitting with PBO_tx = 4 (8 dB power backoff). It is reported in units of 0.1 dB to an accuracy of 0.5 dB within the range of –20 dBm to 5.5 dBm. The number is in offset two's complement notation, with 0.0 dB represented by 0x8000. Implementation of this register is optional.

Change the title and text of 45.2.1.78 to include 25GBASE-T and 40GBASE-T as follows:

45.2.1.78 ~~10~~Multi GBASE-T skew delay register (Registers 1.145 and 1.146)

The skew delay register reports the current skew delay on each of the pair with respect to physical pair A (see Table 45–58). It is reported with ~~1.25 ns~~ resolution ~~to an accuracy of 2.5 ns~~ equal to one symbol period

(see 55.1.3 and 113.1.2) of the PHY (e.g., 1.25 ns for 10GBASE-T) to an accuracy of two symbol periods (e.g., 2.5 ns for 10GBASE-T). The number reported is in two's complement notation with positive values representing delay and negative values representing advance with respect to physical pair A. If the delay exceeds the maximum amount that can be represented by the range (–64 symbols to +63 symbols ~~–80 ns to +78.75 ns~~), the field displays the maximum respective value. The value shall be updated at least once per second.

Change title of Table 45–58 to include 25GBASE-T and 40GBASE-T.

Table 45–58—~~10~~MultiGBASE-T skew delay register bit definitions

Change the title of 45.2.1.79 to include 25GBASE-T and 40GBASE-T as follows:

45.2.1.79 ~~10~~MultiGBASE-T fast retrain status and control register (Register 1.147)

Change the title of Table 45–59 to include 25GBASE-T and 40GBASE-T as follows

Table 45–59—~~10~~MultiGBASE-T fast retrain status and control register bit definitions

45.2.1.79.1 LP fast retrain count (1.147.15:11)

Change the text of 45.2.1.79.1 to include 25GBASE-T and 40GBASE-T as follows:

These bits map to fr_rx_counter as defined in 55.4.5.4~~1~~ for 10GBASE-T and 113.4.5.4 for 25GBASE-T and 40GBASE-T. The counter is a 5-bit count of the number of ~~10GBASE-T~~ fast retrains requested by the link partner. These bits shall be reset to all zeros when read or upon execution of the PMA reset. These bits shall be held at all ones in the case of overflow.

45.2.1.79.2 LD fast retrain count (1.147.10:6)

Change the text of 45.2.1.79.2 to include 25GBASE-T and 40GBASE-T as follows:

These bits map to fr_tx_counter as defined in 55.4.5.4~~1~~ for 10GBASE-T and 113.4.5.4 for 25GBASE-T and 40GBASE-T. The counter is a 5-bit count of the number of ~~10GBASE-T~~ fast retrains requested by the local device. These bits shall be reset to all zeros when read or upon execution of the PMA reset. These bits shall be held at all ones in the case of overflow.

45.2.1.79.5 Fast retrain signal type (1.147.2:1)

Change the text of 45.2.1.79.5 to add reference to 113.3.6.2.2 as follows:

For PHYs that support fast retrain, these bits map to fr_sigtype as defined in 55.3.6.2.2 and 113.3.6.2.2. When Fast retrain signal type is set to 00, the PMA sends IDLE characgters on the receive path during fast retrain. When Fast retrain signal type is set to 01, the PMA sends Local Fault on the receive path during fast retrain. When Fast retrain signal type is set to 10, the PMA sends Link Interruption on the receive path during fast retrain.

45.2.1.79.6 Fast retrain enable (1.147.0)

Change the text of 45.2.1.79.6 to add reference to 113.4.5.1 as follows:

For PHYs that support fast retrain, this bit controls fr_enable as defined in 55.4.5.1 and 113.4.5.1. When PMA reset is executed, this bit is set to one.

NOTE—Setting this bit to zero while a link is up will cause the PHY to stop supporting fast retrain, and the link will drop if the link partner initiates a fast retrain.

45.2.3 PCS registers

Change the names of register addresses 3.20, 3.21, 3.32 and 3.33 in Table 45–119 as follows (unchanged rows not shown):

Table 45–119—PCS registers

Register address	Register name	Subclause
3.20	EEE control and capability 1	45.2.3.9
3.21	~~Reserved~~EEE control and capability 2	45.2.3.9a
3.32	BASE-R and ~~10~~MultiGBASE-T PCS status 1	45.2.3.13
3.33	BASE-R and ~~10~~MultiGBASE-T PCS status 2	45.2.3.14

45.2.3.1 PCS control 1 register (Register 3.0)

45.2.3.1.2 Loopback (3.0.14)

Change the text of 45.2.3.1.2 to include 25GBASE-T and 40GBASE-T as follows:

When ~~the 10GBASE-T~~ any MultiGBASE-T or the 10GBASE-R mode of operation is selected for the PCS using the PCS type selection field (3.7.~~2~~3:0), the PCS shall be placed in a loopback mode of operation when bit 3.0.14 is set to a one. When bit 3.0.14 is set to a one, the 10GBASE-R or ~~10GBASE-T~~ any PCS in the MultiGBASE-T set shall accept data on the transmit path and return it on the receive path. The speed of the loopback is selected by the PCS control 1 (Register 3.0) defined in 45.2.3.1. The specific behavior of the 10GBASE-R PCS during loopback is specified in 49.2. The specific behavior for the 10GBASE-T PCS during loopback is specified in 55.3.7.3. The specific behavior for the 25GBASE-T and 40GBASE-T PCS during loopback is specified in 113.3.7.3. For all other port types, the PCS loopback functionality is not applicable and writes to this bit shall be ignored and reads from this bit shall return a value of zero.

45.2.3.2 PCS status 1 register (Register 3.1)

45.2.3.2.7 PCS receive link status (3.1.2)

Change the text of 45.2.3.2.7 to include 25GBASE-T and 40GBASE-T as follows:

When read as a one, bit 3.1.2 indicates that the PCS receive link is up. When read as a zero, bit 3.1.2 indicates that the PCS receive link is down. When a 10/40/100GBASE-R, 10GBASE-W, or ~~10GBASE-T~~ any MultiGBASE-T mode of operation is selected for the PCS using the PCS type selection field (3.7.~~2~~3:0), this bit is a latching low version of bit 3.32.12. When a 10GBASE-X mode of operation is selected for the PCS using the PCS type selection field (3.7.~~2~~3:0), this bit is a latching low version of bit 3.24.12. The receive link status bit shall be implemented with latching low behavior.

45.2.3.6 PCS control 2 register (Register 3.7)

Change Table 45–123 (as modified by IEEE Std 802.3by-2016) as follows:

Table 45–123—PCS control 2 register bit definitions

Bit(s)	Name	Description	R/W[a]
3.7:15:~~3~~4	Reserved	Value always 0	RO
3.7.~~2~~3:0	PCS type selection	3 2 1 0 1 1 x x= reserved 1 0 1 x= reserved 1 0 0 1= Select 25GBASE-T PCS type 1 0 0 0= reserved 0 1 1 1= Select 25GBASE-R PCS type 0 1 1 0= ~~reserved~~Select 40GBASE-T PCS type 0 1 0 1= Select 100GBASE-R PCS type 0 1 0 0= Select 40GBASE-R PCS type 0 0 1 1= Select 10GBASE-T PCS type 0 0 1 0= Select 10GBASE-W PCS type 0 0 0 1= Select 10GBASE-X PCS type 0 0 0 0= Select 10GBASE-R PCS type	R/W

[a]R/W = Read/Write, RO = Read only

45.2.3.6.1 PCS type selection (3.7.~~2~~3:0)

Change the title and text of 45.2.3.6.1 (as modified by IEEE Std 802.3by-2016) to include revised PCS selection bits as follows:

The PCS type shall be selected using bits ~~2~~3 through 0. The PCS type abilities of the PCS are advertised in bits 3.8.9 and 3.8.7:0. A PCS shall ignore writes to the PCS type selection bits that select PCS types it has not advertised in the PCS status 2 register. It is the responsibility of the STA entity to ensure that mutually acceptable MMD types are applied consistently across all the MMDs on a particular PHY. The PCS type selection defaults to a supported ability.

45.2.3.7 PCS status 2 register (Register 3.8)

Change the reserved row in Table 45–124 (as modified by IEEE Std 802.3by-2016) as shown below, insert a new row for name and description for bit 3.8.9 above the reserved row, and change the reserved row 3.8.6 (as inserted by IEEE Std 802.3by-2016) as follows (unchanged rows not shown):

Table 45–124—PCS status 2 register bit definitions

Bit(s)	Name	Description	R/W[a]
3.8.9	25GBASE-T capable	1 = PCS is able to support 25GBASE-T PCS type 0 = PCS is not able to support 25GBASE-T PCS type	RO
3.8.~~9~~8	Reserved	Value always 0	RO
3.8.6	~~Reserved~~40GBASE-T capable	~~Value always 0~~1 = PCS is able to support 40GBASE-T PCS type 0 = PCS is not able to support 40GBASE-T PCS type	RO

[a]RO = Read only, LH = Latching high

Insert 45.2.3.7.3aa after 45.2.3.7.3 and before 45.2.3.7.3a (as inserted by IEEE Std 802.3by-2016) as follows:

45.2.3.7.3aa 25GBASE-T capable (3.8.9)

When read as a one, bit 3.8.9 indicates that the PCS is able to support the 25GBASE-T PCS type. When read as a zero, bit 3.8.9 indicates that the PCS is not able to support the 25GBASE-T PCS type.

Insert 45.2.3.7.3b after 45.2.3.7.3a (as inserted by IEEE Std 802.3by-2016) as follows:

45.2.3.7.3b 40GBASE-T capable (3.8.6)

When read as a one, bit 3.8.6 indicates that the PCS is able to support the 40GBASE-T PCS type. When read as a zero, bit 3.8.6 indicates that the PCS is not able to support the 40GBASE-T PCS type.

Change the name of Register 3.20 in title and text as shown:

45.2.3.9 EEE control and capability 1 (Register 3.20)

This register is used to indicate the capability of the PCS to support EEE functions for each PHY type. The assignment of bits in the EEE control and capability 1 register is shown in Table 45–125.

Change the title of Table 45-125, and the name and definition of bit 3.20.7 as follows (unchanged rows not shown):

Table 45–125—EEE control and capability 1 register bit definitions

Bit(s)	Name	Description	R/W[a]
3.20.7	~~Reserved~~40GBASE-T EEE	~~Values always 0~~1 = EEE is supported for 40GBASE-T 0 = EEE is not supported for 40GBASE-T	RO

[a] Read/Write, RO = Read only

Insert 45.2.3.9.4a after 45.2.3.9.4 as follows:

45.2.3.9.4a 40GBASE-T EEE supported (3.20.7)

If the device supports EEE operation for 40GBASE-T as defined in 113.1.3.3, this bit shall be set to one.

Insert 45.2.3.9a, Table 45-125a, and 45.2.3.9a.1 after 45.2.3.9 as follows:

45.2.3.9a EEE control and capability 2 (Register 3.21)

Register 3.21 extends register 3.20 and is used to indicate the capability of the PCS to support EEE functions for each PHY type. The assignment of bits in the EEE control and capability 2 register is shown in Table 45–125a.

Table 45–125a—EEE control and capability 2 register bit definitions

Bit(s)	Name	Description	R/W[a]
3.21.15:3	Reserved	Value always 0	RO
3.21.2	25GBASE-T EEE	1 = EEE is supported for 25GBASE-T 0 = EEE is not supported for 25GBASE-T	RO
3.21.1:0	Reserved	Value always 0	RO

[a] Read/Write, RO = Read only

45.2.3.9a.1 25GBASE-T EEE supported (3.21.2)

If the device supports EEE operation for 25GBASE-T as defined in 113.1.3.3, this bit shall be set to one.

45.2.3.13 BASE-R and ~~10~~MultiGBASE-T PCS status 1 register (Register 3.32)

Change the title and text of 45.2.3.13 to include 25GBASE-T and 40GBASE-T as follows:

The assignment of bits in the BASE-R and ~~10~~MultiGBASE-T PCS status 1 register is shown in Table 45–128. All the bits in the BASE-R and ~~10~~MultiGBASE-T PCS status 1 register are read only; a write to the BASE-R and ~~10~~MultiGBASE-T PCS status 1 register shall have no effect. A PCS device that ~~does not~~ implements neither BASE-R ~~and 10GBASE-T~~ nor any member of the MultiGBASE-T set shall return a zero for all bits in the BASE-R and ~~10~~MultiGBASE-T PCS status 1 register. It is the responsibility of the STA management entity to ensure that a port type is supported by all MMDs before interrogating any of its status bits. The

contents of register 3.32 are undefined when the BASE-R ~~PCS~~ or MultiGBASE-T ~~the 10GBASE-T~~ PCS is operating in seed test-pattern mode, PRBS31 test-pattern mode, or PRBS9 test-pattern mode.

Change the title and names and descriptions of bits 3.32.12, 3.3.2.1, and 3.32.0 rows in Table 45–128 as follows (unchanged bits not shown):

Table 45–128—BASE-R and ~~10~~MultiGBASE-T PCS status 1 register bit definitions

Bit(s)	Name	Description	R/W[a]
3.32.12	BASE-R and ~~10~~MultiGBASE-T receive link status	1 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS receive link up 0 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS receive link down	RO
3.32.1	BASE-R and ~~10~~MultiGBASE-T PCS high BER	1 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS reporting a high BER 0 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS not reporting a high BER	RO
3.32.0	BASE-R and ~~10~~MultiGBASE-T PCS block lock	1 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS locked to received blocks 0 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS not locked to received blocks	RO

[a]RO = Read only

45.2.3.13.1 BASE-R and ~~10~~MultiGBASE-T receive link status (3.32.12)

Change the title and text of 45.2.3.13.1 to include 25GBASE-T and 40GBASE-T as follows:

When read as a one, bit 3.32.12 indicates that the PCS is in a fully operational state. When read as a zero, bit 3.32.12 indicates that the PCS is not fully operational. This bit is a reflection of the PCS_status variable defined in 49.2.14.1 for 10GBASE-R, in 55.3.6.1 for 10GBASE-T, in 113.3.7.1 for 25GBASE-T and 40GBASE-T, and in 82.3.1 for 40/100GBASE-R.

45.2.3.13.4 BASE-R and ~~10~~MultiGBASE-T PCS high BER (3.32.1)

Change the title and text of 45.2.3.13.4 to include 25GBASE-T and 40GBASE-T as follows:

For BASE-R, when read as a one, bit 3.32.1 indicates that the 64B/66B receiver is detecting a BER of $\geq 10^{-4}$. When read as a zero, bit 3.32.1 indicates that the 64B/66B receiver is detecting a BER of $< 10^{-4}$. This bit is a direct reflection of the state of the hi_ber variable in the 64B/66B state diagram and is defined in 49.2.13.2.2 for 10GBASE-R and in 82.2.19.2.2 for 40/100GBASE-R.

For ~~10GBASE-T~~ any member of the MultiGBASE-T set, when read as a one, bit 3.32.1 indicates that the 64B/65B receiver is detecting a BER of $\geq 10^{-4}$. When read as a zero, bit 3.32.1 indicates that the 64B/65B receiver is detecting a BER of $< 10^{-4}$. This bit is a direct reflection of the state of the hi_lfer variable in the MultiGBASE-T 64B/65B state diagrams, ~~and is~~ defined in 55.3.6.1 for 10GBASE-T and 113.3.6.2.2 for 25GBASE-T and 40GBASE-T.

Change the title and text of 45.2.3.13.5 to include 25GBASE-T and 40GBASE-T as follows:

45.2.3.13.5 BASE-R and ~~10~~MultiGBASE-T block lock (3.32.0)

When read as a one, bit 3.32.0 indicates that the 64B/66B receiver for BASE-R or the 64B/65B receiver for ~~the 10GBASE-T~~ a member of the MultiGBASE-T set has block lock. When read as a zero, bit 3.32.0 indicates that the 64B/66B receiver for BASE-R or the 64B/65B receiver for the ~~10GBASE-T~~ member of the MultiGBASE-T set has not achieved block lock. This bit is a direct reflection of the state of the block_lock variable in the 64B/66B state diagram and is defined in 49.2.13.2.2 for 10GBASE-R and in 82.2.19.2.2 for 40/100GBASE-R. For the 10GBASE-T PCS the block_lock variable in the 64B/65B state diagram is defined in 55.3.2.3. For both the 25GBASE-T and 40GBASE-T PCS, the block_lock variable in the 64B/65B state diagram is defined in 113.3.6.2.2. For a multi-lane PCS, this bit indicates that the receiver has both block lock and alignment for all lanes and is identical to 3.50.12 (see 45.2.3.21.1).

Change the title and text of 45.2.3.14 to include 25GBASE-T and 40GBASE-T as follows:

45.2.3.14 BASE-R and ~~10~~MultiGBASE-T PCS status 2 register (Register 3.33)

The assignment of bits in the BASE-R and ~~10~~MultiGBASE-T PCS status 2 register is shown in Table 45–129. All the bits in the BASE-R and ~~10~~MultiGBASE-T PCS status 2 register are read only; a write to the BASE-R and ~~10~~MultiGBASE-T PCS status 2 register shall have no effect. A PCS device that ~~does not~~ implements neither BASE-R ~~and 10GBASE-T~~ nor any member of the MultiGBASE-T set shall return a zero for all bits in the BASE-R and ~~10~~MultiGBASE-T PCS status 2 register. It is the responsibility of the STA management entity to ensure that a port type is supported by all MMDs before interrogating any of its status bits. The contents of register 3.33 are undefined when the BASE-R or the ~~10GBASE-T~~MultiGBASE-T set PCS is operating seed test-pattern mode, PRBS31 test-pattern mode, or PRBS9 test-pattern mode.

Change the title and descriptions of bits 3.33.15 and 3.33.14 in Table 45–129 as follows (unchanged bits not shown):

Table 45–129—BASE-R and ~~10~~MultiGBASE-T PCS status 2 register bit definitions

Bit(s)	Name	Description	R/W[a]
3.33.15	Latched block lock	1 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS has block lock 0 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS does not have block lock	RO/LL
3.33.14	Latched high BER	1 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS has reported a high BER 0 = BASE-R or ~~10GBASE-T~~ any MultiGBASE-T PCS has not reported a high BER	RO/LH

[a]RO = Read only, LL = Latching low, LH = Latching high, NR = Non Roll-over

45.2.3.14.1 Latched block lock (3.33.15)

Change the text of 45.2.3.14.1 to include 25GBASE-T and 40GBASE-T as follows:

When read as a one, bit 3.33.15 indicates that the 10/40/100GBASE-R or ~~the 10GBASE-T~~ a member of the MultiGBASE-T set PCS has achieved block lock. When read as a zero, bit 3.33.15 indicates that the 10/40/100GBASE-R or ~~the 10GBASE-T~~ a member of the MultiGBASE-T set PCS has lost block lock.

The latched block lock bit shall be implemented with latching low behavior.

This bit is a latching low version of the 10/40/100GBASE-R and ~~10~~MultiGBASE-T PCS block lock status bit (3.32.0).

45.2.3.14.2 Latched high BER (3.33.14)

Change the text of 45.2.3.14.2 to include 25GBASE-T and 40GBASE-T as follows:

When read as a one, bit 3.33.14 indicates that the 10/40/100GBASE-R or ~~the 10GBASE-T~~ a member of the MultiGBASE-T set PCS has detected a high BER. When read as a zero, bit 3.33.14 indicates that the 10/40/100GBASE-R or ~~the 10GBASE-T~~ a member of the MultiGBASE-T set PCS has not detected a high BER.

The latched high BER bit shall be implemented with latching high behavior.

This bit is a latching high version of the 10/40/100GBASE-R and ~~10~~MultiGBASE-T PCS high BER status bit (3.32.1).

45.2.3.14.3 BER (3.33.13:8)

Change the text of 45.2.3.14.3 to include 25GBASE-T and 40GBASE-T as follows:

The BER counter is a six bit count as defined by the ber_count variable in 49.2.14.2 and 82.2.19.2.4 for 10/40/100GBASE-R and defined by ~~the~~counter lfer_count ~~variable~~ in 55.3.6.2 for 10GBASE-T and in 113.3.6.2.2 for 25GBASE-T and 40GBASE-T. These bits shall be reset to all zeros when the BASE-R and ~~10~~MultiGBASE-T PCS status 2 register is read by the management function or upon execution of the PCS reset. If the BER high order counter, 3.44 (see 45.2.3.23) is not implemented then these bits shall be held at all ones in the case of overflow.

45.2.3.14.4 Errored blocks (3.33.7:0)

Change the text of 45.2.3.14.4 to include 25GBASE-T and 40GBASE-T as follows:

The errored blocks counter is an eight bit count defined by the errored_block_count counter specified in 49.2.14.2 for 10GBASE-R, in 82.3.1 for 40/100GBASE-R and defined by ~~the~~counter errored_block_count ~~variable~~ in 55.3.6.2 for 10GBASE-T and in 113.3.6.2 for 25GBASE-T and 40GBASE-T. These bits shall be reset to all zeros when the errored blocks count is read by the management function or upon execution of the PCS reset. If the Errored blocks high order counter, 3.45 (see 45.2.3.20) is not implemented then these bits shall be held at all ones in the case of overflow.

45.2.7 Auto-Negotiation registers

Change the title and names for register addresses 7.32 and 7.33 to reference MultiGBASE-T (unchanged rows not shown), insert rows for registers 7.64 and 7.65 (to end of table, after 7.63) for MultiGBASE-T AN control 2 and MultiGBASE-T status 2 registers, adjust the reserved row, and add a new reserved row in Table 45–200 as follows:

Table 45–200—Auto-Negotiation MMD registers

Register address	Register name	Subclause
7.32	~~10~~MultiGBASE-T AN control 1	45.2.7.10
7.33	~~10~~MultiGBASE-T AN status 1	45.2.7.11
7.62 through 7.6~~332 767~~	Reserved	
7.64	MultiGBASE-T AN control 2	45.2.7.14a
7.65	MultiGBASE-T AN status 2	45.2.7.14b
7.66 through 7.32 767	Reserved	

Change title and the text of 45.2.7.10 to include 25GBASE-T and 40GBASE-T as follows:

45.2.7.10 ~~10~~MultiGBASE-T AN control 1 register (Register 7.32)

The assignment of bits in the ~~10~~MultiGBASE-T AN control 1 register is shown in Table 45–207. The default values for each bit of the ~~10~~MultiGBASE-T AN control 1 register should be chosen so that the initial state of the device upon power up or reset is a normal operational state without management intervention.

Change the title, the reserved row, and the name and description of bits 7.32.0, 7.32.1, 7.32.2 in Table 45–207 and insert rows for bits 7.32.11, 7.32.10, 7.32.9, and 7.32.3 above and below the reserved row, respectively, as follows (unchanged rows not shown):

Table 45–207—~~10~~MultiGBASE-T AN control 1 register

Bit(s)	Name	Description	R/W[a]
7.32.11	40GBASE-T ability	1 = Advertise PHY as 40GBASE-T capable 0 = Do not advertise the PHY as 40GBASE-T capable	R/W
7.32.10	25GBASE-T ability	1 = Advertise PHY as 25GBASE-T capable 0 = Do not advertise the PHY as 25GBASE-T capable	R/W
7.32.9	25GBASE-T Fast retrain ability	1 = Advertise PHY as 25GBASE-T fast retrain capable 0 = Do not advertise PHY as 25GBASE-T fast retrain capable	R/W
7.32.~~11:3~~8:4	Reserved	Value always 0	RO
7.32.3	40GBASE-T Fast retrain ability	1 = Advertise PHY as 40GBASE-T fast retrain capable 0 = Do not advertise PHY as 40GBASE-T fast retrain capable	R/W
7.32.2	10GBASE-T LD PMA training reset request	1 = Local device requests that link partner reset PMA training PRBS every frame 0 = Local device requests that link partner run PMA training PRBS continuously NOTE — the periodic training sequence request functionality is deprecated. Link partners may ignore a value of one in this bit. It is recommended to always set this bit to zero.	R/W

Table 45–207—~~10~~MultiGBASE-T AN control 1 register *(continued)*

Bit(s)	Name	Description	R/W[a]
7.32.1	10GBASE-T Fast retrain ability	1 = Advertise PHY as 10GBASE-T fast retrain capable 0 = Do not advertise PHY as 10GBASE-T fast retrain capable	R/W
7.32.0	10GBASE-T LD loop timing ability	1 = Advertise PHY as capable of 10GBASE-T loop timing 0 = Do not advertise PHY as 10GBASE-T loop timing	R/W

[a]R/W = Read/Write, RO = Read only

Insert 45.2.7.10.4a through 45.2.7.10.4d after 45.2.7.10.4 as follows:

45.2.7.10.4a 40GBASE-T capability (7.32.11)

Bit 7.32.11 is used to select whether or not Auto-Negotiation advertises the ability to operate as a 40GBASE-T PHY. If bit 7.32.11 is set to one the PHY shall advertise 40GBASE-T PHY capability. If bit 7.32.11 is set to zero the PHY shall not advertise 40GBASE-T PHY capability.

45.2.7.10.4b 25GBASE-T capability (7.32.10)

Bit 7.32.10 is used to select whether or not Auto-Negotiation advertises the ability to operate as a 25GBASE-T PHY. If bit 7.32.10 is set to one the PHY shall advertise 25GBASE-T PHY capability. If bit 7.32.10 is set to zero the PHY shall not advertise 25GBASE-T PHY capability.

45.2.7.10.4c 25GBASE-T Fast retrain ability (7.32.9)

Bit 7.32.9 is used to select whether or not the 25GBASE-T PHY advertises the ability to support 25GBASE-T fast retrain. Fast retrain ability is exchanged during link training, see 113.4.2.5.10. If bit 7.32.9 is set to one, the PHY shall advertise fast retrain ability. If bit 7.32.9 is set to zero, the PHY shall not advertise fast retrain ability.

45.2.7.10.4d 40GBASE-T Fast retrain ability (7.32.3)

Bit 7.32.3 is used to select whether or not the 40GBASE-T PHY advertises the ability to support 40GBASE-T fast retrain. Fast retrain ability is exchanged during link training, see 113.4.2.5.10. If bit 7.32.3 is set to one, the PHY shall advertise fast retrain ability. If bit 7.32.3 is set to zero, the PHY shall not advertise fast retrain ability.

Change the title and the text of 45.2.7.10.5 as shown:

45.2.7.10.5 10GBASE-T LD PMA training reset request (7.32.2)

For 10GBASE-T, i~~I~~f bit 7.32.2 is set to one then the local device expects the link partner to reset the PMA training PRBS for every PMA training frame. If bit 7.32.2 is zero then the local device expects link partner to run PMA training PRBS continuously through every PMA training frame.

The periodic training sequence request functionality is deprecated and may be unsupported by some implementations. The link partner may ignore a request caused by setting this bit to one. It is recommended to always set this bit to zero.

Change the title of 45.2.7.10.6 as follows:

45.2.7.10.6 10GBASE-T Fast retrain ability (7.32.1)

Change the title of 45.2.7.10.7 as follows:

45.2.7.10.7 10GBASE-T LD loop timing ability (7.32.0)

Change the title and the text of 45.2.7.11 to include 40GBASE-T as follows:

45.2.7.11 ~~10~~MultiGBASE-T AN status 1 register (Register 7.33)

The assignment of bits in the ~~10~~MultiGBASE-T AN status 1 register is shown in Table 45–208. All the bits in the ~~10~~MultiGBASE-T AN status 1 register are read only; a write shall have no effect.

Change the title, the reserved row, and the names and descriptions for bits 7.33.9 (and NOTE), 7.33.1, and 7.33.0, insert rows for bit 7.33.8 and 7.33.7 before the reserved row, and bit 7.33.2 after reserved row in Table 45–208 as follows (unchanged rows not shown):

Table 45–208—~~10~~MultiGBASE-T AN status 1 register

Bit(s)	Name	Description	R/W[a]
7.33.9	10GBASE-T LP PMA training reset request	1 = Link partner requests that local device reset PMA training PRBS every frame 0 = Link partner requests that local device run PMA training PRBS continuously See NOTE.	RO
7.33.8	Link partner 40GBASE-T capability	1 = Link partner is able to operate as 40GBASE-T 0 = Link partner is not able to operate as 40GBASE-T	RO
7.33.7	Link partner 25GBASE-T capability	1 = Link partner is able to operate as 25GBASE-T 0 = Link partner is not able to operate as 25GBASE-T	RO
7.33.~~8~~:26:3	Reserved	Value always 0	RO
7.33.2	25GBASE-T Fast retrain ability	1 = Link partner is capable of 25GBASE-T fast retrain 0 = Link partner is not capable of 25GBASE-T fast retrain	RO
7.33.1	10GBASE-T Fast retrain ability	1 = Link partner is capable of 10GBASE-T fast retrain 0 = Link partner is not capable of 10GBASE-T fast retrain	RO
7.33.0	~~Reserved~~40GBASE-T Fast retrain ability	~~Value always 0~~1 = Link partner is capable of 40GBASE-T fast retrain 0 = Link partner is not capable of 40GBASE-T fast retrain	RO
NOTE — The periodic training sequence request functionality is deprecated. Implementations may ignore a value of one in this bit or have it always read as zero.			

[a]RO = Read only, SC = Self-clearing, LH = Latching high

45.2.7.11.1 MASTER-SLAVE configuration fault (7.33.15)

Change the text of 45.2.7.11.1 to include 25GBASE-T and 40GBASE-T as follows:

MASTER-SLAVE configuration fault bit 7.33.15 shall be set in the event that determination of the MASTER-SLAVE cannot be successfully concluded. MASTER-SLAVE configuration fault, as well as the

criteria and method of fault detection, is PHY specific. Additional information regarding the resolution and selection of MASTER-SLAVE and bit 7.33.15 for 10GBASE-T in contained in 55.6.2. Additional information regarding the resolution and selection of MASTER-SLAVE and bit 7.33.15 for 25GBASE-T and 40GBASE-T is contained in 113.6.2. The MASTER-SLAVE configuration fault bit 7.33.15 shall be cleared each time the ~~10~~MultiGBASE-T status register 7.33 is read via the management interface and shall be cleared by a ~~10GBASE-T~~ PMA reset of a PHY in the MultiGBASE-T set. Bit 7.33.15 shall self clear upon Auto-Negotiation enable.

45.2.7.11.2 MASTER-SLAVE configuration resolution (7.33.14)

Change the text of 45.2.7.11.2 to include 25GBASE-T and 40GBASE-T as follows:

Bit 7.33.14 is determined by the 10GBASE-T MASTER-SLAVE configuration resolution function described in 55.6.2 or the 25GBASE-T/40GBASE-T MASTER-SLAVE configuration resolution function described in 113.6.2. If the MASTER-SLAVE configuration resolution bit 7.33.14 is set to one and the Auto-Negotiation complete bit 7.1.5 is set and ~~if~~ MASTER-SLAVE configuration fault bit 7.33.15 in the ~~10~~MultiGBASE-T status register is zero, then MASTER mode of operation has been selected. If the MASTER-SLAVE configuration resolution bit 7.33.14 is set to zero and the Auto-Negotiation complete bit 7.1.5 is set and ~~if~~ MASTER-SLAVE configuration fault bit 7.33.15 in the ~~10~~MultiGBASE-T status register is zero, then SLAVE mode of operation has been selected. In all other cases, neither SLAVE mode nor MASTER mode has been selected.

Change the title and the text of 45.2.7.11.7 as shown:

45.2.7.11.7 10GBASE-T Link partner PMA training reset request (7.33.9)

If bit 7.33.9 is set to one then the link partner is expecting the local device to reset the PMA training PRBS for every PMA training frame. If bit 7.33.9 is zero then the link partner expects the local device to run PMA training PRBS continuously through every PMA training frame. The periodic training sequence request functionality is deprecated. Implementations may ignore a value of one in this bit or have it always read as zero.

Insert 45.2.7.11.7a, 45.2.7.11.7b, and 45.2.7.11.7c after 45.2.7.11.7 as follows:

45.2.7.11.7a Link partner 40GBASE-T capability (7.33.8)

Bit 7.33.8 is only valid when page received bit 7.1.6 is set to one. When read as a one, bit 7.33.8 indicates that the link partner is able to operate as 40GBASE-T. When read as a zero, bit 7.33.8 indicates that the link partner is not able to operate as 40GBASE-T.

45.2.7.11.7b Link partner 25GBASE-T capability (7.33.7)

Bit 7.33.7 is only valid when page received bit 7.1.6 is set to one. When read as a one, bit 7.33.7 indicates that the link partner is able to operate as 25GBASE-T. When read as a zero, bit 7.33.7 indicates that the link partner is not able to operate as 25GBASE-T.

45.2.7.11.7c 25GBASE-T Fast retrain ability (7.33.2)

When read as a one, bit 7.33.2 indicates that the link partner has the ability to support the 25GBASE-T fast retrain capability as specified in 113.4.2.5.16. When read as a zero, bit 7.33.2 indicates that the PHY lacks the ability to support the 25GBASE-T fast retrain capability. This bit is valid only after link is established.

Change the title and the text of 45.2.7.11.8 as follows:

45.2.7.11.8 10GBASE-T Fast retrain ability (7.33.1)

When read as a one, bit 7.33.1 ~~is used to~~ indicates that the link partner has the ability to support the 10GBASE-T fast retrain capability as specified in 55.4.2.5.15. When read as a zero, bit 7.33.1 indicates that the PHY lacks the ability to support the 10GBASE-T fast retrain capability.

Insert 45.2.7.11.9 after 45.2.7.11.8 as follows:

45.2.7.11.9 40GBASE-T Fast retrain ability (7.33.0)

When read as a one, bit 7.33.0 indicates that the link partner has the ability to support the 40GBASE-T fast retrain capability as specified in 113.4.2.5.16. When read as a zero, bit 7.33.0 indicates that the PHY lacks the ability to support the 40GBASE-T fast retrain capability. This bit is valid only after link is established.

45.2.7.13 EEE advertisement (Register 7.60)

Change the text of the first paragraph of 45.2.7.13 to include 25GBASE-T and 40GBASE-T as follows (unchanged paragraphs not shown):

This register defines the EEE advertisement for several device types ~~that is sent~~ Devices that use Clause 28 Auto-Negotiation send EEE advertisement in the Unformatted Next Page following a EEE technology message code as defined in 28C.12 or ~~sent in the unformatted code field of Message Next Page with EEE technology message code as defined in 73A.4 or sent~~ as part of the 10GBASE-T and 1000BASE-T technology message code as defined in 28C.11. Devices that use Clause 73 Auto-Negotiation send EEE advertisement in the unformatted code field of Message Next Page with EEE technology message code as defined in 73A.4. 25GBASE-T and 40GBASE-T EEE advertisement is exchanged in the Infofield during training as defined in 113.4.2.5.10. The assignment of bits in the EEE advertisement register and the correspondence with the bits in the Next Page messages or in the training Infofield are shown in Table 45–210.

Change the name, description, clause reference, and R/W status for bit 7.60.9 and bit 7.60.0 in Table 45–210 as follows: (unchanged rows not shown)

Table 45–210—EEE advertisement register (Register 7.60) bit definitions

Bit(s)	Name	Description	Clause reference; Next Page bit number	R/W[a]
7.60.9	~~Reserved~~40GBASE-T EEE	~~Value always 0~~1 = Advertise that the 40GBASE-T has EEE capability 0 = Do not advertise that the 40GBASE-T has EEE capability	113.4.2.5.10; Infofield Octet 12 bit 7	~~RO~~ R/W
7.60.0	~~Reserved~~25GBASE-T EEE	~~Value always 0~~1 = Advertise that the 25GBASE-T has EEE capability 0 = Do not advertise that the 25GBASE-T has EEE capability	113.4.2.5.10; Infofield Octet 12 bit 7	~~RO~~ R/W

[a]R/W = Read/Write, RO = Read only

Insert 45.2.7.13.4a after 45.2.7.13.4 as follows:

45.2.7.13.4a 40GBASE-T EEE supported (7.60.9)

If the device supports EEE operation for 40GBASE-T as defined in 113.6.1, and EEE operation is desired, this bit shall be set to one.

Insert 45.2.7.13.12a after 45.2.7.13.12 as follows:

45.2.7.13.12a 25GBASE-T EEE supported (7.60.0)

If the device supports EEE operation for 25GBASE-T as defined in 113.6.1, and EEE operation is desired, this bit shall be set to one.

45.2.7.14 EEE link partner ability (Register 7.61)

Change the text of 45.2.7.14 to include 25GBASE-T and 40GBASE-T as follows:

All of the bits in the EEE LP ability register are read-only. A write to the EEE LP ability register shall have no effect. Except for 10GBASE-T, members of the MultiGBASE-T PHY set exchange the EEE ability in the Infofield during link training. For these PHYs, the EEE LP ability register is updated after link is established. For all other PHYs, w~~W~~hen the AN process has been completed, this register shall reflect the contents of the link partner's EEE advertisement register. The assignment of bits in the EEE link partner ability register and the correspondence with the bits in the Next Page messages are shown in Table 45–211.

Change the name, description, clause reference, and Next Page bit number for bits 7.61.9 and 7.61.0 in Table 45–211 as follows (unchanged rows not shown):

Table 45–211—EEE link partner ability (Register 7.61) bit definitions

Bit(s)	Name	Description	Clause reference; Next Page bit number	R/W[a]
7.61.9	~~Reserved~~40GBASE-T EEE	~~Value always 0~~1 = Link partner is advertising EEE capability for 40GBASE-T 0 = Link partner is not advertising EEE capability for 40GBASE-T	113.4.2.5.10; Infofield Octet 12 bit 7	RO
7.61.0	~~Reserved~~25GBASE-T EEE	~~Value always 0~~1 = Link partner is advertising EEE capability for 25GBASE-T 0 = Link partner is not advertising EEE capability for 25GBASE-T	113.4.2.5.10; Infofield Octet 12 bit 7	RO

[a]~~R~~RO = Read only

After 45.2.7.14, insert three new subclauses and Table 45-211a for MultiGBASE-T AN control 2 register and bits as shown:

45.2.7.14a MultiGBASE-T AN control 2 (Register 7.64)

Register 7.64 is a continuation of register 7.32. The assignment of bits in the MultiGBASE-T AN control 2 register is shown in Table 45–211a. The default values of each bit of the MultiGBASE-T AN control

2 register should be chosen so that the initial state of the devices upon power up or reset is a normal operational state without management intervention.

Table 45–211a—MultiGBASE-T AN control 2 (Register 7.64) bit definitions

Bit(s)	Name	Description	R/W[a]
7.64.15:2	Reserved	Value always 0	RO
7.64.1	25GBASE-T THP Bypass Request	1 = Local device requests link partner to initially reset THP during fast retrain 0 = Local device requests link partner not to initially reset THP during fast retrain	R/W
7.64.0	40GBASE-T THP Bypass Request	1 = Local device requests link partner to initially reset THP during fast retrain 0 = Local device requests link partner not to initially reset THP during fast retrain	R/W

[a]R/W = Read/Write, RO = Read only

45.2.7.14a.1 25GBASE-T THP Bypass Request

Bit 7.65 is valid only if 7.32.9 is set to one advertising fast retrain ability, and is used to request the link partner whether to initially reset the THP during fast retrain. THP Bypass Request is exchanged during link training, see 113.4.2.5.10. If bit 7.64.1 is set to zero the local device requests link partner not to reset THP during fast retrain. If bit 7.64.1 is set to one the local device requests link partner to initially reset THP during fast retrain.

45.2.7.14a.2 40GBASE-T THP Bypass Request

Bit 7.64.0 is valid only if 7.32.3 is set to one advertising fast retrain ability, and is used to request the link partner whether to initially reset the THP during fast retrain. THP Bypass Request is exchanged during link training, see 113.4.2.5.10. If bit 7.64.0 is set to zero the local device requests link partner not to reset THP during fast retrain. If bit 7.64.0 is set to one the local device requests link partner to initially reset THP during fast retrain.

Insert three new subclauses and Table 45-211b for MultiGBASE-T AN status 2 register and bits, as follows:

45.2.7.14b MultiGBASE-T AN status 2 (Register 7.65)

Register 7.65 is a continuation of register 7.33. The assignment of bits in the MultiGBASE-T AN status 2 register is shown in Table 45–211b. All the bits in the MultiGBASE-T AN status 2 register are read only; a write shall have no effect.

Table 45–211b—MultiGBASE-T AN status 2 (Register 7.65) bit definitions

Bit(s)	Name	Description	R/W[a]
7.65.15:2	Reserved	Value always 0	RO
7.65.1	25GBASE-T Link Partner THP Bypass Request	1 = Link partner requests local device to initially reset THP during fast retrain 0 = Link partner requests local device not to initially reset THP during fast retrain	RO
7.65.0	40GBASE-T Link Partner THP Bypass Request	1 = Link partner requests local device to initially reset THP during fast retrain 0 = Link partner requests local device not to initially reset THP during fast retrain	RO

[a]RO = Read only

45.2.7.14b.1 25GBASE-T Link Partner THP Bypass Request

Bit 7.65.1 is valid only if 7.33.2 is set to one indicating that the link partner has fast retrain ability. THP Bypass Request is exchanged during link training, see 113.4.2.5.10. Bit 7.65.1 is updated after link is established. When read as a zero, the link partner requests local device not to reset THP during fast retrain. When read as a one, the link partner requests local device to initially reset THP during fast retrain.

45.2.7.14b.2 40GBASE-T Link Partner THP Bypass Request

Bit 7.65.0 is valid only if 7.33.0 is set to one indicating that the link partner has fast retrain ability. THP Bypass Request is exchanged during link training, see 113.4.2.5.10. Bit 7.65.0 is updated after link is established. When read as a zero, the link partner requests local device not to reset THP during fast retrain. When read as a one, the link partner requests local device to initially reset THP during fast retrain.

45.5 Protocol implementation conformance statement (PICS) proforma for Clause 45, Management Data Input/Output (MDIO) interface[5]

45.5.3 PICS proforma tables for the Management Data Input Output (MDIO) interface

45.5.3.2 PMA/PMD MMD options

Insert rows for items *25T and *40T after the row for *10T in 45.5.3.2 as follows (unchanged rows not shown):

Item	Feature	Subclause	Value/Comment	Status	Support
*25T	Implementation of the 25GBASE-T PMA	45.2.1.14b		PMA:O	Yes [ ] No [ ]
*40T	Implementation of the 40GBASE-T PMA	45.2.1.12		PMA:O	Yes [ ] No [ ]

45.5.3.3 PMA/PMD Management functions

Change rows for items MM111 and MM112 to add *25T and *40T in 45.5.3.3 as follows (unchanged rows not shown)

Item	Feature	Subclause	Value/Comment	Status	Support
MM111	Bit set to zero if PMA link_status=FAIL	45.2.1.62.1		PMA*10T:M PMA*25T:M PMA*40T:M	Yes [ ] N/A [ ]
MM112	Skew delay register update rate.	45.2.1.78	At least once per second	PMA*10T:M PMA*25T:M PMA*40T:M	Yes [ ] N/A [ ]

45.5.3.6 PCS options

Change the feature description for item *CT in 45.5.3.6 as follows (unchanged rows not shown):

Item	Feature	Subclause	Value/Comment	Status	Support
*CT	Implementation of the 10GBASE-T PCS or the PCS of any other member of the MultiGBASE-T set.	45.2.3		PCS:O	Yes [ ] No [ ] N/A [ ]

[5]*Copyright release for PICS proformas:* Users of this standard may freely reproduce the PICS proforma in this subclause so that it can be used for its intended purpose and may further publish the completed PICS.

45.5.3.7 PCS management functions

Change the feature descriptions for items RM37, RM38, RM39, and RM40 in 45.5.3.7 as follows (unchanged rows not shown):

Item	Feature	Subclause	Value/Comment	Status	Support
RM37	Writes to BASE-R and ~~10~~MultiGBASE-T PCS status 1 register have no effect	45.2.3.13		CR:M CT:M	Yes [ ] N/A [ ]
RM38	Reads from BASE-R and ~~10~~MultiGBASE-T status 1 register return zero for PCS that does not support 10GBASE-R	45.2.3.13		CR:M CT:M	Yes [ ] N/A [ ]
RM39	Writes to BASE-R and ~~10~~MultiGBASE-T PCS status 2 register have no effect	45.2.3.14		XCR:M CT:M	Yes [ ] N/A [ ]
RM40	Reads from BASE-R and ~~10~~MultiGBASE-T PCS status 2 register return zero for PCS that does not support 10/40/100GBASE-R or ~~10~~any MultiGBASE-T	45.2.3.14		XCR:M CT:M	Yes [ ] N/A [ ]

45.5.3.8 Auto-Negotiation options

Change the feature description for item *AT in 45.5.3.8 as follows (unchanged rows not shown):

Item	Feature	Subclause	Value/Comment	Status	Support
*AT	Implementation of ~~10~~MultiGBASE-T Auto-Negotiation	45.2.7		AN:O	Yes [ ] No [ ] N/A [ ]

45.5.3.9 Auto-Negotiation management functions

Change the feature description for item AM51 and insert rows for items AM61, AM62, AM63, and AM64 below AM60 in 45.5.3.9 as follows (unchanged rows not shown):

Item	Feature	Subclause	Value/Comment	Status	Support
AM51	Bit set to zero by ~~10~~MultiG-BASE-T PMA reset	45.2.7.11.1		AN:M	Yes [ ] N/A [ ]
AM61	Advertise 40GBASE-T PHY capability when bit is set to one	45.2.7.10.4a		AN:M	Yes [ ] N/A [ ]
AM62	40GBASE-T PHY capability not advertised when bit is set to zero	45.2.7.10.4a		AN:M	Yes [ ] N/A [ ]
AM63	Advertise 25GBASE-T PHY capability when bit is set to one	45.2.7.10.4b		AN:M	Yes [ ] N/A [ ]
AM64	25GBASE-T PHY capability not advertised when bit is set to zero	45.2.7.10.4b		AN:M	Yes [ ] N/A [ ]

55. Physical Coding Sublayer (PCS), Physical Medium Attachment (PMA) sublayer and baseband medium, type 10GBASE-T

Change the second paragraph of 55.3.4 as shown:

55.3.4 PMA training side-stream scrambler polynomials

~~Moreover d~~NOTE—During Auto-Negotiation a device may request its link partner to use periodic training sequence initialization. This function is deprecated; devices may ignore this request if it is received, and it is recommended not to send it. A device that receives this request and does not ignore it generates a periodically repeating pattern by reinitializing its scrambler state after every 16384 symbol periods to the 33-bit value ~~each transceiver may request the remote transceiver to reinitialize the values of its scrambler state after every 16384 symbol periods, to generate a periodically repeating pattern with repetition period 16384. The initial 33-bit values of the scrambler state shall be~~ generated by combining 0x39A422 for the 22 MSBs and random value SB10-SB0 from Table 55–15 generated by the local device for the 11 LSBs as shown in Figure 55–13.

55.6 Management interfaces

55.6.1.2 10GBASE-T Auto-Negotiation page use

Change bits U20, U19, and U17 in Table 55-15 as shown [unchanged bits (except U18) not shown]:

Table 55–15—10GBASE-T Base and Next Pages bit assignments

Bit	Name	Description
Extended Next Page (Unformatted Message Code Field)		
U20	10GBASE-T LD PMA training reset request (1 = Local Device requests that Link Partner reset PMA training PRBS every frame 0 = Local Device requests that Link Partner run PMA training PRBS continuously).	Defined in 45.2.7.10.5
U19	10GBASE-T Fast retrain ability (1 = Advertise PHY as supporting fast retrain, 0 = Advertise PHY as not supporting fast retrain)	Defined in 45.2.7.10.6
U18	PHY short reach mode (1 = PHY of Local Device is operating in short reach mode 0 = PHY of Local Device is operating in normal mode)	Defined in 45.2.1.64.2
U17	10GBASE-T LD loop timing ability (1 = Advertise PHY as capable of loop timing and 0 = do not advertise PHY as capable of loop timing)	Defined in 45.2.7.10.7

Change the three technology message code bits in 55.6.2 to accommodate 40GBASE-T as follows:

55.6.2 MASTER-SLAVE configuration resolution

where

U11 is bit 11 of ~~10~~MultiGBASE-T~~/~~ and 1000BASE-T Technology message code,
U12 is bit 12 of ~~10~~MultiGBASE-T~~/~~ and 1000BASE-T Technology message code,
U13 is bit 13 of ~~10~~MultiGBASE-T~~/~~ and 1000BASE-T Technology message code (see Table 55-15).

Delete row PCT19 in 55.12.3 as follows:

55.12.3 Physical Coding Sublayer (PCS)

Item	Feature	Subclause	Status	Support	Value/Comment
~~PCT19~~	~~PMA training scrambler reset~~	~~55.3.4~~	~~M~~	~~Yes [ ]~~	~~If requested by Link Partner during Auto Negotiation~~

78. Energy-Efficient Ethernet (EEE)

78.1 Overview

78.1.3 Reconciliation sublayer operation

Change text in 78.1.3.3.1 (as modified IEEE Std 802.3by-2016) as follows:

78.1.3.3.1 PHY LPI transmit operation

Except for BASE-T, ~~F~~for PHYs with an operating speed of 25 Gb/s or greater that implement the optional EEE capability, two modes of LPI operation may be supported: deep sleep and fast wake. *Deep sleep* refers to the mode for which the transmitter ceases transmission during Low Power Idle (as shown in Figure 78-3) and is equivalent to the only mechanism defined for PHYs with an operating speed of 10 Gb/s or below. Deep sleep support is optional for PHYs with an operating speed of 25 Gb/s or greater that implement EEE with the exception of the PHYs noted in Table 78–1 that do not support deep sleep. *Fast wake* refers to the mode for which the transmitter continues to transmit signals during Low Power Idle so that the receiver can resume operation with a shorter wake time (as shown in Figure 78-4). For transmit, other than the PCS encoding LPI, there is no difference between fast wake and normal operation. Except for BASE-T PHYs, ~~F~~fast wake support is mandatory for PHYs with an operating speed of 25 Gb/s or greater that implement EEE.

78.1.4 PHY types optionally supporting EEE

Insert new rows into Table 78-1 (as modified by IEEE Std 802.3by-2016) after the entry 25GBASE-SR for 25GBASE-T and after the entry 40GBASE-ER4 for 40GBASE-T as follows:

Table 78–1—Clauses associated with each PHY or interface type

PHY or interface type	Clause
25GBASE-T	113
40GBASE-T	113

78.2 LPI mode timing parameters description

Insert new rows into Table 78-2 (as modified by IEEE Std 802.3by-2016) after the entry 25BASE-KR-S/25GBASE-CR-S for 25GBASE-T and after the entry 40GBASE-CR4 for 40GBASE-T as follows:

Table 78–2—Summary of the key EEE parameters for supported PHYs or interfaces

PHY or interface type	T_s (us)		Tq (us)		Tr (us)	
	Min	Max	Min	Max	Min	Max
25GBASE-T	0.768	0.896	15.616	15.616	0.768	0.768
40GBASE-T	0.48	0.56	9.76	9.76	0.48	0.48

78.3 Capabilities Negotiation

Change text in first and second paragraphs of 78.3 as follows:

The EEE capability shall be advertised during the Auto-Negotiation stage, except for PHYs that only support fast wake operation or PHYs that exchange EEE capability during link training. Auto-Negotiation provides a linked device with the capability to detect the abilities (modes of operation) supported by the device at the other end of the link, determine common abilities, and configure for joint operation. Auto-Negotiation is performed at power up, on command from management, due to link failure, or due to user intervention. Fast wake capability shall be advertised using L2 protocol frames as described in 78.4. The EEE capability for 25GBASE-T and 40GBASE-T shall be advertised during link training according to 113.4.2.5.10.

During Auto-Negotiation, both link partners indicate their EEE capabilities. EEE is supported only if during Auto-Negotiation both the local device and link partner advertise the EEE capability for the resolved PHY type. If EEE is not supported, all EEE functionality is disabled and the LPI client does not assert LPI. EEE deep sleep operation shall not be enabled unless both the local device and link partner advertise deep sleep capability during Auto-Negotiation for the resolved PHY type. If EEE is supported by both link partners for the negotiated PHY type, then the EEE function can be used independently in either direction. The same applies to 25GBASE-T and 40GBASE-T except the EEE capabilities are exchanged and resolved during link training instead of during Auto-Negotiation.

78.5 Communication link access latency

Change text in 78.5 as follows:

Case-1 of the ~~10GBASE-T~~ PHY in the MultiGBASE-T set applies when the PHY is requested to transmit the Wake signal before transmission of the Sleep signal to the Link Partner is complete. Case-2 of the ~~10GBASE-T~~ PHY in the MultiGBASE-T set applies when the PHY is requested to transmit the Wake signal after transmission of the Sleep signal to the Link Partner is complete and if the PHY has not indicated LOCAL FAULT at any time during the previous 10 ms.

Insert row into Table 78-4 for 25GBASE-T immediately prior to first 40GBASE PHY type (40GBASE-R, after other 25GBASE parameters inserted by IEEE Std 802.3by-2016) and 40GBASE-T LPI parameters following last 40G PHY type (currently 40GBASE-KR4) as follows:

Table 78–4—Summary of the LPI timing parameters for supported PHYs or interfaces

PHY type	Case	$T_{w\_sys\_tx}$ (min) (us)	T_{w_phy} (min) (us)	$T_{phy\_shrink\_tx}$ (max) (us)	$T_{phy\_shrink\_rx}$ (max) (us)	$T_{w\_sys\_rx}$ (min) (us)
25GBASE-T	Case-1	2.56	2.56	1.792	0	0.768
	Case-2	1.792	1.792	0.64	0	0.768
40GBASE-T	Case-1	1.6	1.6	1.12	0	0.48
	Case-2	1.12	1.12	0.4	0	0.48

80. Introduction to 40 Gb/s and 100 Gb/s networks

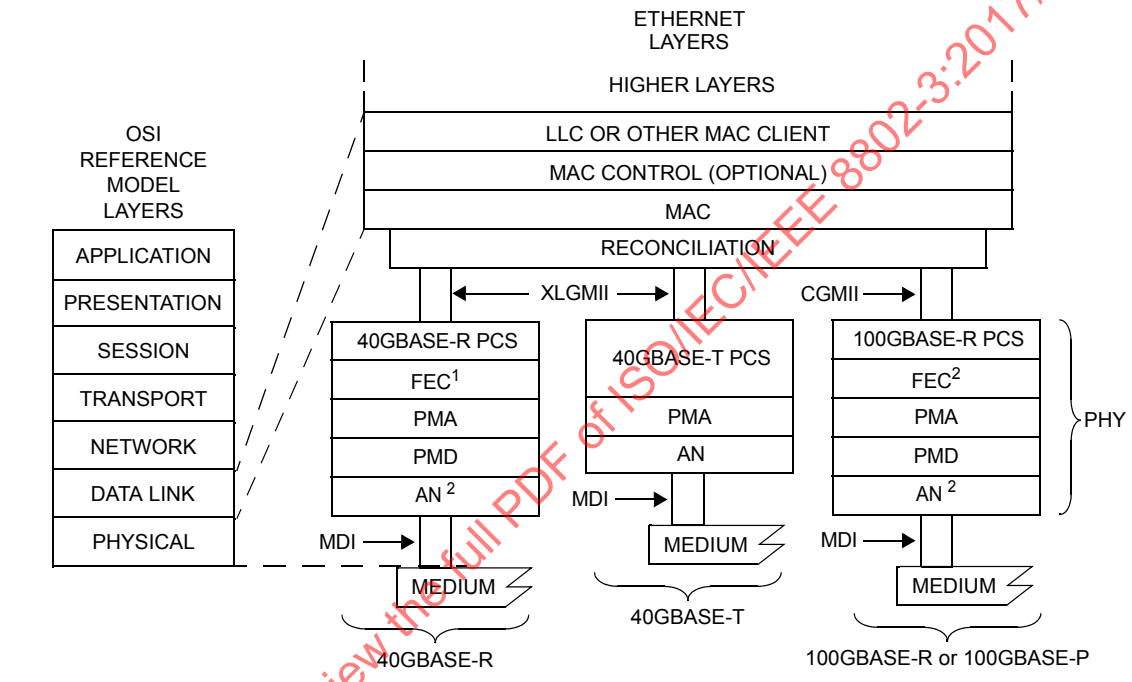
80.1 Overview

80.1.3 Relationship of 40 Gigabit and 100 Gigabit Ethernet to the ISO OSI reference model

Insert new item k) in the second paragraph list of 80.1.3 as follows:

k) The MDI as specified in Clause 113 for 40GBASE-T uses a 4 lane data path.

Replace Figure 80–1 as follows:



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

80.1.4 Nomenclature

Insert the following paragraph after the paragraph on 40GBASE-R and before the paragraph on 100GBASE-R:

40GBASE-T represents Physical Layer devices using Clause 113 Physical Coding Sublayer (PCS), Physical Medium Attachment (PMA) sublayer, and baseband medium, for data communication at 40 Gb/s over a point-to-point 4-pair balanced twisted-pair medium. 40GBASE-T uses a combination of Reed-Solomon-FEC (RS-FEC) and low density parity check (LDPC) FECs in its physical coding sublayer that is mapped to a 128 double-square (DSQ128) constellation for transmission on 4-pair, twisted-pair copper cabling.

Insert the following row between 40GBASE-ER4 and 100GBASE-KR4 in Table 80–1:

Table 80–1—40 Gb/s and 100 Gb/s PHYs

Name	Description
40GBASE-T	40 Gb/s PHY using RS-FEC and LDPC encoding over balanced twisted-pair structured cabling systems (see Clause 113)

80.1.5 Physical Layer signaling systems

Replace Table 80–2 with the following new table as follows:

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	XLAUI	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-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		
40GBASE-T	M			O	M	O												M

[a]O = Optional, M = Mandatory.

80.4 Delay Constraints

Insert row for 40GBASE-T PHY in Table 80-5 Sublayer delay constraints below row for 40GBASE-R PMA as shown (unchanged rows not shown):

Table 80–5—Sublayer delay constraints

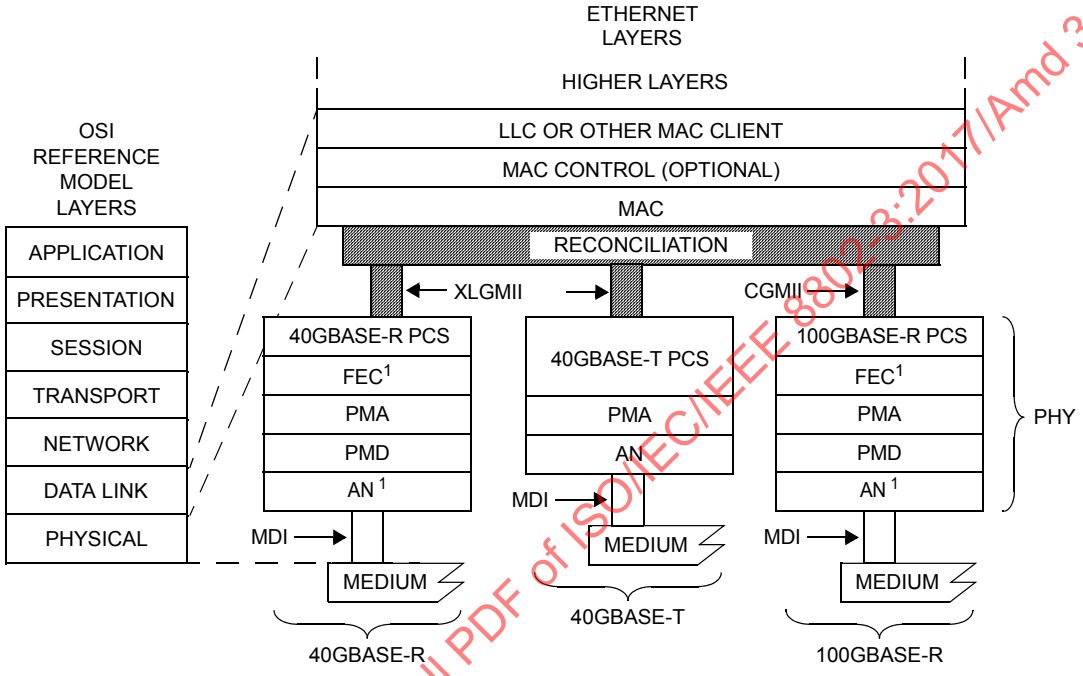
Sublayer	Maximum (bit time)[a]	Maximum (pause_quanta)[b]	Maximum (ns)	Notes[c]
40GBASE-T PHY	25 600	50	640	See 113.11

[a] For 40GBASE-R, 1 bit time (BT) is equal to 25 ps and for 100GBASE-R, 1 bit time (BT) is equal to 10 ps. (See 1.4.117 for the definition of bit time.)
[b] For 40GBASE-R, 1 pause_quantum is equal to 12.8 ns and for 100GBASE-R, 1 pause_quantum is equal to 5.12 ns. (See 31B.2 for the definition of pause_quanta.)
[c] Should there be a discrepancy between this table and the delay requirements of the relevant sublayer clause, the sublayer clause prevails.

81. Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation (XLGMII and CGMII)

81.1 Overview

Replace Figure 81–1 as follows:



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

81.1.7.3 Mapping of PLS_CARRIER.indication

Change 81.1.7.3 as follows:

40 Gb/s and 100 Gb/s operation supports full duplex operation only. The RS only generates this primitive when optional EEE capability or the optional detection of Link Interruption is supported ~~never generates this primitive for PHYs that do not support EEE~~.

~~For PHYs that support EEE capability, CARRIER_STATUS is set in response to LPI_REQUEST as shown in Figure 81-13.~~ CARRIER_STATUS is set to CARRIER_ON if the optional EEE capability is supported and LPI_CARRIER_STATUS is TRUE, or if optional detection of Link Interruption is supported and

link_fault is Link Interruption (see 81.3.4.1). CARRIER_STATUS is otherwise set to CARRIER_OFF. The deferral mechanism based upon the Link Interruption signal may be enabled or disabled by management.

81.3.4 Link fault signaling

Insert a new paragraph after the fourth paragraph in 81.3.4 as follows:

For operation with links that may be temporarily interrupted, optional detection of a third fault condition, Link Interruption, is provided. Link Interruption is indicated by the PHY receive function by continuously sending the Link Interruption ordered set as defined in Table 81–5.

Insert a new row for Link Interruption Description after the Remote Fault Description in Table 81–5 as follows:.

Table 81–5—Sequence ordered sets

Lane 0	Lane 1	Lane 2	Lane 3	Lane 4	Lane 5	Lane 6	Lane 7	Description
Sequence	0x00	0x00	0x00	0x00	0x00	0x00	0x00	Reserved
Sequence	0x00	0x00	0x01	0x00	0x00	0x00	0x00	Local Fault
Sequence	0x00	0x00	0x02	0x00	0x00	0x00	0x00	Remote Fault
Sequence	0x00	0x00	0x03	0x00	0x00	0x00	0x00	Link Interruption

81.3.4.1 Variables and counters

Change definitions of variables: fault_sequence, last_seq_type, link_fault, and seq_type in 81.3.4.1 as follows:

The Link Fault Signaling state diagram uses the following variables and counters:

fault_sequence
: A new column received on RXC<7:0> and RXD<63:0> comprising a Sequence ordered set of eight bytes and consisting of a Sequence control character in lane 0 and a seq_type in lanes 1, 2, 3, 4, 5, 6, and 7 indicating either Local Fault, ~~or~~ Remote Fault, or Link Interruption.

last_seq_type
: The seq_type of the previous Sequence ordered set received
Values:Local Fault; 0x00 in lane 1, 0x00 in lane 2, 0x01 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.
Remote Fault; 0x00 in lane 1, 0x00 in lane 2, 0x02 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.
Link Interruption; 0x00 in lane 1, 0x00 in lane 2, 0x03 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.

link_fault
: An indicator of the fault status.
Values:OK; No fault.
Local Fault; fault detected by the PHY.
Remote Fault; fault detection signaled by the remote RS.
Link Interruption; link temporarily unavailable, signaled by the PHY.

seq_type
: The value received in the current Sequence ordered set

Values:Local Fault; 0x00 in lane 1, 0x00 in lane 2, 0x01 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.
Remote Fault; 0x00 in lane 1, 0x00 in lane 2, 0x02 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.
Link Interruption; 0x00 in lane 1, 0x00 in lane 2, 0x03 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.

81.3.4.2 State diagram

Change second and third paragraphs in 81.3.4.2 to include Link Interruption under conditions for variable link_fault as follows:

The variable link_fault is set to OK following any interval of 128 columns not containing a Remote Fault, ~~or~~ Local Fault, or Link Interruption Sequence ordered set.

The RS output onto TXC<7:0> and TXD<63:0> is controlled by the variable link_fault.

a) link_fault = OK
The RS shall send MAC frames as requested through the PLS service interface. In the absence of MAC frames, the RS shall generate Idle control characters.
b) link_fault = Local Fault
The RS shall continuously generate Remote Fault Sequence ordered sets.
c) link_fault = Remote Fault or link_fault = Link Interruption
The RS shall continuously generate Idle control characters.

81.5 Protocol implementation conformance statement (PICS) proforma for Clause 81, Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation[6]

Insert a new subclause 81.5.3.7 for Link Interruption after 81.5.3.6 as shown:

81.5.3.7 Link Interruption

Item	Feature	Subclause	Value/Comment	Status	Support
LINT1	Detection of Link Interruption	81.3.4	As defined in Table 81–5	O	Yes [ ] No [ ]
LINT2	CARRIER_STATUS response to Link Interruption	81.4.2	Set to CARRIER_ON if link_fault is Link Interruption	LINT1: M	Yes [ ] No [ ]

[6]*Copyright release for PICS proformas:* Users of this standard may freely reproduce the PICS proforma in this subclause so that it can be used for its intended purpose and may further publish the completed PICS.

105. Introduction to 25 Gb/s networks

105.1 Overview

105.1.1 Scope

Change text in first paragraph of 105.1.1 to include 25GBASE-T as follows:

25 Gigabit Ethernet uses the IEEE 802.3 MAC sublayer, connected through a 25 Gigabit Media Independent Interface (25GMII) to Physical Layer entities such as 25GBASE-CR, 25GBASE-CR-S, 25GBASE-KR, 25GBASE-KR-S, ~~and~~ 25GBASE-SR, and 25GBASE-T.

105.1.2 Relationship of 25 Gigabit Ethernet to the ISO OSI reference model

Change Figure 105-1 to include 25GBASE-T as follows:.

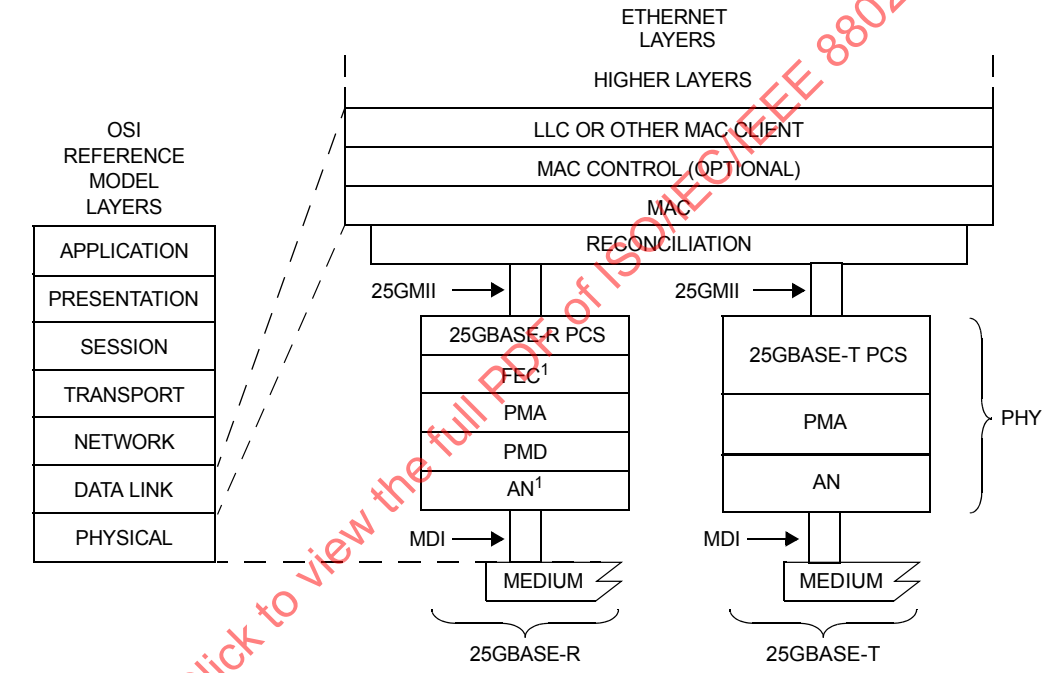


Figure 105–1—Architectural positioning of 25 Gigabit Ethernet

Insert new item d) to add 25GBASE-T 4 lane MDI at the end of the list of exceptions in 105.1.2 (unchanged text not shown) as follows:

d) The MDI as specified in Clause 113 for 25GBASE-T uses a 4 lane data path.

105.1.3 Nomenclature

Insert a new third paragraph as follows:

25GBASE-T represents Physical Layer devices using Clause 113 Physical Coding Sublayer (PCS), Physical Medium Attachment (PMA) sublayer, and baseband medium, for data communication at 25 Gb/s over a point-to-point 4-pair balanced twisted-pair medium. 25GBASE-T uses a combination of Reed-Solomon-FEC (RS-FEC) and low density parity check (LDPC) FECs in its physical coding sublayer that is mapped to a 128 double-square (DSQ128) constellation for transmission on 4-pair, twisted-pair copper cabling.

Insert row in Table 105-1 to add 25GBASE-T after row for 25GBASE-SR (unchanged rows not shown) as follows:.

Table 105–1—25 Gb/s PHYs

Name	Description
25GBASE-T	25 Gb/s PHY using RS-FEC and LDPC encoding over balanced twisted-pair structured cabling systems (see Clause 113).

105.2 Physical Layer signaling systems

Revise title of Table 105-2 and insert row for 25GBASE-T after 25GBASE-SR, column for Clause 28 Auto-Negotiation to the left of Clause 73, and Clause 113 to the right of Clause 112 (all existing entries are blank for the new columns, unchanged rows not shown) as follows:

Table 105–2—Nomenclature and clause correlation, ~~25GBASE-R~~25 Gb/s Ethernet PHYs

Nomenclature	Clause/Annex																	
	28	73	74	78	106		107	108	109	109A	109B	110		111		112	113
	Auto-Negotiation	Auto-Negotiation	BASE-R FEC	EEE	RS	25GMII	25GBASE-R PCS	RS-FEC	PMA	25GAUI C2C	25GAUI C2M	25GBASE-CR PMD	25GBASE-CR-S PMD	25GBASE-KR PMD	25GBASE-KR-S PMD	25GBASE-SR PMD	25GBASE-T PCS/PMA
25GBASE-T	M			O	M	O											M

105.3 Summary of 25 Gigabit Ethernet sublayers

Change third paragraph of 105.3.1 as follows:

105.3.1 Reconciliation Sublayer (RS) and 25 Gigabit Media Independent Interface (25GMII)

While the 25GMII is an optional interface, it is used extensively in this standard as a basis for functional specification and provides a common service interface for ~~the 25GBASE-R PCS (Clause 107)~~a 25 Gb/s PHY.

105.3.6 Auto-Negotiation (AN)

Change text of 105.3.6 to add 25GBASE-T and Clause 28 Auto-Negotiation as follows:

AN provides a linked device with the capability to detect the abilities (modes of operation) supported by the device at the other end of the link, determine common abilities, and configure for joint operation.

Clause 28 AN is used by the 25GBASE-T PHY.

Clause 73 AN is used by the 25GBASE-CR, 25GBASE-CR-S, 25GBASE-KR, and 25GBASE-KR-S PHYs.

105.5 Delay constraints

Insert row for 25GBASE-T after row for 25GBASE-SR in Table 105-3 (unchanged rows not shown) as follows:

Table 105–3—Sublayer delay constraints

Sublayer	Maximum (bit time)[a]	Maximum (pause_quanta)[b]	Maximum (ns)	Notes[c]
25GBASE-T PHY	25 600	50	1024	See 113.11.

[a]1 bit time (BT) is equal to 40 ps. (See 1.4.117 for the definition of bit time.)
[b]1 pause_quantum is equal to 20.48 ns. (See 31B.2 for the definition of pause_quanta.)
[c]Should there be a discrepancy between this table and the delay requirements of the relevant sublayer clause, the sublayer clause prevails.

Insert new Clause 113 and new Annex 113A as follows:

113. Physical Coding Sublayer (PCS), Physical Medium Attachment (PMA) sublayer, and baseband medium, types 25GBASE-T and 40GBASE-T

113.1 Overview

The 25GBASE-T and 40GBASE-T PHYs are members of the 25 Gb/s and 40 Gb/s Ethernet family of high-speed network specifications. The 25GBASE-T PCS, PMA, and baseband medium specifications are intended for users who want 25 Gb/s performance over balanced twisted-pair structured cabling systems. The 40GBASE-T PCS, PMA, and baseband medium specifications are intended for users who want 40 Gb/s performance over balanced twisted-pair structured cabling systems. 25GBASE-T and 40GBASE-T signaling both require four pairs of balanced cabling as specified in ISO/IEC 11801-1:2016 and ANSI/TIA-568-C.2-1-2016.

This clause defines type 25GBASE-T and 40GBASE-T PCS, PMA sublayers, and Medium Dependent Interfaces (MDI). Together, the PCS and PMA sublayers define a Physical Layer (PHY). Functional, electrical, and mechanical specifications for the type 25GBASE-T PMA, 40GBASE-T PMA, and MDI are provided in this clause. This clause also specifies the baseband medium used with 25GBASE-T and 40GBASE-T. Management functions are optionally accessible through the management interface defined in Clause 45, or equivalent. Please refer to Table 105–2 and Table 80–2 for associated sublayers and options for assembling a 25 Gb/s system with the 25GBASE-T PHY and a 40 Gb/s system with the 40GBASE-T PHY, respectively.

This clause also specifies 25GBASE-T and 40GBASE-T Low Power Idle (LPI) as part of Energy-Efficient Ethernet (EEE). This allows the PHY to enter a low power mode of operation during periods of low link utilization as described in Clause 78.

Both 25GBASE-T and 40GBASE-T PHYs may optionally support a fast retrain mechanism. Implementation of the fast retrain option is recommended. Configurations wishing to disable fast retrain on the link may do so by advertising lack of support during Auto-Negotiation, thus preventing the link partner from attempting fast retrain and potentially dropping the link, see 45.2.7.10.

113.1.1 Nomenclature

The 25GBASE-T and 40GBASE-T PHYs described in Clause 113 represent two distinct PHY types that share the same PCS, PMA, and MDI specifications subject to frequency scaling, and differences between the 25GMII and the XLGMII specifications. In order to efficiently describe the two PHYs, the nomenclature 25G/40GBASE-T is used to describe specifications that apply to both the 25GBASE-T and 40GBASE-T PHYs. Additionally, for parameters that scale with the PHYs data rate, the parameter *S* is used for scaling. For 25GBASE-T, $S = 0.625$ and for 40GBASE-T, $S = 1$.

Where a functionality or register refers to any member of the MultiGBASE-T set of PHYs, as defined in 1.4.277a, that nomenclature is used.

Additionally, because the two PHYs specify different media independent interfaces, the nomenclature 25GMII/XLGMII is used as shorthand for referring to 25GMII for 25GBASE-T and XLGMII for 40GBASE-T. Note that 25GMII uses a 4-bit control word and a 32-bit data word, so that when a transfer of 64 bits (8 octets) is referred to on the media independent interface, this results in a reference like “two/one transfer(s) on the 25GMII/XLGMII.”

113.1.2 Relationship of 25GBASE-T and 40GBASE-T to other standards

Relations between the 25GBASE-T and 40GBASE-T PHYs, the ISO Open Systems Interconnection (OSI) reference model, and the IEEE 802.3 Ethernet model are shown in Figure 113–1. The PHY sublayers (shown shaded) in Figure 113–1 connect the IEEE 802.3 MAC to the medium.

The 25GBASE-T PHY service interface is the 25GMII, which is defined in Clause 106. The 40GBASE-T PHY service interface is the XLGMII, which is defined in Clause 81.

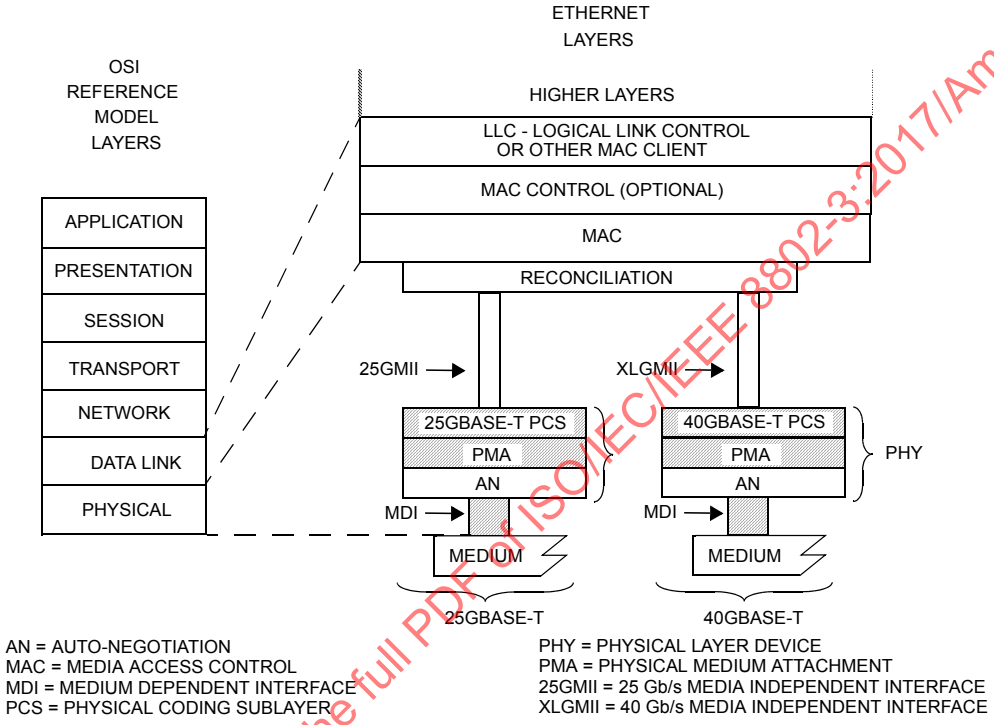


Figure 113–1—Types 25GBASE-T and 40GBASE-T PHY relationship to the ISO Open Systems Interconnection (OSI) reference model and the IEEE 802.3 Ethernet model

113.1.3 Operation of 25GBASE-T and 40GBASE-T

The 25GBASE-T and 40GBASE-T PHY each employ full duplex baseband transmission over four pairs of balanced twisted-pair structured cabling. The aggregate data rates of 25 Gb/s or 40 Gb/s are achieved by transmitting one-quarter of the aggregate data rate in each direction simultaneously on each wire pair, as shown in Figure 113–2. Baseband 16-level PAM signaling with a modulation rate of 2000 MBd for 25GBASE-T and 3200 MBd for 40GBASE-T is used on each of the wire pairs. Ethernet data and control characters are encoded at a rate of 3.125 information bits per PAM16 symbol, along with auxiliary bits. Two consecutively transmitted PAM16 symbols are considered as one two-dimensional (2D) symbol. The 2D symbols are selected from a constrained constellation of 128 maximally spaced 2D symbols, called DSQ128 (double square 128).[7] After link startup, PHY frames consisting of 512 DSQ128 symbols are continuously transmitted. The DSQ128 symbols are determined by seven-bit labels, each comprising 3 RS-FEC (Reed-Solomon FEC) encoded bits and 4 LDPC-encoded bits. The 512 DSQ128 symbols of one PHY frame are

[7]The resulting checkerboard constellation is based on a lattice called RZ^2 in the literature (see Forney [B31]). DSQ constellations have previously been introduced under the name "AMPM" (see Gallagher [B32] for examples of 8 point and 32 point AMPM/DSQ constellations).

transmitted as 4 × 256 PAM16 symbols over the four wire pairs. Data and Control symbols are embedded in a framing scheme that runs continuously after startup of the link. For 25GBASE-T, the modulation symbol rate of 2000 MBd results in a symbol period of 500 ps. For 40GBASE-T, the modulation symbol rate of 3200 MBd results in a symbol period of 312.5 ps.

A 25GBASE-T or 40GBASE-T PHY can be configured either as a MASTER PHY or as a SLAVE PHY. The MASTER-SLAVE relationship between two stations sharing a link segment is established during Auto-Negotiation (see Clause 28, 113.6, Annex 28B, Annex 28C, and Annex 28D). The MASTER PHY uses a local clock to determine the timing of transmitter operations. The MASTER-SLAVE relationship includes loop timing. The SLAVE PHY recovers the clock from the received signal and uses it to determine the timing of transmitter operations, i.e., it performs loop timing, as illustrated in Figure 113–3.

25GBASE-T and 40GBASE-T PHYs optionally provide support for LPI as part of EEE (see Clause 78). This extension allows PHYs to enter an LPI mode when either the local or link partner system requests low power operation. The transmit and receive functions may enter and leave the LPI mode independently so that both symmetric and asymmetric operation is supported. While the PHY is in the LPI mode, the PHY periodically transmits a refresh signal to allow the remote PHY to refresh its receiver state (e.g., timing recovery, adaptive filter coefficients) and thereby track long-term variation in the timing of the link or the underlying link segment characteristics. An easily detectable alert signal is transmitted to signal an end to the LPI mode. The alert signal is followed by a wake signal to enable a rapid transition back to the normal operational mode.

25GBASE-T and 40GBASE-T PHYs may optionally support a fast retrain mechanism. This function allows PHYs to quickly recover from link degradation without a normal two-second retrain.

The PCS and PMA subclauses of this document are summarized in 113.1.3.1 and 113.1.3.2. The EEE capability is summarized in 113.1.3.3. Figure 113–3 shows the functional block diagram.

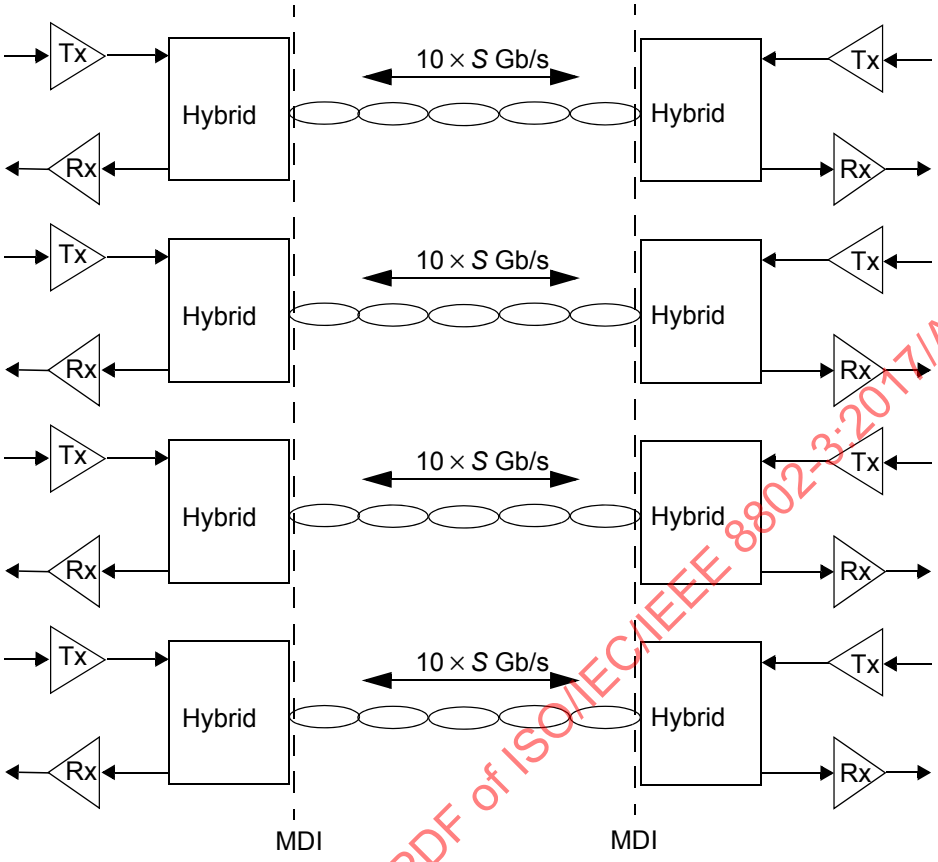
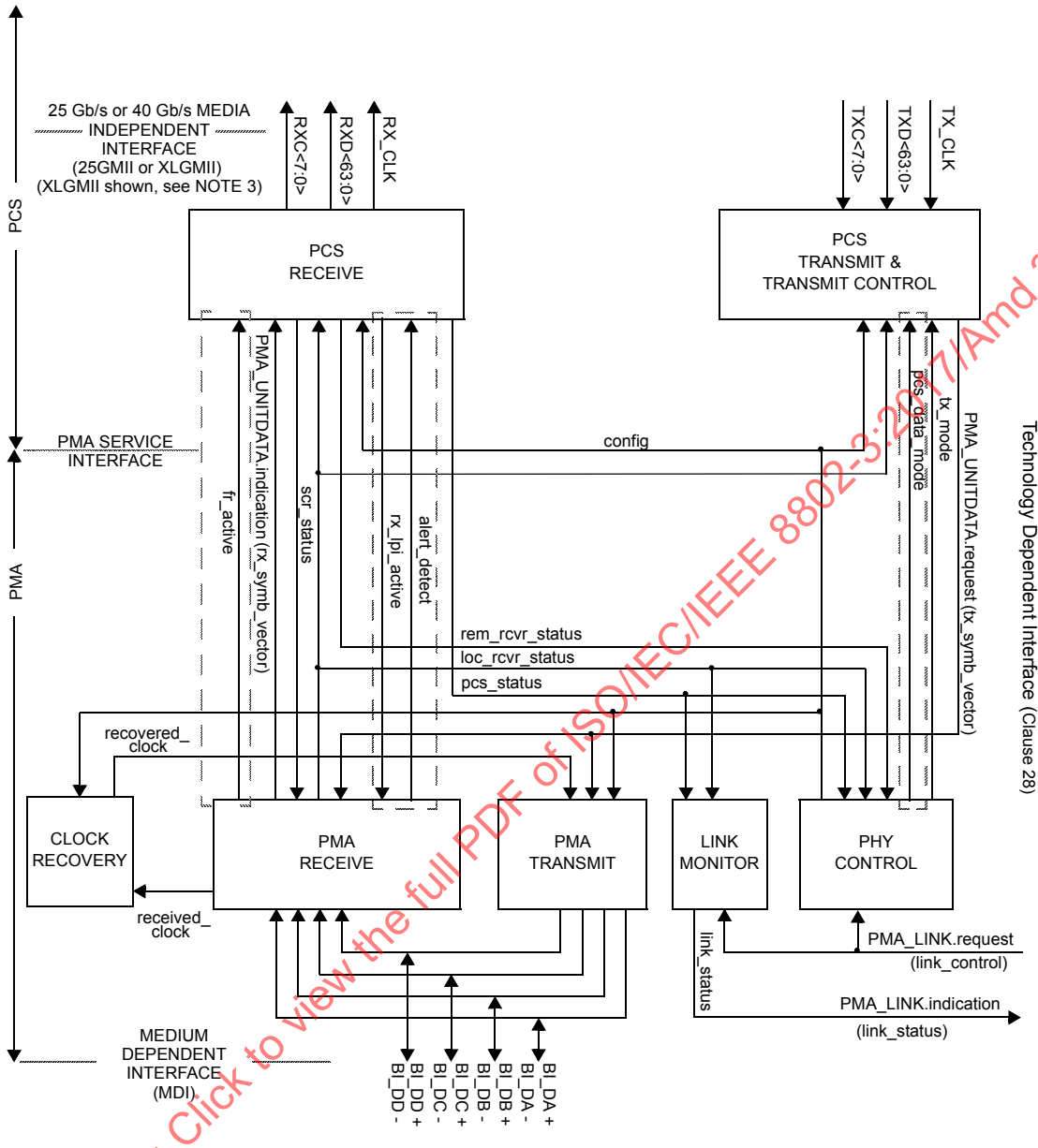


Figure 113–2—25GBASE-T and 40GBASE-T topology



NOTE 1—The recovered_clock arc is shown to indicate delivery of the received clock signal back to PMA TRANSMIT for loop timing.

NOTE 2—pcs_data_mode is required only for the EEE or fast retrain capabilities; alert_detect and rx_lpi_active are only required for the EEE capability; fr_active is only required for the fast retrain capability. Figures and capabilities only required for EEE are noted by dashed boxes.

NOTE 3— Word widths for only the 40 Gb/s Media Independent Interface (XLGMII) are shown for clarity. For the 25 Gb/s Media Independent Interface, a four-bit control word and 32-bit data word are used, i.e., RXC<3:0>, RXD<31:0>, TXC<3:0>, and TXD<31:0>.

Figure 113–3—Functional block diagram

113.1.3.1 Summary of Physical Coding Sublayer (PCS)

The 25GBASE-T PCS couples a 25 Gb/s Media Independent Interface (25GMII), as described in Clause 106, to the 25GBASE-T Physical Medium Attachment (PMA) sublayer.

The 40GBASE-T PCS couples a 40 Gb/s Media Independent Interface (XLGMII), as described in Clause 81, to the 40GBASE-T Physical Medium Attachment (PMA) sublayer.

In addition to the normal mode of operation, the PCS supports a training mode. Furthermore, the PCS contains a management interface.

In the transmit direction, in normal mode, the PCS receives eight 25GMII or XLGMII data octets provided by two transfers on the 25GMII service interface on TXD<31:0>, or one transfer on the XLGMII service interface of TXD<63:0>, and groups them into 64-bit blocks with the 64-bit block boundaries aligned with the boundary of the XLGMII transfer (or two 25GMII transfers). Each group of eight octets along with the data/control indications is transcoded into a 65-bit block. These 65-bit blocks are then aggregated into groups of 50 blocks. The first 48 65-bit blocks are transcoded into six 513-bit blocks, and the subsequent two 65-bit blocks are appended to complete a 50-group block. The resulting bits are then scrambled. This yields an Ethernet payload of $6 \times 513 + 2 \times 65 = 3208$ bits. An auxiliary bit is added after scrambling to obtain a block of 3209 bits.

The 3209 bits are divided into two groups, one of 1486 bits and the other of 1723 bits. To the 1486 bits, among them the auxiliary bit, 2 zero pad bits are added, and these are encoded into 192 eight-bit symbols of an RS-FEC(192, 186, 2^8), code. After substituting 2 random fill bits for the zero pad bits, a total of $1536 = 3 \times 512$ RS-FEC-coded bits are transmitted. The 1723 bits are encoded by a systematic LDPC(1723,2048) encoder, which adds 325 LDPC check bits to form an LDPC codeword of 2048 LDPC-coded bits. The 3×512 RS-FEC-coded bits and the $2048 = 4 \times 512$ LDPC-coded bits are arranged in a frame of 512 seven-bit labels. Each seven-bit label comprises 3 RS-FEC-coded bits and 4 LDPC-coded bits.

The 512 seven-bit labels are mapped into 512 2D modulation symbols selected from a DSQ128 constellation. The DSQ128 symbols are obtained by concatenating two time-adjacent 1D PAM16 symbols and retaining among the 256 possible Cartesian product combinations, 128 maximally spaced 2D symbols.[8]

The DSQ128 constellation is partitioned into eight subsets, each subset containing 16 maximally spaced 2D symbols. The three RS-FEC-coded bits of each seven-bit label select one DSQ128 subset, and the four LDPC-coded-bits of the label select one 2D symbol in this subset.

The obtained PHY frame of 512 DSQ128 symbols is passed on to the PMA as PMA_UNITDATA.request. The PMA transmits the DSQ128 symbols over the four wire pairs in the form of 256 constituent PAM16 symbols per pair.

In the receive direction, in normal mode, the PCS processes code-groups received from the remote PHY via the PMA in 256 4D symbol blocks and maps them to the 25GMII/XLGMII service interface in the receive path. In this receive processing scheme, symbol clock synchronization is done by the PMA Receive function.

For 25GBASE-T, the signals provided by the PCS at the 25GMII conform to the interface requirements of Clause 106. For 40GBASE-T, the signals provided by the PCS at the XLGMII conform to the interface requirements of Clause 81.

[8]The resulting checkerboard constellation is based on a lattice called RZ^2 in the literature (see Forney [B31]).

Details of the PCS functions and state diagrams are covered in 113.3. The interface to the PMA is an abstract message-passing interface specified in 113.2.

113.1.3.2 Summary of Physical Medium Attachment (PMA) sublayer

The PMA couples messages from the PCS service interface onto the balanced cabling physical medium via the Medium Dependent Interface (MDI) and provides the link management and PHY Control functions. The PMA provides full duplex communications at $3200 \times S$ MBd over four pairs of balanced cabling up to 30 m in length.

The PMA Transmit function comprises four transmitters to generate continuous time analog signals on each of the four pairs BI_DA, BI_DB, BI_DC, and BI_DD, as described in 113.4.3.1. In normal mode, each four-dimensional (4D) symbol received from the PCS Transmit function undergoes multiple stages of processing. First the symbol goes through a Tomlinson-Harashima precoder (THP), which maps the PAM16 input (as described in 113.3.2.2.20) in each dimension of the four-dimensional symbol into a quasi-continuous discrete-time value in the range $-16 \le x < 16$. This THP-processed four-dimensional symbol stream may be further processed by a digital transmit filter and is then passed on to four digital to analog converters (DACs). The DAC outputs may be further processed with continuous time filters to roll off the high-frequency spectral response to limit high-frequency emissions and are then applied to each of the four balanced pairs via the MDI port.

The PMA Receive function comprises four independent receivers for pulse-amplitude modulated signals on each of the four pairs BI_DA, BI_DB, BI_DC, and BI_DD, as described in 113.4.3.2. The receivers are responsible for acquiring symbol timing and, when operating in normal mode, for canceling echo, near-end crosstalk, far-end crosstalk, and equalizing the signal. The 4D symbols are provided to the PCS Receive function via the PMA_UNITDATA.indication message. The PMA also contains functions for Link Monitor.

The PMA PHY Control function generates signals that control the PCS and PMA sublayer operations. PHY Control begins following the completion of Auto-Negotiation and provides the startup functions required for successful 25GBASE-T or 40GBASE-T operation. It determines whether the PHY operates in a normal mode, enabling data transmission over the link segment, or in training mode, in which it sends and receives special PAM2 code-groups and data transmission is disabled.

PMA functions and state diagrams are specified in 113.4. PMA electrical specifications are given in 113.5.

The PMA sublayer may also support a fast retrain function. The fast retrain function is specified in 113.4.2.5.16.

113.1.3.3 Summary of EEE capability

A 25GBASE-T or 40GBASE-T PHY may optionally support the EEE capability, as described in 78.1.4. The EEE capability is a mechanism by which 25GBASE-T and 40GBASE-T PHYs are able to reduce power consumption during periods of low link utilization. PHYs can enter this mode of operation after reaching PCS data mode. Each direction of the full duplex link is able to enter and exit the LPI mode independently, supporting symmetric and asymmetric LPI operation. This allows power savings when only one side of the full duplex link is in a period of low utilization. No data frames are lost or corrupted during the transition to or from the LPI mode.

In the transmit direction the transition to the LPI transmit mode begins when the PCS transmit function detects an LPI control character in all eight lanes of one transfer of TXD[63:0] that will be mapped into a single 64B/65B block, according to the position in the 25G/40GBASE-T LDPC frame. Following this event a sleep signal is transmitted by the PMA. The sleep signal is composed of LDPC frames that contain only LP_IDLE 512B/513B and 64B/65B blocks. The sleep signal indicates to the link partner that the transmit function of the PHY is entering the LPI transmit mode. Immediately after the transmission of the sleep

frames, the transmit function of the local PHY enters the LPI transmit mode. While the transmit function is in the LPI mode the PHY may disable data path and control logic to save additional power. Periodically the transmit function of the local PHY transmits refresh frames that are used by the link partner to update adaptive filters and timing circuits in order to maintain link integrity. The LPI mode begins with quiet signaling or with a full refresh period. Partial refreshes (defined as a refresh signal shorter than 6 LDPC frames) that immediately follow the transition to the LPI mode are replaced with quiet signaling. The quiet-refresh cycle continues until the PCS function detects IDLE characters on the 25GMII/XLGMII. These characters signal to the PHY that the LPI transmit mode should end. The PMA Transmit function in the PHY then sends an alert message to the link partner. The alert signal begins on a LDPC frame boundary, but has no fixed relationship to the quiet-refresh cycle. The alert signal wakes the link partner from sleep. The alert signal is followed by a wake signal, composed of LDPC frames containing only IDLE 512B/513B and 64B/65B blocks. After a short recovery time the normal operational mode is resumed.

In the receive direction the transition to the LPI mode is triggered when the PCS Receive function detects LPI control characters within received LDPC frames. This indicates that the link partner is about to enter the LPI transmit mode. Following these frames the link partner ceases transmission and begins quiet-refresh signaling. During the quiet time it is highly recommended that the local receiver power off circuits to reduce power consumption. Periodically the link partner transmits refresh frames that are used by the receiver to update adaptive coefficients and timing circuits. This quiet-refresh cycle continues until the link partner transmits the alert signal, initiating a transition back to the normal operational mode. The alert signal is detected in the PMA and signals that normal data frames will follow. The alert signal is followed by a wake signal that allows the local receiver time to prepare for the normal operational mode. The wake signal is composed of repeated IDLE 512B/513B and 64B/65B blocks. After a short recovery time the normal operational mode is resumed.

Support for the EEE capability is advertised in the Infofield (Octet 12 bit 7) during link startup. Transitions to and from the LPI transmit mode are controlled via 25GMII/XLGMII signaling. Transitions to and from the LPI receive mode are controlled by the link partner using sleep, alert, and wake signaling.

The PCS 64B/65B Transmit state diagram in Figure 113–18a and Figure 113–18b includes additional states for EEE. The PCS 64B/65B Receive state diagram in Figure 113–19a and Figure 113–19b includes additional states for EEE. The EEE Transmit state diagram is contained in the PCS Transmit function and is specified in Figure 113–20.

113.1.4 Signaling

25GBASE-T and 40GBASE-T signaling is performed by the PCS generating continuous code-group sequences that the PMA transmits over each wire pair. The signaling scheme achieves a number of objectives including:

a) Forward error correction (FEC) coded symbol mapping for data.
b) Algorithmic mapping from TXD<31:0> and TXC<3:0> for 25GBASE-T, or TXD<63:0> and TXC<7:0> for 40GBASE-T, to four-dimensional symbols in the transmit path.
c) Algorithmic mapping from the received four-dimensional signals on the MDI port to RXD<31:0> and RXC<3:0> for 25GBASE-T on the 25GMII interface, or RXD<63:0> and RXC<7:0> for 40GBASE-T on the XLGMII interface.
d) Uncorrelated symbols in the transmitted symbol stream.
e) No correlation between symbol streams traveling both directions on any pair combination.
f) No correlation between symbol streams on pairs BI_DA, BI_DB, BI_DC, and BI_DD.
g) Block framing and other control signals.
h) Ability to signal the status of the local receiver to the remote PHY to indicate that the local receiver is not operating reliably and requires retraining.
i) Ability to automatically detect and correct for pair swapping and crossover connections.
j) Ability to automatically detect and correct for incorrect polarity in the connections.

k) Ability to automatically correct for differential delay variations across the wire pairs.
l) Ability to support refresh, quiet and alert signaling during LPI operation.

The PHY operates in two modes—normal mode or training mode. In normal mode, PCS generates a continuous stream of four-dimensional symbols that are transmitted via the PMA at one of eight power levels. In training mode, the PCS is directed to generate only PAM2 symbols for transmission by the PMA, which enable the receiver at the other end to train until it is ready to operate in normal mode. (See Figure 113–5.)

PHYs may also support the EEE capability as described in 113.1.3.3. Transitions to the LPI mode are supported after reaching normal mode.

113.1.5 Interfaces

All 25GBASE-T and 40GBASE-T PHY implementations are compatible at the MDI, and at the 25GMII/XLGMII, if implemented. Implementation of the 25GMII/XLGMII is optional. Designers are free to implement circuitry within the PCS and PMA in an application-dependent manner provided that the MDI and 25GMII/XLGMII (if the 25GMII/XLGMII is implemented) specifications are met. System operation from the perspective of signals at the MDI and management objects are identical whether the 25GMII/XLGMII is implemented or not.

113.1.6 Conventions in this clause

The body of this clause contains state diagrams, including definitions of variables, constants, and functions. Should there be a discrepancy between a state diagram and descriptive text, the state diagram prevails.

The notation used in the state diagrams follows the conventions of 21.5.

Default initializations, unless specified, are left to the implementer.

113.2 25GBASE-T and 40GBASE-T service primitives and interfaces

25GBASE-T transfers data and control information across the following four service interfaces:

a) 25 Gb/s Media Independent Interface (25GMII)
b) Technology Dependent Interface
c) PMA service interface
d) Medium dependent interface (MDI)

The 25GMII is specified in Clause 106; the Technology Dependent Interface is specified in Clause 28. The PMA service interface is defined in 113.2.2 and the MDI is defined in 113.8.

40GBASE-T transfers data and control information across the following four service interfaces:

a1) 40 Gb/s Media Independent Interface (XLGMII)
b1) Technology Dependent Interface
c1) PMA service interface
d1) Medium dependent interface (MDI)

The XLGMII is specified in Clause 81; the Technology Dependent Interface is specified in Clause 28. The PMA service interface is defined in 113.2.2 and the MDI is defined in 113.8.

113.2.1 Technology Dependent Interface

25G/40GBASE-T use the following service primitives to exchange status indications and control signals across the Technology Dependent Interface as specified in Clause 28:

PMA_LINK.request (link_control)

PMA_LINK.indication (link_status)

113.2.1.1 PMA_LINK.request

This primitive allows the Auto-Negotiation algorithm to enable and disable operation of the PMA as specified in 28.2.6.2.

113.2.1.1.1 Semantics of the primitive

PMA_LINK.request (link_control)

The link_control parameter can take on one of three values: SCAN_FOR_CARRIER, DISABLE, or ENABLE.

SCAN_FOR_CARRIER	Used by the Auto-Negotiation algorithm prior to receiving any fast link pulses. During this mode the PMA reports link_status=FAIL. PHY processes are disabled.
DISABLE	Set by the Auto-Negotiation algorithm in the event fast link pulses are detected. PHY processes are disabled. This allows the Auto-Negotiation algorithm to determine how to configure the link.
ENABLE	Used by Auto-Negotiation to turn control over to the PHY for data processing functions.

113.2.1.1.2 When generated

Auto-Negotiation generates this primitive to indicate a change in link_control as described in Clause 28.

113.2.1.1.3 Effect of receipt

This primitive affects operation of the PMA Link Monitor function as defined in 113.4.2.6.

113.2.1.2 PMA_LINK.indication

This primitive is generated by the PMA to indicate the status of the underlying medium as specified in 28.2.6.1. This primitive informs the Auto-Negotiation algorithm about the status of the underlying link.

113.2.1.2.1 Semantics of the primitive

PMA_LINK.indication (link_status)

The link_status parameter can take on one of two values: FAIL or OK.

FAIL	No valid link established.
OK	The Link Monitor function indicates that a valid 25G/40GBASE-T link is established. Reliable reception of signals transmitted from the remote PHY is possible.

113.2.1.2.2 When generated

The PMA generates this primitive to indicate a change in link_status in compliance with the state diagram given in Figure 113–31.

113.2.1.2.3 Effect of receipt

Auto-Negotiation uses this primitive to detect a change in link_status as described in Clause 28.

113.2.2 PMA service interface

25G/40GBASE-T use the following service primitives to exchange symbol vectors, status indications, and control signals across the service interfaces:

PMA_TXMODE.indication (tx_mode)

PMA_CONFIG.indication (config)

PMA_UNITDATA.request (tx_symb_vector)

PMA_UNITDATA.indication (rx_symb_vector)

PMA_SCRSTATUS.request (scr_status)

PMA_PCSSTATUS.request (pcs_status)

PMA_RXSTATUS.indication (loc_rcvr_status)

PMA_REMRXSTATUS.request (rem_rcvr_status)

EEE-capable PHYs additionally support the following service primitives:

PMA_ALERTDETECT.indication (alert_detect)

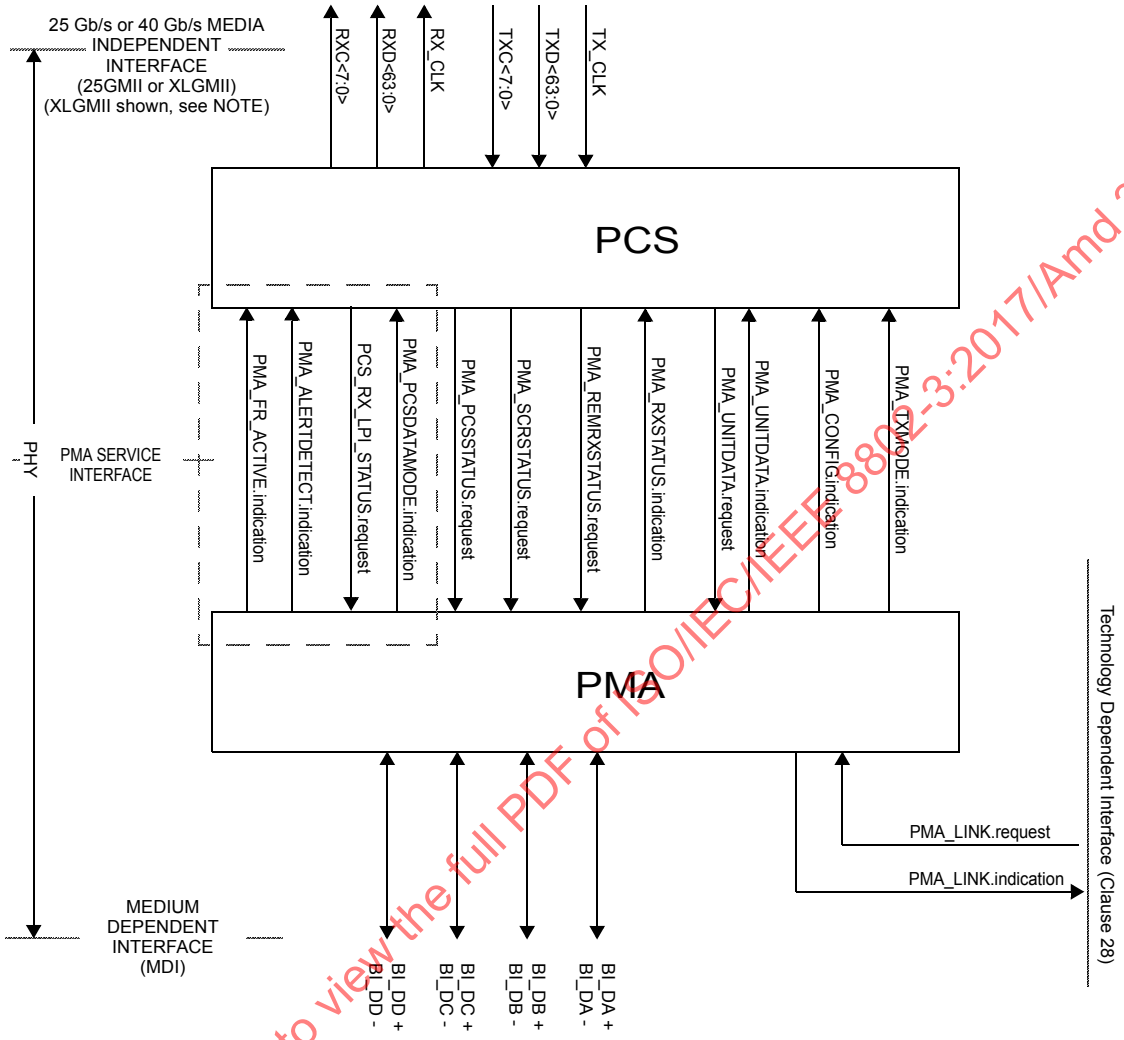
PCS_RX_LPI_STATUS.request (rx_lpi_active)

PMA_PCSDATAMODE.indication (PCS_data_mode)

Fast retrain capable PHYs additionally support the following service primitive:

PMA_FR_ACTIVE.indication (fr_active)

The use of these primitives is illustrated in Figure 113–4. Connections from the management interface (signals MDC and MDIO) to the sublayers are pervasive and are not shown in Figure 113–4.



NOTE—
PMA_PCSDATAMODE.indication is required only for the EEE or fast retrain capabilities.
PMA_ALERTDETECT.indication and PCS_RX_LPI_STATUS.request are only required for the EEE capability.
PMA_FR_ACTIVE.indication is only required for the fast retrain capability.
Word widths for only the 40 Gb/s Media Independent Interface (XLGMII) are shown for clarity. For the 25 Gb/s Media Independent Interface, a four-bit control word and 32-bit data word are used, i.e., RXC<3:0>, RXD<31:0>, TXC<3:0>, and TXD<31:0>.

Figure 113–4—25GBASE-T and 40GBASE-T service interfaces

113.2.2.1 PMA_TXMODE.indication

The transmitter in a 25G/40GBASE-T link normally sends over the four pairs, four-dimensional symbols that represent a 25GMII/XLGMII data stream with framing, scrambling and encoding of data, control information, or idles.

113.2.2.1.1 Semantics of the primitive

PMA_TXMODE.indication (tx_mode)

PMA_TXMODE.indication specifies to PCS Transmit via the parameter tx_mode what sequence of code-groups the PCS should be transmitting. The parameter tx_mode can take on one of the following three values of the form:

SEND_N	This value is continuously asserted when transmission of sequences of four-dimensional symbols representing a 25GMII/XLGMII data stream in normal mode.
SEND_T	This value is continuously asserted in case transmission of sequences of code-groups representing the training mode is to take place.
SEND_Z	This value is continuously asserted in case transmission of zeros is required.

113.2.2.1.2 When generated

The PMA PHY Control function generates PMA_TXMODE.indication messages to indicate a change in tx_mode.

113.2.2.1.3 Effect of receipt

Upon receipt of this primitive, the PCS performs its transmit function as described in 113.3.2.2.

113.2.2.2 PMA_CONFIG.indication

Each PHY in a 25G/40GBASE-T link is capable of operating as a MASTER PHY and as a SLAVE PHY. MASTER-SLAVE configuration is determined during Auto-Negotiation (113.6.1). The result of this negotiation is provided to the PMA.

113.2.2.2.1 Semantics of the primitive

PMA_CONFIG.indication (config)

PMA_CONFIG.indication specifies to PCS and PMA Transmit via the parameter config whether the PHY operates as a MASTER PHY or as a SLAVE PHY. The parameter config can take on one of the following two values of the form:

MASTER	This value is continuously asserted when the PHY operates as a MASTER PHY.
SLAVE	This value is continuously asserted when the PHY operates as a SLAVE PHY.

113.2.2.2.2 When generated

PMA generates PMA_CONFIG.indication messages to indicate a change in config.

113.2.2.2.3 Effect of receipt

PCS and PMA Clock Recovery perform their functions in MASTER or SLAVE configuration according to the value assumed by the parameter config.

113.2.2.3 PMA_UNITDATA.request

This primitive defines the transfer of code-groups in the form of the tx_symb_vector parameter from the PCS to the PMA. The code-groups are obtained in the PCS Transmit function using the encoding rules defined in 113.3.2.2 to represent 25GMII/XLGMII data and control streams or other sequences.

113.2.2.3.1 Semantics of the primitive

PMA_UNITDATA.request (tx_symb_vector)

During transmission, the PMA_UNITDATA.request simultaneously conveys to the PMA via the parameter tx_symb_vector the value of the symbols to be sent over each of the four transmit pairs BI_DA, BI_DB, BI_DC, and BI_DD. For EEE-capable PHYs, the vector also requests the PMA to send the ALERT signal during LPI. The tx_symb_vector parameter takes on the form:

SYMB_4D	A vector of four multi-level symbols, one for each of the four transmit pairs BI_DA, BI_DB, BI_DC, and BI_DD. In normal operation, each symbol takes on one of the values in the set {–15, –13, –11, –9, –7, –5, –3, –1, 1, 3, 5, 7, 9, 11, 13, 15}. The symbols additionally take the value 0 when zeros are to be transmitted in the following two cases: 1) when PMA_TXMODE.indication is SEND_Z during PMA training, and 2) after data mode is reached, the transmit function is in the LPI transmit mode and lpi_tx_mode is QUIET
ALERT	A vector used to indicate that the PMA should transmit the alert sequence. ALERT is asserted for a time equal to 4 LDPC frames.

The symbols that are elements of tx_symb_vector are called, according to the pair on which each is transmitted, tx_symb_vector[BI_DA], tx_symb_vector[BI_DB], tx_symb_vector[BI_DC], and tx_symb_vector[BI_DD].

113.2.2.3.2 When generated

The PCS generates PMA_UNITDATA.request synchronously with every transmit clock cycle.

113.2.2.3.3 Effect of receipt

Upon receipt of this primitive the PMA transmits on the MDI the signals corresponding to the indicated symbols after processing with the THP, the transmit filter and other specified PMA Transmit processing. The parameter tx_symb_vector is also used by the PMA Receive function to process the signals received on pairs BI_DA, BI_DB, BI_DC, and BI_DD for cancelling the echo and near-end crosstalk (NEXT).

113.2.2.4 PMA_UNITDATA.indication

This primitive defines the transfer of code-groups in the form of the rx_symb_vector parameter from the PMA to the PCS.

113.2.2.4.1 Semantics of the primitive

PMA_UNITDATA.indication (rx_symb_vector)

During reception the PMA_UNITDATA.indication simultaneously conveys to the PCS via the parameter rx_symb_vector the values of the symbols detected on each of the four receive pairs BI_DA, BI_DB, BI_DC, and BI_DD. The rx_symb_vector parameter takes on the form:

SYMB_4D	A vector of the four 1D symbols that is the receiver's best estimate of the symbols that were sent by the remote transmitter across the four pairs with reliability measures.

113.2.2.4.2 When generated

The PMA generates PMA_UNITDATA.indication (SYMB_4D) messages synchronously every four symbols received at the MDI. The nominal rate of the PMA_UNITDATA.indication primitive is $3200 \times S$ MHz, as governed by the recovered clock.

113.2.2.4.3 Effect of receipt

The effect of receipt of this primitive is unspecified.

113.2.2.5 PMA_SCRSTATUS.request

This primitive is generated by PCS Receive to communicate the status of the descrambler for the local PHY. The parameter scr_status conveys to the PMA Receive function the information that the training mode descrambler has achieved synchronization.

113.2.2.5.1 Semantics of the primitive

PMA_SCRSTATUS.request (scr_status)

The scr_status parameter can take on one of two values of the form:

OK	The training mode descrambler has achieved synchronization.
NOT_OK	The training mode descrambler is not synchronized.

113.2.2.5.2 When generated

PCS Receive generates PMA_SCRSTATUS.request messages to indicate a change in scr_status.

113.2.2.5.3 Effect of receipt

The effect of receipt of this primitive is specified in 113.4.2.4, 113.4.2.5, and 113.4.6.1.

113.2.2.6 PMA_PCSSTATUS.request

This primitive is generated by PCS Receive to indicate the fully operational state of the PCS for the local PHY. The parameter pcs_status conveys to the PMA Receive function the information that the PCS is operating reliably in data mode.

113.2.2.6.1 Semantics of the primitive

PMA_PCSSTATUS.request (pcs_status)

The pcs_status parameter can take on one of two values of the form:

OK	The PCS is operating reliably in data mode.
NOT_OK	The PCS is not operating reliably in data mode.

113.2.2.6.2 When generated

PCS Receive generates PMA_PCSSTATUS.request messages to indicate a change in pcs_status.

113.2.2.6.3 Effect of receipt

The effect of receipt of this primitive is specified in 113.4.6.

113.2.2.7 PMA_RXSTATUS.indication

This primitive is generated by PMA Receive to indicate the status of the receive link at the local PHY. The parameter loc_rcvr_status conveys to the PCS Transmit, PCS Receive, PMA PHY Control function, and Link Monitor the information on whether the status of the overall receive link is satisfactory or not. Note that loc_rcvr_status is used by the PCS Receive decoding functions. The criterion for setting the parameter loc_rcvr_status is left to the implementer. It can be based, for example, on observing the mean-square error at the decision point of the receiver and detecting errors during reception of symbol streams.

113.2.2.7.1 Semantics of the primitive

PMA_RXSTATUS.indication (loc_rcvr_status)

The loc_rcvr_status parameter can take on one of two values of the form:

OK	This value is asserted and remains true during reliable operation of the receive link for the local PHY.
NOT_OK	This value is asserted whenever operation of the link for the local PHY is unreliable.

113.2.2.7.2 When generated

PMA Receive generates PMA_RXSTATUS.indication messages to indicate a change in loc_rcvr_status on the basis of signals received at the MDI.

113.2.2.7.3 Effect of receipt

The effect of receipt of this primitive is specified in Figure 113–28 and in 113.2 and 113.4.6.3.

113.2.2.8 PMA_REMRXSTATUS.request

This primitive is generated by PCS Receive to indicate the status of the receive link at the remote PHY as communicated by the remote PHY via its encoding of its loc_rcvr_status parameter. The parameter rem_rcvr_status conveys to the PMA PHY Control function the information on whether reliable operation of the remote PHY is detected or not. The criterion for setting the parameter rem_rcvr_status is left to the implementer. It can be based, for example, on asserting rem_rcvr_status is NOT_OK until loc_rcvr_status is OK and then asserting the detected value of rem_rcvr_status after proper PCS Receive decoding is achieved.

113.2.2.8.1 Semantics of the primitive

PMA_REMRXSTATUS.request (rem_rcvr_status)

The rem_rcvr_status parameter can take on one of two values of the form:

OK	The receive link for the remote PHY is operating reliably.
NOT_OK	Reliable operation of the receive link for the remote PHY is not detected.

113.2.2.8.2 When generated

The PCS generates PMA_REMRXSTATUS.request messages to indicate a change in rem_rcvr_status on the basis of signals received at the MDI.

113.2.2.8.3 Effect of receipt

The effect of receipt of this primitive is specified in Figure 113–28.

113.2.2.9 PMA_ALERTDETECT.indication

This primitive is generated by PMA Receive to indicate the status of the receive link at the local PHY when rx_lpi_active is TRUE. The parameter alert_detect conveys to the PCS receive function information regarding the detection of the LPI alert signal by the PMA receive function. The criterion for setting the parameter alert_detect is left to the implementer.

113.2.2.9.1 Semantics of the primitive

PMA_ALERTDETECT.indication (alert_detect)

The alert_detect parameter can take on one of two values of the form:

TRUE	The alert signal has been reliably detected at the local receiver.
FALSE	The alert signal at the local receiver has not been detected.

113.2.2.9.2 When generated

The PMA generates PMA_ALERTDETECT.indication messages to indicate a change in the alert_detect status.

113.2.2.9.3 Effect of receipt

The effect of receipt of this primitive is specified in 113.3.2.3, Figure 113–19a, and Figure 113–19b.

113.2.2.10 PCS_RX_LPI_STATUS.request

When the PHY supports the EEE capability this primitive is generated by the PCS receive function to indicate the status of the receive link at the local PHY. The parameter PCS_RX_LPI_STATUS.request conveys to the PCS transmit and PMA receive functions information regarding whether the receive function is in the LPI receive mode. The parameter is generated by the Receive 64B/65B state diagram in Figure 113–19a.

113.2.2.10.1 Semantics of the primitive

PCS_RX_LPI_STATUS.request (rx_lpi_active)

The rx_lpi_active parameter can take on one of two values of the form:

TRUE	The receive function is in the LPI receive mode.
FALSE	The receive function is not in the LPI receive mode.

113.2.2.10.2 When generated

The PCS generates PCS_RX_LPI_STATUS.request messages to indicate a change in the rx_lpi_active variable as determined by the receive state diagram in Figure 113–19a.

113.2.2.10.3 Effect of receipt

The effect of receipt of this primitive is specified in 113.3.2.3 and Figure 113–32.

113.2.2.11 PMA_PCSDATAMODE.indication

This primitive indicates whether or not the PCS state diagrams are able to transition from their initialization states. The pcs_data_mode variable is generated by the PMA PHY Control function. It is passed to the PCS Control function via the PMA_PCSDATAMODE.indication primitive.

113.2.2.11.1 Semantics of the primitive

PMA_PCSDATAMODE.indication (pcs_data_mode)

The pcs_data_mode parameter can take on one of two values of the form:

TRUE	PHY is in state PCS_Data (see Figure 113–28)
FALSE	PCS is not in state PCS_Data (see Figure 113–28).

113.2.2.11.2 When generated

The PMA PHY Control function generates PMA_PCSDATAMODE.indication messages continuously.

113.2.2.11.3 Effect of receipt

Upon receipt of this primitive, the PCS performs its transmit function as described in 113.3.2.2.

113.2.2.12 PMA_FR_ACTIVE.indication

This primitive indicates whether or not the PMA is currently performing a fast retrain. The fr_active variable is generated by the PMA PHY Control function. It is passed to the PCS Receive Control function via the PMA_FR_ACTIVE.indication primitive. This primitive is only supported by PHYs with the fast retrain capability.

113.2.2.12.1 Semantics of the primitive

PMA_FR_ACTIVE.indication (fr_active)

The fr_active parameter can take on one of two values of the form:

TRUE	PHY is currently performing a fast retrain
FALSE	PCS is not currently performing a fast retrain.

113.2.2.12.2 When generated

The PMA PHY Control function generates PMA_FR_ACTIVE.indication messages continuously.

113.2.2.12.3 Effect of receipt

The effect of receipt of this primitive is specified in Figure 113–19a.

113.3 Physical Coding Sublayer (PCS)

113.3.1 PCS service interface (25GMII/XLGMII)

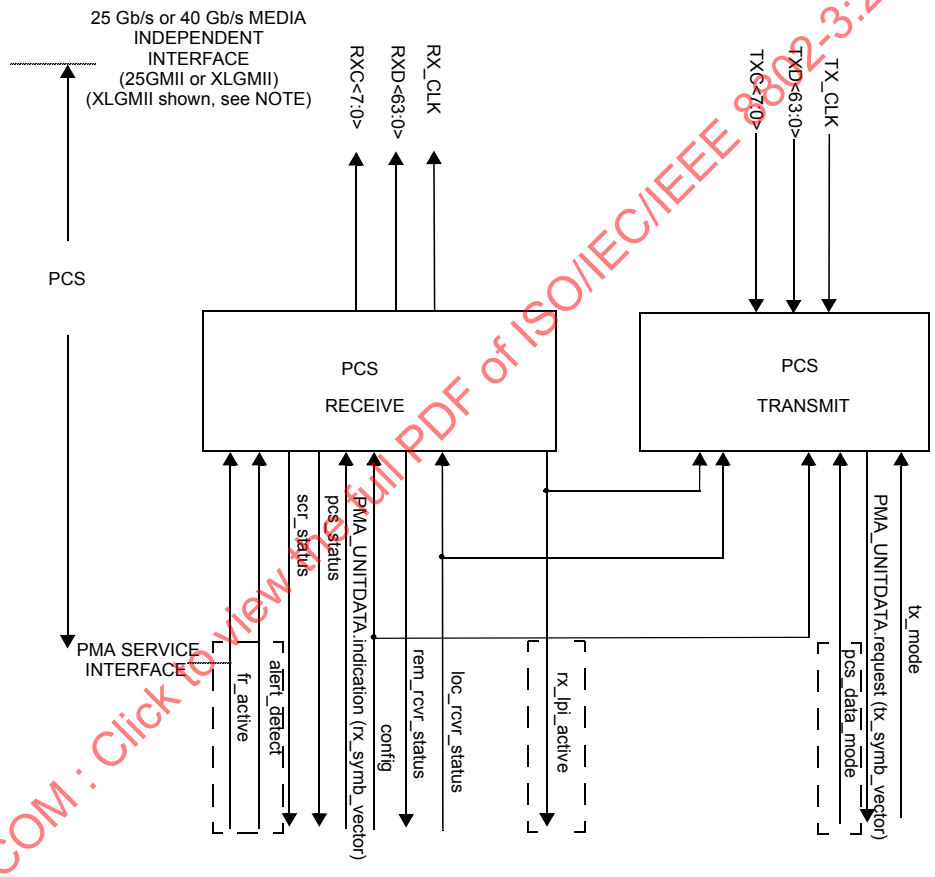
The PCS service interface allows the 25G/40GBASE-T PCS to transfer information to and from a PCS client. The PCS Interface is precisely defined for 25GBASE-T as the 25 Gb/s Media Independent Interface

(25GMII) in Clause 106, and, for 40GBASE-T, the 40 Gb/s Media Independent Interface (XLGMII) in Clause 81.

113.3.2 PCS functions

The PCS comprises one PCS Reset function and two simultaneous and asynchronous operating functions. The PCS operating functions are: PCS Transmit and PCS Receive. All operating functions start immediately after the successful completion of the PCS Reset function.

The PCS reference diagram, Figure 113–5, shows how the two operating functions relate to the messages of the PCS-PMA interface. Connections from the management interface (signals MDC and MDIO) to other layers are pervasive and are not shown in Figure 113–5.



NOTE—
pcs_data_mode is required only for the EEE or fast retrain capabilities.
alert_detect and rx_lpi_active are only required for the EEE capability.
fr_active is only required for the fast retrain capability.
Word widths for only the 40 Gb/s Media Independent Interface (XLGMII) are shown for clarity. For the 25 Gb/s Media Independent Interface, a four-bit control word and 32-bit data word are used, i.e., RXC<3:0>, RXD<31:0>, TXC<3:0>, and TXD<31:0>.

Figure 113–5—PCS reference diagram

113.3.2.1 PCS Reset function

PCS Reset initializes all PCS functions. The PCS Reset function shall be executed whenever one of the following conditions occur:

a) Power on (see 113.3.6.2.2).
b) The receipt of a request for reset from the management entity.

PCS Reset sets pcs_reset = true while any of the above reset conditions hold true. All state diagrams take the open-ended pcs_reset branch upon execution of PCS Reset. The reference diagrams do not explicitly show the PCS Reset function.

113.3.2.2 PCS Transmit function

The PCS Transmit function shall conform to the PCS 64B/65B Transmit state diagram in Figure 113–18a and Figure 113–18b, and to the PCS Transmit bit ordering in Figure 113–6 and Figure 113–8.

Dashed rectangles in Figure 113–18a and Figure 113–18b are used to indicate states and state transitions in the transmit process state diagram that shall be supported by PHYs with the EEE capability. PHYs without the EEE capability do not support these transitions.

When communicating with the 25GMII, the 25GBASE-T PCS uses a four octet-wide, synchronous data path, with packet delimiting being provided by transmit control signals and receive control signals. When communicating with the XLGMII, the 40GBASE-T PCS uses a eight octet-wide, synchronous data path, with packet delimiting being provided by transmit control signals and receive control signals. Alignment to 64B/65B is performed in the PCS. The PMA sublayer operates independently of block and packet boundaries. The PCS provides the functions necessary to map packets between the 25GMII/XLGMII format and the PMA service interface format.

When the transmitter is in normal mode, the PCS Transmit process continuously generates 65B blocks based upon the TXD <31:0> and TXC <3:0> signals on the 25GMII for 25GBASE-T, or the TXD <63:0> and TXC <7:0> signals on the XLGMII for 40GBASE-T. The subsequent functions of the PCS Transmit process then transcode the first 96 25GMII transfers for 25GBASE-T, or 48 XLGMII transfers for 40GBASE-T into 512B/513B blocks, append the subsequent four 25GMII transfers (25GBASE-T), or two XLGMII transfers (40GBASE-T) as (non-transcoded) 64B/65B blocks, scramble the bits, pack the resulting blocks, appending an unscrambled auxiliary bit, and split the bits into two sets. The first set is encoded by a Reed-Solomon encoder, and the second set is processed by a low density parity check (LDPC) encoder and then the two sets are joint mapped into a transmit LDPC frame of DSQ128 symbols. Transmit data-units are sent to the PMA service interface via the PMA_UNITDATA.request primitive.

In each symbol period, when communicating with the PMA, the PCS Transmit generates a code-group (A_n, B_n, C_n, D_n) that is transferred to the PMA via the PMA_UNITDATA.request primitive. The PMA transmits symbols A_n, B_n, C_n, D_n over wire pairs BI_DA, BI_DB, BI_DC, and BI_DD, respectively. The integer, n, is a time index that is introduced to establish a temporal relationship between different symbol periods. A symbol period, T, is 312.5/S ps.

If a PMA_TXMODE.indication message has the value SEND_Z, PCS Transmit shall pass a vector of zeros at each symbol period to the PMA via the PMA_UNITDATA.request primitive.

If a PMA_TXMODE.indication message has the value SEND_T, PCS Transmit shall generate sequences of code-groups (TA_n, TB_n, TC_n, TD_n) defined in 113.3.4.2 to the PMA via the PMA_UNITDATA.request primitive. These code-groups are used for training mode and only transmit the values {–9, 9} to keep the transmit power in the training mode the same as the transmit power in normal mode.

During training mode an Infofield is transmitted at regular intervals containing messages for startup operation. By this mechanism, a PHY indicates the status of its own receiver to the link partner and makes requests for remote transmitter settings. (See 113.4.2.5.)

If a PMA_TXMODE.indication message has the value SEND_N, the PCS is in the normal mode of operation, and the PCS Transmit function shall use a 65B coding technique, transcoded to a mixed 513B-65B-RS-FEC-LDPC encoding to generate at each symbol period code-groups that represent data or control. During transmission, the six blocks of 513B transcoded bits and the two blocks of 65B encoded bits are scrambled by the PCS using a PCS scrambler and an auxiliary bit is added, then frames are encoded into a code-group of four-dimensional symbols and transferred to the PMA. During data encoding, PCS Transmit utilizes a Reed Solomon encoder to generate 3x512 bits and an LDPC frame encoder for the remaining 4x512 bits.

After reaching the normal mode of operation, EEE-capable PHYs may enter the LPI transmit mode under the control of the MAC via the 25GMII/XLGMII. The EEE Transmit state diagram is contained within the PCS Transmit function. The EEE capability is described in 113.3.2.2.23.

113.3.2.2.1 Use of blocks

The PCS maps 25GMII/XLGMII signals into 65-bit blocks inserted into an LDPC frame, and vice versa, using a 65B-LDPC coding scheme. The PAM2 PMA training frame synchronization allow establishment of LDPC frame and 65B boundaries by the PCS Synchronization process. Blocks and frames are unobservable and have no meaning outside the PCS. During the LPI mode, LDPC frame boundaries delimit sleep, wake, refresh, quiet, and alert cycles. The PCS functions ENCODE and DECODE generate, manipulate, and interpret blocks and frames as provided by the rules in 113.3.2.2.2.

113.3.2.2.2 65B-LDPC transmission code

The PCS uses a transmission code to improve the transmission characteristics of information to be transferred across the link and to support transmission of control and data characters. In addition, the code enables the receiver to achieve PCS synchronization alignment on the incoming PHY bit stream.

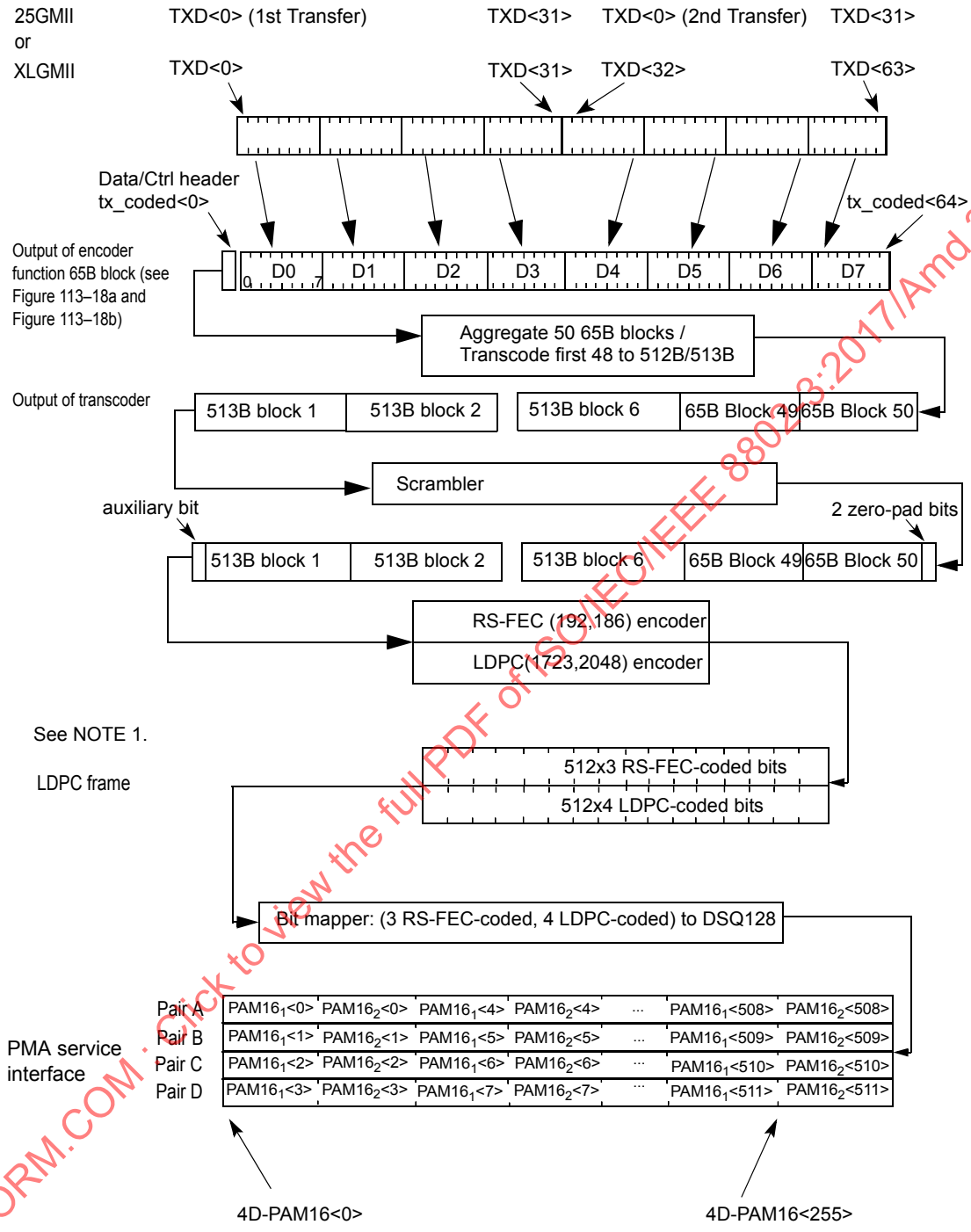
The relationship of block bit positions to 25GMII/XLGMII, PMA, and other PCS constructs is illustrated in Figure 113–6 for transmit and Figure 113–7 for receive. These figures illustrate the processing of a multiplicity of blocks containing 8 data octets. See 113.3.2.2.5 for information on how blocks containing control characters are mapped.

113.3.2.2.3 Notation conventions

For values shown as binary, the leftmost bit is the first transmitted bit.

64B/65B encodes 8 data octets or control characters into a block. Blocks containing control characters also contain a block type field. Data octets are labeled D_0 to D_7. Control characters other than /O/, /S/ and /T/ are labeled C_0 to C_7. The control character for ordered set is labeled as O_0 or O_4 since it is only valid on the first octet of the 25GMII/XLGMII. The control character for start is labeled as S_0 or S_4 for the same reason. The control character for terminate is labeled as T_0 to T_7.

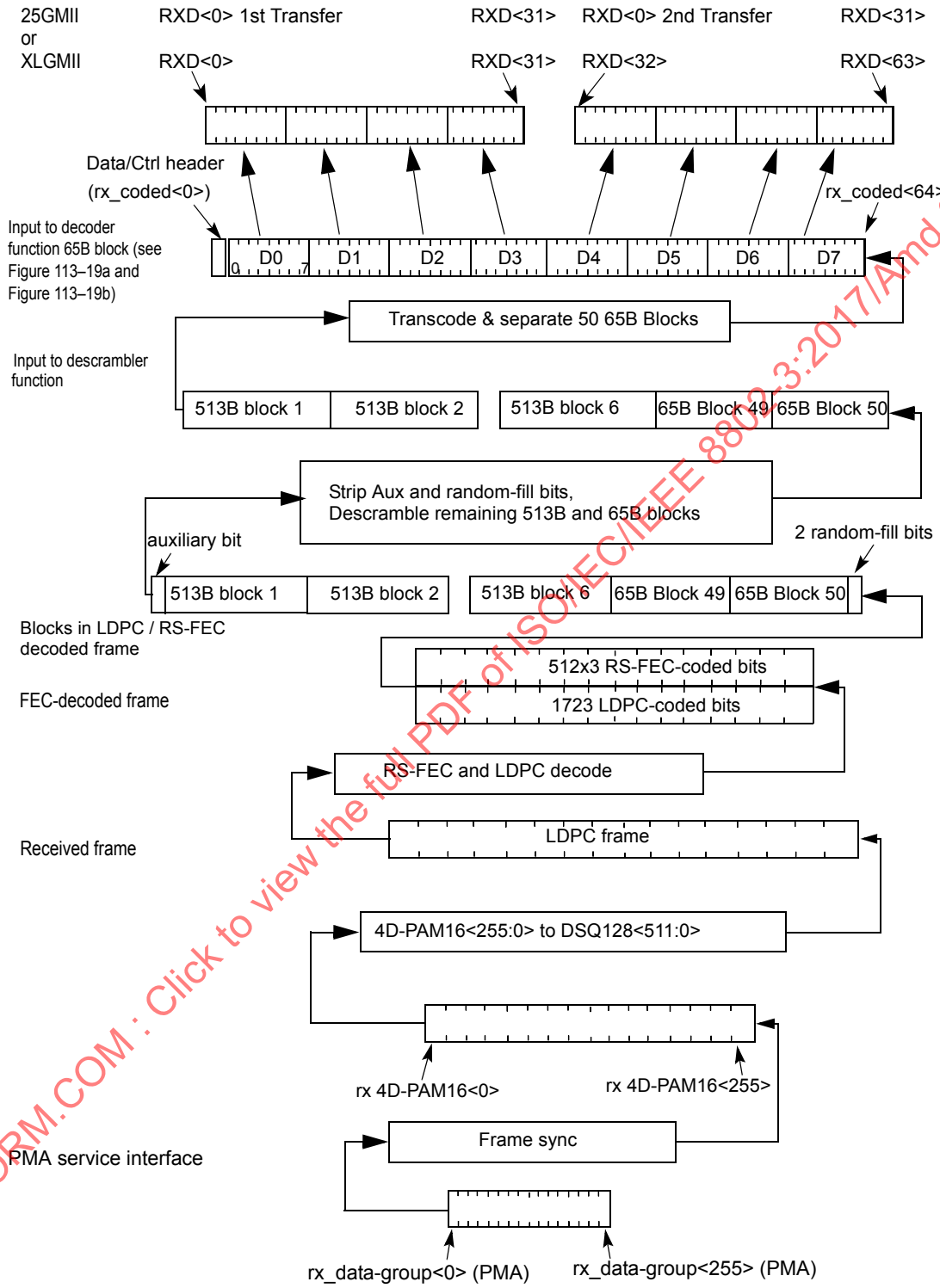
For 25GBASE-T, two 25GMII transfers or, for 40GBASE-T, a single XLGMII transfer provide eight characters that are encoded into one 65-bit transmission block. The subscript in the above labels indicates the position of the character in the eight characters from the 25GMII/XLGMII transfer(s).



NOTE 1—Zero-pad replaced with random bits for transmission.

NOTE 2—This figure shows the mapping from the 25GMII/XLGMII to a 64B/65B block for a block containing eight data characters.

Figure 113–6—PCS Transmit bit ordering



NOTE—This figure shows the mapping from a 64B/65B block a block containing eight data characters to the 25GMII/XLGMII.

Figure 113–7—PCS Receive bit ordering

Contents of block type fields, data octets, and control characters are shown as hexadecimal values. The LSB of the hexadecimal value represents the first transmitted bit. For instance, the block type field 0x1E is sent from left to right as 01111000. The bits of a transmitted or received block are labeled TxB<31:0> and RxB<31:0> for 25GBASE-T, or TxB<64:0> and RxB<64:0> for 40GBASE-T, respectively where TxB<0> and RxB<0> represent the first transmitted bit. The value of the data/ctrl header is shown as a binary value. Binary values are shown with the first transmitted bit (the LSB) on the left.

113.3.2.2.4 Transmission order

The PCS Transmit bit ordering is shown in Figure 113–6 and Figure 113–8. Note that Figure 113–6 shows the mapping from the 25GMII/XLGMII to a 64B/65B block for a block containing eight data characters.

113.3.2.2.5 Block structure

Blocks consist of 65 bits. The first bit of a block is the data/ctrl header. Blocks are either data blocks or control blocks. The data/ctrl header is 0 for data blocks and 1 for control blocks. The remainder of the block contains the payload.

Data blocks contain eight data characters. Control blocks begin with an eight-bit block type field that indicates the format of the remainder of the block. For control blocks containing a Start or Terminate character, that character is implied by the block type field. Other control characters are encoded in a seven-bit control code or a four-bit O Code. Each control block contains eight characters.

The format of the blocks for 25GBASE-T is as shown in Figure 113–9. The format of the blocks for 40GBASE-T is as shown in Figure 113–10. In the figure, the column labeled Input Data shows, in abbreviated form, the eight characters used to create the 65-bit block. These characters are either data characters or control characters and, when transferred across the 25GMII/XLGMII interface, the corresponding TXC or RXC bit is set accordingly. Within the Input Data column, D_0 through D_7 are data octets and are transferred with the corresponding TXC or RXC bit set to zero. All other characters are control octets and are transferred with the corresponding TXC or RXC bit set to one. The single bit fields (thin rectangles with no label in the figure) are sent as zero and ignored upon receipt..

Bits and field positions are shown with the least significant bit on the left. Hexadecimal numbers are shown prepended with ‘0x’, and with the least significant digit on the right. For example the block type field 0x1E is sent as 01111000 representing bits 1 through 8 of the 65-bit block. The least significant bit for each field is placed in the lowest numbered position of the field.

All unused values of block type field are reserved.[9]

113.3.2.2.6 Control codes

The same set of control characters are supported by the 25GMII/XLGMII and the 25G/40GBASE-T PCS. The representations of the control characters are the control codes. The 25GMII/XLGMII encodes a control character into an octet (an eight-bit value). The 25G/40GBASE-T PCS encodes the start and terminate control characters implicitly by the block type field. The 25G/40GBASE-T PCS encodes the ordered set control codes using a combination of the block type field and a four-bit O code for each ordered set. The 25G/40GBASE-T PCS encodes each of the other control characters into a seven-bit C code.

[9]The block type field values have been chosen to have a four-bit Hamming distance between them. The only unused value that maintains the Hamming distance is 0x00.

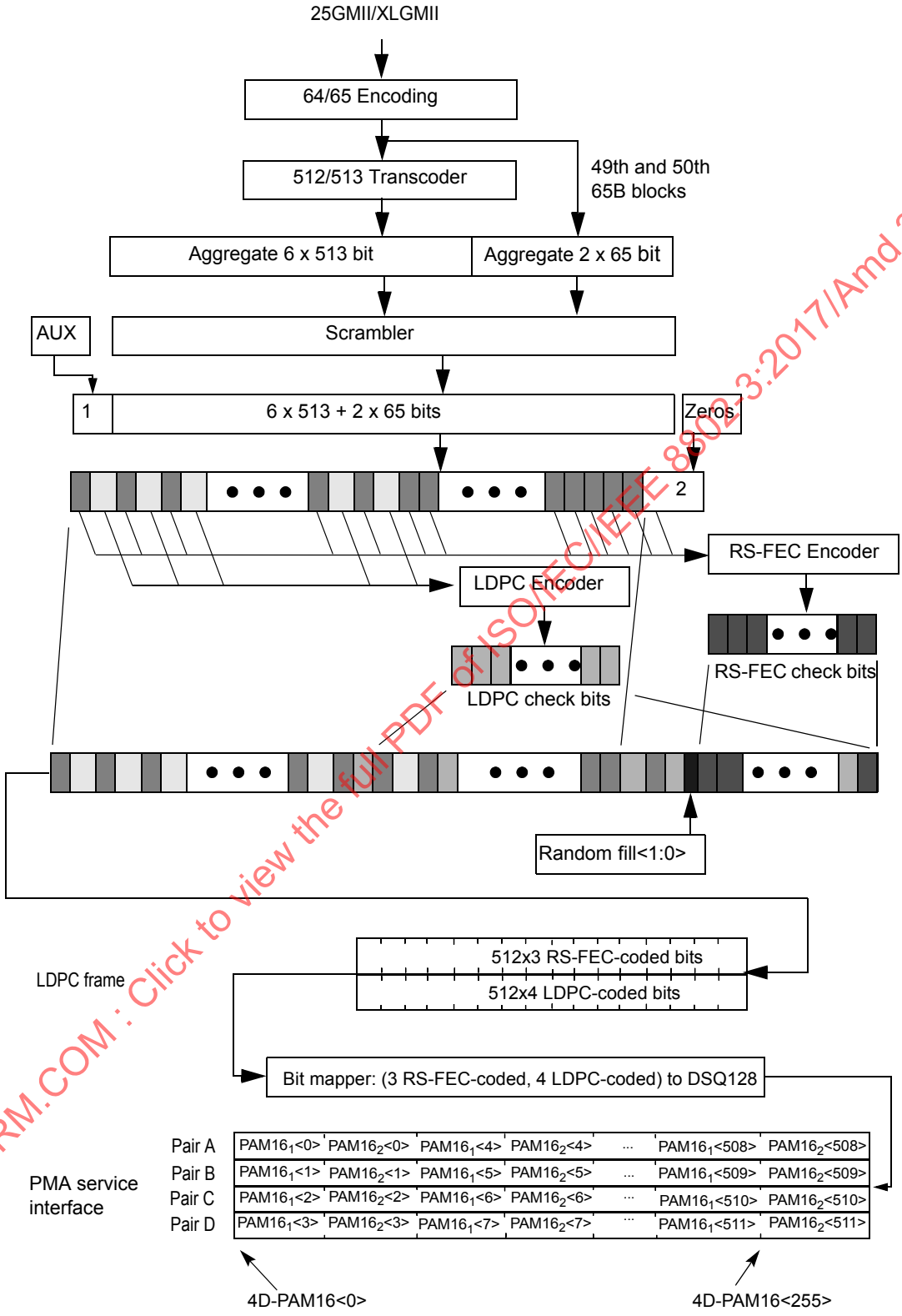


Figure 113–8—PCS detailed transmit bit ordering

Input Data	data ctrl header	Block Payload							
Bit Position: / Data Block Format:	0	1							64
D_0 D_1 D_2 D_3/D_4 D_5 D_6 D_7	0	D_0	D_1	D_2	D_3	D_4	D_5	D_6	D_7
Control Block Formats:		Block							
C_0 C_1 C_2 C_3/C_4 C_5 C_6 C_7	1	0x1E	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7
C_0 C_1 C_2 C_3/O_4 D_5 D_6 D_7	1	0x2D	C_0	C_1	C_2	C_3	O_4	D_5	D_6	D_7
C_0 C_1 C_2 C_3/S_4 D_5 D_6 D_7	1	0x33	C_0	C_1	C_2	C_3		D_5	D_6	D_7
O_0 D_1 D_2 D_3/S_4 D_5 D_6 D_7	1	0x66	D_1	D_2	D_3	O_0		D_5	D_6	D_7
O_0 D_1 D_2 D_3/O_4 D_5 D_6 D_7	1	0x55	D_1	D_2	D_3	O_0	O_4	D_5	D_6	D_7
S_0 D_1 D_2 D_3/D_4 D_5 D_6 D_7	1	0x78	D_1	D_2	D_3	D_4		D_5	D_6	D_7
O_0 D_1 D_2 D_3/C_4 C_5 C_6 C_7	1	0x4B	D_1	D_2	D_3	O_0	C_4	C_5	C_6	C_7
T_0 C_1 C_2 C_3/C_4 C_5 C_6 C_7	1	0x87		C_1	C_2	C_3	C_4	C_5	C_6	C_7
D_0 T_1 C_2 C_3/C_4 C_5 C_6 C_7	1	0x99	D_0		C_2	C_3	C_4	C_5	C_6	C_7
D_0 D_1 T_2 C_3/C_4 C_5 C_6 C_7	1	0xAA	D_0	D_1		C_3	C_4	C_5	C_6	C_7
D_0 D_1 D_2 T_3/C_4 C_5 C_6 C_7	1	0xB4	D_0	D_1	D_2		C_4	C_5	C_6	C_7
D_0 D_1 D_2 D_3/T_4 C_5 C_6 C_7	1	0xCC	D_0	D_1	D_2	D_3		C_5	C_6	C_7
D_0 D_1 D_2 D_3/D_4 T_5 C_6 C_7	1	0xD2	D_0	D_1	D_2	D_3	D_4		C_6	C_7
D_0 D_1 D_2 D_3/D_4 D_5 T_6 C_7	1	0xE1	D_0	D_1	D_2	D_3	D_4	D_5		C_7
D_0 D_1 D_2 D_3/D_4 D_5 D_6 T_7	1	0xFF	D_0	D_1	D_2	D_3	D_4	D_5	D_6	

Figure 113–9—64B/65B block formats for 25GBASE-T

The control characters and their mappings to 25G/40GBASE-T control codes and 25GMII/XLGMII control codes are specified in Table 113–1 for 25GBASE-T and Table 113–2 for 40GBASE-T. All 25GMII/XLGMII and 25G/40GBASE-T control code values that do not appear in the table shall not be transmitted and shall be treated as an error if received.

Input Data	data ctrl header	Block Payload							
Bit Position: / Data Block Format:	0	1							64
$D_0\ D_1\ D_2\ D_3/D_4\ D_5\ D_6\ D_7$	0	D_0	D_1	D_2	D_3	D_4	D_5	D_6	D_7
Control Block Formats:		Block							
$C_0\ C_1\ C_2\ C_3/C_4\ C_5\ C_6\ C_7$	1	0x1E	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7
$S_0\ D_1\ D_2\ D_3/D_4\ D_5\ D_6\ D_7$	1	0x78	D_1	D_2	D_3	D_4	D_5	D_6	D_7
$O_0\ D_1\ D_2\ D_3/Z_4\ Z_5\ Z_6\ Z_7$	1	0x4B	D_1	D_2	D_3	O_0	0x000_0000		
$T_0\ C_1\ C_2\ C_3/C_4\ C_5\ C_6\ C_7$	1	0x87		C_1	C_2	C_3	C_4	C_5	C_6	C_7
$D_0\ T_1\ C_2\ C_3/C_4\ C_5\ C_6\ C_7$	1	0x99	D_0		C_2	C_3	C_4	C_5	C_6	C_7
$D_0\ D_1\ T_2\ C_3/C_4\ C_5\ C_6\ C_7$	1	0xAA	D_0	D_1		C_3	C_4	C_5	C_6	C_7
$D_0\ D_1\ D_2\ T_3/C_4\ C_5\ C_6\ C_7$	1	0xB4	D_0	D_1	D_2		C_4	C_5	C_6	C_7
$D_0\ D_1\ D_2\ D_3/T_4\ C_5\ C_6\ C_7$	1	0xCC	D_0	D_1	D_2	D_3		C_5	C_6	C_7
$D_0\ D_1\ D_2\ D_3/D_4\ T_5\ C_6\ C_7$	1	0xD2	D_0	D_1	D_2	D_3	D_4		C_6	C_7
$D_0\ D_1\ D_2\ D_3/D_4\ D_5\ T_6\ C_7$	1	0xE1	D_0	D_1	D_2	D_3	D_4	D_5		C_7
$D_0\ D_1\ D_2\ D_3/D_4\ D_5\ D_6\ T_7$	1	0xFF	D_0	D_1	D_2	D_3	D_4	D_5	D_6	

Figure 113–10—64B/65B block formats for 40GBASE-T

Table 113–1—Control codes for 25GBASE-T

Control character	Notation	25GMII control codes	25GBASE-T control codes	25GBASE-T O code
idle	/I/	0x07	0x00	
LPI	/LI/	0x06	0x06	
start	/S/	0xFB	Encoded by block type field	
terminate	/T/	0xFD	Encoded by block type field	
error	/E/	0xFE	0x1E	
Sequence ordered set	/Q/	0x9C	Encoded by block type field plus O code	0x0
reserved0		0x1C	0x2D	reserved0
reserved1		0x3C	0x33	reserved1
reserved2		0x7C	0x4B	reserved2
reserved3		0xBC	0x55	reserved3
reserved4		0xDC	0x66	reserved4
reserved5		0xF7	0x78	reserved5
Signal ordered set[a]	/Fsig/	0x5C	Encoded by block type field plus O code	0xF

[a]Reserved for INCITS TA Fibre Channel use.

Table 113–2—Control codes for 40GBASE-T

Control character	Notation	XLGMII control codes	40GBASE-T control codes	40GBASE-T O code
idle	/I/	0x07	0x00	
LPI[a]	/LI/	0x06	0x06	
start	/S/	0xFB	Encoded by block type field	
terminate	/T/	0xFD	Encoded by block type field	

Table 113–2—Control codes for 40GBASE-T *(continued)*

Control character	Notation	XLGMII control codes	40GBASE-T control codes	40GBASE-T O code
error	/E/	0xFE	0x1E	
Sequence ordered set	/Q/	0x9C	Encoded by block type field plus O code	0x0
Signal ordered set[b]	/Fsig/	0x5C	Encoded by block type field plus O code	0xF

[a]Use of idle and LPI ordered sets per 81.3.
[b]Reserved for INCITS T11 Fibre Channel use.

113.3.2.2.7 Ordered sets

Ordered sets are used to extend the ability to send control and status information over the link such as remote fault and local fault status. Ordered sets consist of a control character followed by three data characters. Ordered sets always begin on the first octet of the 25GMII or XLGMII. 25 Gigabit and 40 Gigabit Ethernet use one kind of ordered set: the sequence ordered set (see 81.3.4). The sequence ordered set control character is denoted /Q/. An additional ordered set, the signal ordered set, has been reserved and it begins with another control code. The four-bit O field encodes the control code. See Table 113–1 for the mappings for 25GBASE-T, and Table 113–2 for the mappings for 40GBASE-T.

113.3.2.2.8 Idle (/I/)

Idle control characters (/I/) are transmitted when idle control characters are received from the 25GMII/XLGMII. Idle characters may be added or deleted by the PCS to adapt between clock rates. /I/ insertion and deletion shall occur in groups of 4. /I/s may be added following idle or ordered sets. They shall not be added while data is being received. When deleting /I/s, the first four characters after a /T/ shall not be deleted.

113.3.2.2.9 LPI (/LI/)

Low power idle (LPI) control characters (/LI/) on the 25GMII/XLGMII indicate that the LPI client is requesting operation in the LPI transmit mode. A continuous stream of LPI control characters (/LI/) is used to maintain a link in the LPI transmit mode. Idle control characters (/I/) are used to transition from the LPI transmit mode to the normal mode. PHYs that support EEE respond to the LPI 25GMII/XLGMII control characters using the procedure outlined in 113.1.3.3. LPI characters may be added or deleted by the PCS to adapt between clock rates. /LI/ insertion and deletion shall occur in groups of four. /LI/s may be added following low power idle characters. They shall not be added while data is being received.

If EEE is not supported, then /LI/ is not a valid control character.

113.3.2.2.10 Start (/S/)

The start control character (/S/) indicates the start of a packet. This delimiter is only valid on the first octet of the 25GMII (TXD<7:0> and RXD<7:0>), or XLGMII (TXD<7:0> and RXD<7:0>). Receipt of an /S/ on any other octet of TXD indicates an error. For 25 Gb/s and 40 Gb/s transmission, block type field values implicitly encode an /S/ as the first character of the block. These are the only characters of a block on which a start can occur.

113.3.2.2.11 Terminate (/T/)

The terminate control character (/T/) indicates the end of a packet. Since packets may be any length, the /T/ can occur on any octet of the 25GMII/XLGMII interface and within any character of the block. The location of the /T/ in the block is implicitly encoded in the block type field. A valid end of packet occurs when a block containing a /T/ is followed by a control block that does not contain a /T/.

113.3.2.2.12 ordered set (/O/)

The ordered set control characters (/O/) indicate the start of an ordered set. There are two kinds of ordered sets: the sequence ordered set and the signal ordered set (which is reserved). When it is necessary to designate the control character for the sequence ordered set specifically, /Q/ is used. /O/ is only valid on the first octet of the 25GMII/XLGMII. Receipt of an /O/ on any other octet of TXD indicates an error. For 40 Gb/s transmission, block type field values implicitly encode an /O/ as the first character of the block. The four-bit O code encodes the specific /O/ character for the ordered set.

Sequence ordered sets may be deleted by the PCS to adapt between clock rates. Such deletion shall only occur when two consecutive sequence ordered sets have been received and shall delete only one of the two. Only Idles may be inserted for clock compensation. Signal ordered sets are not deleted for clock compensation.

113.3.2.2.13 Error (/E/)

The /E/ is sent whenever an /E/ is received. The /E/ allows physical sublayers such as the PCS to propagate received errors. See R_BLOCK_TYPE and T_BLOCK_TYPE function definitions in 113.3.6.2.4 for further information.

113.3.2.2.14 Transmit process

The transmit process generates blocks based upon the TXD and TXC signals received from the 25GMII/XLGMII. 100 25GMII or 50 XLGMII data transfers are encoded into an LDPC frame. It takes 256 PMA_UNITDATA transfers to send an LDPC frame of data. Therefore, for 25GBASE-T, if the PCS is connected to a 25GMII and PMA sublayer where the ratio of their transfer rates is exactly 25:64, or, for 40GBASE-T, an XLGMII and PMA sublayer where the ratio of their transfer rates is exactly 25:128, then the transmit process does not need to perform rate adaptation. Where the 25GMII or XLGMII and PMA sublayer data rates are not synchronized to that ratio, the transmit process needs to insert idles, delete idles, or delete sequence ordered sets to adapt between the rates.

The transmit process generates blocks as specified in the PCS 64B/65B Transmit state diagram (see Figure 113–18a and Figure 113–18b). The contents of each block are contained in a vector tx_coded<64:0>, which is passed to the transcoder/ scrambler. tx_coded<0> contains the data/ctrl header and the remainder of the bits contain the block payload.

113.3.2.2.15 64B/65B to 512B/513B Transcoder

The 513B transcoder constructs a 513-bit block tx_xcoded<512:0>, from a group of eight 65-bit blocks, tx_coded_j<64:0> where j=0 to 7. For each group of eight 65-bit blocks, j=7 corresponds to the most recently received block. Bit 0 in each 65-bit block is the first bit received and corresponds to the data control header.

If for all j=0 to 7, tx_coded_j<0> = 0, tx_xcoded<512:0> shall be constructed as follows:

a) tx_xcoded<0>=1

b) tx_xcoded<(64j+64):(64j+1)>=tx_coded_j<64:1> for j=0 to 7.

If for any *j*=0 to 7, tx_coded_*j*<0>=1, tx_xcoded<512:0> shall be constructed as follows:

a1) tx_xcoded<0>=0

b1) Within the group of eight 65-bit blocks, let *C* be the set of *k* integers corresponding to the values of *j* that have tx_coded_*j*<0> = 1, and *U* be the set of 8–*k* integers corresponding to the values of *j* that have tx_coded_*j*<0> = 0, where the integers that comprise both *C* and *U* are arranged in ascending order. For instance, if tx_coded_1<0>=1 and tx_coded_4<0>=1, *C* = {1,4}, and *U* = {0,2,3,5,6,7}.

c1) For the allowed set of 25GBASE-T 64B/65B or 40GBASE-T 64B/65B control codes (see Figure 113–9 and Figure 113–10), let the four-bit code shown in the rightmost column of Table 113–4 represent the Block Type for transcoding purposes.

Given this, a 513-bit block can be constructed consisting of a leading 1, followed by all of the transcoded control blocks in *C*, followed by all of the data blocks in *U*.

Table 113–3 shows the transcoding scheme. The first column represents a 513B control byte that replaces the eight-bit block type found in the 64B/65B control code, and contains 3 fields:

a2) A continuation flag (FC) that if set to 1 indicates that another control block is to follow, and if set to 0 indicates that this is the last control block in the group of 8 transcoded 65B blocks, followed by

b2) Three position bits (Position) used to indicate which of the eight 65B blocks in the transcoding group the control block came from, followed by

c2) Four block-type bits (BlockType), which contain the transcoded block type, as shown in the right-most column of Table 113–4.

The resulting translation can be represented as the 513B control byte = {FC, Position<2:0>, BlockType<3:0>}.

Table 113–3—513B Transcoded Blocks including Control blocks (without leading 0)

513B control byte	64/65B block payload
513B Control Byte C_0 (1, Position<2:0>, Block Type<3:0>)	tx_coded_C_0<9:64>
513B Control Byte C_1 (1, Position<2:0>, Block Type<3:0>)	tx_coded_C_1<9:64>
:	:
513B Control Byte C_{k-1}(0, Position <2:0>, Block Type<3:0>)	tx_coded_C_{k-1}<9:64>
tx_coded_U_0<1:8>	tx_coded_U_0<9:64>
tx_coded_U_1<1:8>	tx_coded_U_1<9:64>
:	:
tx_coded_U_{7-k}<1:8>	tx_coded_U_{7-k}<9:64>

The following examples are provided for clarity, shown in Figure 113–11.

Example #1: *C* = {1,4}, and *U* = {0,2,3,5,6,7}, with the first control block being 0x1E, and the second being 0x78. Thus:

1.1) 65B control words are present, so the 513B control flag bit gets set to 0

Table 113–4—64B/65B Block field to BlockType<3:0> translation

Block field (see Figure 113–10)	BlockType<3:0>
0x1E	1000
0x2D	0100
0x33	1110
0x66	1101
0x55	1011
0x78	0111
0x4B	0001
0x87	1100
0x99	1010
0xAA	1001
0xB4	0101
0xCC	0011
0xD2	0110
0xE1	0000
0xFF	1111

1.2) The first control word is C_0 where Position = 0x1, and BlockType = 0x8. Since this is not the last control word the continuation flag FC = 1. Thus the 513B control word for this block is:
C_0 Control Word = {1,0x1, 0x8} = 1 100 0001 in bit order of transmission

1.3) The second control word is C_4 where Position = 0x4, and BlockType = 0x7. Since this is the last control word the continuation flag FC = 0. Thus the 513B control word for this block is:
C_4 Control Word = {0,0x4, 0x7} = 0 001 1110 in bit order of transmission

1.4) After this the payload of the remaining data blocks is placed

Example #2: C = {7}, and U = {0,1,2,3,4,5,6}, with the control block being 0xB4. Thus:

2.1) 65B control words are present, so the 513B control flag bit gets set to 0

2.2) The first and only control word is C_0 where Position = 0x7, and BlockType = 0x5. Since this is also the last control word the continuation flag FC = 0. Thus the 513B control word for this block is:
C_0 Control Word = {0,0x7, 0x5} = 0 111 1010 in bit order of transmission

2.3) After this the payload of the remaining data blocks is placed

Example #3: C = {}, and U = {0,1,2,3,4,5,6,7}. Thus:

3.1) No 65B control words are present, so the 513B control flag bit gets set to 1

3.2) After this the payload of all of the data blocks is placed

0	1100 000	C0	C1	C2	C3	C4	C5	C6	C7	65B #1
	0001 111	D1	D2	D3	D4	D5	D6	D7		65B #4
	D0	D1	D2	D3	D4	D5	D6	D7		65B #0
	D0	D1	D2	D3	D4	D5	D6	D7		65B #2
	D0	D1	D2	D3	D4	D5	D6	D7		65B #3
	D0	D1	D2	D3	D4	D5	D6	D7		65B #5
	D0	D1	D2	D3	D4	D5	D6	D7		65B #6
	D0	D1	D2	D3	D4	D5	D6	D7		65B #7

Example 1: *C*={1,4}, *U*={0,2,3,5,6,7}, 1st Control block = 0x1E, 2nd = 0x78

0	0111 101	C0	C1	C2	C3	C4	C5	C6	C7	65B #7
	D0	D1	D2	D3	D4	D5	D6	D7		65B #0
	D0	D1	D2	D3	D4	D5	D6	D7		65B #1
	D0	D1	D2	D3	D4	D5	D6	D7		65B #2
	D0	D1	D2	D3	D4	D5	D6	D7		65B #3
	D0	D1	D2	D3	D4	D5	D6	D7		65B #4
	D0	D1	D2	D3	D4	D5	D6	D7		65B #5
	D0	D1	D2	D3	D4	D5	D6	D7		65B #6

Example 2: Control block 0xB4, *C*={7}, *U*={0,1,2,3,4,5,6}

1	D0	D1	D2	D3	D4	D5	D6	D7	65B #0
	D0	D1	D2	D3	D4	D5	D6	D7	65B #1
	D0	D1	D2	D3	D4	D5	D6	D7	65B #2
	D0	D1	D2	D3	D4	D5	D6	D7	65B #3
	D0	D1	D2	D3	D4	D5	D6	D7	65B #4
	D0	D1	D2	D3	D4	D5	D6	D7	65B #5
	D0	D1	D2	D3	D4	D5	D6	D7	65B #6
	D0	D1	D2	D3	D4	D5	D6	D7	65B #7

Example 3: All data blocks

Figure 113–11—Examples of the construction of tx_xcoded

113.3.2.2.16 Aggregation

The six 512B/513B transcoded blocks and two 64B/65B blocks are concatenated to form tx_aggregated<3207:0>. The tx_xcoded_*j*<512:0> is mapped into tx_aggregated<513*j*+512:513*j*> where j=0 to 5. tx_xcoded_0<512:0> is the first to be transmitted. The two 64B/65B blocks are mapped into tx_aggregated<3142:3078> and tx_aggregated<3207:3143>.

113.3.2.2.17 PCS Scrambler

The payload of the PCS PHY frame tx_aggregated<3207:0> is scrambled to tx_scrambled<3207:0> with a self-synchronizing scrambler. The scrambler for the MASTER shall produce the same result as the

implementation shown in Figure 113–12. This implements the scrambler polynomial[10] as shown in Equation (113–1):

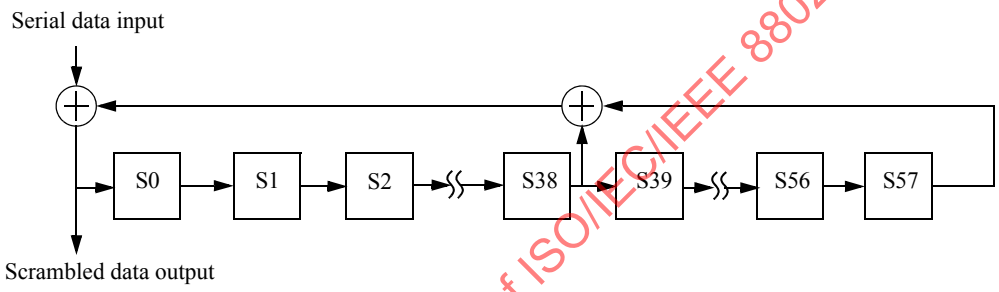
$$G(x) = 1 + x^{39} + x^{58} \quad (113–1)$$

The scrambler for the SLAVE shall produce the same result as the implementation shown in Figure 113–12. This implements the scrambler polynomial as shown in Equation (113–2):

$$G(x) = 1 + x^{19} + x^{58} \quad (113–2)$$

The initial seed values for the MASTER and SLAVE are left to the implementer. The scrambler is run continuously on all payload bits.

PCS scrambler employed by the MASTER



PCS scrambler employed by the SLAVE

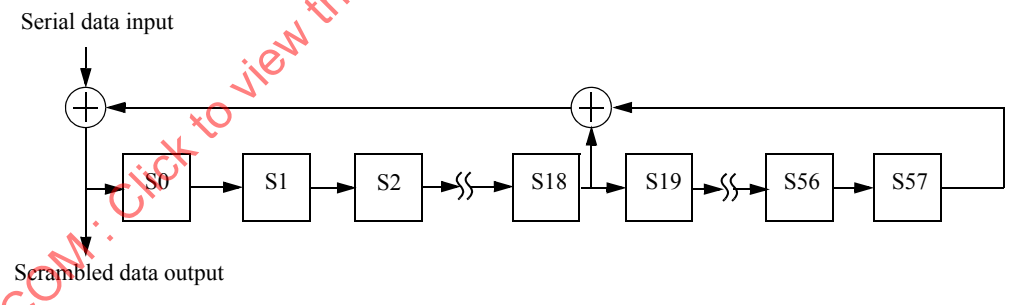


Figure 113–12—MASTER and SLAVE PCS scramblers

[10]The convention here, which considers the most recent bit into the scrambler to be the lowest order term, is consistent with most references and with other scramblers shown in this standard. Some references consider the most recent bit into the scrambler to be the highest order term and would therefore identify this as the inverse of the polynomial in Equation (113–1). In case of doubt, note that the conformance requirement is based on the representation of the scrambler in the figure rather than the polynomial equation.

113.3.2.2.18 LDPC framing and LDPC encoder

The resulting payload of scrambled six 513B transcoded blocks plus two 65B blocks, followed by two random fill bits, and six eight-bit Reed Solomon parity symbols and preceded by 1 auxiliary bit results in a total payload of 6 × 513 + 2 × 65 + 2+ 6× 8+ 1 = 3259 bits. The use of the auxiliary bit for vendor-specific communication is outside the scope of this document. It is highly recommended that the auxiliary bit be randomized. For the purposes of this standard it is ignored by the link partner, as are the random fill bits. From the total payload of 3259 bits, 1536 (3 bits for each of the 512 DSQ128 symbols) are RS-FEC-coded bits and the remaining 1723 shall be encoded by the LDPC(1723, 2048) generator matrix G. G is described in Annex 55A. See Figure 113–8 and 113.3.2.2.19 for details on PCS bit ordering and RS-FEC encoding.

The LPDC encoding takes the 1723-bit input code vector x = [x_0 x_1 x_2 ... x_{1722}], and shall generate the 2048-bit codeword c represented by the matrix multiplication c = x × G. For both x and c the encoder shall follow the notation described in 113.3.2.2.3 where the LSB (leftmost element of the vectors x and c) is the first bit into the LDPC encoder and the first transmitted bit.

113.3.2.2.19 Reed Solomon encoder

The group of 1536 bits are encoded using a Reed-Solomon encoder operating over the Galois Field GF(2^8) where the symbol size is 8 bits. The encoder processes k=186 eight-bit RS-FEC message symbols to generate 2t=6 eight-bit RS-FEC parity symbols, which are then appended to the message to produce a codeword of n=k+2t=192 eight-bit RS-FEC symbols. For the purposes of this clause, the Reed-Solomon code is denoted RS-FEC(n,k), and the particular Reed-Solomon code is RS-FEC(192,186).

The code is based on the generating polynomial given by Equation (113–3):

$$g(x) = \prod_{j=0}^{5} (x - \alpha^j) = g_6x^6 + g_5x^5 + g_4x^4 + g_3x^3 + g_2x^2 + g_1x + g_0 \quad (113–3)$$

In Equation (113–3), α, is a primitive element of the finite field defined by the primitive polynomial 0x11D = $x^8+x^4+x^3+x^2+1$.

Equation (113–4) defines the message polynomial m(x) whose coefficients are the message symbols m_{185} to m_0:

$$m(x) = m_{185}x^{191} + m_{184}x^{190} + \ldots + m_1x^7 + m_0x^6 \quad (113–4)$$

Each message symbol m_i is the bit vector ($m_{i,7}$, $m_{i,6}$, ... $m_{i,1}$,$m_{i,0}$) that is identified with the element of the finite field. The message symbols are composed of the bits in tx_RSmessage<1487:0> where

$m_{i,j}$=tx_RSmessage<(185-i)8 + j>, i=0 to 185, j=0 to 7.

tx_RSmessage<1487:0> is formed as follows:

tx_RSmessage<0> = Auxiliary bit
tx_RSmessage<2:1>=tx_scrambled<1:0>
tx_RSmessage<1292:3> = tx_RSmessage<3j+5:3j+3> = tx_scrambled<8+7j:6+7j> where j = 429 to 0
tx_RSmessage<1485:1293> = tx_scrambled<3207:3015>
tx_RSmessage<1487:1486> = 00 (It is highly recommended that two random fill bits be transmitted instead of zeros, and then this information is discarded upon receipt)

The first symbol input to the encoder is m_{185}.

tx_scrambled<3207:0> is defined in 113.3.2.2.17.

Equation (113–5) defines the parity polynomial *p(x)* whose coefficients are the parity symbols p_5 to p_0:

$$p(x) = p_5x^5 + p_4x^4 + p_3x^3 + p_2x^2 + p_1x + p_0 \quad (113–5)$$

The parity polynomial is the remainder from the division of m(x) by g(x). This may be computed using the shift register implementation illustrated in Figure 113–13. The outputs of the delay elements are initialized to zero prior to the computation of the parity for a given message. After the last message symbol, m_0, is processed by the encoder, the outputs of the delay elements are the parity symbols for that message.

The codeword polynomial c(x) is then the sum of m(x) and p(x) where the coefficient of the highest power of x, $c_{191} = m_{185}$ is transmitted first and the coefficient of the lowest power of *x*, $c_0 = p_0$ is transmitted last. The first bit transmitted from each symbol is bit 0.

The coefficients of the generator polynomial for the code are presented in Table 113–5.

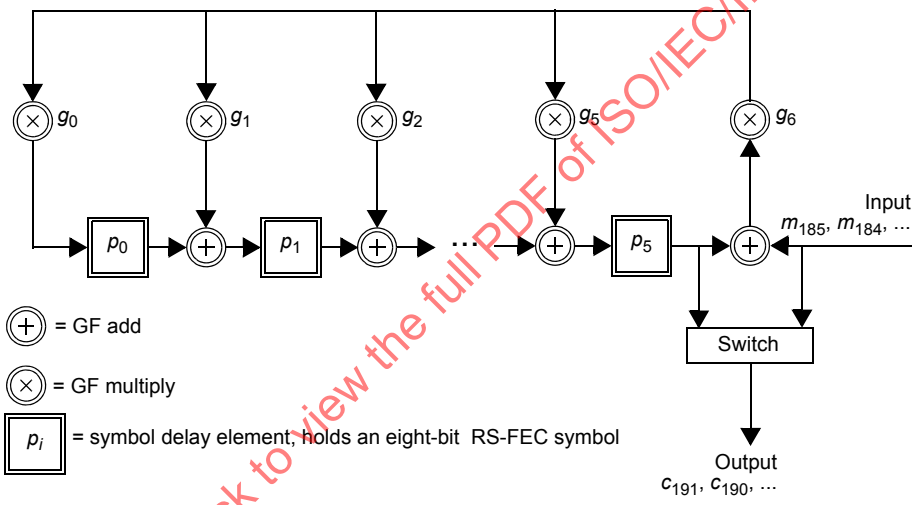


Figure 113–13—Reed-Solomon encoder functional model

113.3.2.2.20 DSQ128 bit mapping

DSQ128 refers to a two-dimensional constellation with 128 possible 2D values, such that the combined 2D symbol carries $\log_2(128)$ or 7 bits. The PHY frame of 1536 RS-FEC-coded bits and 2048 (LDPC output) LDPC-coded bits described in Figure 113–6 shall be partitioned into seven-bit groups of (3 RS-FEC-coded, 4 LDPC-coded) as described in Figure 113–8. The bit partition is as follows, the first 1290 (3 × 430) RS-FEC-coded bits are paired with the first 1720 (4 × 430) LDPC input bits. The following 3 RS-FEC-coded bits are paired with the last 3 LDPC input bits and the first LDPC parity bit. Finally the remaining 243 (3 × 81) RS-FEC-coded bits are paired with the remaining 324 (4 × 81) LDPC parity bits.

Table 113–5—Coefficients of the generator polynomial g_i (decimal)

i	RS-FEC(192, 186)
0	38
1	227
2	32
3	218
4	1
5	63
6	1

Each 2D-DSQ128 value has two PAM16 components denoted $PAM16_1$ and $PAM16_2$ respectively. The DSQ128 can be constructed by pruning the 256 values of a 2D-PAM16 where every other point in 2D is discarded (like the black or white squares in a checkerboard). The PAM16 components $PAM16_1$ and $PAM16_2$ can each take any of the values from the set {–15, –13, –11, –9, –7, –5, –3, –1, 1, 3, 5, 7, 9, 11, 13, 15}. The mapping from 7 bits where u_0 u_1 u_2 denote the 3 RS-FEC-coded bits and c_0 c_1 c_2 c_3 denote the 4 LDPC-coded bits to the DSQ128 is described by the following four steps (the bits from the scrambler output shall be read LSB first):

Step 1:

$x_{13} = (!u_0)\ \&\ u_2$

$x_{12} = u_0\ \text{XOR}\ u_2$

$x_{11} = c_0$

$x_{10} = c_0\ \text{XOR}\ c_1$

$x_{23} = (u_1\ \&\ u_2)\ \text{OR}\ (u_0\ \&\ !u_1)$

$x_{22} = u_1\ \text{XOR}\ u_2$

$x_{21} = c_2$

$x_{20} = c_2\ \text{XOR}\ c_3$

Step 2:

$x_1 = 8x_{13} + 4x_{12} + 2x_{11} + x_{10}$

$x_2 = 8x_{23} + 4x_{22} + 2x_{21} + x_{20}$

Step 3:

$y_1 = (x_1 + x_2)\ \text{mod}16$

$y_2 = (-x_1 + x_2)\ \text{mod}16$

Step 4:

$$PAM16_1 = 2y_1 - 15$$

$$PAM16_2 = 2y_2 - 15$$

where '*n* mod16' for an integer *n*, is defined as the integer value *p* in the range 0 to 15 (both inclusive) such that '$p = n + 16m$', for some integer *m*.

113.3.2.2.21 DSQ128 to 4D-PAM16

The DSQ mapper generates 512 2D-DSQ128 symbols per LDPC frame that are mapped onto 256 4D-PAM16 symbols prior to sending to the PMA via PMA_UNITDATA.request. The mapping of DSQ128 to 4D-PAM16 is illustrated in Figure 113–6. As shown in Figure 113–6, the two PAM16 components of each DSQ128 symbol are mapped onto two consecutive time periods on the same wire pair.

113.3.2.2.22 Block-LDPC framer

The Block-LDPC framer adapts between the mixed 513B and 65B blocks and the 4D-PAM16 width of the PMA. When the transmitter is operating in normal mode, the 65B-LDPC sends four PAM16 of transmit data at a time via PMA_UNITDATA.request primitives. The PMA_UNITDATA.request primitives are fully packed with bits.

113.3.2.2.23 EEE capability

The optional 25G/40GBASE-T EEE capability allows compliant PHYs to transition to an LPI mode of operation when link utilization is low.

PHYs that support EEE shall implement the EEE transmit state diagram, shown in Figure 113–20, within the PCS.

When PCS_Reset is asserted or pcs_data_mode is not asserted, the state diagram enters the TX_NORMAL state.

When a complete 64B/65B block of LPI characters is generated by the PCS transmit function, the PHY transmits the sleep signal to indicate to the link partner that it is transitioning to the LPI transmit mode. If the sleep signal begins on an LDPC frame boundary, then it contains 6 full LDPC frames each composed entirely of RS-FEC and LDPC encoded LP_IDLE blocks. If the sleep signal does not begin on an LDPC frame boundary, then it contains one LDPC frame partially composed of LP_IDLE blocks followed by 6 LDPC frames fully composed of LP_IDLE blocks.

Following the transmission of the sleep signal, quiet-refresh signaling begins, as described in 113.3.5.

After the sleep signal is transmitted LPI control characters shall be input to the PCS scrambler continuously until the PCS Transmit Function exits the LPI transmit mode.

While the PMA asserts SEND_N, the lpi_tx_mode variable shall control the transmit signal through the PMA_UNITDATA.request primitive described as follows:

When the PHY is not in the PCS_Data state, the lpi_tx_mode variable is ignored.

When the lpi_tx_mode variable takes the value NORMAL and the PMA asserts SEND_N, the PCS passes coded data to the PMA via the PMA_UNITDATA.request primitive as described in 113.3.2.2.

When the lpi_tx_mode variable takes the value QUIET and the PMA asserts SEND_N, the PCS passes zeros to the PMA through the PMA_UNITDATA.request primitive.

When the lpi_tx_mode variable takes the value REFRESH_A and the PMA asserts SEND_N, the PCS passes the PMA training signal to the PMA on pair A, to allow both the local and remote PHY to refresh adaptive filters and timing loops. The PCS passes zeros to all other pairs in this condition. REFRESH_B, REFRESH_C and REFRESH_D operate in an analogous manner for the other pairs.

When the lpi_tx_mode variable takes the value ALERT and the PMA asserts SEND_N, the PCS passes the ALERT vector to the PMA.

The quiet-refresh cycle is repeated until codewords other than LP_IDLE are detected at the 25GMII/XLGMII. These codewords indicate that the local system is requesting a transition back to the normal operational mode. Following this event, the PMA_UNITDATA.request parameter tx_symb_vector is set to the value ALERT. The alert signal is not synchronized with respect to the quiet-refresh cycle but shall be synchronized so that the alert signal from the PMA begins on a LDPC frame boundary.

The PHY will also transition back to the normal operation mode if an error condition occurs. This error condition is defined as the detection of any characters other than LPI or IDLE at the 25GMII/XLGMII.

After the alert signal the PCS completes the transition from LPI mode to normal mode by sending a wake signal containing lpi_wake_time LDPC frames composed of IDLE 64B/65B blocks.

lpi_wake_time is a fixed parameter that is defined as 9 LDPC frames as shown in Table 113–6. The maximum PHY wake time when wake is requested before sleep has been completely transmitted is 1.6 µs (lpi_wake_timer=T_{w_phy} as defined by Clause 78). The maximum PHY wake time when wake is requested after sleep has been completely transmitted is 1.12 µs.

Table 113–6—LPI wake time

lpi_wake_time	lpi_wake_timer when wake starts before sleep signal is complete		lpi_wake_timer when wake starts after sleep signal is complete	
(frames)	(frames)	(µs)	(frames)	(µs)
9	20	1.6	14	1.12

113.3.2.3 PCS Receive function

The PCS Receive function shall conform to the PCS 64B/65B receive state diagram in Figure 113–19a and Figure 113–19b and the PCS Receive bit ordering in Figure 113–7 including compliance with the associated state variables as specified in 113.3.6.

The PCS Receive function accepts received code-groups provided by the PMA Receive function via the parameter rx_symb_vector. The PCS receiver uses knowledge of the encoding rules to correctly align the 65B-LDPC frames. The received 65B-LDPC frames are decoded with error correction on both the RS-FEC and LDPC codes, and framing is checked; the six 512B/513B blocks are transcoded to 64B/65B. This process generates the 64B/65B block vector rx_coded<64:0> that is then decoded to form the 25GMII signals RXD<31:0> and RXC<3:0> for 25GBASE-T or the XLGMII signals RXD<63:0> and RXC<7:0> for 40GBASE-T, as specified in the PCS 64B/65B Receive state diagram (see Figure 113–19a and Figure 113–19b). 100 25GMII or 50 XLGMII data transfers are decoded from one LDPC frame. Where the 25GMII or XLGMII and PMA sublayer data rates are not synchronized in a 25:64 ratio or 25:128 ratio,

respectively, the receive process inserts idles, deletes idles, or deletes sequence ordered sets to adapt between rates.

During PMA training mode, PCS Receive checks the received PAM2 framing and signals the reliable acquisition of the descrambler state by setting the scr_status parameter of the PMA_SCRSTATUS.request primitive to OK.

When the PCS Synchronization process has obtained synchronization, the LDPC frame error ratio (LFER) monitor process monitors the signal quality asserting hi_lfer if excessive LDPC frame errors are detected (LDPC parity error or uncorrectable RS-FEC error). If 40 consecutive LDPC frame errors are detected, the block_lock flag is de-asserted. When block_lock is asserted and hi_lfer is de-asserted, the pcs_status parameter of the PMA_PCSSTATUS.request primitive is set to OK, and the PCS Receive process continuously accepts blocks. The PCS Receive process monitors these blocks and generates RXD and RXC on the 25GMII/XLGMII.

When the receiver is in training mode, the PCS Synchronization process continuously monitors PMA_RXSTATUS.indication (loc_rcvr_status). When loc_rcvr_status indicates OK, then the PCS Synchronization process accepts data-units via the PMA_UNITDATA.request primitive. It attains frame and block synchronization based on the PMA training frames and conveys received blocks to the PCS Receive process. The PCS Synchronization process sets the block_lock flag to indicate whether the PCS has obtained synchronization. The PMA training sequence includes one-bit pattern on pair A every 256 PAM2 symbols, which is aligned with the PCS PHY frame boundary. When the PCS Synchronization process is synchronized to this pattern, block_lock is asserted.

PHYs with the EEE capability support transition to the LPI mode when the PHY has successfully completed training and pcs_data_mode is TRUE. Transitions to and from the LPI mode are allowed to occur independently in the transmit and receive functions. The PCS receive function is responsible for detecting transitions to and from the LPI receive mode and indicating these transitions using signals defined in 113.2.2.

The link partner signals a transition to the LPI mode of operation by transmitting 6 LDPC frames composed entirely of 64B/65B blocks of /LI/. When blocks of /LI/ are detected at the output of the 64B/65B decoder, rx_lpi_active is asserted by the PCS receive function and the /LI/ character is continuously asserted at the receive 25GMII/XLGMII. These frames may be preceded by a frame composed partially of /LI/ characters. After these frames the link partner begins transmitting zeros, and it is recommended that the receiver power down receive circuits to reduce power consumption. The receive function uses LDPC frame counters to maintain synchronization with the remote PHY and receives periodic refresh signals that are used to update coefficients, so that the integrity of adaptive filters and timing loops in the PMA is maintained. LPI signaling is defined in 113.3.5. The quiet-refresh cycle continues until the PMA asserts alert_detect to indicate that the alert signal has been reliably detected. After the alert signal the link partner transmits repeated /I/ characters, representing a wake signal. The PHY receive function sends /I/ to the 25GMII/XLGMII for 9 LDPC frame periods and then resumes normal operation.

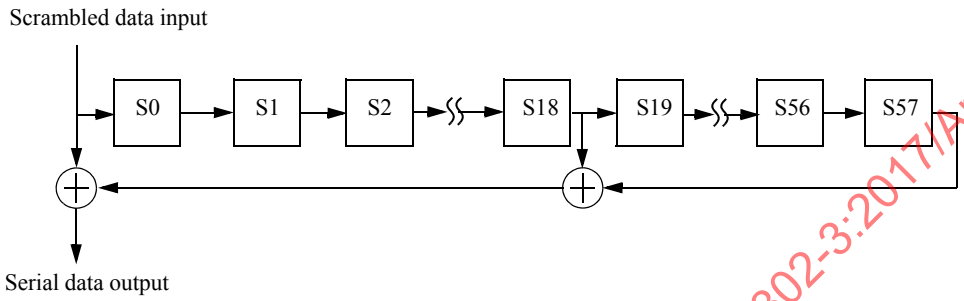
113.3.2.3.1 Frame and block synchronization

When the receiver is operating in normal mode, the frame and block synchronization function receives data via 4D-PAM16 PMA_UNITDATA.request primitives. It shall form a 4D-PAM16 stream from the primitives by concatenating requests with the PAM16s of each primitive in order from rx_data-group<0> to rx_data-group<255> (see Figure 113–7). It obtains block_lock to the LDPC frames during the PAM2 training pattern using synchronization bits provided on pair A. The 65-bit blocks are extracted based on their location in the LDPC frame.

113.3.2.3.2 PCS descrambler

The descrambler processes the payload to reverse the effect of the scrambler using the same polynomial. It shall produce the same result as the implementations shown in Figure 113–14 for the MASTER and the SLAVE.

PCS descrambler employed by the MASTER



PCS descrambler employed by the SLAVE

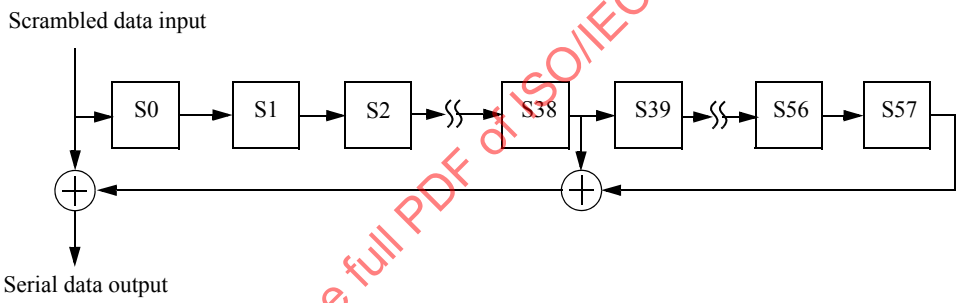


Figure 113–14—MASTER and SLAVE PCS descramblers

113.3.2.3.3 Invalid blocks

A block is invalid if any of the following conditions exists:

a) The block type field contains a reserved value.
b) For 25GBASE-T, any control character contains a value not in Table 113–1, or Table 113–2 for 40GBASE-T.
c) For 25GBASE-T, any O code contains a value not in Table 113–1, or Table 113–2 for 40GBASE-T.
d) The block contains information from the payload of an invalid received PHY frame or the first 64B/65B block following an invalid received PHY frame.

The PCS Receive function shall check the integrity of the LDPC and RS-FEC parity bits defined in 113.3.2.2.18 and 113.3.2.2.19, respectively. If either check fails the PHY frame is invalid.

R_BLOCK_TYPE of an invalid block is set to E.

113.3.3 Test-pattern generators

The test-pattern generator mode is provided for enabling joint testing of the local transmitter, link segment, and remote receiver. When the transmit PCS is operating in test-pattern mode it shall transmit continuously

as illustrated in Figure 113–6, with the input to the scrambler set to zero and the initial condition of the scrambler set to any non-zero value. When the receiver PCS is operating in test-pattern mode it shall receive continuously as illustrated in Figure 113–7. After acquiring the self-synchronizing scrambler state, the output of the received scrambled values should ideally be zero. Any nonzero values correspond to receiver bit errors. This mode is further described as test mode 7 in 113.5.2.

113.3.4 PMA training side-stream scrambler polynomials

The PCS Transmit function employs side-stream scrambling for generating 2-level PAM PMA training sequences as shown in Figure 113–15. An implementation of MASTER and SLAVE PHY side-stream scramblers is shown in the "Main PN sequence" box. The bits stored in the shift register delay line at time n are denoted by $Scr_n[32:0]$. At each symbol period, the shift register is advanced by one bit, and one new bit represented by $Scr_n[0]$ is generated. The transmitter side-stream scrambler is reset upon execution of the PCS Reset function. If PCS Reset is executed, all bits of the 33-bit vector representing the side-stream scrambler state are arbitrarily set. The initialization of the scrambler state is left to the implementer. In no case shall the scrambler state be initialized to all zeros.

113.3.4.1 Generation of bits Sa_n, Sb_n, Sc_n, Sd_n

PMA training signal encoding rules are based on the generation, at time n, of the four bits Sa_n, Sb_n, Sc_n, Sd_n. These four bits are generated in a systematic fashion using the bits in $Scr_n[32:0]$, and an auxiliary generating polynomial. For both MASTER and SLAVE PHYs, they are obtained by the same linear combinations of bits stored in the transmit scrambler shift register delay line. These four bits are derived from elements of the same maximum-length shift register sequence of length $2^{33}–1$ as $Scr_n[0]$, but shifted in time. The associated delays are all large and different so that there is no short-term correlation among the bits Sa_n, Sb_n, Sc_n, Sd_n. The four bits are generated using the bit $Scr_n[0]$ and the equations in Figure 113–15 in the "Derived sequences" box.

113.3.4.2 Generation of 4D symbols TA_n, TB_n, TC_n, TD_n

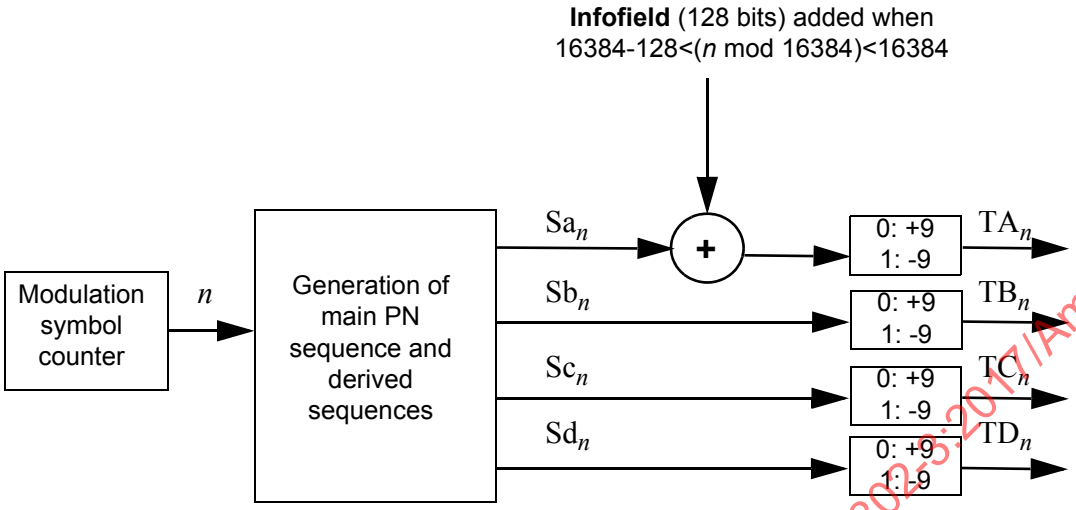
The four bits Sa_n, Sb_n, Sc_n, Sd_n are mapped to a 4D symbol (TA_n, TB_n, TC_n, TD_n) as shown in Figure 113–15.

The inversion on pair A at 256 intervals ($n = k \times 256$, $k = 0, 1, 2, \ldots$) defines the LDPC boundary during data mode.

Notice that over the repeating time intervals of 16384 and of length 128, $m \times 16384 - 128 \le n < m \times 16384$, $m = 1, 2, 3, \ldots$, the PMA training pattern in pair A is XOR'ed with the Infofield. Thus, pair A transmits the Infofield, which communicates to the remote transceiver settings of THP and power backoff and other control information.

113.3.4.3 PMA training mode descrambler polynomials

The PHY shall acquire descrambler state synchronization to the PAM2 training sequence and report success through scr_status. For side-stream descrambling, the MASTER PHY shall employ the receiver descrambler generator polynomial $g'_M(x) = 1 + x^{20} + x^{33}$ and the SLAVE PHY shall employ the receiver descrambler generator polynomial $g'_S(x) = 1 + x^{13} + x^{33}$.



Main PN sequence:

$$Scr_n[32:1] = Scr_{n-1}[31:0]$$

$$Scr_n[0] = \begin{cases} Scr_{n-1}[12]+Scr_{n-1}[32] \text{ if PMA\_CONFIG=MASTER} \\ Scr_{n-1}[19]+Scr_{n-1}[32] \text{ if PMA\_CONFIG = SLAVE} \end{cases}$$

Derived sequences:

$$Sa_n = \begin{cases} Scr_n[0] \oplus 1 & \text{if } n \bmod 256 = 0 \\ Scr_n[0] & \text{otherwise} \end{cases}$$

$$Sb_n = Scr_n[3] \oplus Scr_n[8]$$

$$Sc_n = Scr_n[6] \oplus Scr_n[16]$$

$$Sd_n = Scr_n[9] \oplus Scr_n([14] \oplus Scr_n([19] \oplus Scr_n[24]))$$

Figure 113–15—A realization of PMA training PAM2 sequences

113.3.5 LPI signaling

PHYs with EEE capability have transmit and receive functions that can enter and leave the LPI mode independently. The PHY can transition to the LPI mode when the PHY has successfully completed training and pcs_data_mode is TRUE. The transmit function of the PHY initiates a transition to the LPI transmit mode when it generates 64B/65B blocks composed entirely of LPI control characters, as described in 113.3.2.2.23. The transmit function of the link partner signals the transition using the sleep signal. When the transmitter begins to send the sleep signal, it asserts tx_lpi_active and the transmit function enters the LPI transmit mode.

Within the LPI mode PHYs use a repeating quiet-refresh cycle (see Figure 113–16). The first part of this cycle is known as the quiet period and lasts for a time lpi_quiet_time equal to 122 LDPC frame periods. The quiet period is defined in 113.3.5.2. The second part of this cycle is known as the refresh period and lasts for a time lpi_refresh_time equal to 6 LDPC frame periods. The refresh period is defined in 113.3.5.3. A cycle composed of one quiet period and one refresh period is known as a *single pair LPI cycle* and lasts for a time lpi_qr_time equal to 128 LDPC frame periods. The time taken to complete a quiet-refresh cycle for all four pairs is known as a *complete LPI cycle*.

lpi_offset, lpi_quiet_time, lpi_refresh_time, lpi_qr_time, and lpi_allpairs_qr_time are timing parameters that are integer multiples of the LDPC frame period. lpi_offset is a fixed value equal to lpi_qr_time/2 that is used to ensure refresh signals are appropriately offset by the link partners.

PHYs begin the transition from the LPI receive mode when the alert signal is detected by the PMA as defined in 113.4.2.4.

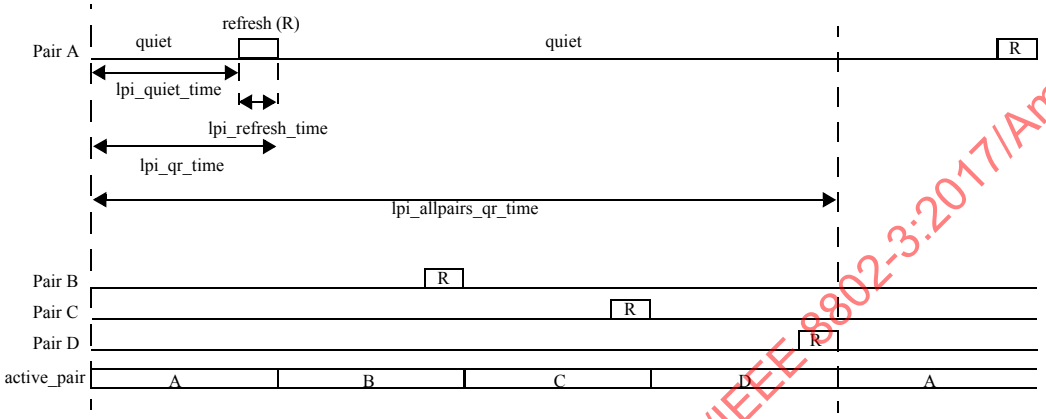


Figure 113–16—Timing periods for LPI signals

113.3.5.1 LPI Synchronization

To maximize power savings, maintain link integrity, and ensure interoperability, EEE-capable PHYs synchronize refresh intervals during the LPI mode. The transition to PCS_Test is used as a fixed timing reference for the link partners. Refresh signaling is derived by counting LDPC frames from the transition to PCS_Test. An EEE-capable PHY shall support loop timing and loop timing shall be enabled on the slave PHY.

In initial training, normal retraining, and fast retraining, with or without the EEE capability being supported, the master and slave signal when they will transition to PCS_Test using the transition counter following the procedure described in 113.4.2.5.15.

An EEE-capable PHY in slave mode is responsible for synchronizing its PMA training frame to the master's PMA training frame during the transition to PMA_Training_Init_S. The slave shall ensure that its PMA training frames are synchronized to the master's PMA training frames within 1 LDPC frame, measured at the slave MDI on pair A. In addition, the slave shall initialize its transition counter so that it transitions to PCS_Test within 1 LDPC frame of the master PHY's transition to PCS_Test, measured at the slave PHY's MDI on pair A. This mechanism ensures that the refresh offset is bounded to a small value at both MDI interfaces, thus ensuring there is no overlap of master and slave signals when both transmit and receive are in the LPI mode.

Following the transition to PCS_Test, the PCS counts transmitted and received LDPC frames, and uses these counters to generate refresh and pair control signals for the transmit and receive functions. The transmitted LDPC frame count is named tx_ldpc_frame_cnt. The received LDPC frame count is named rx_ldpc_frame_cnt.

The master and slave shall derive the active pair and refresh_active signals from the LDPC frame counters as shown in Table 113–7 and Table 113–8.

Table 113–7—Synchronization logic derived from slave signal LDPC frame count

Slave-side variable	Master-side variable	for master u=rx_ldpc_frame_cnt for slave u=tx_ldpc_frame_cnt
tx_refresh_active=true	rx_refresh_active=true	lpi_offset – lpi_refresh_time ≤ mod(u,lpi_qr_time) < lpi_offset
tx_lpi_full_refresh=true	N/A	lpi_offset – lpi_refresh_time = mod(u,lpi_qr_time)
tx_active_pair=PAIR_A	rx_active_pair=PAIR_A	lpi_offset + lpi_qr_time £ u < lpi_offset + 2 × lpi_qr_time
tx_active_pair=PAIR_B	rx_active_pair=PAIR_B	lpi_offset + 2 × lpi_qr_time ≤ u < lpi_offset + 3 × lpi_qr_time
tx_active_pair=PAIR_C	rx_active_pair=PAIR_C	lpi_offset + 3 × lpi_qr_time ≤ u < 4 × lpi_qr_time OR 0 ≤ u < lpi_offset
tx_active_pair=PAIR_D	rx_active_pair=PAIR_D	lpi_offset ≤ u < lpi_offset + lpi_qr_time

Table 113–8—Synchronization logic derived from master signal LDPC frame count

Slave-side variable	Master-side variable	for master v=tx_ldpc_frame_cnt for slave v=rx_ldpc_frame_cnt
rx_refresh_active=true	tx_refresh_active=true	lpi_quiet_time ≤ mod(v,lpi_qr_time)
N/A	tx_lpi_full_refresh=true	lpi_quiet_time = mod(v,lpi_qr_time)
rx_active_pair=PAIR_A	tx_active_pair=PAIR_A	0 ≤ v < lpi_qr_time
rx_active_pair=PAIR_B	tx_active_pair=PAIR_B	lpi_qr_time ≤ v < 2 × lpi_qr_time
rx_active_pair=PAIR_C	tx_active_pair=PAIR_C	2 × lpi_qr_time ≤ v < 3 × lpi_qr_time
rx_active_pair=PAIR_D	tx_active_pair=PAIR_D	3 × lpi_qr_time ≤ v < 4 × lpi_qr_time

113.3.5.2 Quiet period signaling

During the quiet period the transmitters on all four pairs should be turned off. Average launch power (as measured from 28 LDPC frames after a refresh period to 28 LDPC frames before the next refresh period on the same lane) for each Transmitter shall be less than –41 dBm. This requirement does not apply to the periods when the alert signal is transmitted as defined in 113.4.2.2.1.

113.3.5.3 Refresh period signaling

During the LPI mode 25G/40GBASE-T PHYs use staggered, out-of-phase refresh signaling to maximize power savings. Two-level PAM refresh symbols are generated using the PMA side-stream scrambler polynomials described in 113.3.4 and exactly as is shown in Figure 113–15 with the exception that the Infofield consists of a sequence of 128 zeros. The training sequence shall be used during the LPI mode.

Refresh signals shall be sent using the THP filter as described in 113.4.3.1. At the start of each refresh signal the THP feedback delay line shall be initialized with zeros.

While a transmit function is in the LPI transmit mode only one of the transmit pairs is active during a refresh period. tx_symb_vector for all transmit pairs that are not active shall be set to zero.

When tx_symb_vector has the value ALERT and the PHY is master, the transmitter on pair A shall be active and all other pairs shall be quiet. When tx_symb_vector has the value ALERT and the PHY is slave, the transmitter on pair C shall be active and all other pairs shall be quiet. If lpi_tx_mode=REFRESH_A on a MASTER PHY or lpi_tx_mode=REFRESH_C on a SLAVE PHY, and tx_symb_vector has the value ALERT, then the alert signalling shall be transmitted in place of the refresh signalling where the signals overlap.

113.3.6 Detailed functions and state diagrams

113.3.6.1 State diagram conventions

The body of this subclause is comprised of state diagrams, including the associated definitions of variables, constants, and functions. Should there be a discrepancy between a state diagram and descriptive text, the state diagram prevails.

The notation used in the state diagrams follows the conventions of 21.5. State diagram timers follow the conventions of 14.2.3.2. The notation ++ after a counter or integer variable indicates that its value is to be incremented.

113.3.6.2 State diagram parameters

113.3.6.2.1 Constants

EBLOCK_R<71:0>
: 72-bit vector to be sent to the 25GMII/XLGMII interface containing /E/ in all the eight character locations.

EBLOCK_T<64:0>
: 65-bit vector to be sent to the 512B/513B transcoder and block-LDPC framer containing /E/ in all the eight character locations.

LBLOCK_R<71:0>
: 72-bit vector to be sent to the 25GMII/XLGMII interface containing a Local Fault ordered set. The Local Fault ordered set is defined in 81.3.4.

LBLOCK_T<64:0>
: 65-bit vector to be sent to the 512B/513B transcoder and block-LDPC framer containing a Local Fault ordered set.

LPBLOCK_R<71:0>
: 72-bit vector to be sent to the 25GMII/XLGMII containing /LI/ in all the eight character locations.

LPBLOCK_T<64:0>
: 65-bit vector to be sent to the 512B/513B transcoder and block-LDPC framer containing /LI/ in all the eight character locations.

IBLOCK_R<71:0>
: 72-bit vector to be sent to the 25GMII/XLGMII containing /I/ in all the eight character locations.

IBLOCK_T<64:0>
: 65-bit vector to be sent to the 512B/513B transcoder and block-LDPC framer containing /I/ in all the eight character locations.

UBLOCK_R<71:0>

72-bit vector to be sent to the 25GMII/XLGMII containing a Link Interruption ordered set. The Link Interruption ordered set is defined in 81.3.4.

113.3.6.2.2 Variables

lfer_test_lf
: Boolean variable that is set true when a new LDPC frame is available for testing and false when LFER_TEST_LF state is entered. A new LDPC frame is available for testing when the Block Sync process has accumulated enough symbols from the PMA to evaluate the next LDPC frame.

block_lock
: Boolean variable that is set true when receiver acquires block delineation.

hi_lfer
: Boolean variable that is asserted true when the lfer_cnt reaches 16 errors in one lfer_timer interval.

pcs_reset
: Boolean variable that controls the resetting of the PCS. It is true whenever a reset is necessary including when reset is initiated from the MDIO, during power on, and when the MDIO has put the PCS into low-power mode.

rx_coded<64:0>
: Vector containing the input to the 64B/65B decoder. The format for this vector is shown in Figure 113–10. The leftmost bit in the figure is rx_coded<0> and the rightmost bit is rx_coded<64>.

rx_raw<71:0>
: Vector containing two successive 25GMII transfers or a single XLGMII output transfer. For 25GBASE-T, RXC<0> through RXC<3> for the first transfer are taken from rx_raw<0> through rx_raw<3>, respectively. RXC<0> through RXC<3> for the second transfer are taken from rx_raw<4> through rx_raw<7>, respectively. RXD<0> through RXD<31> for the first transfer are taken from rx_raw<8> through rx_raw<39>, respectively. RXD<0> through RXD<31> for the second transfer are taken from rx_raw<40> through rx_raw<71>, respectively. For 40GBASE-T, RXC<0> through RXC<7> for the transfer are taken from rx_raw<0> through rx_raw<7>, respectively. RXD<0> through RXD<63> for the transfer are taken from rx_raw<8> through rx_raw<71>, respectively.

lf_valid
: Boolean indication that is set true if received LDPC frame is valid. LDPC frame is valid if both:
  a. All parity checks of the LDPC-coded bits are satisfied
  b. The RS-FEC-coded bits, after decoding, form a valid RS-FEC codeword.

tx_coded<64:0>
: Vector containing the output from the 64B/65B encoder. The format for this vector is shown in Figure 113–10. The leftmost bit in the figure is tx_coded<0> and the rightmost bit is tx_coded<64>.

tx_raw<71:0>
: Vector containing two successive 25GMII transfers or a single XLGMII transfer. For 25GBASE-T, TXC<0> through TXC<3> for the first transfer are placed in tx_raw<0> through tx_raw<3>, respectively. TXC<0> through TXC<3> for the second transfer are placed in tx_raw<4> through tx_raw<7>, respectively. TXD<0> through TXD<31> for the first transfer are placed in tx_raw<8> through tx_raw<39>, respectively. TXD<0> through TXD<31> for the second transfer are placed in tx_raw<40> through tx_raw<71>, respectively. For 40GBASE-T, TXC<0> through TXC<7> for the transfer are placed in tx_raw<0> through tx_raw<7>, respectively. TXD<0> through TXD<63> for the transfer are placed in tx_raw<8> through tx_raw<71>, respectively.

The following variables are required for PHYs that support the EEE capability:

tx_lpi_active
: A Boolean variable that is set true when the PHY transmit function is operating in the LPI transmit

mode and during transitions to and from the LPI transmit mode (i.e., at any time when the PHY is transmitting sleep, alert, wake, or quiet-refresh signaling). It is set false otherwise.

tx_lpi_qr_active
: A Boolean variable that is set true during the LPI transmit mode, when the PHY is transmitting quiet-refresh signaling. Set false otherwise.

rx_lpi_active
: A Boolean variable that is set true when the PHY receive function is operating in the LPI receive mode and set false otherwise. The LPI receive mode begins when the sleep signal is detected and lasts until the alert signal is detected. When the EEE capability is not supported, rx_lpi_active is set false.

tx_lpi_req
: A Boolean variable that is set true when the LPI client indicates that it is requesting operation in the LPI transmit mode via the 25GMII/XLGMII and set false otherwise.

alert_detect
: Indicates that an alert signal from the link partner has been received at the MDI as indicated by PMA_ALERTDETECT.indication(alert_detect).

tx_lpi_alert_active
: A Boolean variable that is set true when the PHY is transmitting ALERT signaling. Set false otherwise.

rx_lpi_wake
: A Boolean variable that is set true when the PHY receiver is in the WAKE state and sending IDLE to the 25GMII/XLGMII. Set false otherwise. When the EEE capability is not supported, rx_lpi_wake is set false.

tx_active_pair
: A variable indicating the transmit active pair during the LPI transmit mode. The variable may take the values PAIR_A, PAIR_B, PAIR_C, PAIR_D. This variable is defined in 113.3.5.1.

lpi_tx_mode
: A variable indicating the signaling to be used from the PCS to the PMA across the PMA_UNITDATA.request (tx_symb_vector) interface.
lpi_tx_mode controls tx_symb_vector only when tx_mode is set to SEND_N.
The variable is set to NORMAL when (!tx_lpi_qr_active * !tx_lpi_alert_active), indicating that the PCS is in the normal mode of operation and will encode code-groups as described in Figure 113–18a and Figure 113–18b.
The variable is set to REFRESH_A when (tx_lpi_qr_active * (tx_active_pair=PAIR_A) * tx_refresh active).
The variable is set to REFRESH_B when (tx_lpi_qr_active * (tx_active_pair=PAIR_B) * tx_refresh active).
The variable is set to REFRESH_C when (tx_lpi_qr_active * (tx_active_pair=PAIR_C) * tx_refresh active).
The variable is set to REFRESH_D when (tx_lpi_qr_active * (tx_active_pair=PAIR_D) * tx_refresh active).
The variable is set to QUIET when (tx_lpi_qr_active * (!tx_refresh_active + tx_lpi_initial_quiet))
The variable is set to ALERT when (tx_lpi_alert_active)

tx_refresh_active
: A Boolean value. This variable is set true following the logic described in 113.3.5.1.

tx_lpi_full_refresh
: A Boolean value. This variable is set true following the logic described in 113.3.5.1.

tx_lpi_initial_quiet
: A Boolean value. This variable is set true when the transmit function enters the LPI transmit mode and a partial refresh will be replaced by quiet signaling.

ldpc_frame_done
: A Boolean value. This variable is set true when the final symbol of each LDPC frame is transmitted and is set false otherwise.

The following variable is only required for PHYs that support the fast retrain capability:

fr_sigtype

If fast retrain is supported, this variable controls the block type the PMA sends on the receive path during fast retrain. If MDIO is supported, this variable is set based on the value in 1.147.2:1 as follows:

00 IBLOCK_R
01 LBLOCK_R
10 UBLOCK_R
11 Reserved.

If MDIO is not supported, an equivalent method of controlling fast retrain functionality should be provided.

113.3.6.2.3 Timers

State diagram timers follow the conventions described in 14.2.3.2.

lfer_timer
: Timer that is triggered every 125/(4 × *S*) μs +1%, –25%. When the timer reaches its terminal count, lfer_timer_done = TRUE

The following timers are required for PHYs that support the EEE capability:

lpi_tx_sleep_timer
: This timer defines the time the local transmitter sends the sleep signal to the link partner.
Values: The condition lpi_tx_sleep_timer_done becomes true upon timer expiration.
Duration: This timer shall have a period equal to 6 LDPC frame periods.

lpi_tx_alert_timer
: This timer defines the time the local transmitter transmits the alert signal.
Values: The condition lpi_tx_alert_timer_done becomes true upon timer expiration.
Duration: This timer shall have a period equal to 4 LDPC frame periods.

lpi_tx_wake_timer
: This timer defines the time the local transmitter transmits the wake signal.
Values: The condition lpi_tx_wake_timer_done becomes true upon timer expiration.
Duration: This timer shall have a period equal to lpi_wake_time LDPC frame periods.

lpi_rx_wake_timer
: This timer defines the time the receiver sends IDLE blocks to the 25GMII/XLGMII after the alert signal is detected.
Values: The condition lpi_rx_wake_timer_done becomes true upon timer expiration.
Duration: This timer shall have a period equal to lpi_wake_time LDPC frame periods.

113.3.6.2.4 Functions

DECODE(rx_symb_vector<64:0>)
: In the PCS Receive process, this function takes as its argument 65-bit rx_coded<64:0> from the LDPC decoder and decodes the 65B-LDPC bit vector returning a vector rx_raw<71:0>, which is sent to the 25GMII/XLGMII. The DECODE function shall decode the block based on code specified in 113.3.2.2.2.

ENCODE(tx_raw<71:0>)
: Encodes the 72-bit vector received from the 25GMII/XLGMII, returning 65-bit vector tx_coded. The ENCODE function shall encode the block as specified in 113.3.2.2.2.

R_BLOCK_TYPE = {C, S, T, D, E, I, LI, LII}

When the EEE capability is not supported, this function classifies each 65-bit rx_coded vector as belonging to one of the five types {C, S, T, D, E} depending on its contents.

When the EEE capability is supported, this function classifies each 65-bit rx_coded vector as belonging to the eight types depending on its contents. A vector may simultaneously belong to the C and I types when it contains eight valid control characters that are all /I/, but in every other case the vector belongs to only one type.

Values: C; The vector contains a data/ctrl header of 1 and one of the following:

a) A block type field of 0x1E and eight valid control characters other than /E/ and /LI/;
b) a block type field 0x4B, a valid O code, and zeros;
c) for 25GBASE-T only, a block type field 0x2D, a valid O code, and four valid control characters;
d) for 25GBASE-T only, a blcok type field of 0x55 and two valid O codes.

S; The vector contains a data/ctrl header of 1 and one of the following:

a) A block type field of 0x78;
b) for 25GBASE-T only, a block type field of 0x66 and a valid O code;
c) for 25GBASE-T only, a block type field of 0x33 and four valid control characters.

T; The vector contains a data/ctrl header of 1, a block type field of 0x87, 0x99, 0xAA, 0xB4, 0xCC, 0xD2, 0xE1, or 0xFF and all control characters are valid.

D; The vector contains a data/ctrl header of 0.

I; If the optional EEE capability is supported, then the I type is a special case of the C type where the vector contains a data/ctrl header of 1, a block type field of 0x1e, and eight control characters of /I/.

LI: If the optional EEE capability is supported, then the LI type occurs when the vector contains a data/ctrl header of 1, a block type field of 0x1e, and eight control characters of /LI/.

LII: If the optional EEE capability is supported, then the LII type occurs when the vector contains a data/ctrl header of 1, a block type field of 0x1E, and one of the following:

a) four control characters of /LI/ followed by four control characters of /I/;
b) four control characters of /I/ followed by four control characters of /LI/

E; The vector does not meet the criteria for any other value.

For 25GBASE-T, a valid control character is one containing a 25GBASE-T control code specified in Table 113–1. A valid O code is one containing an O code specified in Table 113–1,

For 40GBASE-T, a valid control character is one containing a 40GBASE-T control code specified in Table 113–2. A valid O code is one containing an O code specified in Table 113–2.

R_TYPE(rx_coded<64:0>)

Returns the R_BLOCK_TYPE of the rx_coded<64:0> bit vector.

R_TYPE_NEXT

Presient end of packet check function. It returns the R_BLOCK_TYPE of the rx_coded vector immediately following the current rx_coded vector.

T_BLOCK_TYPE = {C, S, T, D, E, I, LI, LII}

When the EEE capability is not supported, this function classifies each 72-bit tx_raw vector as belonging to one of the five types {C, S, T, D, E} depending on its contents.

When the EEE capability is supported, this function classifies each 72-bit tx_raw vector as belonging to the eight types depending on its contents. A vector may simultaneously belong to the C and I types when it contains eight valid control characters that are all /I/, but in every other case the vector belongs to only one type.

Values: C; The vector contains one of the following:

a) eight valid control characters other than /O/, /S/, /T/, /E/, and /LI/;
b) one valid ordered set followed by four data bytes and zeros for block code 0x4B;
c) for 25GBASE-T only, one valid ordered set followed by four valid control characters other than /O/, /S/ and /T/;

d) for 25GBASE-T only, two valid ordered sets.

S; The vector contains an /S/ in its first character, or, for 25GBASE-T, its first or fifth character. Any characters before the S character are valid control characters other than /O/, /S/ and /T/ or form a valid ordered set, and all characters following the /S/ are data characters.

T; The vector contains a /T/ in one of its characters, all characters before the /T/ are data characters, and all characters following the /T/ are valid control characters other than /O/, /S/ and /T/.

D; The vector contains eight data characters.

I; If the optional EEE capability is supported, then the I type is a special case of the C type where the vector contains eight control characters of /I/.

LI: If the optional EEE capability is supported, then the LI type occurs when the vector contains eight control characters of /LI/.

LII: If the optional EEE capability is supported, then the LII type occurs when the vector contains one of the following:

a) four control characters of /LI/ followed by four control characters of /I/;
b) four control characters of /I/ followed by four control characters of /LI/.

E; The vector does not meet the criteria for any other value.

A tx_raw character is a control character if its associated TXC bit is asserted. For 25GBASE-T, a valid control character is one containing a 25GMII control code specified in Table 113–1. A valid ordered set consists of a valid /O/ character in the first or fifth characters and data characters in the three characters following the /O/. A valid /O/ is any character with a value for O code in Table 113–1. For 40GBASE-T, a valid control character is one containing an XLGMII control code specified in Table 113–2. A valid ordered set consists of a valid /O/ character in the first character and data characters in the three characters following the /O/. A valid /O/ is any character with a value for O code in Table 113–2.

T_TYPE(tx_raw<71:0>)
: Returns the T_BLOCK_TYPE of the tx_raw<71:0> bit vector.

T_TYPE_NEXT
: Prescient end of packet check function. It returns the FRAME_TYPE of the tx_raw vector immediately following the current tx_raw vector.

113.3.6.2.5 Counters

lfer_cnt
: Count up to a maximum of 16 of the number of invalid LDPC frames within the current lfer_timer period.

The following counters are required for PHYs that support the EEE capability:

tx_ldpc_frame_cnt
: An integer value that counts transmit LDPC frame periods. The counter is reset when the first symbol of the first LDPC frame crosses the MDI on pair A in the transmit direction after normal training or fast retraining. It is incremented after the last symbol of each transmitted LDPC frame. tx_ldpc_frame_cnt is reset to 0 when tx_ldpc_frame_cnt = lpi_qr_time × 4.

rx_ldpc_frame_cnt
: An integer value that counts receive LDPC frame periods. The counter is reset when the first symbol of the first LDPC frame crosses the MDI on pair A in the receive direction after normal training or fast retraining. It is incremented after the last symbol of each received LDPC frame. rx_ldpc_frame_cnt is reset to 0 when rx_ldpc_frame_cnt = lpi_qr_time × 4.

lpi_rxw_err_cnt
: An integer value that counts the number of receive wake on error conditions. lpi_rxw_err_cnt is reset to zero during PCS_Test. The counter is reflected in register 3.22 (see 45.2.3.10).

113.3.6.3 State diagrams

The LFER Monitor state diagram shown in Figure 113–17 monitors the received signal for high LDPC frame error ratio.

The 64B/65B Transmit state diagram shown in Figure 113–18a controls the encoding of 65B transmitted blocks. It makes exactly one transition for each 65B transmit block processed. Though the Transmit state diagram sends Local Fault ordered sets when reset is asserted, the scrambler and 65B-LDPC are not guaranteed to be operational during reset. Thus, the Local Fault ordered sets are not guaranteed to appear on the PMA service interface.

The 64B/65B Receive state diagram shown in Figure 113–19a controls the decoding of 65B received blocks. It makes exactly one transition for each receive block processed except for the transition from RX_WE to RX_E, which occurs immediately after the RX_WE processes are complete.

The PCS shall perform the functions of LFER Monitor, Transmit, and Receive as specified in these state diagrams. The PCS shall not perform the LFER Monitor function during LPI receive operation from the time that the PCS 64B/65B Receiver enters the state RX_L, until the state RX_W is exited.

Transitions surrounded by dashed rectangles indicate requirements for 25G/40GBASE-T EEE-capable implementations.

113.3.7 PCS management

The following objects apply to PCS management. If an MDIO Interface is provided (see Clause 45), they are accessed via that interface. If not, it is recommended that an equivalent access be provided.

113.3.7.1 Status

pcs_status
: Indicates whether the PCS is in a fully operational state. It is only true if block_lock is true and hi_lfer is false. This status is reflected in MDIO register 3.32.12. A latch low view of this status is reflected in MDIO register 3.1.2 and a latch high of the inverse of this status, Receive fault, is reflected in MDIO register 3.8.10.

block_lock
: Indicates the state of the block_lock variable. This status is reflected in MDIO register 3.32.0. A latch low view of this status is reflected in MDIO register 3.33.15.

hi_lfer
: Indicates the state of the hi_lfer variable. This status is reflected in MDIO register 3.32.1. A latch high view of this status is reflected in MDIO register 3.33.14.

Rx LPI indication
: For EEE capability, this variable indicates the current state of the receive LPI function. This flag is set to TRUE (register bit set to one) when the PCS 64B/65B Receive state diagram (Figure 113–19b) is in the RX_L or RX_W states. This status is reflected in MDIO register 3.1.8. A latch high view of this status is reflected in MDIO register 3.1.10 (Rx LPI received).

Tx LPI indication
: For EEE capability, this variable indicates the current state of the transmit LPI function. This flag is set to TRUE (register bit set to one) when the PCS 64B/65B Transmit state diagram (Figure 113–18b) is in the TX_L or TX_W states. This status is reflected in MDIO register 3.1.9. A latch high view of this status is reflected in MDIO register 3.1.11 (Tx LPI received).

113.3.7.2 Counters

The following counters are reset to zero upon read and upon reset of the PCS. When they reach all ones, they stop counting. Their purpose is to help monitor the quality of the link.

lfer_count
: Six-bit counter that counts each time LFER_BAD_LF state is entered. This counter is reflected in MDIO register bits 3.33.13:8. The counter is reset when register 3.33 is read by management. Note that this counter counts a maximum of 16 counts per lfer_timer period since the LFER_BAD_LF can be entered a maximum of 16 times per lfer_timer window.

errored_block_count
: Eight-bit counter. When the receiver is in normal mode, errored_block_count counts once for each time RX_E state is entered. This counter is reflected in MDIO register bits 3.33.7:0

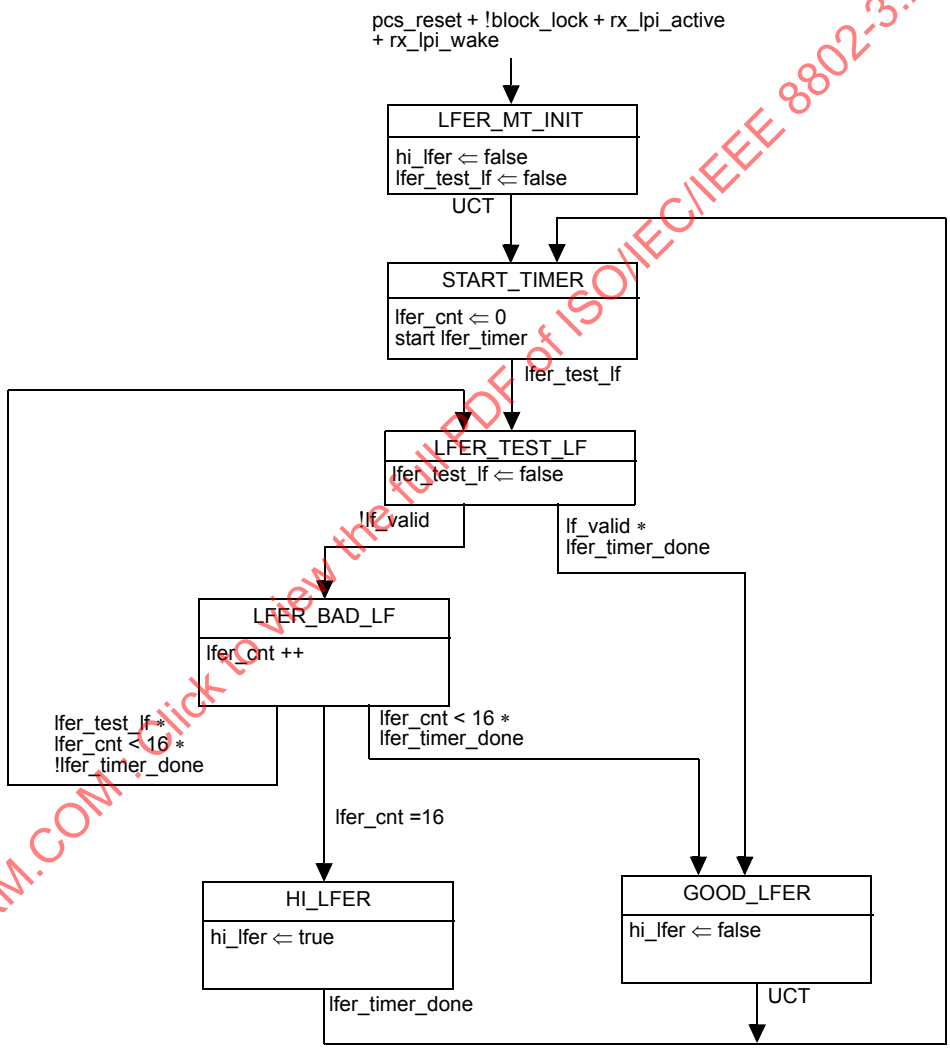
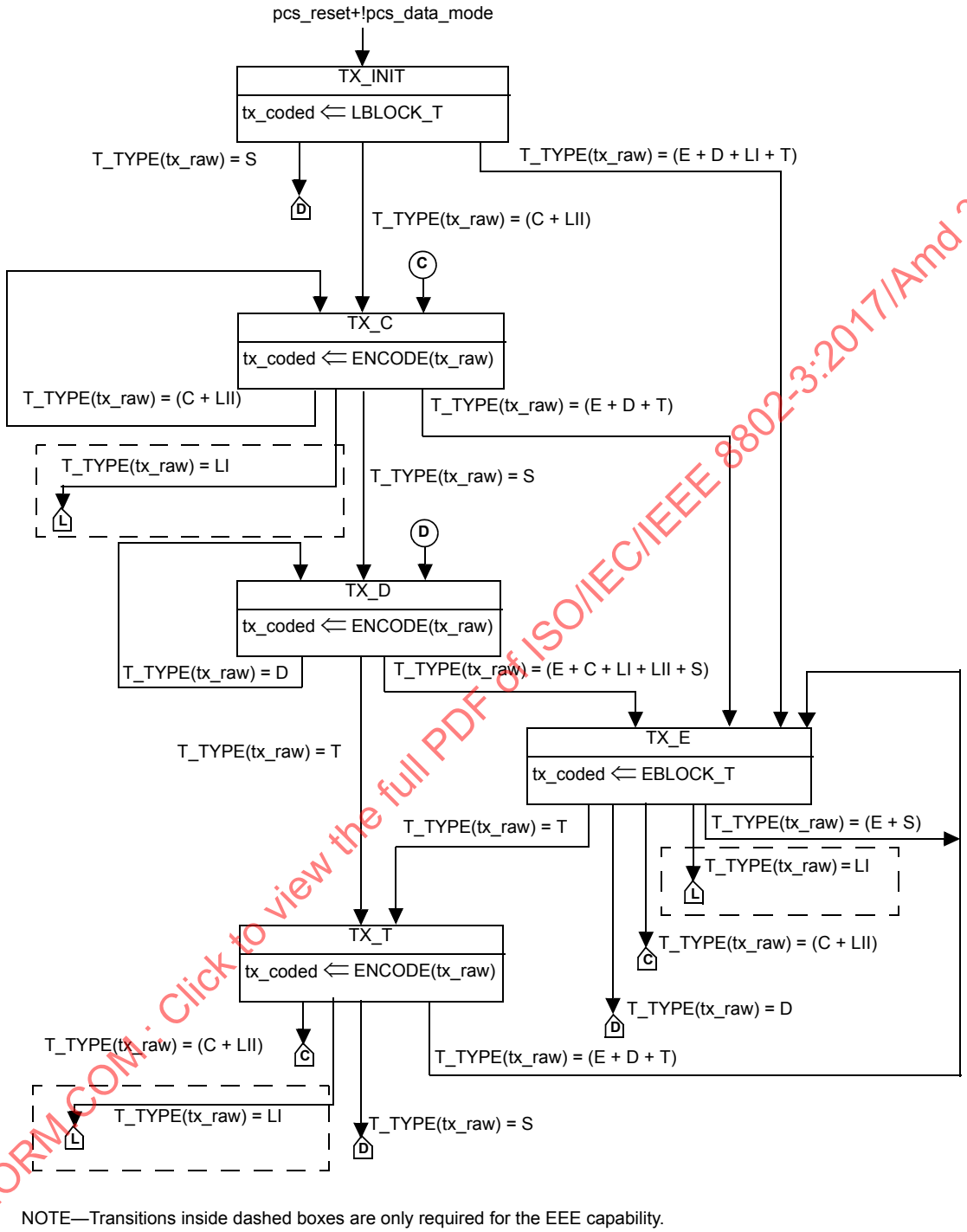
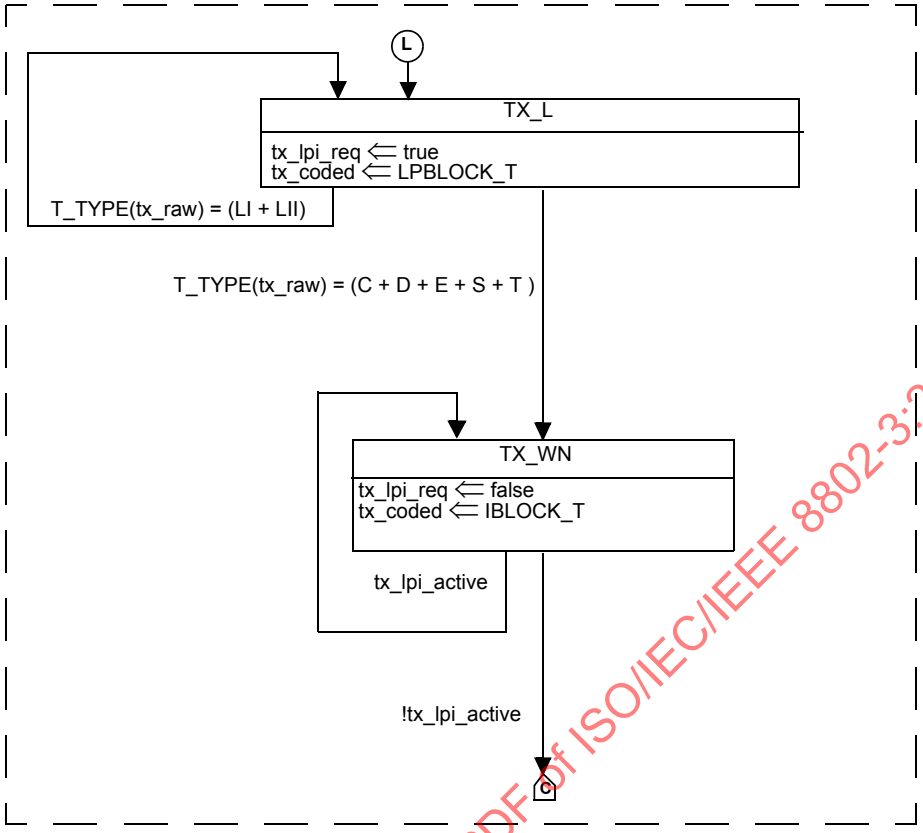


Figure 113–17—LFER monitor state diagram



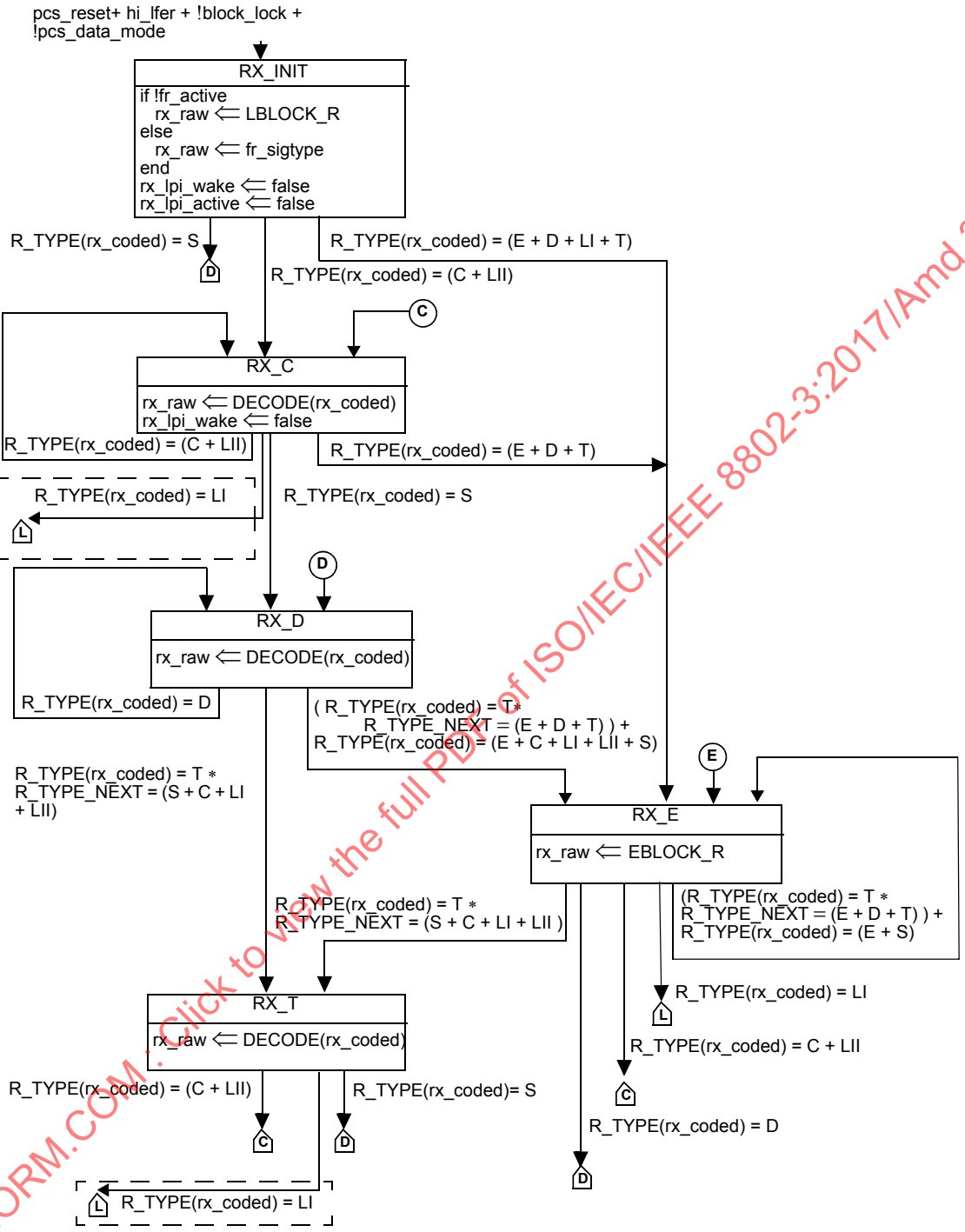
NOTE—Transitions inside dashed boxes are only required for the EEE capability.

Figure 113–18a—PCS 64B/65B Transmit state diagram, part a



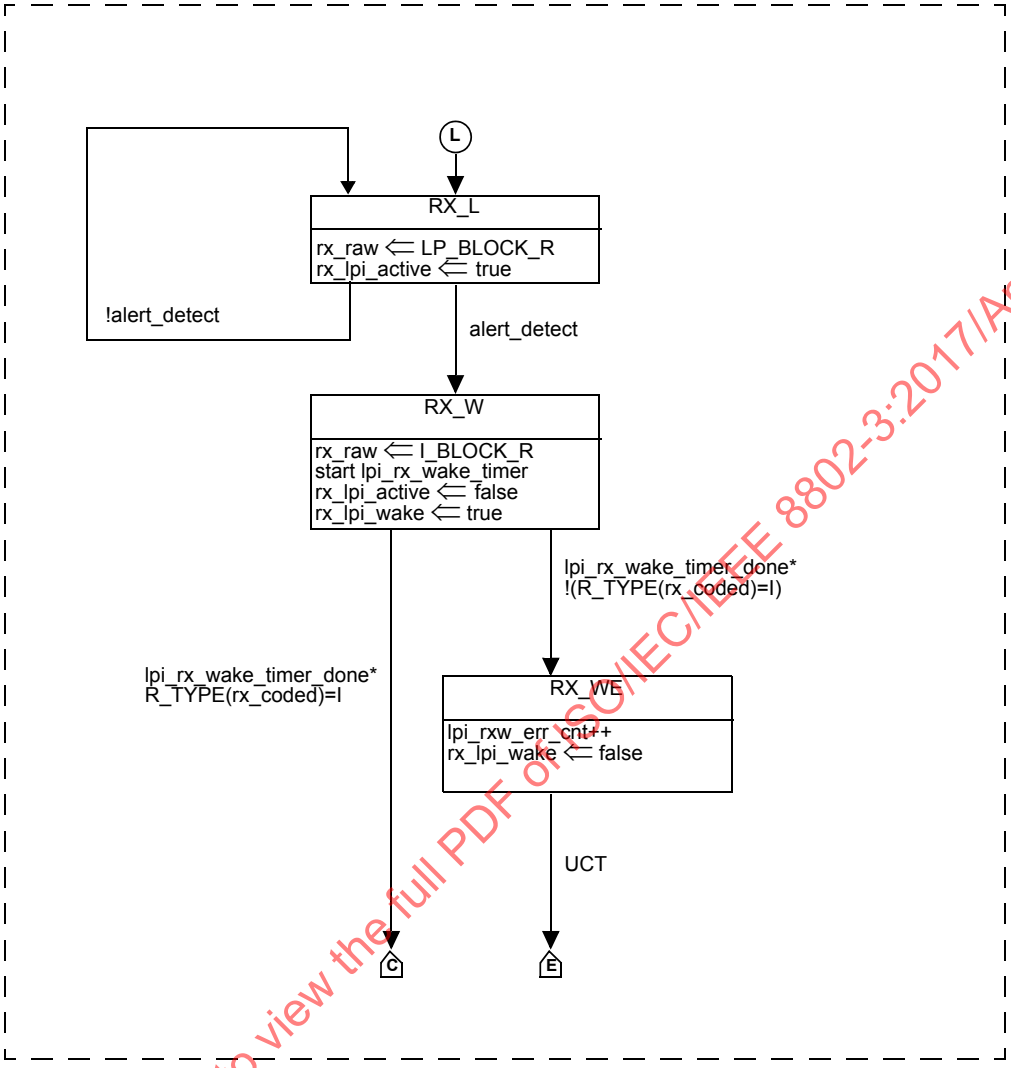
NOTE—This figure is mandatory for PHYs with the EEE capability.

Figure 113–18b—PCS 64B/65B Transmit state diagram, part b



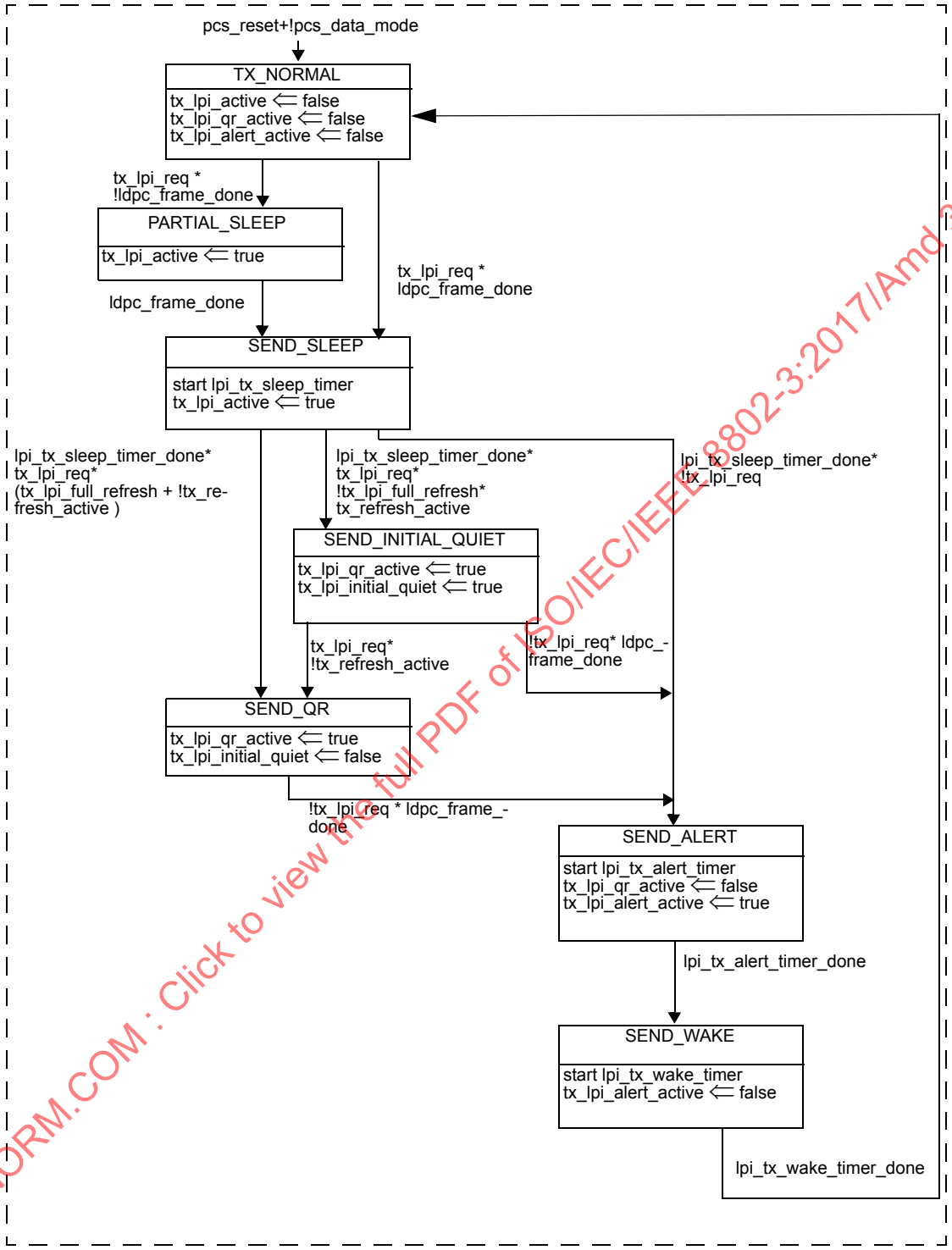
NOTE—Signals and functions shown with dashed lines are only required for the EEE capability.

Figure 113–19a—PCS 64B/65B Receive state diagram, part a



NOTE—This figure is mandatory for PHYs with the EEE capability.

Figure 113–19b—PCS 64B/65B Receive state diagram, part b



NOTE—This figure is mandatory for PHYs with the EEE capability.

Figure 113–20—EEE transmit state diagram

113.3.7.3 Loopback

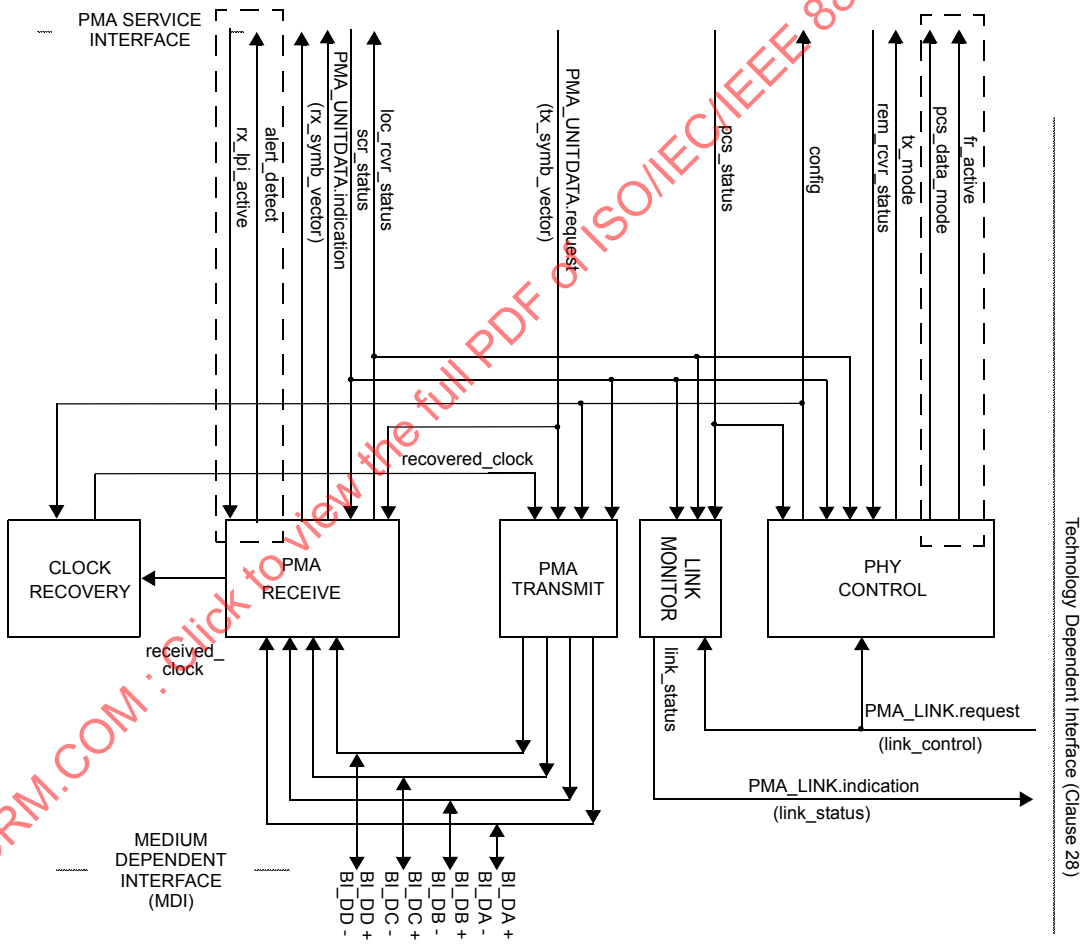
The PCS shall be placed in loopback mode when the loopback bit in MDIO register 3.0.14 is set to a one. In this mode, the PCS shall accept data on the transmit path from the 25GMII/XLGMII and return it on the receive path to the 25GMII/XLGMII. In addition, the PCS shall transmit a continuous stream of 65B-LDPC encoded 4D-PAM16 symbols to the PMA sublayer, and shall ignore all data presented to it by the PMA sublayer.

113.4 Physical Medium Attachment (PMA) sublayer

113.4.1 PMA functional specifications

The PMA couples messages from a PMA service interface specified in 113.2.2 to the 25G/40GBASE-T baseband medium, specified in 113.7.

The interface between PMA and the baseband medium is the Medium Dependent Interface (MDI), which is specified in 113.8.



NOTE 1—The recovered_clock arc is shown to indicate delivery of the recovered clock signal back to PMA TRANSMIT for loop timing.

NOTE 2—pcs_data_mode is required only for the EEE or fast retrain capabilities alert_detect and rx_lpi_active are only required for the EEE capability fr_active is only required for the fast retrain capability.

Figure 113–21—PMA reference diagram

113.4.2 PMA functions

The PMA sublayer comprises one PMA Reset function and five simultaneous and asynchronous operating functions. The PMA operating functions are PHY Control, PMA Transmit, PMA Receive, Link Monitor, and Clock Recovery. All operating functions are started immediately after the successful completion of the PMA Reset function.

The PMA reference diagram, Figure 113–21, shows how the operating functions relate to the messages of the PMA Service interface and the signals of the MDI. Connections from the management interface, comprising the signals MDC and MDIO, to other layers are pervasive and are not shown in Figure 113–21.

113.4.2.1 PMA Reset function

The PMA Reset function shall be executed whenever one of the two following conditions occur:

a) Power on (see 113.3.6.2.2)
b) The receipt of a request for reset from the management entity

All state diagrams take the open-ended pma_reset branch upon execution of PMA Reset. The reference diagrams do not explicitly show the PMA Reset function.

113.4.2.2 PMA Transmit function

The PMA Transmit function comprises four synchronous transmitters to generate four pulse-amplitude modulated signals on each of the four pairs BI_DA, BI_DB, BI_DC, and BI_DD. While send_fail is FALSE and ALERT is not indicated by tx_symb_vector, PMA Transmit shall continuously transmit onto the MDI pulses modulated by the symbols given by tx_symb_vector[BI_DA], tx_symb_vector[BI_DB], tx_symb_vector[BI_DC], and tx_symb_vector[BI_DD], respectively, after processing with the THP, optional transmit filtering, digital to analog conversion (DAC), and subsequent analog filtering. When ALERT is indicated by tx_symb_vector, the alert signal is transmitted as specified in 113.4.2.2.1. When send_fail is TRUE, the link failure signal is transmitted as specified in 113.4.2.2.2. The four transmitters shall be driven by the same transmit clock, TX_TCLK. The signals generated by PMA Transmit shall follow the mathematical description given in 113.4.3.1 and shall comply with the electrical specifications given in 113.5.

When the PMA_CONFIG.indication parameter config is MASTER, for both normal and LPI operation, the PMA Transmit function shall source TX_TCLK from a local clock source while meeting the transmit jitter requirements of 113.5.3.3. The MASTER/SLAVE relationship includes loop timing. If the PMA_CONFIG.indication parameter config is SLAVE, the PMA Transmit function shall source TX_TCLK from the recovered clock of 113.4.2.8 while meeting the jitter requirements of 113.5.3.3.

The PMA Transmit fault function is optional. The faults detected by this function are implementation specific. If the MDIO interface is implemented, then this function shall be mapped to the transmit fault bit as specified in 45.2.1.7.4.

EEE-capable PHYs shall implement a PMA Transmit function that generates the alert signal as defined in 113.4.2.2.1. PHYs that support the fast retrain capability shall implement a PMA Transmit function that generates the link failure signal as defined in 113.4.2.2.2. If ALERT is indicated by tx_symb_vector at the same time as send_fail is TRUE, then link failure signaling is transmitted.

113.4.2.2.1 Alert signal

PHYs that support the optional EEE capability transmit the following PAM2 sequence when the PMA_UNITDATA.request parameter is set to ALERT. The alert signal is sent for a total of 4 LDPC frame

periods and begins on a LDPC frame boundary. The alert signal is transmitted without THP filtering. The alert signal is transmitted on pair A when the PHY operates as a MASTER. The alert signal is transmitted on pair C when the PHY operates as a SLAVE. All other pairs transmit quiet as described in 113.3.5.

When the PMA_CONFIG.indication(config) is MASTER, the alert signal is composed of 7 repetitions of the following 128 symbol PAM2 sequence, followed by 128 zero symbols.

xpr_master =

9 9 –9 –9 –9 –9 –9 –9 9 9 –9 –9 9 9 9 9

9 9 9 9 –9 –9 9 9 9 9 –9 –9 9 9 –9 –9

–9 –9 –9 –9 –9 –9 9 9 –9 –9 –9 –9 –9 –9 9 9

–9 –9 –9 –9 –9 –9 –9 –9 9 9 –9 –9 9 9 –9 –9

–9 –9 9 9 9 9 9 9 9 9 9 9 –9 –9 –9 –9

9 9 –9 –9 –9 –9 9 9 9 9 –9 –9 9 9 –9 –9

–9 –9 –9 –9 –9 –9 –9 –9 9 9 9 9 –9 –9 9 9

9 9 –9 –9 9 9 –9 –9 9 9 9 9 –9 –9 9 –9

When the PMA_CONFIG.indication(config) is SLAVE, the alert signal is composed of 7 repetitions of the following 128 symbol PAM2 sequence, followed by 128 zero symbols:

xpr_slave =

–9 –9 –9 –9 9 9 9 9 –9 –9 9 9 –9 –9 9 9

9 9 –9 –9 9 9 9 9 –9 –9 –9 –9 –9 –9 –9 –9

–9 –9 9 9 –9 –9 9 9 9 9 –9 –9 –9 –9 9 9

–9 –9 –9 –9 9 9 9 9 9 9 9 9 9 9 –9 –9

–9 –9 9 9 –9 –9 9 9 –9 –9 –9 –9 –9 –9 –9 –9

9 9 –9 –9 –9 –9 –9 –9 9 9 –9 –9 –9 –9 –9 –9

–9 –9 9 9 –9 –9 9 9 9 9 –9 –9 9 9 9 9

9 9 9 9 –9 –9 9 9 –9 –9 –9 –9 –9 –9 9 9

The alert signal is followed by a wake signal composed of repeated IDLE characters encoded using the mixed 512B/513B 64B/65B RS-FEC and LDPC encoding used in normal data mode. At the start of the wake signal all THP feedback delay lines are initialized with zeros.

113.4.2.2.2 Link failure signal

PHYs that support the fast retrain capability transmit the link failure signal under the control of the Fast Retrain state diagram. The link failure signal indicates to the link partner that a link failure has been detected and that the link partners should begin the fast retrain procedure.

The link failure signal is sent for 4 LDPC frames and begins on a LDPC frame boundary. The link failure signal is transmitted without THP filtering. The link failure signal is transmitted on pair A when the PHY operates as a MASTER. The link failure signal is transmitted on pair C when the PHY operates as a SLAVE. All other pairs transmit quiet as described in 113.3.5.

When the PMA_CONFIG.indication(config) is MASTER, the link failure signal is composed of 7 repetitions of the following 128 symbol PAM2 sequence, followed by 128 zero symbols.

xfr_master = xpr_master × (–1)

When the PMA_CONFIG.indication(config) is SLAVE, the link failure signal is composed of 7 repetitions of the following 128 symbol PAM2 sequence, followed by 128 zero symbols.

xfr_slave = xpr_slave × (–1)

113.4.2.3 PMA transmit disable function

113.4.2.3.1 Global PMA transmit disable function

The Global_PMA_transmit_disable function allows all of the transmitters to be disabled, when either:

a) When a Global_PMA_transmit_disable variable is set to TRUE, this function shall turn off all of the transmitters so that the each transmitter Average Launch Power of the OFF Transmitter is less than –53 dBm.
b) If a PMA_transmit_fault is detected, then the PMA may set the Global_PMA_transmit_disable to TRUE, turning off the transmitter on each pair.

113.4.2.3.2 PMA pair by pair transmit disable function

The PMA_transmit_disable function allows the transmitters on each pair to be selectively disabled.

When a PMA_transmit_disable_N variable is set to TRUE, this function shall turn off the transmitter associated with that variable so that the transmitter Average Launch Power of the OFF Transmitter is less than –53 dBm.

113.4.2.3.3 PMA MDIO function mapping

The MDIO capability described in Clause 45 defines several variables that provide control and status information for and about the PMA. Mapping of MDIO control variables to PMA control variables is shown in Table 113–9. Mapping of MDIO status variables to PMA status variables is shown in Table 113–10.

Table 113–9—MDIO/PMA control variable mapping

MDIO control variable	PMA register name	Register/bit number	PMA control variable
Reset	Control register 1	1.0.15	PMA_reset
Global transmit disable	Transmit disable register	1.9.0	Global_PMA_transmit_disable
Transmit disable pair D	Transmit disable register	1.9.4	PMA_transmit_disable_D
Transmit disable pair C	Transmit disable register	1.9.3	PMA_transmit_disable_C
Transmit disable pair B	Transmit disable register	1.9.2	PMA_transmit_disable_B
Transmit disable pair A	Transmit disable register	1.9.1	PMA_transmit_disable_A

Table 113–10—MDIO/PMA status variable mapping

MDIO status variable	PMA register name	Register/bit number	PMA status variable
Fault	Status register 1	1.1.7	PMA_fault
Transmit fault	Status register 2	1.8.11	PMA_transmit_fault
Receive fault	Status register 2	1.8.10	PMA_receive_fault

113.4.2.4 PMA Receive function

The PMA Receive function comprises four independent receivers for pulse-amplitude modulated signals on each of the four pairs BI_DA, BI_DB, BI_DC, and BI_DD. The PMA Receive function contains the circuits necessary to both detect symbol sequences from the signals received at the MDI over receive pairs BI_DA, BI_DB, BI_DC, and BI_DD and to present these sequences to the PCS Receive function. The signals received at the MDI are described mathematically in 113.4.3.2. The PMA translates the signals received on pairs BI_DA, BI_DB, BI_DC, and BI_DD into the PMA_UNITDATA.indication parameter rx_symb_vector. The quality of these symbols shall allow an LFER of less than 3.2×10^{-9} after RS-FEC and LDPC decoding, over a link segment meeting the requirements of 113.7. The receiver shall correct for differential delay variations of up to 17 ns across the wire pairs. The delay skew is removed by computing the relative received delay of the four known transmit patterns described in 113.3.4.

To achieve the indicated performance, it is highly recommended that PMA Receive include the functions of signal equalization, echo, and crosstalk cancellation. The sequence of code-groups assigned to rx_symb_vector is needed to perform echo and self near-end crosstalk cancellation.

The PMA Receive function uses the scr_status parameter and the state of the equalization, cancellation, estimation, and LPI functions to determine the quality of the receiver performance, and generates the loc_rcvr_status variable accordingly. The precise algorithm for generation of loc_rcvr_status is implementation dependent.

The receiver uses the sequence of symbols during the training sequence to detect and correct for pair swaps and crossovers. The receiver pairs BI_DA, BI_DB, BI_DC, and BI_DD may be connected in any manner described in 113.4.4 to the corresponding transmit pairs. The receiver also detects and corrects for polarity mismatches on any pairs and corrects for differential delay variations across the wire pairs.

The PMA Receive fault function is optional. The PMA Receive fault function is the logical OR of the link_status = FAIL and any implementation specific fault. If the MDIO interface is implemented, then this function shall contribute to the receive fault bit specified in 45.2.1.7.5.

PMA receive functions that support the optional EEE capability shall generate alert_detect when the alert signal is detected at the receiver. The PMA receive function asserts alert_detect after the entire alert signal (3.5 LDPC frame periods of the xpr_master or xpr_slave sequence and 0.5 frames of silence) has been detected. The alert signal is specified in 113.4.2.2.1. The criterion used to generate alert_detect is left to the implementer.

PHYs that support the fast retrain capability shall set link_fail_detect to TRUE when the link failure signal is reliably detected at the receiver. The PMA receive function asserts link_fail_detect after the entire link failure signal (3.5 LDPC frame periods of the xfr_master or xfr_slave sequence and 0.5 frames of silence) has been detected. The link failure signal is specified in 113.4.2.2.2. The criterion used to generate link_fail_detect is left to the implementer. It is highly recommended that the generation of link_fail_detect is qualified with repeated errored frames at the LDPC decoder output.

113.4.2.5 PHY Control function

PHY Control generates the control actions that are needed to bring the PHY into a mode of operation during which frames can be exchanged with the link partner. PHY Control shall comply with the state diagram description given in Figure 113–28.

During PMA training (includes PMA_Training_Init_M, PMA_Training_Init_S, PMA_PBO_Exch, PMA_Coeff_Exch, and PMA_Fine_Adjust states in Figure 113–28), PHY Control information is exchanged between link partners with a 16 octet Infofield, which is XOR'ed with the last 128 bits of the PMA 16384 PAM2 frame on pair A (see Figure 113–15). The link partner is not required to decode every Infofield transmitted but is required to decode Infofields at a rate that enables the correct actions to timer expiration times, transition counter values, etc., described in Figure 113–28, Figure 113–29, and Figure 113–30.

The 16-octet Infofield shall include the fields in 113.4.2.5.2 through 113.4.2.5.14, also shown in the overview Figure 113–22, and the more detailed Figure 113–23, Figure 113–24, and Figure 113–25.

Start of Frame Delimiter 0xBBA70000	3 Transmitter Settings	Message Field	SNR Margin	Message Field Dependent	Message Field Dependent	CRC16
4 octets	3 octets	1 octet	4 bits	1.5 octets	4 octets	2 octets

Figure 113–22—Infofield format

Start of Frame Delimiter 0xBBA70000	3 Transmitter Settings	Mes-sage Field	SNR Margin	Reser-ved	Transition Counter	Reser-ved/ Ability	Vendor Specific	CRC16
4 octets	3 octets	1 octet	4 bits	2 bits	10 bits	2 octets	2 octets	2 octets

Figure 113–23—Infofield transition counter format

Start of Frame Delimiter 0xBBA70000	3 Transmitter Settings	Message Field	SNR Margin	Coefficient Exchange	Coefficient Field	CRC16
4 octets	3 octets	1 octet	4 bits	1.5 octets	4 octets	2 octets

Figure 113–24—Infofield coefficient exchange format

Start of Frame Delimiter 0xBBA70000	3 Transmitter Settings	Mes-sage Field	SNR Margin	Reserved	Reserved /Ability	Vendor Specific	CRC16
4 octets	3 octets	1 octet	4 bits	1.5 octets	2 octets	2 octets	2 octets

Figure 113–25—Infofield not transition counter and not coefficient exchange format

113.4.2.5.1 Infofield notation

For all the Infofield notation in the following subclauses, Reserved<bit location> represents any unused values and shall be set to zero and ignored by the link partner. For all PBO Infofield values in the following subclauses, the PBO<6:4> are unsigned three-bit values 000, 001, 010, 011, 100, 101, 110, and 111 shall indicate power backoffs of 0 dB, 2 dB, 4 dB, 6 dB, 8 dB, 10 dB, 12 dB, and 14 dB, respectively. The Infofield is transmitted following the notation described in 113.3.2.2.3 where the LSB of each octet is sent first and the octets are sent in increasing number order (that is, the LSB of Octet 1 is sent first).

113.4.2.5.2 Start of Frame Delimiter

The start of Frame Delimiter consist of 4 octets [Octet 1<7:0>, Octet 2<7:0>, Octet 3<7:0>, Octet 4<7:0>] and shall use the hexadecimal value 0xBBA70000. 0xBB corresponds to Octet 1<7:0> and so forth.

113.4.2.5.3 Current transmitter settings

Current transmitter setting (1 octet). Represented by Octet 5{Valid<7>, PBO<6:4>, Reserved<3:0>} and shown in Figure 113–26. Used to announce the current fixed PBO setting during PMA_Training_Init_M, PMA_Training_Init_S, and PMA_PBO_Exch, and the current programmable PBO setting during PMA_Coeff_Exch. For every other state this octet is set to zero and ignored by the link partner. The bit Valid

shall be set to one if the corresponding octet information is valid and shall be set to zero if it the octet information is not valid. If Valid is set to zero, the octet is ignored by the link partner.

Single transmitter setting detail (one for current, next or requested)

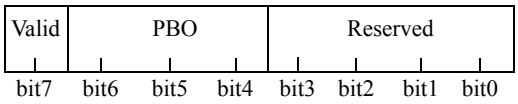


Figure 113–26—Infofield transmitter setting format

113.4.2.5.4 Next transmitter settings

Next transmitter setting (1 octet). Represented by Octet 6{Valid<7>, PBO<6:4>, Reserved<3:0>} and shown in Figure 113–26. Used to announce the next programmable PBO setting during PMA_PBO_Exch that takes effect upon entering PMA_Coeff_Exch state. For every other state, this octet is set to zero and ignored by the link partner. The bit Valid shall be set to one if the corresponding octet information is valid and shall be set to zero if it the octet information is not valid. If Valid is set to zero, the octet is ignored by the link partner.

113.4.2.5.5 Requested transmitter settings

Requested remote transmitter setting (1 octet). Represented by Octet 7{Valid<7>, PBO<6:4>, Reserved<3:0>} and shown in Figure 113–26. Used to request the remote transmitter programmable PBO setting during PMA_PBO_Exch that takes effect upon entering PMA_Coeff_Exch state. For every other state, this octet is set to zero and ignored by the link partner. The bit Valid shall be set to one if the corresponding octet information is valid and shall be set to zero if it the octet information is not valid. If Valid is set to zero, the octet is ignored by the link partner.

113.4.2.5.6 Message Field

Message Field (1 octet). For the MASTER, this field is represented by Octet 8{PMA_state<7:6>, loc_rcvr_status<5>, en_slave_tx<4>, trans_to_Coeff_Exch<3>, Coeff_exchange<2>, trans_to_Fine_Adjust<1>, trans_to_PCS_Test<0>}. For the SLAVE, this field is represented by Octet 8{PMA_state<7:6>, loc_rcvr_status<5>, timing_lock_OK<4>, trans_to_Coeff_Exch<3>, Coeff_exchange<2>, trans_to_Fine_Adjust <1>, trans_to_PCS_Test<0>}.

The two state-indicator bits PMA_state<7:6> shall communicate the state of the transmitting transceiver to the link partner. PMA_state<7:6>=00 indicates PMA_Training_Init_M or PMA_Training_Init_S, PMA_state<7:6>=01 indicates PMA_PBO_Exch, PMA_state<7:6>=10 indicates PMA_Coeff_Exch, and PMA_state<7:6>=11 indicates PMA_Fine_Adjust.

All possible Message Field settings are listed in Table 113–11 for the MASTER and Table 113–12 for the SLAVE. No other value shall be transmitted, and all other values shall be ignored at the receiver. The Message Field setting for the first transmitted PMA frame shall be the first row of Table 113–11 for the MASTER and the first row of Table 113–12 for the SLAVE. Moreover, for a given Message Field setting, the following Message Field setting shall be the same Message Field setting or the Message Field setting corresponding to a row below the current setting. When loc_rcvr_status=OK the Infofield variable is set to loc_rcvr_status<5>=1 and set to 0 otherwise.

Table 113–11—Infofield message field valid MASTER settings

PMA_state<7:6>	loc_rcvr_ status	en_slave_tx	trans_to_ Coeff_Exch	Coeff_ exchange	trans_to_ Fine_Adjust	trans_to_ PCS_Test
00	0	0	0	0	0	0
00	0	1	0	0	0	0
01	0	1	0	0	0	0
01	0	1	1	0	0	0
10	0	1	0	0	0	0
10	0	1	0	1	0	0
10	0	1	0	0	0	0
10	0	1	0	0	1	0
11	0/1	1	0	0	0	0
11	1	1	0	0	0	1

Table 113–12—Infofield message field valid SLAVE settings

PMA_state<7:6>	loc_rcvr_ status	timing_lock _OK	trans_to_ Coeff_Exch	Coeff_ exchange	trans_to_ Fine_Adjust	trans_to_ PCS_Test
00	0	0	0	0	0	0
00	0	0/1	0	0	0	0
01	0	1	0	0	0	0
01	0	1	1	0	0	0
10	0	0/1	0	0	0	0
10	0	1	0	1	0	0
10	0	1	0	0	0	0
10	0	1	0	0	1	0
11	0	0/1	0	0	0	0
11	0/1	1	0	0	0	0
11	1	1	0	0	0	1

113.4.2.5.7 SNR_margin

SNR_margin (4 bits). Represented by Octet 9<7:4>, which reports received decision point SNR margin in 1/2 dB steps. SNR_margin is relative to the SNR required for reception of LDPC-coded DSQ128 at an LDPC frame error ratio of less than 3.2×10^{-9}. The SNR_margin<7:4> four-bit values, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110 shall indicate the decision point SNR margin

values of –1.5, –1, –0.5, 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5 dB, respectively. The value 0001 shall indicate a margin of –2 dB or less, and the value 1111 shall indicate 5 dB or more. Finally the value 0000 shall indicate that the SNR margin value is unknown.

113.4.2.5.8 Transition counter

Transition counter (10 bits). Represented by the 1.25 octets [Octet 9<1:0>, Octet 10<7:0>]. When configured as Transition counter (Coeff_exchange<2>=0 and a transition is announced to PMA_Coeff_Exch, PMA_Fine_Adjust or PCS_Test), this field is used as a 10-bit counter that counts the number of remaining frames until the next transition (PMA_Coeff_Exch, PMA_Fine_Adjust, PCS_Test).

113.4.2.5.9 Coefficient exchange handshake

Coefficient exchange handshake (12 bits). Represented by the 1.5 octets [Octet 9<3:0>, Octet 10<7:0>]. If Coeff_exchange<2>=1, this field is configured as a Coefficient exchange handshake and is used as a handshake control channel during programmable THP coefficient exchange. The details of the coefficient exchange are described in 113.4.2.5.15.

113.4.2.5.10 Ability Fields

Ability field (1 octet). Represented by Octet 12{EEE Ability<7>, THP Bypass Request<6>, Fast Retrain<5>, Reserved<4:0>}. Used to advertise the abilities of the PHY during the PMA_PBO_Exch state when Message<7:6> = 01.

For every other state, this octet is set to zero and ignored by the link partner. The Ability bits are defined as follows:

Octet 12<4:0> = Reserved

Octet 12<5> = Fast Retrain

0 = Fast Retrain not supported

1 = Fast Retrain supported

Octet 12<6> = THP Bypass Request in PMA_Coeff_Exchstate

0 = Local device requests link partner not to bypass THP during fast retrain

1 = Local device requests link partner to bypass THP during fast retrain

Octet 12<7> = EEE Ability

0 = EEE not supported

1 = EEE supported.

113.4.2.5.11 Reserved fields

All Infofield fields denoted Reserved in Figure 113–23, Figure 113–24, and Figure 113–25 are reserved for future use. This includes Octet 11 and Octet 12 when Coeff_exchange<2>=0 and Message<7:6>!=01, Octet 9<3:2> when transition counter is announced and [Octet 9<3:0>, Octet 10<7:0>] when no transition is announced and no coefficients are exchanged.

113.4.2.5.12 Vendor-specific field

If Coeff_exchange<2>=0 Octet 13 and Octet 14 are vendor-specific fields. If during Auto-Negotiation both transceivers agree on the use of the two vendor-specific octets, they may be used as a PHY communication

channel; otherwise they are set to zero and ignored by the link partner. Represented by Octet 13<7:0> and Octet 14<7:0>.

113.4.2.5.13 Coefficient Field

Coefficient Field (4 octets). Represented by Octet 11<7:0>, Octet 12<7:0>, Octet 13<7:0>, and Octet 14<7:0>. When Coeff_exchange<2>=1, this field is used to exchange programmable THP coefficients. It transmits four 8-bit THP coefficients out of the total of 64 (16 coefficients over each of the 4 pairs). The order is pair A, coefficients 0:3, followed by coefficients 4:7, followed by 8:11 and 12:15. For all cases the first coefficient (indices 0, 4, 8 and 12) is mapped to Octet 11, the second coefficient (indices 1, 5, 9, 13) is mapped to Octet 12 and so on. The same coefficient order is followed to transmit the coefficients for pair B, followed by pair C, and finally pair D. The details of the coefficient exchange are described in 113.4.2.5.15.

113.4.2.5.14 CRC16

CRC16 (2 octets). Shall implement the CRC16 polynomial $(x+1)(x^{15}+x+1)$ of the previous 10 octets, Octet 5<7:0>, Octet 6<7:0>, Octet 7<7:0>, Octet 8<7:0>, Octet 9<7:0>, Octet 10<7:0>, Octet 11<7:0>, Octet 12<7:0>, Octet 13<7:0>, and Octet 14<7:0>. The CRC16 shall produce the same result as the implementation shown in Figure 113–27. In Figure 113–27 the 16 delay elements S0,..., S15, shall be initialized to zero. Afterwards Octet 5 through Octet 14 are used to compute the CRC16 with the switch connected, which is setting CRCgen in Figure 113–27. After all the 10 octets have been processed, the switch is disconnected (setting CRCout) and the 16 values stored in the delay elements are transmitted in the order illustrated, first S15, followed by S14, and so on, until the final value S0.

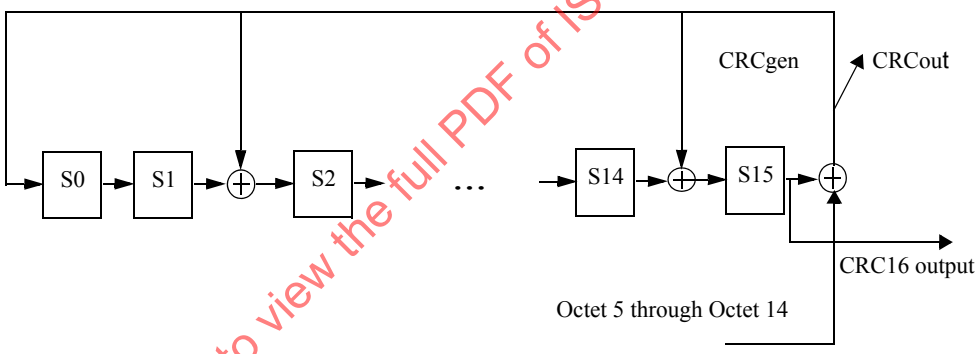


Figure 113–27—CRC16

113.4.2.5.15 Startup sequence

The startup sequence shall comply with the state diagram description given in Figure 113–28 and the transition counter state diagrams Figure 113–29 and Figure 113–30.

During Auto-Negotiation, PHY Control is in the DISABLE_25G/40GBASE-T_TRANSMITTER state and the transmitters are disabled. During normal training, prior to enabling the transmitter, the THP coefficients are set to zero.

When the Auto-Negotiation process asserts link_control=ENABLE, PHY Control enters the INIT_MAXWAIT_TIMER state. Upon entering this state, the maxwait_timer is started and PHY Control enters the SILENT state, which starts the minwait_timer and forces transmission of zeros by setting tx_mode=SEND_Z.

In MASTER mode, after expiration of the minwait_timer, PHY Control transitions to the PMA_Training_Init_M state.

Upon entering the PMA_Training_Init_M and PMA_Training_Init_S states, the PHY Control forces transmission into the training mode by asserting tx_mode=SEND_T, which includes the transmission of Infofields.

Upon entering state PMA_Training_Init_M, the MASTER starts transmission with a fixed transmit power level, PBO=4 (corresponding to a power backoff of 8 dB). The PBO variable is communicated to the link partner via the current transmitter octet of the Infofield.

Initially the MASTER is not ready for the SLAVE to respond and sets en_slave_tx=0, which is communicated to the link partner via the Infofield. After the MASTER has sufficiently converged the necessary circuitry, the MASTER sets en_slave_tx=1 to allow the SLAVE to transition to PMA_Training_Init_S.

In SLAVE mode, PHY Control transitions to the PMA_Training_Init_S state only after the SLAVE PHY acquires timing, converges its equalizers, acquires its descrambler state, and sets loc_SNR_margin=OK. The SLAVE shall respond using the fixed PBO transmit power level, PBO=4 (corresponding to a power backoff of 8 dB). For PHYs with the EEE capability, further requirements for this transition are described in 113.3.5.1.

While in states PMA_Training_Init_S, PMA_PBO_Exch, or PMA_Coeff_Exch, whenever a SLAVE operating in loop timing mode loses the MASTER timing reference (for example, after transmit power level transitions) it sets timing_lock_OK=0, which is communicated to the link partner via the Infofield. Otherwise, timing_lock_OK is set to one.

In MASTER mode, PHY Control enters the PMA_PBO_Exch state after loc_SNR_margin=OK and in SLAVE mode PHY Control enters the PMA_PBO_Exch state after the loc_SNR_margin=OK and minwait_timer expires. In the PMA_PBO_Exch state while Infofield Message<7:6> = 01, the PHY advertises EEE and Fast Retrain capability in octet 12 of the Infofield. When both the local device and remote device advertise EEE capability, then EEE is supported. When both the local device and remote device advertise Fast Retrain capability then Fast Retrain is supported. In the PMA_PBO_Exch state, after the MASTER has computed the final desired programmable PBO level, it shall request a PBO change using the requested transmitter setting in the Infofield (octet 7). In SLAVE mode, after the MASTER has requested the desired PBO level, the SLAVE shall request a desired PBO level that is within two levels (within 4 dB) of the requested MASTER PBO level. Both MASTER and SLAVE shall use the lower of the two PBO levels (i.e., that providing the larger transmit power).

Following PBO exchange for both transceivers, each PHY shall announce the next PBO setting using the next transmitter setting (octet 6). Afterwards, each PHY announces a transition to the PMA_Coeff_Exch state using the trans_to_Coeff_Exch=1 and transition_count as described in 113.4.5.1. MASTER initiates the transition to PMA_Coeff_Exch count with the trans_to_Coeff_Exch=1 flag and a transition counter value of 2^9. The SLAVE responds prior to the MASTER transition counter reaching 2^6 by setting trans_to_Coeff_Exch=1 flag and a transition counter value matching the MASTER. The PMA frame after each transceiver transition_count reaches zero, the PHYs shall enter the PMA_Coeff_Exch state and enable the requested PBO. Therefore, both PHYs will enter the PMA_Coeff_Exch state within one PMA frame.

While both MASTER and SLAVE are in state PMA_Coeff_Exch, when either end has computed the programmable THP settings, the programmable THP coefficient exchange process can begin, using the 1.5 octet Coefficient exchange handshake and the 4 octet Coefficient Field as follows:

a) During PMA_Coeff_Exch each PHY begins a coefficient exchange by setting the Coeff_Exchange flag to 1 in the Message Field.

b) During coefficient exchange, the transition counter bits are used as the Coefficient Exchange Handshake
   1) Octet 9{Reserved<3:0>}: unused
   2) Coefficient Pair Received, Octet 10<7:6>: 01 for local transmitter pair A, 10 for B, 11 for C, and 00 for D (default). This is the handshake to tell the remote unit the last coefficients received.
   3) Coefficient Group Received, Octet 10<5:4>: 01 for coefficients 0:3, 10 for 4:7, 11 for 8:11, and 00 for 12:15 (default). This is the handshake to tell the remote unit the last coefficients received.
   4) Coefficient Pair Sent, Octet 10<3:2>: 01 for remote transmitter pair A, 10 for B, 11 for C, and 00 for D (default). This is the handshake to tell the remote unit the current coefficients being sent.
   5) Coefficient Group Sent, Octet 10<1:0>: 01 for 0:3, 10 for 4:7, 11 for 8:11, and 00 for 12:15 (default). This is the handshake to tell the remote unit the current coefficients being sent.

c) The Coefficient Field is used to send four 8-bit coefficients in each frame designated by the Coefficient Pair Sent and Coefficient Group Sent bits. The coefficient format is:
   1) 8 bits per coefficient. Use one octet per coefficient in twos complement notation
   2) Coefficient range is –2.0 to 1.984375 in steps of 0.015625
   3) The sign of the coefficients shall be consistent with Equation (113–6)

d) Each PHY begins the exchange by sending pair A coefficients 0:3 with Coefficient Pair Sent=01 and Coefficient Group Sent=01.

e) The remote unit acknowledges by setting Coefficient Pair Received=01 and Coefficient Group Received=01.

f) Following each acknowledgement, the PHY increments through the Coefficient Group and then Coefficient Pair settings until Coefficient Pair Sent=00 and Coefficient Group Sent=00 and Coefficient Pair Received=00 and Coefficient Group Received=00. At this time, coefficient exchange is done and both PHYs set Coeff_Exchange=0.

Following coefficient exchange for both transceivers, each PHY announces a transition to the PMA_Fine_Adjust state (trans_to_Fine_Adjust=1) and starts the transition_count as described in 113.4.5.1. During the first PMA frame after the transition_count reaches zero, the PHYs enter the PMA_Fine_Adjust state and enable the THP precoders with the requested coefficients. At the closure of the THP feedback loop, the initial state of the THP feedback filters shall be the last 16 symbols from the state PMA_Coeff_Exch.

The THP coefficients and PBO setting are not changed during PMA_Fine_Adjust. The final convergence of the adaptive filter parameters is completed in the PMA_Fine_Adjust state.

After the PHY completes successful training and establishes proper receiver operations, PCS Transmit conveys this information to the link partner via transmission of the parameter Infofield value loc_rcvr_status. The link partner's value for loc_rcvr_status is stored in the local device parameter rem_rcvr_status. When the condition loc_rcvr_status=OK and rem_rcvr_status=OK is satisfied, each PHY announces a transition to the PCS_Test state (trans_to_PCS_Test=1) and start the transition counter as described in 113.4.5.1. For PHYs with the EEE capability, further requirements for this transition are described in 113.3.5.1.

The normal mode of operation corresponds to the PCS_Data state, where PHY Control asserts tx_mode=SEND_N and transmission of data over the link can take place.

PHY Control may force the transmit scrambler state to be initialized to an arbitrary value by requesting the execution of the PCS Reset function defined in 113.3.2.1.

The operation of the maxwait_timer requires that the PHY complete the startup sequence from state SILENT to PMA_Fine_Adjust in the PHY Control state diagram (Figure 113–28) in less than 2000 ms to

avoid link_status being changed to FAIL by the Link Monitor state diagram (Figure 113–31). To ensure interoperability the timing in Table 113–13 should be observed.

After reaching the PCS_Data state PHYs with the EEE capability can transition to the LPI receive mode under the control of the link partner and to the LPI transmit mode under control of the local LPI client.

Table 113–13—Recommended startup sequence timing

Master	Recommended maximum time (ms)	Recommended average time (ms)	Slave
SILENT plus (PMA_Training_Init_M state AND en_slave_tx = 0)	350	315	SILENT
(PMA_Training_Init_M state AND en_slave_tx = 1) plus PMA_PBO_Exch state	480	432	PMA_Training_Init_S state plus PMA_PBO_Exch state
PMA_Coeff_Exch state	100	90	PMA_Coeff_Exch state with timing_lock_OK=0
	520	468	Total for PMA Coeff Exch state
PMA_Fine_Adjust state	650	585	PMA_Fine_Adjust state
Total	2000	1800	

113.4.2.5.16 Fast retrain function

PHYs that support the fast retrain capability shall implement the fast retrain state diagram shown in Figure 113–33. PHYs may request a fast retrain by setting the variable loc_fr_req to TRUE. This causes the transmission of an easily detected link failure signal specified in 113.4.2.2.2. After completing the link failure signal the PHY shall transition to the PMA_INIT_FR state followed immediately by the PMA_Coeff_Exch state. If the link partner requested THP bypass for fast retrain the PHY shall bypass the THP (or set THP coefficients to zero). Otherwise the PHY shall keep its THP turned on with its previously exchanged coefficients, and send PAM2 signaling within a time period equivalent to 9 LDPC frame periods.

After the detection of the link failure signal, a PHY shall transition to the PMA_Coeff_Exch state and respond with PAM2 signaling within a time period equivalent to 9 LDPC frame periods after receiving the link failure signal.

The PAM2 symbols are generated using the PMA sidestream scrambler polynomials shown in Figure 113–15. The training sequence in 113.3.4 shall be used during fast retraining.

Note that reliable traffic on the transmitter may be interrupted when the local receiver requests a fast retrain.

Following the link failure signal, the two link partners transition back to the PMA_Coeff_Exch state and follow the training procedure described in 113.4.2.5.15, with the exception that the initial infofield countdown values are reduced as indicated in Figure 113–29 and Figure 113–30.

To ensure interoperability the training times in Table 113–14 should be observed during the fast retrain.

Table 113–14—Recommended fast retrain sequence timing

State	Recommended maximum time (ms)
PMA_Coeff_Exch state	20
PMA_Fine_Adjust state	10

113.4.2.6 Link Monitor function

Link Monitor determines the status of the receiver and communicates it via the variable link_status. Failure of the receiver typically causes the PMA's clients to suspend normal operation.

The Link Monitor function shall comply with the state diagram of Figure 113–31.

Upon power on, reset, or release from power down, the Auto-Negotiation algorithm sets link_control=SCAN_FOR_CARRIER and, during this period, sends fast link pulses to signal its presence to a remote station. If the presence of a remote station is sensed through reception of fast link pulses, the Auto-Negotiation algorithm sets link_control=DISABLE and exchanges Auto-Negotiation information with the remote station. During this period, link_status=FAIL is asserted. If the presence of a remote 25G/40GBASE-T station is established, the Auto-Negotiation algorithm permits full operation by setting link_control=ENABLE. As soon as reliable transmission is achieved, the variable link_status=OK is asserted, upon which further PHY operations can take place.

113.4.2.7 Refresh Monitor function

The Refresh Monitor is required for PHYs that support the EEE capability. The Refresh Monitor operates when the PHY is in the LPI receive mode. The Refresh Monitor shall comply with the state diagram of Figure 113–19b. The function forces a link retrain if a refresh signal is not reliably detected within a moving time window equivalent to 50 complete quiet-refresh cycles (nominally equal to 2.048 ms), when the PHY is in the lower power receive mode.

113.4.2.8 Clock Recovery function

The Clock Recovery function couples to all four receive pairs. It may provide independent clock phases for sampling the signals on each of the four pairs.

The Clock Recovery function shall provide clocks suitable for signal sampling on each line so that the LDPC FER indicated in 113.4.2.4 is achieved. The received clock signal should be stable and ready for use when training has been completed (loc_rcvr_status=OK). The received clock signal is supplied to the PMA Transmit function by received_clock.

113.4.3 MDI

Communication through the MDI is summarized in 113.4.3.1 and 113.4.3.2.

113.4.3.1 MDI signals transmitted by the PHY

The symbols to be transmitted by the PMA on the four pairs BI_DA, BI_DB, BI_DC, and BI_DD are denoted by tx_symb_vector[BI_DA], tx_symb_vector[BI_DB], tx_symb_vector[BI_DC], and tx_symb_vector[BI_DD], respectively. The modulation scheme used over each pair is PAM16. PMA Transmit generates a pulse-amplitude modulated signal on each pair in the following form:

$$x_n = M\left(a_n - \sum_{k=1}^{16} x_{n-k} c_k\right) = a_n + 32m_n - \sum_{k=1}^{16} x_{n-k} c_k \quad (113\text{–}6)$$

$$s(t) = \sum_{n=0}^{\infty} x_n h_T(t - nT) \quad (113\text{–}7)$$

In Equation (113–6), a_n is the PAM16 modulation symbol from the set {–15, –13, –11, –9, –7, –5, –3, –1, 1, 3, 5, 7, 9, 11, 13, 15} to be transmitted at time nT. Each of the 16 THP coefficients c_1, c_2,..., c_{16} per wire pair is represented in two's complement form by 8 bits described in 113.4.2.5. The nonlinear THP operation given by $M(\alpha) = (\alpha + 16) mod 32 - 16$ corresponds to changing the modulation symbol a_n to an augmented modulation symbol $\tilde{a}_n = a_n + 32m_n$ with the integer m_n chosen such that the THP output lies in the interval $-16 \le x_n < 16$. Equation (113–7) describes the convolution of the THP output signals with the transmitter symbol response $h_T(t)$ to obtain the transmit signal $s(t)$ at the MDI. The values of the programmable THP coefficients are exchanged in the Infofield during PMA_Coeff_Exch. The THP filter coefficients shall be fixed after startup.

The nominal power (denoted Ptx) and the symbol response of the PMA transmitted signal $s(t)$ shall comply with the electrical specifications given in 113.5. When the link segment does not experience the maximum insertion loss (IL), each transceiver indicates to the link partner that the link partner PMA Transmit signal shall be reduced in increments of 2 dB. The minimum power backoff level requested shall comply with the power backoff schedule in Table 113–15. If a given receiver has sufficient decision point SNR margin, it may choose to request from the link partner larger power backoff (up to 14 dB) than shown in Table 113–15. Additionally, the Slave shall select a PBO level as described in the PMA_PBO_Exch state of 113.4.2.5.15. The PMA Transmit shall be capable of eight power backoff settings in approximately 2 dB steps. The difference between each consecutive power setting shall be 2 ± 0.25 dB, and each step shall be centered at $2 \times n$ dB (n = 0 to 7) reduction from nominal, with a maximum error of ± 1 dB.

The received signal power at the MDI, P (dBm), in Table 113–15, should be the estimate of the average received power across all four pairs from the remote transmitter when the link partner PMA Transmit is at nominal power (after accounting for local transmitter power). If the remote transmitter is not at nominal power during the measurement, the estimate of the received power should be incremented by the amount of power backoff of the link partner transmitter during the measurement. Nominal power refers to the transmit power without any power backoff and is specified in 113.5.3.4. The estimate of the received signal power is stored in registers 1.141 to 1.144 as described in 45.2.1. The values in the length, L (m), column in Table 113–15 are for reference only (not required for power backoff evaluation).

Table 113–15—Power backoff schedule

Received signal power at MDI, P (dBm)	Length L(m) (reference)	Minimum power backoff (dB)
$-7.2 < P$	$0 \le L < 13$	6
$-9.4 < P \le 7.2$	$13 \le L < 19$	4
$-11.4 < P \le -9.4$	$19 \le L < 25$	2
$P \le -11.4$	$25 \le L$	0

113.4.3.2 Signals received at the MDI

Signals received at the MDI can be expressed for each pair as pulse-amplitude modulated signals that are corrupted by noise as follows:

$$r(t) = \sum_{n=0}^{\infty} \tilde{a}_n h_R(t - nT) + w(t) \quad (113\text{–}8)$$

In Equation (113–8), $\tilde{a}_n$ are the augmented PAM16 modulation symbols described in 113.4.3.1, $h_R(t)$ denotes the symbol response of the overall signal path from the THP precoder to the MDI at the receiver, and $w(t)$ represents the contribution of various noise sources including uncancelled crosstalk. The four signals received on pairs BI_DA, BI_DB, BI_DC, and BI_DD are processed within the PMA Receive function to yield the received symbols rx_symb_vector.

113.4.4 Automatic MDI/MDI-X configuration

Automatic MDI/MDI-X configuration is intended to eliminate the need for crossover cables between similar devices. Automatic MDI/MDI-X configuration is required for 25G/40GBASE-T devices and shall comply with 40.4.4.1 and 40.4.4.2.

Having established MDI/MDI-X configuration, the receiver shall detect and correct for several configurations of pair swaps and crossovers and arbitrary polarity swaps. The receiver pairs BI_DA, BI_DB, BI_DC, and BI_DD might be connected to the corresponding transmit pairs in any of the following ways with arbitrary polarity:

a) No crossover
b) A/B crossover only
c) C/D crossover only
d) A/B crossover and C/D crossover

For EEE-capable PHYs, the MDI/MDIX function configuration shall apply to refresh and alert signaling. For PHYs with the fast retrain capability, the MDI/MDIX function configuration shall apply to link failure signaling.

113.4.5 State variables

113.4.5.1 State diagram variables

coeff_exchange_done
: This variable reports that both transceivers have received the corresponding coefficients from the link partner.
Values: TRUE: The coefficient exchange has completed.
FALSE: The coefficient exchange has not completed.

config
: The PMA shall generate this variable continuously and pass it to the PCS via the PMA_CONFIG.indication primitive.
Values: MASTER or SLAVE

link_control
: The link_control parameter generated by Auto-Negotiation and passed to the PMA via the PMA_LINK.request primitive (see 113.2.1.1)

link_status
: The link_status parameter set by PMA Link Monitor state diagram and communicated through the PMA_LINK.indication primitive.
Values: OK or FAIL

loc_rcvr_status
: Variable set by the PMA Receive function to indicate correct or incorrect operation of the receive link for the local PHY.
Values: OK: The receive link for the local PHY is operating reliably.
NOT_OK: Operation of the receive link for the local PHY is unreliable.

loc_SNR_margin

This variable reports whether the local device has sufficient SNR margin to continue to the next state. The criterion for setting the parameter loc_SNR_margin is left to the implementer.

Values: OK: The local device has sufficient SNR margin.
NOT_OK: The local device does not have sufficient SNR margin.

master_transition_counter

This variable reports the current value of the MASTER's transition counter reported in the Infofield defined in 113.4.2.5.

Values: 0 to 2^9

MessageField_IF

This variable reports that a receiver has successfully received and decoded the Infofield from the remote device. This variable takes on the value contained in the Message Field. If the Message Field cannot be decoded or no explicit action is outstanding the value Null is returned.

Values: trans_to_Coeff_Exch, trans_to_Fine_Adjust, trans_to_PCS_Test or Null

PBO

PBO is a variable that can take any integer value from 0 to 7 and indicates the power backoff level. Denoting Ptx as the maximum nominal power, the PBO values are:

Values: 0, 1, 2, 3, 4, 5, 6, 7, which correspond to transmit power levels of *Ptx*, *Ptx*–2 dB, *Ptx*–4 dB, *Ptx*–6 dB, *Ptx*–8 dB, *Ptx*–10 dB, *Ptx*–12 dB, *Ptx*–14 dB, respectively

PBO_next

PBO_next is a variable that can take any integer value from 0 to 7 and indicates the next power backoff level to be used at the local transmitter. The value is taken from the fixed set of values during PMA_Training_Init_M and PMA_Training_Init_S as described in 113.4.2.5. The value is taken from the decoded value of the link partner Infofield during PMA_PBO_Exch

Values: 0, 1, 2, 3, 4, 5, 6, 7, which correspond to transmit power levels of *Ptx*, *Ptx*–2 dB, *Ptx*–4 dB, *Ptx*–6 dB, *Ptx*–8 dB, *Ptx*–10 dB, *Ptx*–12 dB, *Ptx*–14 dB, respectively

PBO_tx

PBO_tx is a variable that can take any integer value from 0 to 7 and indicates the power backoff level currently used at the local transmitter.

Values: 0, 1, 2, 3, 4, 5, 6, 7, which correspond to transmit power levels of *Ptx*, *Ptx*–2 dB, *Ptx*–4 dB, *Ptx*–6 dB, *Ptx*–8 dB, *Ptx*–10 dB, *Ptx*–12 dB, *Ptx*–14 dB, respectively

PBO_exchange_done

This variable reports that both transceivers have received the corresponding PBO levels from the link partner.

Values: TRUE: The PBO exchange has completed.
FALSE: The PBO exchange has not completed.

pcs_status

The pcs_status parameter generated by the PCS and passed to the PMA via the PMA_PCSTATUS.request primitive (see 113.2.2.6).

pma_reset

Allows reset of the PHY Control and Link Monitor state diagrams.

Values: ON or OFF

rem_rcvr_status

Variable set by the PCS Receive function to indicate whether correct operation of the receive link for the remote PHY is detected or not.

Values: OK: The receive link for the remote PHY is operating reliably.
NOT_OK: Reliable operation of the receive link for the remote PHY is not detected.

THP_next

THP_next is a variable that contains sixteen eight-bit values and describes the next transmitter setting of the THP coefficients. It refers to the programmable THP coefficients selected during coefficient exchange described in 113.4.2.5.

Values: 16 coefficients of 8-bit values each. Range is –2.0 to 1.984375 in steps of 0.015625

THP_tx

THP_tx is a variable that contains sixteen 8-bit values and describes the current transmitter setting of the THP coefficients. It refers to the programmable THP coefficients selected during coefficient exchange described in 113.4.2.5.

Values: 16 coefficients of eight-bit values each. Range is –2.0 to 1.984375 in steps of 0.015625

trans_to_Coeff_Exch

Message field variable defined in 113.4.2.5 that flags a transition by the local device to the PMA_Coeff_Exch state.

Values: 1: The local device transitions to the PMA_Coeff_Exch state on expiration of the transition counter.

0: The local device does not transition to the PMA_Coeff_Exch state.

trans_to_Fine_Adjust

Message field variable defined in 113.4.2.5 that flags a transition by the local device to the PMA_Fine_Adjust state.

Values: 1: The local device transitions to the PMA_Fine_Adjust state on expiration of the transition counter.

0: The local device does not transition to the PMA_Fine_Adjust state.

trans_to_PCS_Test

Message field variable defined in 113.4.2.5 that flags a transition by the local device to the PCS_Test state.

Values: 1: The local device transitions to the PCS_Test state on expiration of the transition counter.

0: The local device does not transition to the PCS_Test state.

transition_count

This variable reports the value of the transition counter contained in the Infofield sent to the remote device. Transition_count must comply with the state diagram description given in 113.4.6.2. When the message field contains a flag for a state transition, the transition counter denotes the remaining number of Infofield until the next state transition. MASTER initiates the transition to PMA_Coeff_Exch count with the trans_to_Coeff_Exch=1 flag and a counter value of 2^9. The SLAVE responds prior to the counter reaching 2^6 with the same flag and a count value matching the MASTER. Then both PHYs transition to PMA_Coeff_Exch within one PMA frame. The same sequence is performed in the transition to PMA_Fine_Adjust state and PCS_Test state using the trans_to_Fine_Adjust=1 and trans_to_PCS_Test=1 flags, respectively. In EEE-capable PHYs, synchronization of the PMA frames is tightly controlled as described in 113.3.5.1. When the message field does not contain a flag for a state transition, the transition counter is set to zero and ignored by the receiver.

Values: 0 to 2^9

tx_mode

PCS Transmit sends code-groups according to the value assumed by this variable.

Values: SEND_N: This value is continuously asserted when transmission of sequences of code-groups representing a 25GMII/XLGMII data stream take place.

SEND_T: This value is continuously asserted when transmission of sequences of code-groups representing the sequences of code-groups (TA_n, TB_n, TC_n, TD_n) defined in 113.3.4.2 is to take place.

SEND_Z: This value is asserted when transmission of zero code-groups is to take place.

The following variables are required only for PHYs that support the EEE capability:

lpi_refresh_detect
: Set TRUE when the receiver has reliably detected refresh signaling and FALSE otherwise. The exact criteria left to the implementer.

pcs_data_mode
: Generated by the PMA PHY Control function and indicates whether or not the local PHY may transition its PCS state diagrams out of their initialization states. The current value of the pcs_data_mode is passed to the PCS via the PMA_PCSDATAMODE.indicate primitive. In the absence of the optional EEE and fast retrain capabilities, the PHY operates as if the value of this variable is TRUE.

mtc
: mtc is the transition count for a MASTER PHY during normal training and fast retraining. mtc shall be equal to 2^9 for normal training and 2^5 for fast retrain.

stc
: stc is the transition count for a SLAVE PHY during normal training and fast retraining. stc shall be equal to 2^6 for normal training and 2^4 for fast retrain.

The following six variables are required only for PHYs that support the fast retrain capability:

fr_enable
: This variable is set to TRUE if fast retrain is supported. The variable is set to FALSE otherwise. If MDIO is supported, this variable is based on the value of 1.147.0. with the value of TRUE corresponding to 1.147.0 set to 1. If MDIO is not supported, an equivalent method of controlling fast retrain functionality should be provided.

loc_fr_req
: Set TRUE when the receiver has detected a link failure condition and is requesting a fast retrain; set FALSE otherwise.

loc_fr_detect
: Set TRUE when the receiver has reliably detected the link failure signal. It is highly recommended that loc_fr_detect is qualified with the reception of errored blocks at the LDPC decoder output. Set FALSE when the link failure signal is not detected.

send_link_fail
: When TRUE indicates that the PMA should send the link failure signal. When FALSE the variable has no effect.

fr_active
: Set TRUE when the PHY is performing a fast retrain and set FALSE otherwise.

fast_retrain_flag
: Set TRUE after the PHY generates or detects a link failure signal and set FALSE otherwise.

113.4.5.2 Timers

All timers operate in the manner described in 14.2.3.2.

maxwait_timer
: A timer used to limit the amount of time during which a receiver dwells in the SILENT and TRAINING states. The timer shall expire 2000 ms ± 10 ms after being started. This timer is used jointly in the PHY Control and Link Monitor state diagrams. The maxwait_timer is tested by the Link Monitor to force link_status to be set to FAIL if the timer expires and loc_rcvr_status is NOT_OK. See Figure 113–28 and Figure 113–31.

minwait_timer
: A timer used to determine the minimum amount of time the PHY Control stays in the SILENT, PMA_Training_Init_S, PCS_Test and PCS_Data states. The timer shall expire 1 ms ± 0.1 ms after being started.

The following timer is required only for PHYs that support the EEE capability:

lpi_refresh_rx_timer
: This timer is used to monitor link quality during the LPI receive mode. If the PHY does not reliably detect reliable refresh signaling before this timer expires then a full retrain is performed.
Values: The condition lpi_refresh_rx_timer_done becomes true upon timer expiration.
Duration: This timer shall have a period equal to 50 complete quiet-refresh signal periods, equivalent to 2.048× *S* ms.

The following two timers are required only for PHYs that support the fast retrain capability:

link_fail_sig_timer
: Determines the period of time the PHY sends the link failure signal.
Values: The condition link_fail_sig_timer_done becomes true upon timer expiration.
Duration: This timer shall have a period equal to 4 LDPC frame periods.

fr_maxwait_timer
: Determines the period of time the PHY has to transition its PCS Control State to PCS_Test following a fast retrain before the fast retrain is aborted and a full retrain performed.
Values: The condition fr_maxwait_timer_done becomes true upon timer expiration.
Duration: This timer shall have a period equal to 30 ms.

113.4.5.3 Functions

Exchange_Final_PBO
: This function transmits and receives the final PBO settings using the Infofield as described in 113.4.2.5.

Exchange_THP_Coefficients
: This function compiles and sends to the link partner and receives from the link partner the desired programmable THP coefficients using the Infofield as described in 113.4.2.5.

113.4.5.4 Counters

The following two counters are required only for PHYs that support the fast retrain capability:

fr_tx_counter
: Counts the number of times the PHY initiates a fast link retrain by transmitting the link failure signal. This counter is reflected in MDIO register 1.147.10:6 specified in 45.2.1.79.2.

fr_rx_counter
: Counts the number of times the PHY begins a fast link retrain in response to the detection of link failure signalling from the link partner. This counter is reflected in MDIO register 1.147.15:11 specified in 45.2.1.79.1.

113.4.6 State diagrams

113.4.6.1 PHY Control state diagram

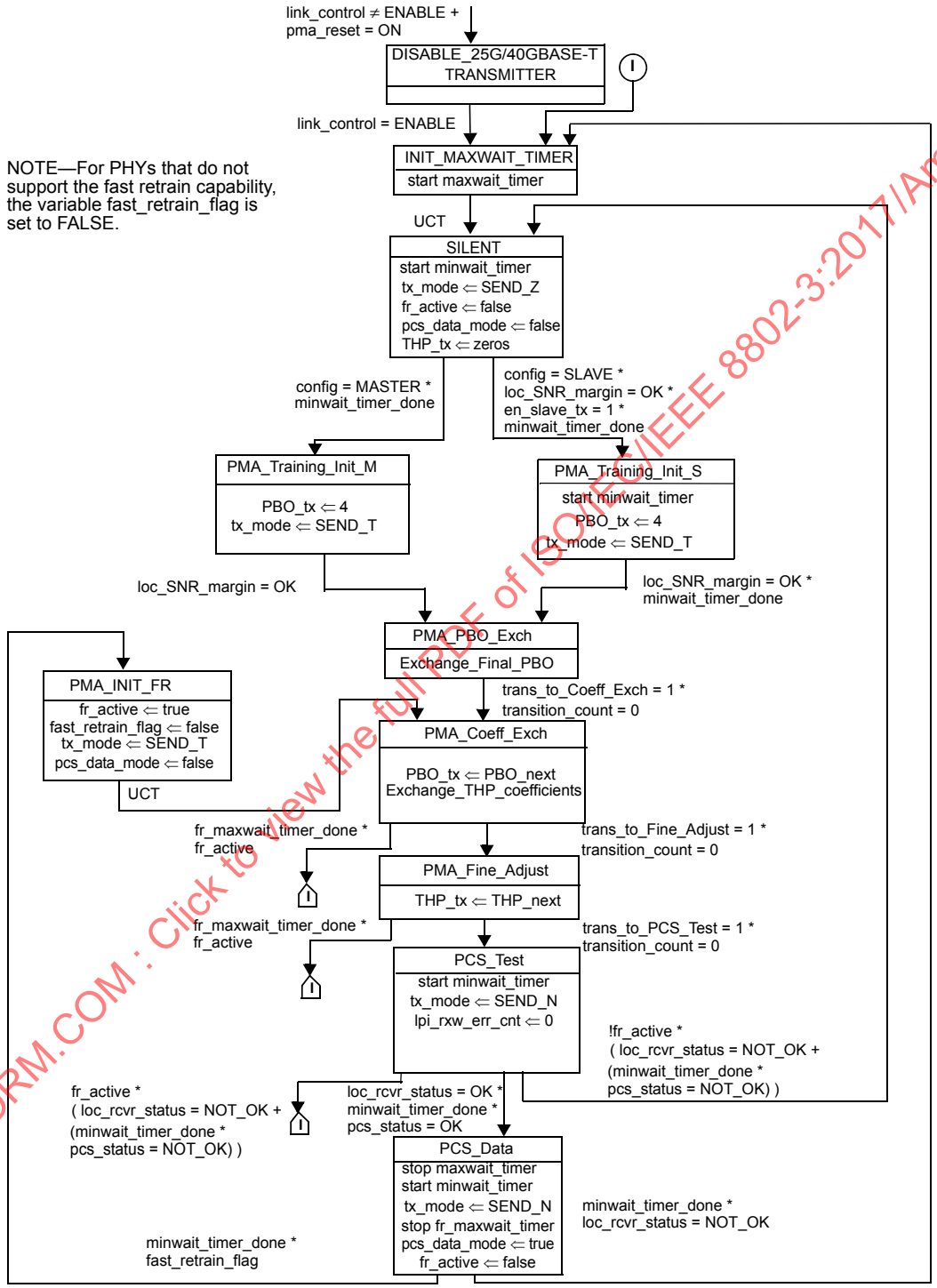


Figure 113–28—PHY Control state diagram

113.4.6.2 Transition counter state diagrams

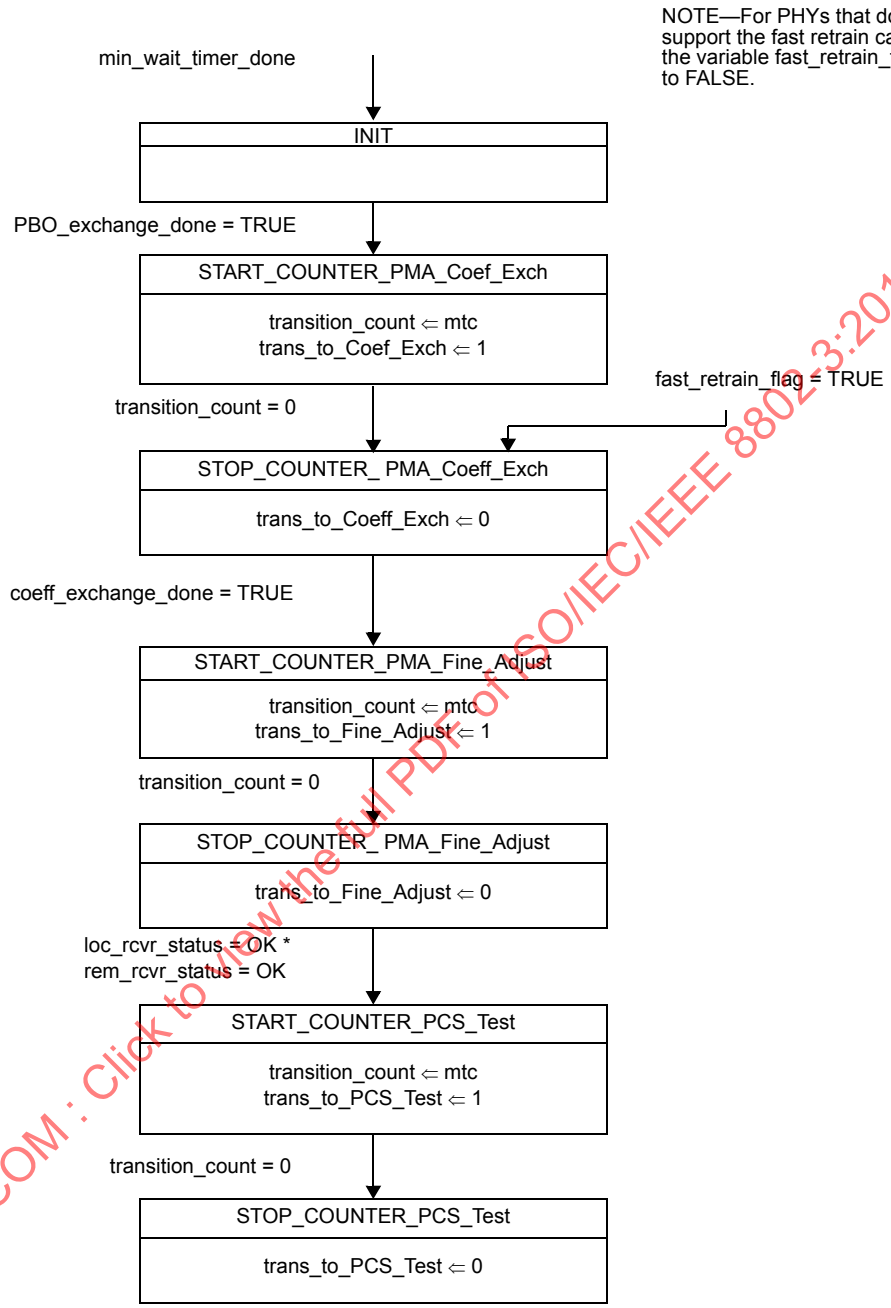


Figure 113–29—MASTER transition counter state diagram

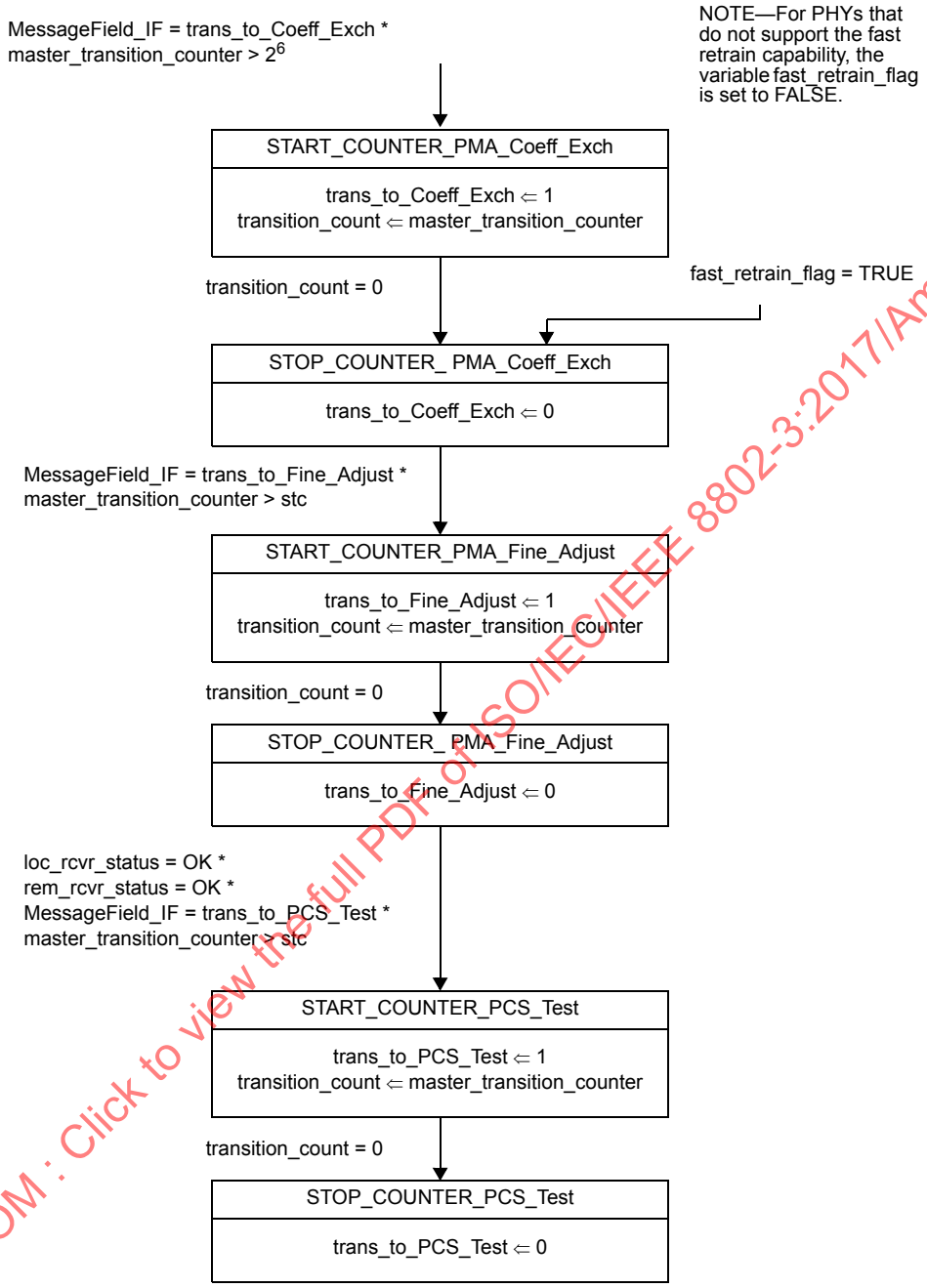
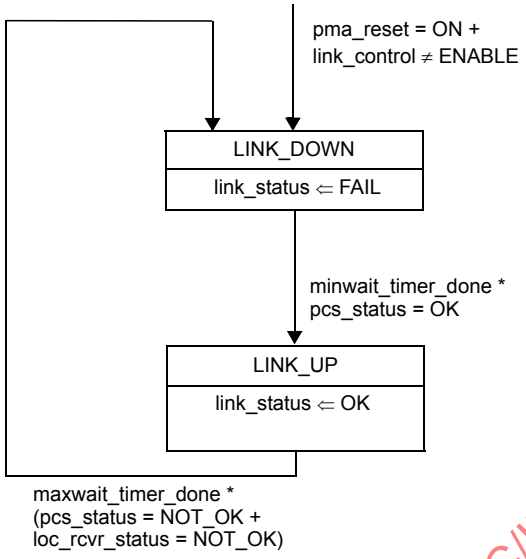


Figure 113–30—SLAVE transition counter state diagram

113.4.6.3 Link Monitor state diagram

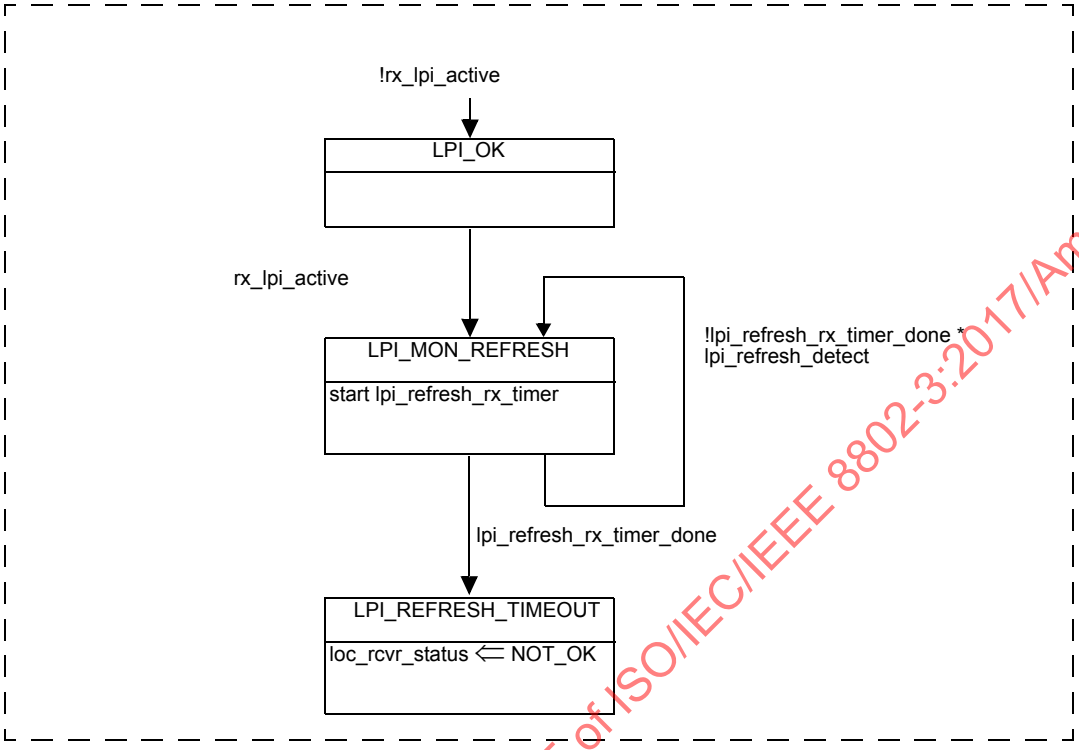


NOTE 1—maxwait_timer is started in PHY Control state diagram (see Figure 113–28).

NOTE 2—The variables link_control and link_status are designated as link_control_25GigT and link_status_25GigT (for 25GBASE-T) or link_control_40GigT and link_status_40GigT (for 40GBASE-T), respectively, by the Auto-Negotiation Arbitration state diagram (Figure 28-16).

Figure 113–31—Link Monitor state diagram

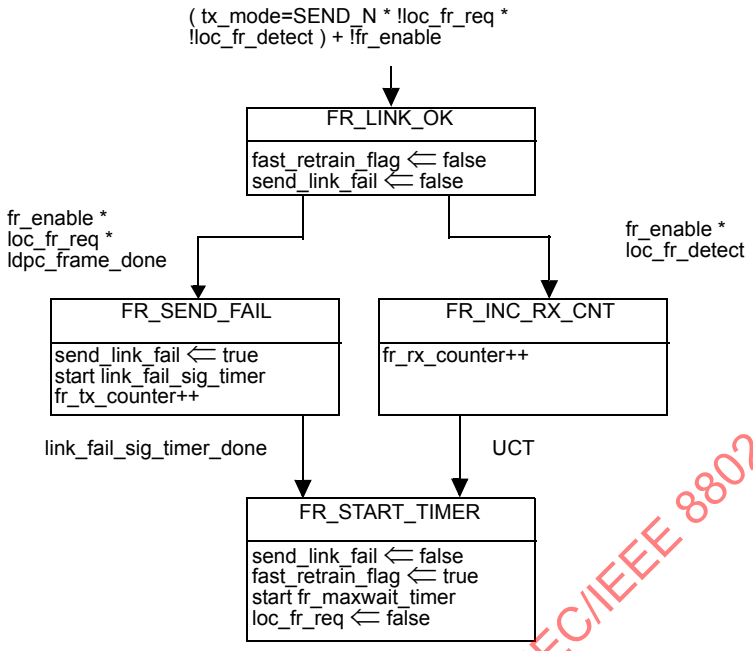
113.4.6.4 EEE Refresh monitor state diagram



NOTE—This state diagram is only required when the PHY supports the EEE capability.

Figure 113–32—EEE Refresh monitor state diagram

113.4.6.5 Fast retrain state diagram



NOTE—This state diagram is only required when the PHY supports the fast retrain capability.

Figure 113–33—Fast retrain control state diagram

113.5 PMA electrical specifications

This subclause defines the electrical characteristics of the PMA and specifies PMA-to-MDI interface tests.

113.5.1 Isolation requirement

The PHY shall provide electrical isolation between the port device circuits, including frame ground (if any) and all MDI leads. This electrical isolation shall withstand at least one of the following electrical strength tests:

a) 1500 V rms at 50 Hz to 60 Hz for 60 s, applied as specified in 5.2.2 of IEC 60950-1:2001.
b) 2250 V dc for 60 s, applied as specified in 5.2.2 of IEC 60950-1:2001.
c) A sequence of ten 2400 V impulses of alternating polarity, applied at intervals of not less than 1 s. The shape of the impulses is 1.2/50 µs (1.2 µs virtual front time, 50 µs virtual time or half value), as defined in Annex N of IEC 60950-1:2001.

There shall be no insulation breakdown, as defined in 5.2.2 of IEC 60950-1:2001, during the test. The resistance after the test shall be at least 2 MΩ, measured at 500 V dc.

113.5.2 Test modes

The test modes described below shall be provided to allow for testing of the transmitter waveform, transmitter distortion, transmitted jitter, transmitter droop, and BER testing.

For a PHY with an MDIO management interface, these modes shall be enabled by setting bits 1.132.15:13 (MultiGBASE-T test mode register) of the MDIO Management register set as shown in Table 113–16. These test modes shall only change the data symbols provided to the transmitter circuitry and shall not alter the electrical and jitter characteristics of the transmitter and receiver from those of normal (non-test mode) operation. PHYs without a MDIO shall provide a means to enable these modes for conformance testing.

Table 113–16—MDIO management register settings for test modes

1.132.15	1.132.14	1.132.13	Mode
0	0	0	Normal operation.
0	0	1	Test mode 1—Setting of MASTER transmitter required by SLAVE for transmit jitter test in SLAVE mode.
0	1	0	Test mode 2—Transmit jitter test in MASTER mode.
0	1	1	Test mode 3—Transmit jitter test in SLAVE mode.
1	0	0	Test mode 4—Transmit distortion test.
1	0	1	Test mode 5—Normal operation with no power backoff. This is for the PSD mask and power level test.
1	1	0	Test mode 6—Transmitter droop test mode.
1	1	1	Test mode 7—Pseudo random test mode for BER Monitor.

Test mode 1 is a mode provided for enabling testing of timing jitter on a SLAVE transmitter. When test mode 1 is enabled, the PHY shall transmit the PMA training pattern (PRBS 33) continually from all four transmitters with the THP turned off, with no power backoff and with the transmitted symbols timed from its local clock source.

Test mode 2 is for transmitter jitter testing when transmitter is in MASTER timing mode. When test mode 2 is enabled, the PHY shall transmit {two +16 symbols followed by two –16 symbols} continually from all four transmitters with the THP turned off, with no power backoff and with the transmitted symbols timed from its local clock source. The transmitter output is a 800 × *S* MHz signal.

When test mode 3 is enabled on a PHY, the PHY shall transmit, with THP turned off, the data symbol sequence {two +16 symbols followed by two –16 symbols} repeatedly on pair D with the symbols timed from its recovered receive data clock in SLAVE timing mode. A PHY operates in test mode 3 when there is no input signal on pair D. The transmitter output is a 800 × *S* MHz signal on pair D and shall be silence on pairs A, B, and C.

Test mode 4 is for transmitter nonlinear distortion testing. When test mode 4 is enabled, the PHY shall transmit, with the THP turned off, transmitted symbols, timed from a transmit clock (as specified in 113.5.3.5) in the MASTER timing mode, defined by the bits1.132.12:10 and Table 113–17.

Table 113–17—MDIO management register settings for transmit frequencies in test mode 4

1.132.12	1.132.11	1.132.10	Output waveform frequencies in MHz
			Two tone frequency pairs
0	0	0	Reserved
0	1	1	Reserved
1	1	1	Reserved
0	0	1	(3200 × *S*/1024) × 47, (3200 × *S*/1024) × 53
0	1	0	(3200 × *S*/1024) × 101, (3200 × *S*/1024) × 103
1	0	0	(3200 × *S*/1024) × 179, (3200 × *S*/1024) × 181
1	0	1	(3200 × *S*/1024) × 277, (3200 × *S*/1024) × 281
1	1	0	(3200 × *S*/1024) × 397, (3200 × *S*/1024) × 401

The peak-to-peak levels used in this test shall correspond to the ± 16 symbol levels and the relative amplitudes of the tones in a two-tone pair shall be within 0.5 dB of each other.

Test mode 5 is for checking whether the transmitter is compliant with the transmit PSD mask and the transmit power level. When test mode 5 is enabled, the PHY shall transmit as in normal operation but with the power backoff disabled.

Test mode 6 is for testing transmitter droop. When test mode 6 is enabled, the PHY shall transmit the following sequence of data symbols A_n, B_n, C_n, D_n, of 113.4.3.1 continually from all four transmitters, with the THP turned off:

{One hundred twenty eight +16 followed by one hundred twenty eight –16 symbols}.

Test mode 7 is for enabling measurement of the bit error ratio of the link including the LDPC encoder/decoder, the transmit and receive analog front ends of the PHY and a cable connecting two PHYs. This mode shall reuse the 25G/40GBASE-T scrambler and is defined in detail in 113.3.3.

113.5.2.1 Test fixtures

The following fixtures (illustrated by Figure 113–34, Figure 113–35, and Figure 113–36), or their functional equivalents, can be used for measuring the transmitter specifications described in 113.5.3.

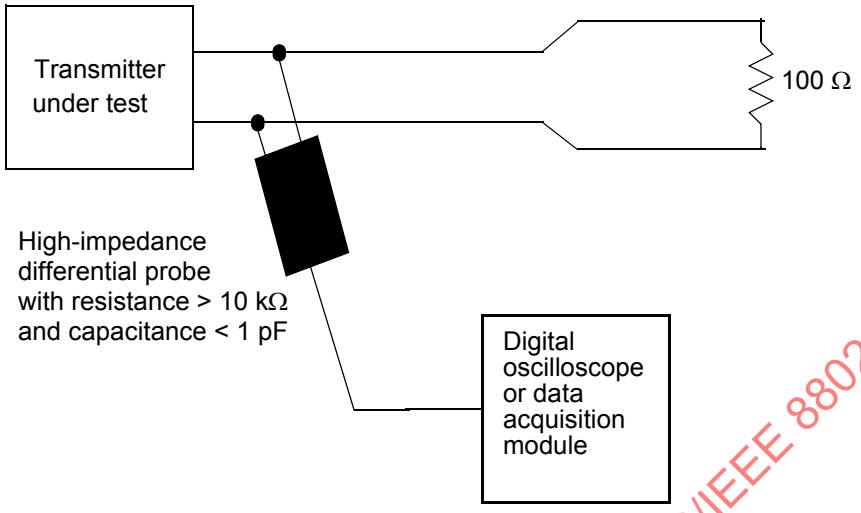


Figure 113–34—Transmitter test fixture 1 for transmitter droop measurement

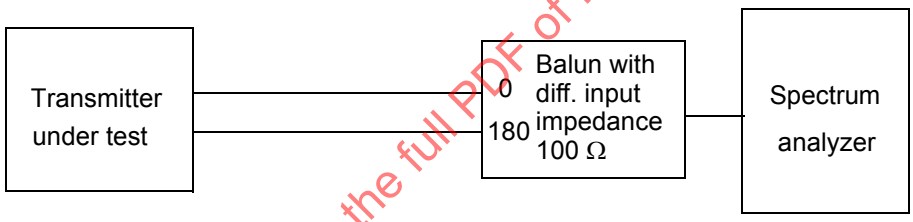


Figure 113–35—Transmitter test fixture 2 for linearity measurement, power spectral density measurement, and transmit power level measurement

The high-impedance probe shown in Figure 113–34 in transmitter test fixture 1 has resistance > 10 KΩ and capacitance < 1 pF over the frequency range of 1 MHz to 1600 MHz. Figure 113–35 includes a power summer or balun device to couple the 100 Ω differential output of the transmitter to the 50 Ω single-ended input typically found in a spectrum analyzer input. The center frequency (F_c) of the band pass filter show in Figure 113–36 is $800 \times S$ MHz ± 200 kHz and the band pass filter noise bandwidth (B_n) is 2 MHz ± 200 kHz.

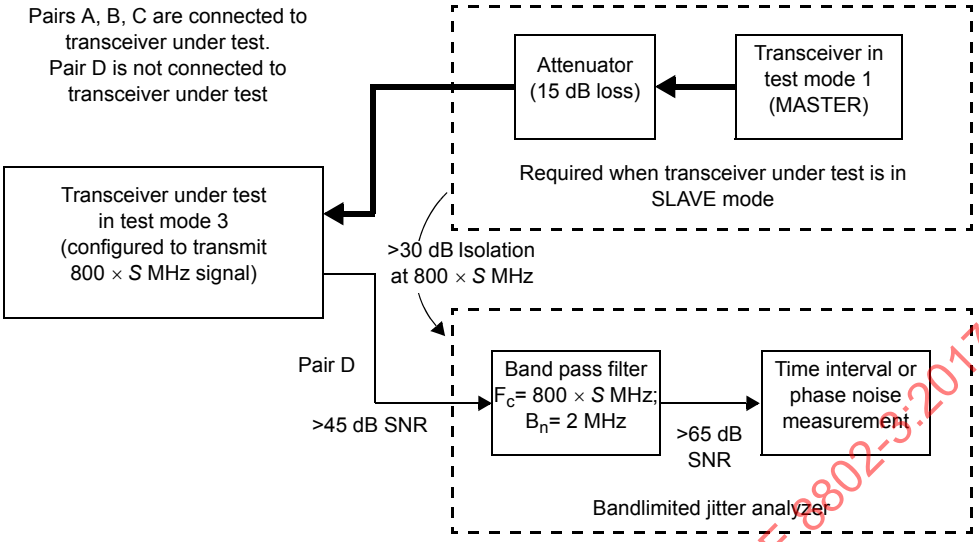


Figure 113–36—Transmitter test fixture 3 for transmitter jitter measurement

113.5.3 Transmitter electrical specifications

The PMA provides the Transmit function specified in 113.4.2.2 in accordance with the electrical specifications of this clause. The PMA shall operate with AC-coupling to the MDI.

Where a load is not specified, the transmitter shall meet the requirements of this clause with a 100 Ω resistive differential load connected to each transmitter output.

113.5.3.1 Maximum output droop

With the transmitter in test mode 6 and using the transmitter test fixture 1, the magnitude of both the positive and negative droop shall be less than 10%, measured with respect to an initial value at 2.5/*S* ns after the zero crossing and a final value at 22.5/*S* ns after the zero crossing.

113.5.3.2 Transmitter nonlinear distortion

When in test mode 4 and observing the spectrum of the differential signal output at the MDI using transmitter test fixture 2, for each pair, with no intervening cable, the transmitter nonlinear distortion mask is defined as follows:

The SFDR of the transmitter, with dual tone inputs as specified in test mode 4, shall meet the requirement shown in Equation (113–9) that:

$$\text{SFDR} \geq \quad 2.5 + \ \min\{52, 58 - 20 \times \log_{10}(f/(100 \times S))\} \tag{113–9}$$

where f is the maximum frequency of the two test tones in MHz, and SFDR is the ratio in dB of the minimum RMS value of either input tone to the RMS value of the worst intermodulation product in the frequency range of 1 MHz to 1600 × S MHz.

This specification on transmit linearity is derived from the requirement for interoperability with the far-end device.

113.5.3.3 Transmitter timing jitter

When in test mode 2, the PHY transmits {two +16 symbols followed by two –16 symbols} continually with the THP turned off and with no power backoff. In this mode, the transmitter output should be a 800 × S MHz signal and the RMS period jitter measured at the PHY MDI output shall be less than 1.3/S ps. The RMS period jitter is measured as per the test configuration shown in Figure 113–36 over an integration time interval of 1 ms ± 10%.

The SLAVE mode RMS period jitter test is measured using the test configuration shown in Figure 113–36. For this test, the MASTER PHY is in test mode 1 and the SLAVE PHY is in test mode 3. The MASTER is transmitting the PMA training pattern (PRBS 33) to the SLAVE PHY on pairs A, B, and C. The SLAVE PHY is in loop timing mode, synchronizing its transmit clock to the signals received from the MASTER PHY on pairs A, B, and C. In this configuration, the transmitter output on pair D should be a 800 × S MHz signal and the RMS period jitter measured at the SLAVE PHY MDI output shall be less than 1.3/S ps. The RMS period jitter is measured over an integration time interval of 1 ms ± 10%.

RMS period jitter over an integration time interval of 1 ms ± 10% is defined as the root mean square period difference from the average period ($T - T_{avg}$), accumulated over a sample size of 200 000 ± 20 000, as shown in Equation (113–10):

$$\text{RMS period jitter} = \sqrt{\frac{\sum[(T - T_{avg})^2]}{\text{Sample size}}} \tag{113–10}$$

113.5.3.4 Transmitter power spectral density (PSD) and power level

In test mode 5 (normal operation with no power backoff), the transmit power shall be in the range –1.0 dBm to 1.0 dBm and the power spectral density of the transmitter, measured into a 100 Ω, load using the test fixture shown in Figure 113–35 shall be between the upper and lower masks specified in Equation (113–11) and Equation (113–12). The masks are shown graphically in Figure 113–37.

$$\text{Upper PSD}(f) \le \begin{cases} -88.5 - 10\log_{10}S \text{ dBm/Hz} & 0 < \frac{f}{S} \le 280 \\ -88.5 - 10\log_{10}S - \left(\frac{\frac{f}{S} - 280}{320}\right) \text{dBm/Hz} & 280 < \frac{f}{S} \le 600 \\ -89.5 - 10\log_{10}S - \left(\frac{\frac{f}{S} - 600}{232}\right) \text{dBm/Hz} & 600 < \frac{f}{S} \le 2920 \\ -89.5 - 10\log_{10}S - \left(\frac{\frac{f}{S} - 1320}{160}\right) \text{dBm/Hz} & 2920 < \frac{f}{S} \le 7160 \\ -126 \text{ dBm/Hz} & 7160 < \frac{f}{S} \le \frac{12000}{S} \end{cases} \tag{113–11}$$

and

$$\text{Lower PSD}\,(f) \geq \begin{cases} -93 \text{ dBm/Hz} & 20 \leq f \leq 200 \\ -93 - \left(\dfrac{\dfrac{f}{S} - 200}{200}\right) \text{dBm/Hz} & 200 < \dfrac{f}{S} \leq 800 \\ -96 - \left(\dfrac{\dfrac{f}{S} - 800}{100}\right) \text{dBm/Hz} & 800 < \dfrac{f}{S} \leq \dfrac{1600}{S} \end{cases} \quad (113\text{–}12)$$

where f is in MHz.

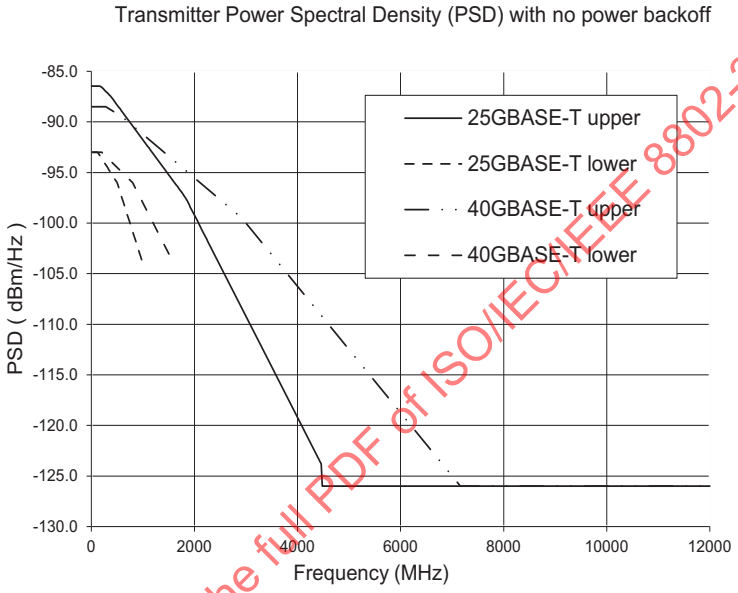


Figure 113–37—Transmitter power spectral density masks

113.5.3.5 Transmit clock frequency

The symbol transmission rate on each pair of the MASTER PHY shall be within the range 3 200× S MBd ± 50 ppm.

For a MASTER PHY, when the transmitter is in the LPI transmit mode or when the receiver is in the LPI receive mode the transmitter clock short-term rate of frequency variation shall be less than 0.1 ppm/second. The short-term frequency variation limit shall also apply when switching to and from the LPI mode.

113.5.4 Receiver electrical specifications

The PMA provides the Receive function specified in 113.4.2.4 in accordance with the electrical specifications of this clause using cabling that is within the limits specified in 113.7.

113.5.4.1 Receiver differential input signals

Differential signals received at the MDI that were transmitted from a remote transmitter within the specifications of 113.5.3 and have passed through a link specified in 113.7 shall be received with a BER less

than 10^{-12} after LDPC and RS-FEC decoding, and sent to the 25GMII/XLGMII after link reset completion. This specification can be verified by a frame error ratio less than 9.6×10^{-9} for 800 octet frames with minimum IPG or greater than 799 octet IPG.

113.5.4.2 Receiver frequency tolerance

The receive feature shall properly receive incoming data, per the requirements of 113.5.4.1, with a symbol rate within the range 3200× *S* MBd ± 50 ppm.

113.5.4.3 Rejection of External EM Fields

When the cabling system is subjected to electromagnetic fields, currents are generated in the shield that can be converted to interference. This specification is provided to limit the sensitivity of the PMA receiver to external EM fields picked up by the cabling and interconnect system. It provides an assessment method of the electromagnetic performance of the link segment and the PHY, including the MDI.

An 80 MHz to 2000 MHz test can be made based on the cable clamp test described in Annex 113A, a 30 m plug-terminated cabling span that meets the requirements of 113.7, and suitable broadband ferrites. All components that are exposed to the induced fields should remain over the ground reference plane. A sine wave with the amplitude held constant over the whole frequency range from 80 MHz to 2000 MHz, with the amplitude calibrated so that the signal power measured at the output of the clamp does not exceed 6 dBm, is used to generate the external electromagnetic field and corresponding shield current.

A system integrating a 25G/40GBASE-T PHY may perform this test to evaluate anticipated performance in regulatory test environments. Operational requirements of the transceiver during the test are determined by the manufacturer.

NOTE—The 6 dBm limit includes the 10% frequency-dependent variation mentioned in Annex 113A.3.

113.5.4.4 Alien crosstalk noise rejection

While receiving data from a transmitter compliant with specifications in 113.5.3, through a 30 m link segment compliant with the specifications in 113.7, a receiver shall operate with an Ethernet frame error ratio less than 9.6×10^{-9} for 800 octet frames with either a minimum IPG or greater than 799 octet IPG with four noise sources at the specified levels representing alien crosstalk, one connected to each of the four pairs. Independent noise sources should be injected into each MDI input using couplers that do not significantly alter the link segment characteristics. Each noise source shall have a flat noise spectrum, with 3 dB bandwidth at least 10 MHz to 2000 MHz and a power spectral density such that at the MDI port of the device under test the power spectral density of the injected noise is –154 dBm/Hz. A flat noise source is chosen to model the sum of all alien noise sources. See Figure 113–38.

The structure shown for injecting the noise in Figure 113–38 is illustrative and alternative approaches are possible. The loss of the coupling structure shown in Figure 113–38, which consists of two baluns and a coupler, is approximately 2.5 dB. The overall insertion loss of the link segment together with the insertion loss of the coupling structure should be adjusted to match the insertion loss specified in 113.7.2.1 to within ±0.5 dB. The balun-coupler-balun structure shown in Figure 113–38 can be replaced by resistively coupling a balanced noise source to the twisted-pair using 500 Ω resistors. In either case, calibration of the test setup is required to confirm the overall insertion loss and the injected noise power at the MDI of the receiver under test.

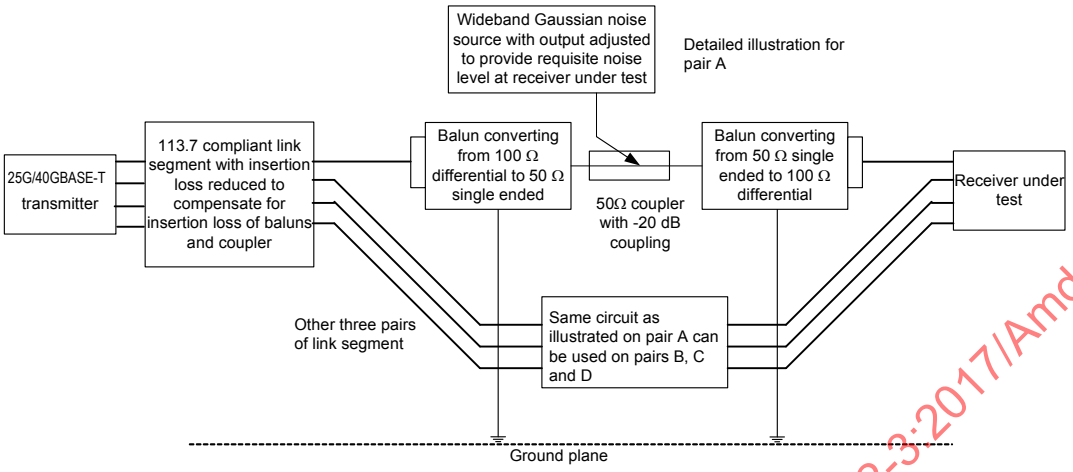


Figure 113–38—Alien crosstalk noise rejection test

113.5.4.5 Short reach mode

The PHY short reach register setting 1.131.0 indicates whether the PHY is operating in the short reach mode.

In short reach mode (indicating operation over a short reach link segment) while receiving data from a transmitter compliant with specifications in 113.5.3 (whether or not in short reach mode), through a short reach link segment meeting the requirements of 113.7.4, a receiver shall operate with a frame error ratio less than 9.6×10^{-9} for 800 octet frames with minimum IPG or greater than 799 octet IPG (e.g., operate with a BER less than 10^{-12}). When operating in short reach mode, only operation over the direct attach link segment specified in 113.7.4 is required.

113.6 Management interfaces

25G/40GBASE-T makes extensive use of the management functions that may be provided by the MDIO (Clause 45), and the communication and self-configuration functions provided by Auto-Negotiation (Clause 28). Additional Auto-Negotiation requirements are set forth within this subclause.

113.6.1 Support for Auto-Negotiation

All 25GBASE-T and 40GBASE-T PHYs shall provide support for Auto-Negotiation (Clause 28) and shall be capable of operating as MASTER or SLAVE. All 25GBASE-T and 40GBASE-T PHYs shall provide support for Extended Next Pages as defined in 28.2.3.4.2 and shall support and use optimized FLP Burst to FLP burst timing as defined in 28.2.1.1.1, and nlp_link_test_min_timer and link_fail_inhibit_timer as defined in 28.3.2.

Auto-Negotiation is performed as part of the initial setup of the link and allows the PHYs at each end to advertise their capabilities (speed, PHY type, half or full duplex) and to automatically select the operating mode for communication on the link. Auto-Negotiation signaling is used for the following primary purposes for 25G/40GBASE-T:

a) To negotiate that the PHY is capable of supporting 25GBASE-T or 40GBASE-T transmission.
b) To determine the MASTER-SLAVE relationship between the PHYs at each end of the link.

113.6.1.1 25G/40GBASE-T use of registers during Auto-Negotiation

When Clause 45 registers are implemented, a 25G/40GBASE-T PHY shall use the management register definitions and values specified in Table 113–18.

Table 113–18—25G/40GBASE-T registers

Register	Bit	Name	Description	Type [a]
7.0	7.0.15:0	AN control register	Defined in 45.2.7.1	R/W
7.1	7.1.15:0	AN status register	Defined in 45.2.7.2	RO
7.2, 7.3	7.2.15:0, 7.3.15:0	AN device identifier registers	Defined in 45.2.7.3	R/W
7.5, 7.6	7.5.15:0, 7.6.15:0	AN devices in package registers	Defined in 45.2.7.4	R/W
7.14, 7.15	7.14.15:0, 7.15.15:0	AN package identifier registers	Defined in 45.2.7.5	R/W
7.16	7.16.15:0	AN advertisement register	Defined in 45.2.7.6	R/W
7.19	7.19.15:0	AN LP Base Page ability register	Defined in 45.2.7.7	RO
7.22, 7.23, 7.24	7.22.15:0, 7.23.15:0, 7.24.15:0	AN XNP transmit register	Defined in 45.2.7.8	R/W
7.25, 7.26, 7.27	7.25.15:0, 7.26.15:0, 7.27.15:0	AN LP XNP ability register	Defined in 45.2.7.9	RO
7.32	7.32.15:0	MultiGBASE-T AN control 1 register	Defined in 45.2.7.10	R/W
7.33	7.33.15:0	MultiGBASE-T AN status 1 register	Defined in 45.2.7.11	RO
7.32	7.32.15:0	MultiGBASE-T AN control 2 register	Defined in 45.2.7.14a	R/W
7.33	7.33.15:0	MultiGBASE-T AN status 2 register	Defined in 45.2.7.14b	RO

[a] R/W = Read/Write, RO = Read only

113.6.1.2 25G/40GBASE-T Auto-Negotiation page use

25G/40GBASE-T PHYs shall exchange a MultiGBASE-T and 1000BASE-T formatted Extended Next Page, as specified in Table 113–19, immediately following the exchange of the Base Page.

Note that the Acknowledge 2 bit is not utilized and has no meaning when used for the 25G/40GBASE-T message page exchange.

113.6.1.3 Sending Next Pages

Implementers who do not wish to send additional Extended Next Pages (i.e., Extended Next Pages in addition to those required to perform PHY configuration as defined in this clause) can use Auto-Negotiation as defined in Clause 28. Implementers who wish to send additional Extended Next Pages may do so using the AN XNP transmit registers. See 45.2.7.8.

Table 113–19—25G/40GBASE-T Base and Next Pages bit assignments

Bit	Name	Description
Base Page		
D15	Next Page	Defined in 28.2.1.6
D14	Acknowledge	Defined in 28.2.1.5
D13	Remote Fault	Defined in 28.2.1.4
D12	Extended Next Page	Defined in 28.2.1.3
D11:D5	Technology Ability Field	Defined in 28.2.1.2
D4:D0	Selector Field	Defined in 28.2.1.1
Extended Next Page (Message Code Field and Flags Field)		
M10:M0	Next Page message code	Defined in Annex 28C
T	Toggle	Defined in 28.2.3.4.7
Ack2	Acknowledge 2	Defined in 28.2.3.4.6
MP	Message Page	Defined in 28.2.3.4.5
Ack	Acknowledge	Defined in 28.2.3.4.4
NP	Next Page	Defined in 28.2.3.4
Extended Next Page (Unformatted Message Code Field)		
U31:U27	Reserved, transmit as 0	
U26	40GBASE-T ability (1 = support of 40GBASE-T and 0 = no support)	Defined in 45.2.7.10.4a
U25	25GBASE-T ability (1 = support of 25GBASE-T and 0 = no support)	Defined in 45.2.7.10.4b
U24	10GBASE-T EEE (1 = Advertise EEE capability for 10GBASE-T 0 = Do not advertise EEE capability for 10GBASE-T)	Defined in 45.2.7.13.4
U23	1000BASE-T EEE (1 = Advertise EEE capability for 1000BASE-T 0 = Do not advertise EEE capability for 1000BASE-T)	Defined in 45.2.7.13.5
U22	100BASE-TX EEE (1 = Advertise EEE capability for100BASE-TX 0 = Do not advertise EEE capability for 100BASE-TX)	Defined in 45.2.7.13.6
U21	Reserved	Value always 0
U20	10GBASE-T LD PMA training reset request 0 = Local Device requests that Link Partner run PMA training PRBS continuously) This bit is not defined for 10GBASE-T but reserved for future use.	Defined in 45.2.7.10.5
U19	10GBASE-T Fast retrain ability (1 = Advertise PHY as supporting fast retrain, 0 = Advertise PHY as not supporting fast retrain)	Defined in 45.2.7.10.6 This bit is not defined for 10GBASE-T but reserved for future use.
U18	PHY short reach mode (1 = PHY of Local Device is operating in short reach mode 0 = PHY of Local Device is operating in normal mode)	Defined in 45.2.1.64.2
U17	10GBASE-T LD loop timing ability (1 = Advertise PHY as capable of loop timing (mandatory for 25G/40GBASE-T) and 0 = do not advertise PHY as capable of loop timing)	Defined in 45.2.7.10.7

Table 113–19—25G/40GBASE-T Base and Next Pages bit assignments *(continued)*

Bit	Name	Description
U16	10GBASE-T ability (1 = support of 10GBASE-T and 0 = no support)	Defined in 45.2.7.10.4
U15	1000BASE-T half duplex (1 = half duplex and 0 = no half duplex)	
U14	1000BASE-T full duplex (1 = full duplex and 0 = no full duplex)	
U13	Port type bit (1 = multiport device and 0 = single-port device)	Defined in 45.2.7.10.3
U12	MultiGBASE-T MASTER-SLAVE config value (1 = MASTER and 0 = SLAVE) This bit is ignored if 7.32.15=0.	Defined in 45.2.7.10.1
U11	MultiGBASE-T MASTER-SLAVE manual config enable (1 = manual configuration enable) This bit is intended to be used for manual selection in a particular MASTER-SLAVE mode and is to be used in conjunction with bit 7.32.14.	Defined in 45.2.7.10.2
U10	MASTER-SLAVE seed Bit 10 (SB10) (MSB)	
U9	MASTER-SLAVE seed Bit 9 (SB9)	
U8	MASTER-SLAVE seed Bit 8 (SB8)	
U7	MASTER-SLAVE seed Bit 7 (SB7)	
U6	MASTER-SLAVE seed Bit 6 (SB6)	
U5	MASTER-SLAVE seed Bit 5 (SB5)	
U4	MASTER-SLAVE seed Bit 4 (SB4)	
U3	MASTER-SLAVE seed Bit 3 (SB3)	
U2	MASTER-SLAVE seed Bit 2 (SB2)	
U1	MASTER-SLAVE seed Bit 1 (SB1)	
U0	MASTER-SLAVE seed Bit 0 (SB0)	

113.6.2 MASTER-SLAVE configuration resolution

Since both PHYs that share a link segment are capable of being MASTER or SLAVE, a prioritization scheme exists to ensure that the correct mode is chosen. The MASTER-SLAVE relationship shall be determined during Auto-Negotiation using Table 113–20 with the 25G/40GBASE-T Technology Ability Next Page bit values specified in Table 113–19 and information received from the link partner. This process is conducted at the entrance to the FLP LINK GOOD CHECK state shown in the Arbitration state diagram (Figure 28-18.)

The following four equations are used to determine these relationships:

manual_MASTER = U11 * U12
manual_SLAVE = U11 * !U12
single-port device = !U11 * !U13
multiport device = !U11 * U13

where

U11 is bit 11 of MultiGBASE-T and 1000BASE-T Technology message code,
U12 is bit 12 of MultiGBASE-T 1000BASE-T Technology message code,
U13 is bit 13 of MultiGBASE-T and 1000BASE-T Technology message code (see Table 113–19).