

(12) **United States Patent**
Sun et al.

(10) **Patent No.:** US 11,032,111 B2
(45) **Date of Patent:** Jun. 8, 2021

(54) **SERDES PRE-EQUALIZER HAVING ADAPTABLE PRESET COEFFICIENT REGISTERS**

(71) Applicant: **CREDO TECHNOLOGY GROUP LIMITED**, Grand Cayman (KY)

(72) Inventors: **Junqing Sun**, Fremont, CA (US); **Haoli Qian**, Fremont, CA (US)

(73) Assignee: **CREDO TECHNOLOGY GROUP LIMITED**, Grand Cayman (KY)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **16/552,927**

(22) Filed: **Aug. 27, 2019**

(65) **Prior Publication Data**

US 2020/0076651 A1 Mar. 5, 2020

Related U.S. Application Data

(60) Provisional application No. 62/723,701, filed on Aug. 28, 2018.

(51) **Int. Cl.**
H04L 25/03 (2006.01)
H04L 7/00 (2006.01)
(Continued)

(52) **U.S. Cl.**
CPC **H04L 25/03019** (2013.01); **G06N 20/00** (2019.01); **H04B 1/40** (2013.01); **H04L 7/0087** (2013.01); **H04L 2025/03808** (2013.01)

(58) **Field of Classification Search**
CPC . H04L 25/03019; H04L 7/0087; H04L 27/01; H04L 25/03885; G06N 20/00; H04B 1/40
(Continued)

(56) **References Cited**

U.S. PATENT DOCUMENTS

8,301,036 B2 10/2012 He
9,071,479 B2 6/2015 Qian et al.
(Continued)

OTHER PUBLICATIONS

Information Technology—Open Systems Interconnection—Basic Reference Model: The Basic Model, 1994.

(Continued)

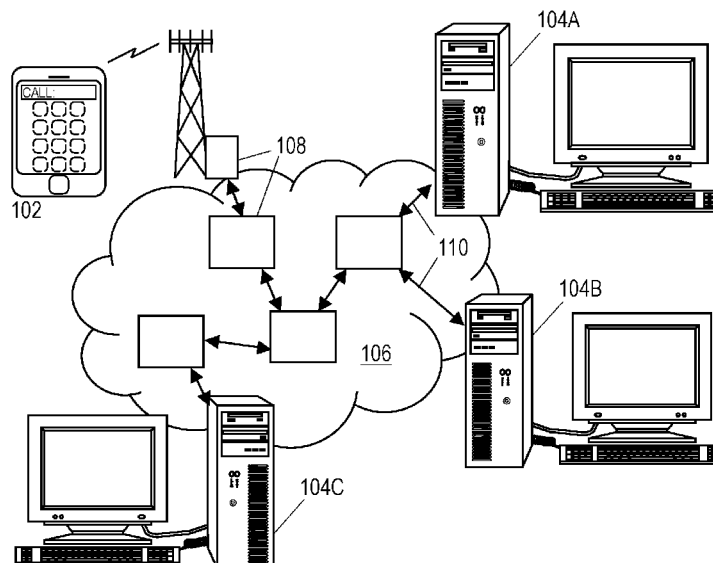
Primary Examiner — Leila Malek

(74) *Attorney, Agent, or Firm* — Ramey & Schwaller LLP; Daniel J. Krueger

(57) **ABSTRACT**

An illustrative SerDes (serializer-deserializer) communications method embodiment may include a transceiver: selecting one of multiple registers to specify initial pre-equalizer coefficient values; updating the initial pre-equalizer coefficient values during a training phase; and using the updated pre-equalizer coefficient values to convey a transmit data stream. In an illustrative embodiment of a chip-to-module communications link, a port connector couples a port transceiver to a pluggable module transceiver, the pluggable module transceiver including: one or more transmit filters to each pre-equalize a corresponding serial symbol stream being transmitted to the port transceiver; and a controller having multiple registers, each of the multiple registers containing a set of initial coefficient values, the controller using one of the registers to set initial coefficient values for the one or more transmit filters.

19 Claims, 7 Drawing Sheets



(51) **Int. Cl.**

G06N 20/00 (2019.01)

H04B 1/40 (2015.01)

(58) **Field of Classification Search**

USPC 375/219

See application file for complete search history.

(56)

References Cited

U.S. PATENT DOCUMENTS

9,385,859	B2	7/2016	Kuan et al.	
9,935,800	B1	4/2018	He	
10,212,260	B2	2/2019	Sun	
2010/0232493	A1 *	9/2010	Thirumoorthy	 H04L 27/0014 375/232
2013/0201316	A1 *	8/2013	Binder	 H04L 67/12 348/77
2013/0223506	A1 *	8/2013	Kolze	 H04L 27/01 375/232
2013/0343400	A1	12/2013	Lusted et al.	
2014/0086264	A1	3/2014	Lusted et al.	
2014/0146833	A1	5/2014	Lusted et al.	
2015/0003505	A1	1/2015	Lusted et al.	
2016/0337114	A1	11/2016	Baden et al.	
2020/0076651	A1	3/2020	Sun et al.	

OTHER PUBLICATIONS

International Search Report dated Dec. 6, 2017 in International Application No. PCT/CN2017/075961.

Non-Final Office Action dated Jan. 24, 2018 in U.S. Appl. No. 15/487,045.

Reed-Solomon Forward Error Correction (RS-FEC) Sublayer for 100GBASE-R PHYs.

Notice of Allowance dated Jun. 8, 2018 in U.S. Appl. No. 15/487,045.

"Physical Coding Sublayer (PCS) for 64B/66B, type 50GBASE-R" Draft Amendment to IEEE Std 802.3-2015, Feb. 3, 2017, pp. 132-141, 10 pgs.

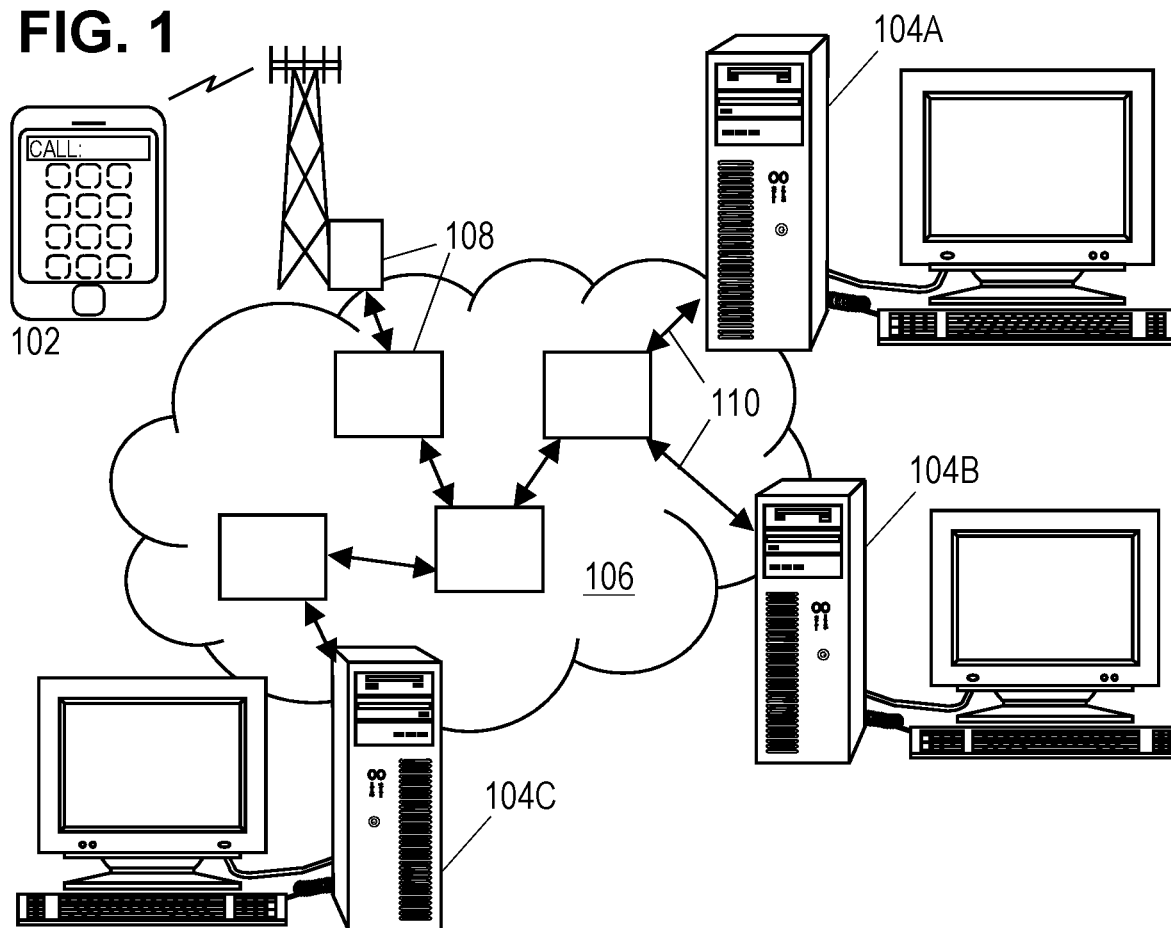
"Physical Medium Attachment (PMA) sublayer, Physical Medium Dependent (PMD) sublayer, and baseband medium, type 1000GBASE-KP4," IEEE Standard for Ethernet, Section Six, pp. 481-519, 39 pgs.

"Physical Medium Dependent (PMD) sublayer and baseband medium, type 50GBASE-CR, 100GBASE-CR2, and 200GBASE-CR4," Draft Amendment to IEEE Std 802.3-2015, Feb. 3, 2017, pp. 184-215, 32 pgs.

"Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 50GBASE-R PHYs," Draft Amendment to IEEE Std 802.3-2015, Feb. 3, 2017, pp. 142-159, 18 pgs.

* cited by examiner

FIG. 1



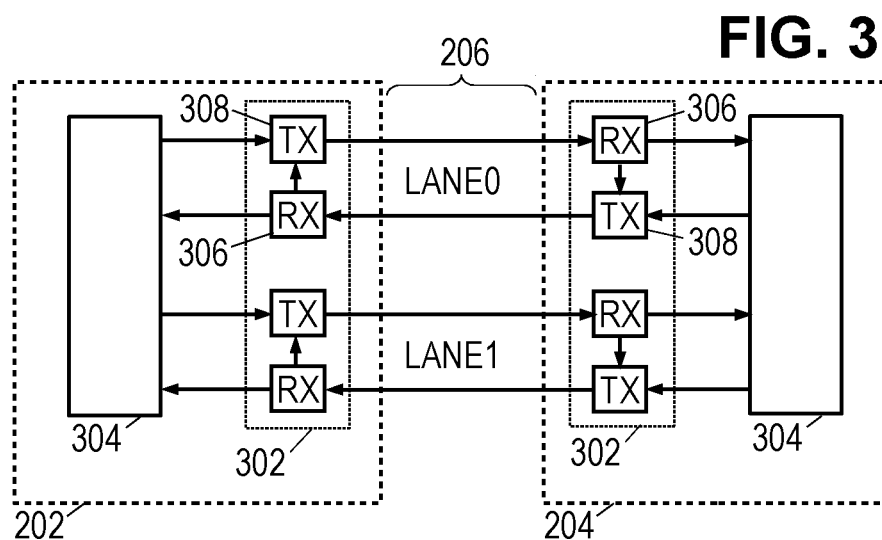
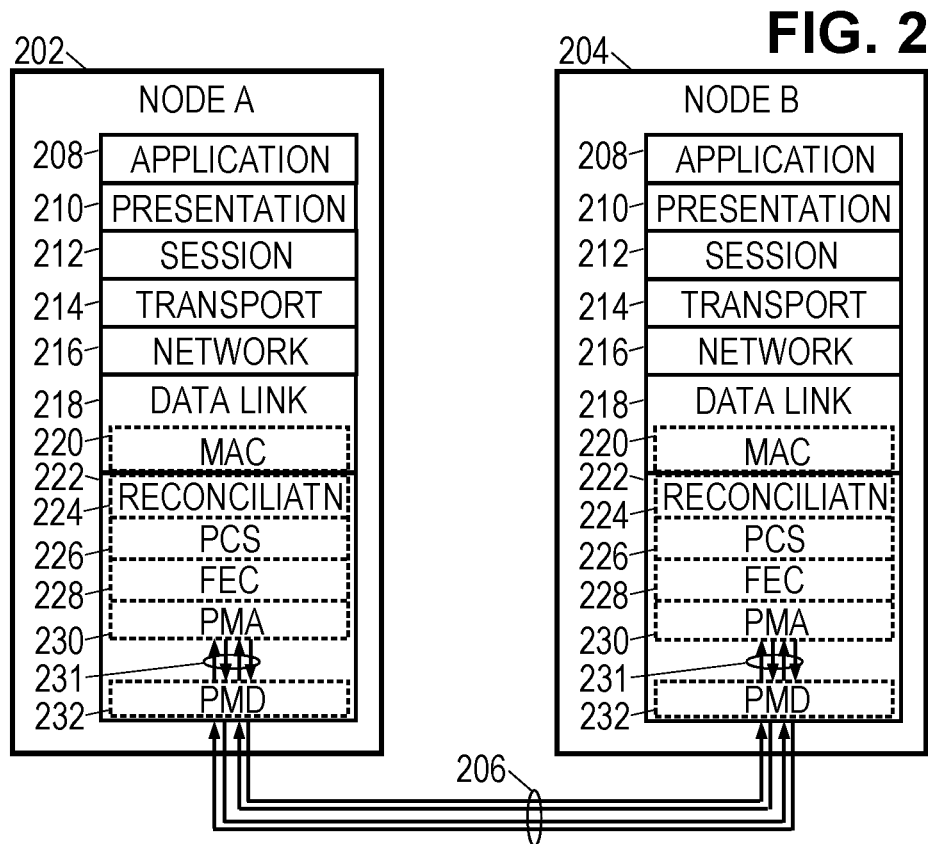
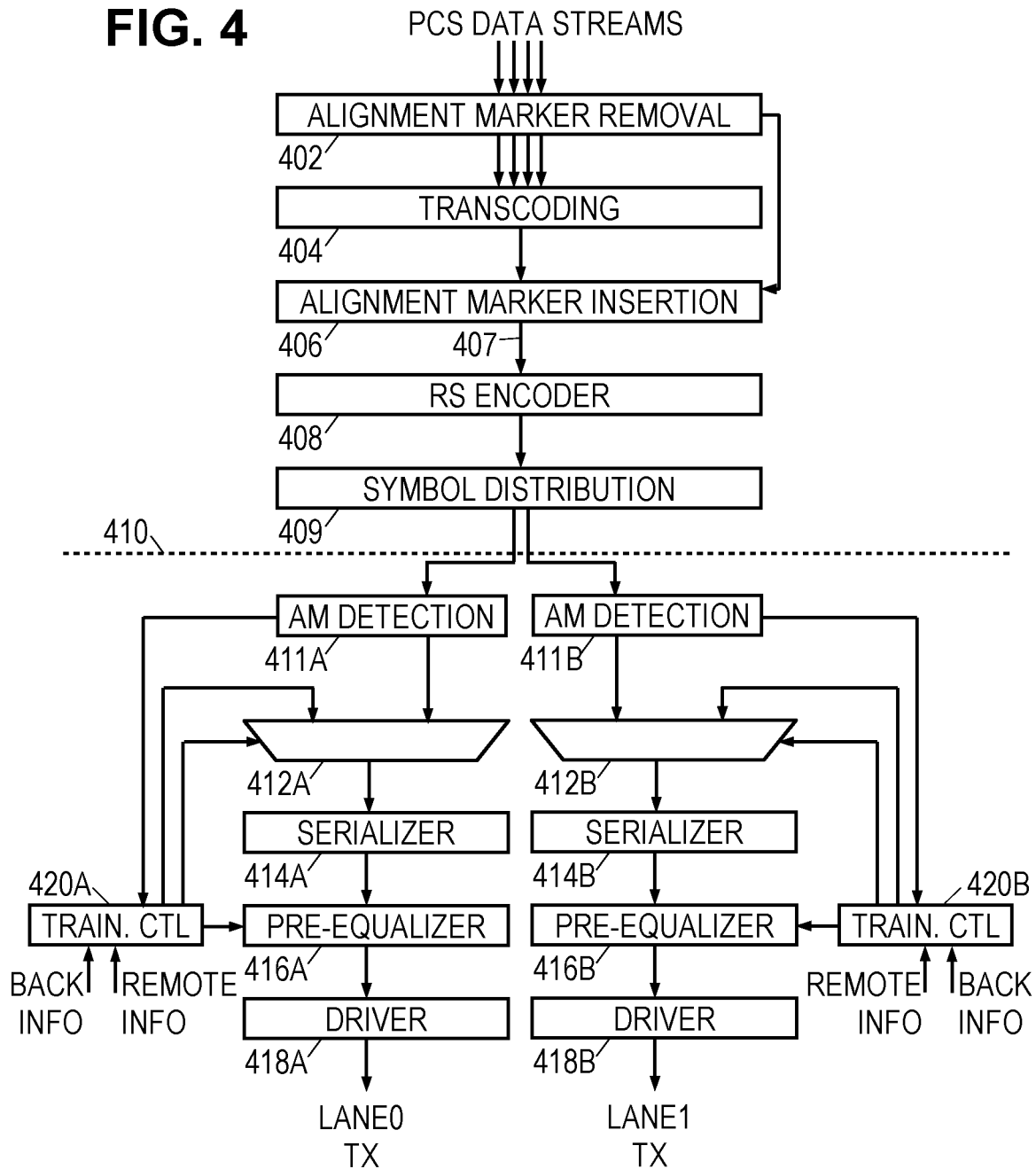


FIG. 4

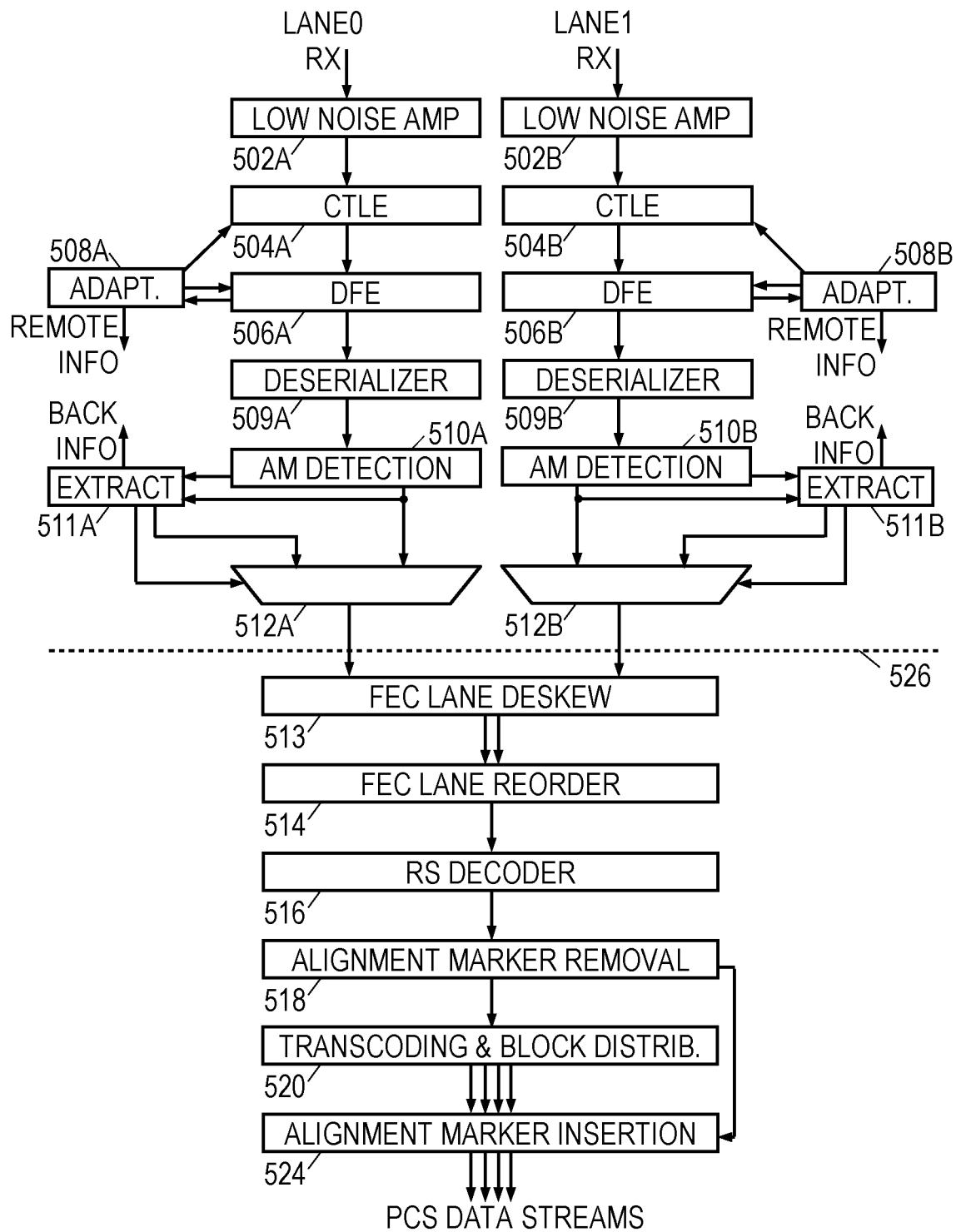
**FIG. 5**

FIG. 6A

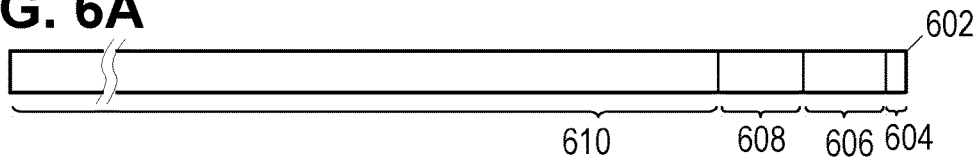


FIG. 6B

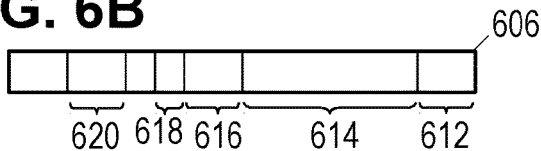


FIG. 6C

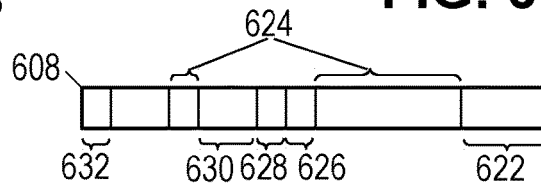


FIG. 7A

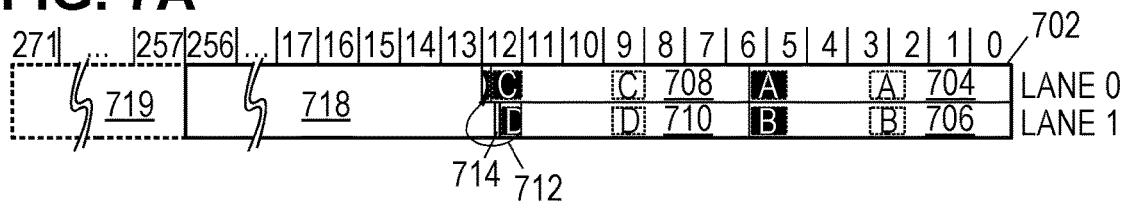


FIG. 7B

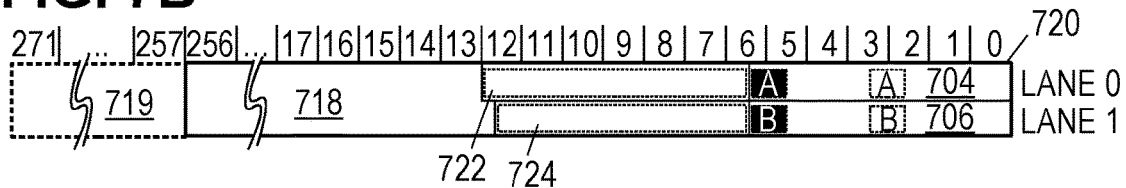


FIG. 7C

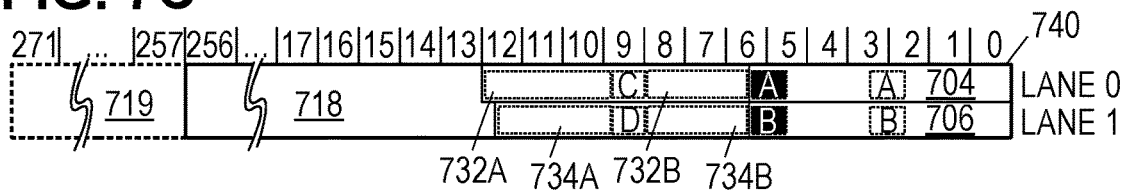


FIG. 7D

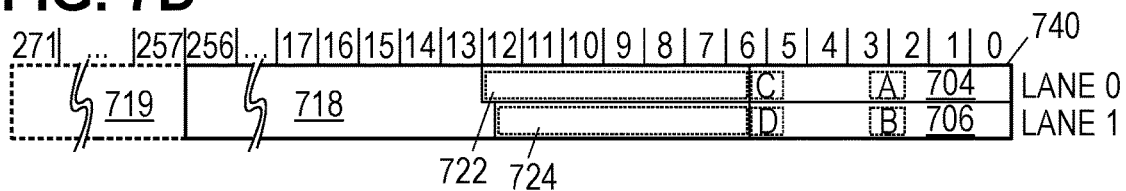


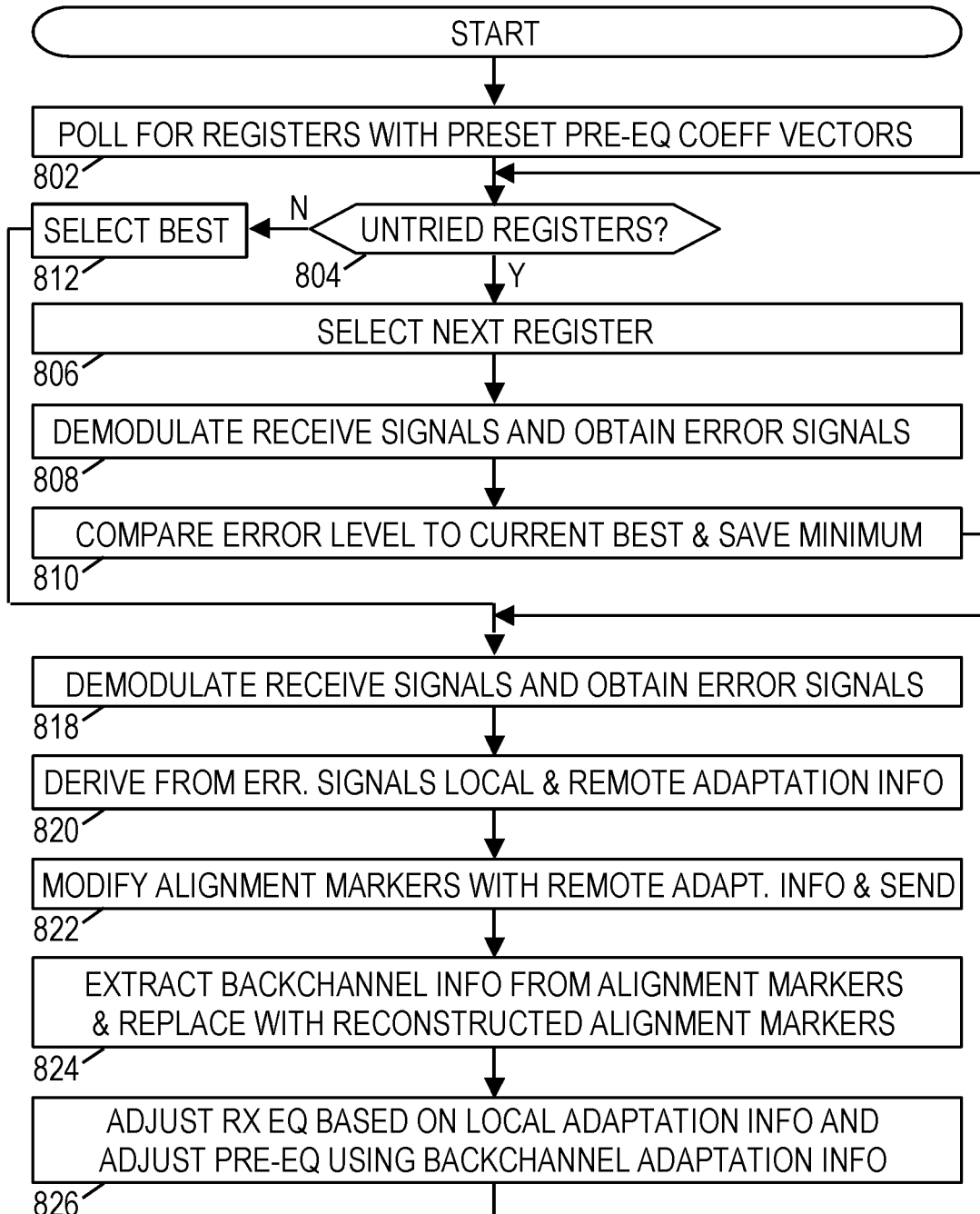
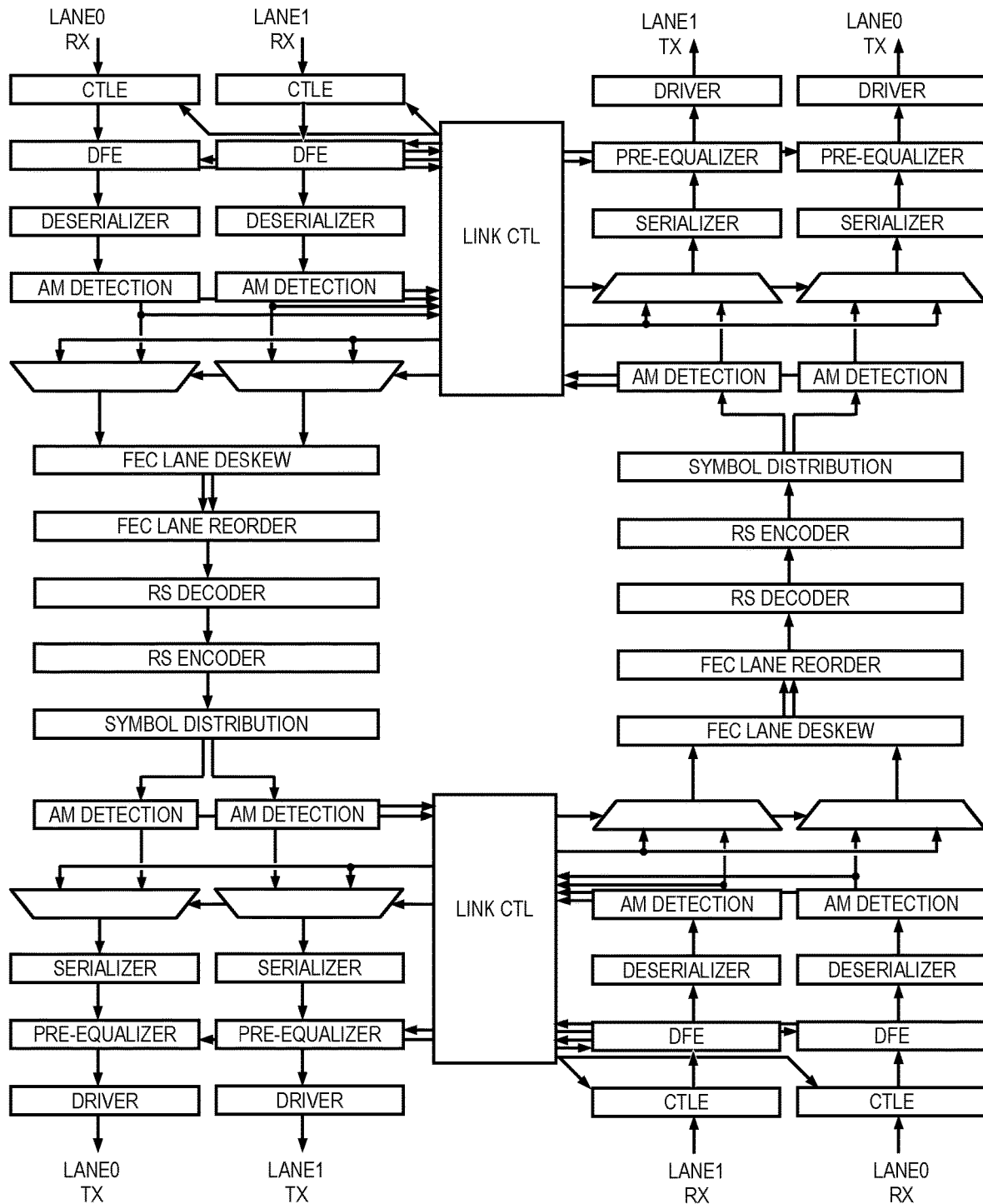
FIG. 8

FIG. 9



1

SERDES PRE-EQUALIZER HAVING ADAPTABLE PRESET COEFFICIENT REGISTERS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to Provisional Application US 62/723,701, filed Aug. 28, 2018 and titled “SerDes Pre-Equalizer Having Efficient Adaptation” by inventors J. Sun and H. Qian, which is hereby incorporated herein by reference. This application further relates to Application U.S. Ser. No. 15/654,446, filed Jul. 19, 2018 and titled “SerDes Architecture with a Hidden Backchannel Protocol” by inventor J. Sun.

BACKGROUND

The Institute of Electrical and Electronics Engineers (IEEE) Standards Association publishes an IEEE Standard for Ethernet, IEEE Std 802.3-2015, which will be familiar to those of ordinary skill in the art to which this application pertains. This standard provides a common media access control specification for local area network (LAN) operations at selected speeds from 1 Mb/s to 100 Gb/s over coaxial cable, twisted wire pair cable, fiber optic cable, and electrical backplanes, with various channel signal constellations. As demand continues for ever-higher data rates, extensions to the standard must account for increased channel attenuation and dispersion even as the equalizers are forced to operate at faster symbol rates. One potential course of action employs enhanced equalization techniques, but such techniques should be implemented in a manner that does not limit performance or introduce undesirable hardware complexity or cost.

SUMMARY

Accordingly, there are disclosed herein a serializer-deserializer (SerDes) architecture and method that implements an efficient adaptation interface for transmitter pre-equalizers in a standards-suitable fashion. An illustrative SerDes (serializer-deserializer) communications method embodiment may include a transceiver: selecting one of multiple registers to specify initial pre-equalizer coefficient values; updating the initial pre-equalizer coefficient values during a training phase; and using the updated pre-equalizer coefficient values to convey a transmit data stream.

In an illustrative embodiment of a chip-to-module communications link, a port connector couples a port transceiver to a pluggable module transceiver, the pluggable module transceiver including: one or more transmit filters to each pre-equalize a corresponding serial symbol stream being transmitted to the port transceiver; and a controller having multiple registers, each of the multiple registers containing a set of initial coefficient values, the controller using one of the registers to set initial coefficient values for the one or more transmit filters.

Each of the foregoing embodiments may be implemented individually or in combination, and may be implemented with any one or more of the following features in any suitable combination: 1. each of the multiple registers corresponds to a different channel model. 2. at least some of the different channel models are for chip-to-module (C2M) channels with different insertion loss and package loss characteristics. 3. at least some of the different channel models presume different types of receiver equalization. 4.

2

said selecting includes using the initial pre-equalizer coefficient values to determine a performance characteristic for each of the multiple registers. 5. the performance characteristic is an error signal energy. 6. the performance characteristic is a bit error rate. 7. updating the selected register with the updated pre-equalizer coefficient values. 8. one of the registers is selected by the port transceiver to specify the initial coefficient values. 9. the port transceiver iteratively selects different ones of the multiple registers and determines a performance characteristic for each. 10. the port transceiver generates updates for the coefficient values of the one or more transmit filters. 11. the controller saves updated coefficient values in one of the registers. 12. performing simultaneous updates of multiple coefficients values. 13. performing coarse updates before performing fine updates of coefficient values.

BRIEF DESCRIPTION OF THE DRAWINGS

In the drawings:

FIG. 1 shows an illustrative communications network.

FIG. 2 is a diagram of an illustrative multi-lane communications link model.

FIG. 3 is a block diagram of an illustrative communications link.

FIG. 4 is a block diagram of a transmit chain in an illustrative multi-lane transceiver.

FIG. 5 is a block diagram of a receive chain in an illustrative multi-lane transceiver.

FIGS. 6A-6C show fields of an illustrative training frame.

FIG. 7A shows sets of grouped alignment markers in an illustrative multi-lane data stream.

FIG. 7B shows a first data stream embodiment having hidden backchannel fields.

FIG. 7C shows a second data stream embodiment having hidden backchannel fields.

FIG. 7D shows a third data stream embodiment having hidden backchannel fields.

FIG. 8 is a flowchart of an illustrative multi-lane communications method.

FIG. 9 is a block diagram of an illustrative pluggable module.

It should be understood, however, that the specific embodiments given in the drawings and detailed description do not limit the disclosure. On the contrary, they provide the foundation for one of ordinary skill to discern the alternative forms, equivalents, and modifications that are encompassed in the scope of the appended claims.

DETAILED DESCRIPTION

The disclosed apparatus and methods are best understood in the context of the larger environments in which they operate. Accordingly, FIG. 1 shows an illustrative communications network 100 including mobile devices 102 and computer systems 104A-C coupled via a routing network 106. The routing network 106 may be or include, for example, the Internet, a wide area network, or a local area network. In FIG. 1, the routing network 106 includes a network of equipment items 108, such as switches, routers, and the like. The equipment items 108 are connected to one another, and to the computer systems 104A-C, via point-to-point communication links 110 that transport data between the various network components. At least some of the links 110 in network 106 are high-speed multi-lane links such as Ethernet links operating in compliance with the IEEE Std 802.3-2015 (or later) at 10 Gb/s or more.

FIG. 2 shows a communications link between two nodes **202**, **204**, that operate in accordance with the ISO/IEC Model for Open Systems Interconnection (See ISO/IEC 7498-1:1994.1) to communicate over a physical medium represented by transmit and receive channels **206**. The interconnection reference model employs a hierarchy of layers with defined functions and interfaces to facilitate the design and implementation of compatible systems by different teams or vendors. While it is not a requirement, it is expected that the higher layers in the hierarchy will be implemented primarily by software or firmware operating on programmable processors while the lower layers may be implemented as application-specific hardware.

The Application Layer **208** is the uppermost layer in the model, and it represents the user applications or other software operating on different systems (e.g., equipment **108**), which need a facility for communicating messages or data. The Presentation Layer **210** provides such applications with a set of application programming interfaces (APIs) that provide formal syntax, along with services for data transformations (e.g., compression), establishing communication sessions, connectionless communication mode, and negotiation to enable the application software to identify the available service options and select therefrom. The Session Layer **212** provides services for coordinating data exchange including: session synchronization, token management, full- or half-duplex mode implementation, and establishing, managing, and releasing a session connection. In connectionless mode, the Session Layer may merely map between session addresses and transport addresses.

The Transport Layer **214** provides services for multiplexing, end-to-end sequence control, error detection, segmenting, blocking, concatenation, flow control on individual connections (including suspend/resume), and implementing end-to-end service quality specifications. The focus of the Transport Layer **214** is end-to-end performance/behavior. The Network Layer **216** provides a routing service, determining the links used to make the end-to-end connection and when necessary acting as a relay service to couple together such links. The Data link layer **218** serves as the interface to physical connections, providing delimiting, synchronization, sequence and flow control across the physical connection. It may also detect and optionally correct errors that occur across the physical connection. The Physical layer **222** provides the mechanical, electrical, functional, and procedural means to activate, maintain, and deactivate channels **206**, and to use the channels **206** for transmission of bits across the physical media.

The Data Link Layer **218** and Physical Layer **222** are subdivided and modified slightly by IEEE Std 802.3-2015, which provides a Media Access Control (MAC) Sublayer **220** in the Data Link Layer **218** to define the interface with the Physical Layer **222**, including a frame structure and transfer syntax. Within the Physical Layer **222**, the standard provides a variety of possible subdivisions such as the one illustrated in FIG. 2, which includes an optional Reconciliation Sublayer **224**, a Physical Coding Sublayer (PCS) **226**, a Forward Error Correction (FEC) Sublayer **228**, a Physical Media Attachment (PMA) Sublayer **230**, and a Physical Medium Dependent (PMD) Sublayer **232**.

The optional Reconciliation Sublayer **224** merely maps between interfaces defined for the MAC Sublayer **220** and the PCS Sublayer **226**. The PCS Sublayer **226** provides scrambling/descrambling, data encoding/decoding (with a transmission code that enables clock recovery and bit error detection), block and symbol redistribution, PCS alignment marker insertion/removal, and block-level lane synchroni-

zation and deskew. To enable bit error rate estimation by components of the Physical Layer **222**, the PCS alignment markers typically include Bit-Interleaved-Parity (BIP) values derived from the preceding bits in the lane up to and including the preceding PCS alignment marker.

The FEC Sublayer **228** provides, e.g., Reed-Solomon coding/decoding that distributes data blocks with controlled redundancy across the lanes to enable error correction. In some embodiments (e.g., in accordance with Article 91 or proposed Article 134 for the IEEE Std 802.3), the FEC Sublayer **228** modifies the number of lanes. For example, under proposed Article 134, a four-lane outgoing data stream (including PCS alignment markers) may be converted into a two-lane transmit data stream. Conversely, the FEC Sublayer **228** may convert a two-lane receive data stream into a four-lane incoming data stream. In both directions, the PCS alignment markers may be preserved, yielding pairs (or more generally, "sets") of grouped PCS alignment markers in the multi-lane data streams being communicated to and from the PMA Sublayer **230**. (Article 91 provides for a 20-to-4 lane conversion, yielding sets of 5 grouped PCS alignment markers in each lane of the data streams communicated between the FEC and PMA sublayers.)

The PMA Sublayer **230** provides lane remapping, symbol encoding/decoding, framing, and octet/symbol synchronization. In at least some contemplated embodiments, the multi-lane data stream **231** communicated to and from the PMA Sublayer **230** complies with the attachment unit interface (AUI) requirements of the IEEE standard. Several varieties of the AUI requirements exist and are named based on the nominal bit rate and the number of lanes. For example, Annexes 120E and 135G to the IEEE standard provide requirements for 50GAUI-1, 100GAUI-2, 200GAUI-4, and 400GAUI-8, which respectively employ 1, 2, 4, and 8 lanes of NRZ signaling at a per-lane symbol rate of 53.125 GBd. Recently proposed extensions provide requirements for 100GAUI-1, 200GAUI-2, 400GAUI-4, and 800GAUI-8, which respectively employ 1, 2, 4, and 8 lanes of PAM4 signaling at per-lane symbol rates of 53.125 GBd.

The PMD Sublayer **232** specifies the transceiver conversions between transmitted/received channel signals and the corresponding AUI-compliant bit (or digital symbol) streams. The PMD Sublayer also implements an initial start-up of the communications channels **206**, optionally conducting an auto-negotiation phase and a link-training phase before entering a normal operating phase. The auto-negotiation phase enables the end nodes to exchange information about their capabilities, and the training phase enables the end nodes to adapt both transmit-side and receive-side equalization filters in a fashion that combats the channel non-idealities.

More information regarding the operation of the sublayers, as well as the electrical and physical specifications of the connections between the nodes and the communications medium (e.g., pin layouts, line impedances, signal voltages & timing), arrangements of the communications medium (e.g., the network topology), and the electrical and physical specifications for the communications medium itself (e.g., conductor arrangements in copper or fiber optic cable, limitations on attenuation, propagation delay, signal skew), can be found in the standard, and any such details should be considered to be well within the knowledge of those having ordinary skill in the art. The discussion below focuses on modifications specific to the present disclosure.

FIG. 3 provides a block diagram of those portions of nodes **202**, **204**, that may implement the Data Link Layer **218** and Physical Layer **222**. Transceivers **302** for multiple

5

transmit and receive channels are coupled to a host interface **304**. The transceivers **302** and host interface **304** may be, e.g., part of a network interface peripheral coupled to the I/O bus of a personal computer, server, network switch, or other network-connected electronic system. The host interface **304** may implement the MAC Sublayer, optional Reconciliation Sublayer, and PCS Sublayer, and elements of the FEC, PMA, and PMD Sublayers, as application-specific integrated circuitry to enable high-rate processing and data transmission. The illustrated transceivers **302**, preferably embodied as application-specific integrated circuitry for very high-rate serial data transmission and reception, include multiple pairs of a receiver **306** and transmitter **308**, each pair coupled to two unidirectional channels (a receive channel and a transmit channel) to implement one lane of a multi-lane physical connection. The physical connection thus accepts from each node transmit channel signals representing a multi-lane transmit stream and conveys the signals to the other node, delivering them as receive channel signals representing a multi-lane receive data stream. The receive signals may be degraded due to the physical channel's introduction of noise, attenuation, and signal dispersion.

Note that while the following discussion glosses over the traces conveying AUI-compliant multi-lane stream **231** between the PMA Sublayer **230** and the PMD Sublayer **232**, focusing instead on communication channels **206**, the disclosed principles are also applicable to this so-called "chip-to-module" (C2M) link, which will receive more detailed consideration after the description of FIGS. 4-5.

FIGS. 4 and 5 provide a more detailed block diagram of illustrative transmit and receive chains in the sublayers below the PCS. The transmit chain in FIG. 4 accepts a four-lane data stream from the PCS. Pursuant to the standard, the PCS data stream is already encoded with a transmission code that provides DC balance and enables timing recovery. The PCS data stream lanes further include PCS alignment markers for synchronizing the lanes with each other. Once the data stream lanes are aligned, an alignment marker removal module **402** removes the alignment markers from each lane, passing them to a downstream alignment marker insertion module **406**. A transcoding module **404** modifies the transmission code from a 64b/66b code to a 256b/257b code more appropriate for use with the Reed-Solomon encoder. By repeatedly transcoding four 66-bit blocks taken in parallel from the four incoming lanes into individual 257-bit blocks, the transcoding module may essentially convert the four lanes into a single lane data stream.

The previously-mentioned alignment marker insertion module **406** accepts the PCS alignment marker information from removal module **402** and the single-lane data stream from transcoding module **404**. The insertion module **406** combines the alignment marker information from the four lanes in a manner discussed further below to form a set of grouped alignment markers in a 257-bit block and, accounting for the operation of the transcoding module **404**, inserts the alignment block in a fashion that preserves its location relative to the other data in the data stream **407**. As described further with respect to FIGS. 7A-7C, the alignment block is designed to account for the operation of the encoder module **408** and symbol distribution modules **409** such that the alignment markers appear essentially intact and in order in the two-lane transmit data stream crossing boundary **410**.

A Reed-Solomon (RS) encoder module **408** operates on 10-bit "symbol" blocks of the data stream **407** from the insertion module **406**, adding redundancy to enable down-

6

stream correction of symbol errors. Typically, the encoder module **408** operates to preserve the original data stream content while appending so-called "parity" information, e.g., 30 parity symbol blocks appended to 514 data symbol blocks to form a complete code word block. Thus the alignment blocks inserted by module **406** will remain present in the output data-stream from the encoder module. A symbol distribution module **409** distributes code word symbols across multiple transmission lanes in round-robin fashion, directing each lane to a corresponding transmitter. Boundary **410** may be considered as the boundary between the FEC sublayer **228** and the PMA sublayer **230**. Where it is desired to maintain this boundary as strongly as possible, the PMA sublayer may include an alignment marker detection module **411A** (and **411B**) for each lane of the transmit data stream to detect the alignment markers inserted by module **406** with suitable data buffering. Alternatively, this boundary can be relaxed and the alignment marker detection modules **411A**, **411B** omitted in favor of appropriate direct signaling from the alignment marker insertion module **406**. In either case, the training control modules **420A**, **420B** control the multiplexers **412A**, **412B** in each lane, based at least in part on detection signals indicating the presence of the alignment markers.

Multiplexers **412A**, **412B** forward the encoded data streams to serializer modules **414A**, **414B** during normal operations and in the absence of alignment markers. During auto-negotiation and training phases, the multiplexers supply negotiation and training data streams from the training control modules **420A**, **420B** to the serializers. During normal operations in the presence of alignment markers, the multiplexers **412A**, **412B** (acting as alignment marker replacement modules) supply the serializer modules with modified alignment markers as discussed further below. The serializers **414A**, **414B**, each accept a stream of transmit data blocks and convert the stream of blocks into a (higher-rate) stream of channel symbols. Where, for example, a 4-PAM signal constellation is used, each serializer may produce a stream of two-bit symbols.

Each stream of channel symbols is filtered by a pre-equalizer module **416A**, **416B** to produce a transmit signal, which is amplified and supplied to the transmit channel by a driver **418A**, **418B**. The pre-equalizer modules compensate for at least some of the channel dispersion, reducing or eliminating the need for receiver-side equalization. Such pre-equalization may be advantageous in that it avoids the noise enhancement often associated with receiver-side equalization and enables digital filtering with a reduced bit-width. The bit width reduction directly reduces power consumption by requiring a less complex filter, but may further reduce power consumption by obviating the parallelization that a more complex filter might require to operate at the required bandwidth. However, pre-equalization generally requires knowledge of the channel.

One or more training controllers **420A**, **420B**, operate to characterize the channel after conducting an initial auto-negotiation phase. During the auto-negotiation phase, at least one training controller generates a sequence of auto-negotiation frames conveying capabilities of the local node to the remote node and negotiating to select a combination of features to be used for subsequent communications. When the auto-negotiation phase is complete, each training controller generates a sequence of training frames, so that training is carried out independently on each of the lanes.

An illustrative training frame **602** is now described with reference to FIGS. 6A-6C. The training frame **602** begins with a frame marker **604** indicating the start of the training

7

frame. The frame marker **604** is followed by a coefficient update field **606**, a status report field **608**, and a training pattern **610**. Unlike the training pattern, which provides a spectrally-dense channel symbol sequence to facilitate training, the preceding fields are sent using differential Manchester encoding to facilitate timing recovery and ensure reliable communication even with untrained equalizers.

FIG. 6B shows an illustrative coefficient update field **606**, having a two-bit request field **612** to indicate whether a selected coefficient should be incremented, decremented, maintained at the present value, or disabled; a six-bit selection field **614** to select one of up to 64 pre-equalizer coefficients; a two-bit modulation field **616** to select a desired modulation scheme (e.g., 2-PAM, 4-PAM, with or without precoding); a one-bit filter length field **618** to indicate whether the pre-equalization filter should be short (no more than four coefficients) or long (more than four coefficients); and a two-bit initialization request field **620** to select a pre-programmed set of pre-equalization coefficient values. The unlabeled fields may be reserved for future use. The training controller(s) may use the coefficient update field **606** to convey backchannel information, i.e., adjustments to the pre-equalization filter coefficients of the remote transmitter.

FIG. 6C shows an illustrative status report field **608**, having a three-bit coefficient status field **622** to report a successful update to a coefficient, an unsuccessful update, or an error condition (e.g., coefficient not supported, maximum voltage or coefficient limit reached); a one-bit and five-bit (totaling six bit) coefficient selection echo field **624** to echo back the coefficient selection from the remote node; a one bit initial condition status field **626** to report successful or unsuccessful setting of the coefficient values to an initial set of preprogrammed values; a one-bit lock status field **628** to indicate whether the receiver has achieved frame lock; a modulation status field **630** to report which modulation scheme is being employed by the transmitter; and a one-bit receiver ready field **632** to indicate whether the local node has finished training and is ready to begin data transmission. The training controller(s) may use the status report field **608** to convey confirmations and other local status information to the remote node.

Returning to FIG. 4, the one or more training controllers **420A, B**, receive backchannel information extracted by the receiver from the received data stream and use the backchannel information to adjust the coefficients of the pre-equalization filters. The controllers further receive “remote info”, which includes locally-generated information for adapting the coefficients of the pre-equalization filter in the remote node. Based on this information the controllers populate the coefficient selection field **614** and request field **612** of the training frames to provide backchannel information to the remote node. As training frames are employed only during the training phase, and as it may be desirable to continue updating the pre-equalization filter during normal operations, the training controller(s) **420** may include similar backchannel information in or with the modified alignment markers supplied via multiplexers **412A, 412B** during normal operations.

Though shown as a single lane, the 10-bit symbol blocks of data stream **407** can conceptually be arranged into two lanes in anticipation of the operation of symbol distribution module **409**, which places adjacent symbol blocks in alternate lanes. FIG. 7A shows a 514 symbol message word **702** (corresponding to the length of twenty 257-bit blocks from transcoding module **404**) in this fashion, with broken outline for the 30 parity symbols **719** to be appended by the

8

Reed-Solomon encoder **408** to form a complete 544 symbol code word. The spacing of the alignment markers in the four-lane PCS data stream is such that when a set of grouped alignment markers are inserted in the data stream **407**, the grouped alignment markers form the first 257 bits of a message word **702** (and also the first 257 bits of the resulting Reed-Solomon codeword) while data blocks **718** occupy the remainder of the message word. Within these 257 bits, the information from the four PCS data stream alignment markers are arranged in alternating 10-bit symbol blocks so that within each lane the information appears, in order, and as essentially intact PCS alignment markers **704, 706, 708**, and **710**. As the four 64-bit alignment markers fall one bit short of the 257-bit length of blocks put out by the transcoding module, a one-bit pad **714** is appended to the grouped alignment markers. Moreover, within these first 257 bits of the message word **702**, the alignment markers don’t align well with the 10-bit symbol block boundaries. To fill out a full 10-bit symbol in lane 0, a 2-bit portion of marker **710** is excerpted from position **712**.

The alignment markers in the four lane PCS data stream carry bit-interleaved parity values (A, B, C, D in inverted and non-inverted versions) and employ unique marker (UM) patterns that are lane specific. The alignment insertion module **406** modifies the pattern of alignment marker **706** to agree with that of alignment marker **704**, so that they share a common marker (CM) pattern to facilitate detection and synchronization of the two lanes at the receiver. The BIP values are preserved. While alignment markers **708, 710** could in theory be used to make the synchronization process more robust, their presence is only necessary for identifying lanes at the beginning of the communications process and thereafter only to convey the contents of their BIP fields. Once the lanes have been identified, the lane-specific patterns for markers **708, 710** are redundant and these patterns may be omitted to make room for backchannel information. Thus multiplexers **412A, 412B**, optionally with the support of alignment marker detection modules **411A, 411B**, may replace the sets of grouped alignment markers with modified versions as shown in FIGS. 7B-7D.

In some embodiments, the replacements begin only after a “settling period” (e.g., for one second after the multi-lane communications begin) or after a predetermined number (e.g., 100) of grouped alignment markers have been sent, with all sets of grouped alignment markers being modified thereafter. In other embodiments, the replacements occur periodically, with 1 out of every 2 sets being replaced, or 2 out of 3, or (N-1) out of N, with N potentially ranging as high as 10^5 to provide at least one unmodified set on the order of once per second. In still other embodiments, the replacement may occur on an as-needed basis, e.g., when degradation or drift is observed in the performance of the equalizers. In these as-needed embodiments, the replacements may begin upon detection of the reduced performance and continue until training convergence is achieved.

In FIG. 7B the UM alignment markers **708** and **710** are replaced in their entirety by a backchannel adaptation field **722** having 66 bits and a backchannel status field **724** having 63 bits (including the former pad bit **714**). At the receiver, the backchannel information is extracted and the alignment markers **708, 710** restored with placeholders for the lost BIP (inverted and non-inverted) values. With SerDes systems employing forward error correction, the BIP values are no longer essential for determining bit error rates, and in this implementation the BIP values are sacrificed for additional

backchannel bandwidth that will not substantively impact the operation or throughput of the higher sublayers and layers.

Nevertheless, different implementations may be employed to prevent the loss of BIP values in the restored alignment markers. In FIG. 7C, the UM patterns and inverted BIP values are dropped from alignment markers **708**, **710** while the non-inverted BIP values are left in place. The UM patterns and inverted BIP values are replaced by an adaptation field **732A**, **732B** (for marker **708**) and a status field **734A**, **734B**. This approach preserves alignment markers **704** and **706** unchanged from the markers inserted by module **406** while reallocating 56 bits from each of markers **708**, **710** for hidden backchannel fields **732**, **734**. At the receiver, the backchannel information is extracted and the alignment markers **708**, **710** restored (and the BIP values preserved) without impact on the operation or throughput of the higher sublayers and layers.

In FIG. 7D, alignment markers **704**, **706** are modified to replace their inverted BIP values with the non-inverted BIP values from alignment markers **708**, **710**. Their CM patterns remain unaffected so that the standard synchronization process still functions. Meanwhile, additional space is made available for backchannel information. As in FIG. 7B, backchannel adaptation field **722** has 66 bits while backchannel status field **724** has 63 bits (including the former pad bit **714**). At the receiver, the backchannel information is extracted and the alignment markers **704**, **706**, **708**, **710** are fully restored (and the BIP values preserved) without impact on the operation or throughput of the higher sublayers and layers.

As with the training frame, the backchannel information field(s) of the modified alignment markers may provide a six-bit field for selecting pre-equalization filter coefficients; a two-bit field for specifying that the selected coefficient should be incremented, decremented, maintained, or disabled; and perhaps a one-bit field for indicating whether the backchannel information field contains a valid command or should be ignored. The backchannel information field(s) may further provide a status report, with a six-bit field for echoing a filter coefficient selection; a two-bit field for indicating whether the coefficient has been updated, not updated, is at its limit, or is not supported; and perhaps a one-bit field for indicating that the receiver has obtained a lock on the lane alignments.

In an alternative embodiment, each coefficient of the pre-equalization filter may have a bit pair allocated within the modified alignment markers to indicate (in the adaptation field) whether that coefficient should be incremented, decremented, maintained, or disabled; or (in the status field) to indicate whether that coefficient is maximized, minimized, updated, or not updated.

As with the training frame, the backchannel information fields of the modified alignment markers may be encoded using differential Manchester encoding to ensure reliable delivery. Additionally or alternatively, reliability of the backchannel information field delivery may be enhanced using other techniques. For example, the backchannel information field may be sent redundantly, using multiple copies of a field to enable error detection and (for three or more copies) voting-based error correction. (Notably, a standards-compliant alignment marker includes a bitwise-inverted duplication of each element, which can be extended to include the backchannel information field. If an odd number of copies is desired, the last field may be partly inverted.) A

parity check or short FEC code may also provide controlled redundancy enabling the detection and possible correction of bit errors.

Having discussed the transmit chain and the fields (including hidden fields **722-724** or **732-734**) potentially employed to communicate backchannel information during the training and normal operations phases, we turn now to the operation of an illustrative receive chain such as that shown in FIG. 5. The receive chain obtains analog electrical signals from different receive channels (indicated by Lane0-rx and Lane1-rx). These may be obtained directly from electrical conductors, if the physical medium is an electrical bus or cable, or indirectly via transducers if the physical medium is wireless. Low noise amplifiers (LNA) **502A,B**, each provide a high input impedance to minimize channel loading and amplifies the receive signal to drive the input of a continuous time linear equalizer (CTLE) filter **504A,B**.

CTLE filters **504A,B** provide continuous time filtering to shape the receive signal spectrum in an adaptive fashion to reduce the length of the channel impulse response while minimizing leading inter-symbol interference (ISI). Decision feedback equalizers (DFE) **506A,B** operate on the filtered signals to correct for trailing ISI and detect each transmitted channel bit or symbol, thereby producing a demodulated digital data stream. Some embodiments employ oversampling. Clock recovery and adaptation modules **508A,B** derive a sampling clock signal from the input and/or output of the DFE's decision element and supply it back to the DFE to control timing of the symbol detection. The adaptation modules **508A,B** further derive an error signal of the DFE decision element's input relative to the output or (during the training phase) to a known training pattern, and use the error signal to adapt the DFE coefficient(s) and the response of the CTLE filters. The adaptation modules still further use the error signal to generate "remote info", i.e., adaptation information for the remote pre-equalizers. This remote info is supplied to the training controller(s) **420** (FIG. 4).

Deserializers **509A,B** group the digital receive data stream bits or symbols into blocks to enable the use of lower clock rates for subsequent on-chip operations. Alignment marker detection modules **510A**, **510B** monitor the receive data stream to detect the CM pattern of the alignment markers and achieve alignment marker lock during normal operations, or during training operations to detect the training frame markers and achieve lock thereto. The backchannel information extraction modules **511A**, **511B** extract the backchannel information from the appropriate portions of the training frames and alignment markers, providing the pre-equalizer adaptation information and status report information to the training controller(s) **420**. During normal operations, the multiplexers **512A**, **512B**, operating under control of the extraction modules **511A**, **511B**, replace the modified alignment markers with sets of grouped PCS alignment markers, thereby hiding the backchannel information fields from the higher layers. As with the transmit chain, the receive chain may impose a hard boundary **526** between the PMA sublayer and the FEC sublayer, or alternatively, the alignment marker detection information may be communicated to the FEC lane deskew module **513**.

During normal operations, the receive data streams from the deserializers are aligned by an FEC lane deskew module **513**. If the FEC lanes have somehow gotten switched, an FEC lane reordering module **514** detects the switch based on the contents of the alignment markers and compensates for the switch as the two lanes are multiplexed on a symbol-block basis into a single lane. An RS decoder module **516**

11

operates on the single lane data stream to detect and correct any symbol errors, removing the FEC coding redundancy (parity symbols) during the decoding process.

A transcoding module **520** converts the 256b/257b transmission code words into blocks of four 64b/66b transmission code words distributing the 66-bit code word blocks across four PCS lanes. An alignment marker insertion module **524** converts the sets of grouped alignment markers into individual alignment markers having lane-specific UM patterns, and inserts the individual alignment markers at appropriate positions in the four lanes, accounting for the operation of the transcoding module **520**. The four lane PCS data stream is provided to the higher hierarchy layers of the node for eventual communication of the conveyed data to the destination application.

With a sufficiently long pre-equalization filter, the DFE feedback filter may be kept quite short, at say 1 or 2 taps, and even when the channel has a relatively lengthy channel response, the pre-equalization filter may obviate any need for a digital feed forward equalizer (FFE) filter. In this case, the power savings associated with this architecture is expected to be substantial. With temperature changes and general evolution of the channel, however, ongoing adaptation of the pre-equalization filter (i.e., during the normal operations phase) is expected to be necessary.

Returning now to the C2M link, the PMD Sublayer **232** may often be implemented in the form of a pluggable module (e.g., a small form-factor pluggable (SFP) or one of its many variations including SFP+, cSFP, SFP-DD, QSFP, QSFP+, QSFP14, QSFP28, QSFP56, etc.) that is inserted into a port in a host such as a switch or a router. The pluggable module may be essentially a concatenation of the PMA Sublayer and transceiver components set forth in FIGS. 4-5 to support bi-directional communications between the AUI lanes and the communication channels with or without the accompanying FEC encoding/decoding elements, as shown in FIG. 9. (The “Link Ctl” elements are each an aggregation of blocks **420**, **508**, and **511**.) The stream **231** being communicated between the PMA and PMD sublayers may be transmitted from a packaged integrated circuit mounted to a printed circuit board (PCB) in the host, through traces on the PCB, through a “mated connector” (i.e., a receptacle with plug inserted to provide electrical contact between respective pins), through a capacitance, through traces on the PCB in the module, to a packaged integrated circuit mounted to the PCB module. The stream also includes signals traveling the opposite direction. At the contemplated signaling rates (26.5625 GBd or 53.125 GBd), the C2M link model includes up to 20 dB of signal attenuation (up to 4 dB host package, up to 11.5 dB host side trace, about 2.5 db module side trace, and about 2 dB module package).

The pluggable module should be usable with the widest possible range of host ports. Even though each port may be advertised as being compliant with the standard, different host manufacturers may employ different integrated circuit packages and different PCB layouts to connect the packaged ICs to the ports, causing the equivalent C2M link model to vary significantly between manufacturers or even between different products of a given manufacturer. To a limited extent, the host transceivers could be equipped with transmit-side and receive-side equalization that at least partly compensates for such variation; however there are severe limitations on the area, complexity, and power available on the host side of the link. A better option is to provide the pluggable module with flexibility to adapt to the different C2M link characteristics.

12

Accordingly, the manufacturer of the pluggable module may formulate a range of C2M link models using different assumptions based on known or predicted layouts and link characteristics, including “reference” transceivers. For example, the host-side receivers might be presumed to employ a continuous time linear equalizer (CTLE) in combination with any one of: (a) a four-tap DFE with a maximum first post-cursor tap coefficient of 0.5; (b) a four-tap DFE with a maximum first post-cursor tap coefficient of 0.2; (c) a four-tap DFE with a maximum first post-cursor tap coefficient of 0.0; (d) a five-tap linear equalizer with a one-tap DFE; (e) a five-tap linear equalizer with no DFE. These receiver assumptions, combined with different host-side trace lengths and package losses, may give rise to a reasonable number of potential C2M link models categories.

With the foregoing context, we now turn to certain variations offering more efficient training and adaptation of the pre-equalization implemented by the transmitter-side feed forward equalizer (FFE), suitable for a standard interface. As an initial matter, we may specify a maximum number of FFE coefficients for the pre-equalizer, e.g., 64 coefficients. We may further specify a minimum range for each coefficient and a minimum resolution such as, e.g., $\frac{1}{64}$ or $\frac{1}{128}$ of full swing. The drivers **418** (FIG. 4) operate as digital to analog converters, which we may require to provide a minimum equivalent number of bits (ENOB), e.g., 5 bits. The resulting combinatorial range of potential FFE coefficient values may be impractically large for a naive training approach.

However, there are certain constraints, such as the reasonable number of C2M link model categories, that may be taken into account to improve the efficiency of the training and adaptation process. For the communication channels, we may require the channels to satisfy certain minimum qualification thresholds for, e.g., copper wire cabling or printed circuit board traces, which can be modeled. These models enable the design of “default” pre-equalizer coefficients for each case, e.g., minimum-quality channels, typical quality channels, and predictable variations thereof. The models, or at least the minimum qualification thresholds, may be chosen to require that the bulk of the pre-equalizer coefficient energy be concentrated at three or four major taps near the cursor (i.e., the FFE tap with nominally zero delay). The remainder of the coefficient taps may then be employed for reflection (“echo”) reduction.

As another constraint, the transmitter FFE is expected to be limited in terms of power and/or maximum permitted voltage swing, such that the sum total of the coefficient values (including the sum of their absolute values) can be no greater than a predetermined limit. Once the limit is reached, the increase in magnitude of one coefficient value may need to be offset by the reduction in magnitude of at least one other coefficient value.

With the foregoing constraints, the “reasonable number” of channel model categories may be on the order of 16 or 32. Accordingly, we may provide a set of, e.g., 16 or 32 registers for the training controllers **420** (FIG. 4) or link controllers (FIG. 9) to store default coefficient values. A first register may store a vector of coefficient values suitable for a first channel response (e.g., a minimum-quality channel); a second register may store a vector of coefficient values suitable for a second channel response (e.g., a near-minimum quality channel with reduced dispersion); and so on. In some contemplated embodiments, the default coefficient values for each register, or at least some of the registers, would be set by the standard based on different channel loss profiles, while others of the registers may be reserved for configu-

13

ration by customers of the original equipment manufacturer (OEM). In any case, at least some of the default coefficient values are preferably capable of being overwritten by the system controller or as part of the operating protocol to remember previous solutions and provide faster start-ups or failure recoveries. As part of the negotiating phase or training process, the coefficient values in each register may be communicated between the transmitter and receiver.

As previously mentioned, the channel qualification requirements are expected to enable pre-equalizer coefficient energy to be concentrated at the first three to four FFE taps, and thus in at least some contemplated embodiments the default coefficient values specified by the standard may be potentially nonzero for only the first four taps and may be zero for all higher taps. (In embodiments permitting the default coefficient values to be overwritten, this constraint may be omitted at least with respect to the replacement values.) To allow for fine-tuning of the default coefficient values, the standard may constrain the sum of the absolute value of the default coefficients to no more than certain percentage, e.g. 80%, of the predetermined limit for pre-equalizer. At least one of the default coefficient value registers is expected to provide a good starting point for training/adaptation, i.e., by providing an open decision eye at the input to the decision element if the receiver design is close to the reference receiver provided by the standard. With this starting point, the percentage constraint leaves room for convex optimization via known filter adaptation procedures such as, e.g., least mean squares (LMS) adaptation, or recursive least squares (RLS) adaptation. (The coefficient adaptations can be conveyed via the backchannel or via the management interface as discussed previously.) The reserved percentage (e.g., 20%) can thus be used to cancel reflections or increase signal swing as provided by the adaptation procedure.

Preferably, the predetermined limit on the sum of coefficient values is enforced (at least for the default coefficient values) by either the transmitter's training controller, or by the adaptation procedure which may be implemented at the receiver. If the limit is exceeded, the coefficient values for at least the major FFE taps, which may alternatively characterized as the first four taps or the largest coefficient magnitudes (though these are not necessarily the same), may be proportionally reduced to make room for further optimization.

Thus the pre-equalizer training/adaption process may proceed in stages. In a first stage, the default coefficient values are determined and stored in a set of registers. The default values may be set by the standard, and overwriting of the values by manufacturers, by system vendors, or alternatively by system controllers in the course of normal operations, may be permitted. In a second stage, the default coefficient values are communicated between the transmitter to the receiver, e.g., as the result of a register scan requested by the receiver during the auto-negotiation and/or training process when the channel is enabled or is recovering from failure. The receiver may select the starting point for the pre-equalizer optimization process by selecting one of the registers. In the third stage, as the training phase progresses, the receiver may communicate pre-equalization coefficient updates to the transmitter to reduce the bit error rate (BER). Such updates may also be communicated via the hidden backchannel discussed previously, or via a management interface, during the ongoing operation of the channel.

At least some system embodiments will support simultaneous updating of multiple pre-equalizer coefficient values. In at least some implementations, these systems will also

14

support coarse adaptation and fine adaptation steps to accelerate the training process. The training controller may maintain status registers with fields for the sum of absolute values of current coefficient values, the total margin (i.e., difference between the sum and the predetermined limit), and the margin for each coefficient (i.e., how far the current value is from the maximum or minimum value, whichever is nearest). The fields may be bit-limited without rollover, i.e., the maximum value also represents all larger values. For example the total margin field may have four bits, while the margin field for each coefficient might have only two bits.

FIG. 8 is a flow diagram of the an illustrative communications method. As the link is first established, which may happen, e.g., during automated testing as part of the manufacturing process, when a module is inserted into a port, and/or when a cable connector is inserted into a module the training controller or link controller (hereafter, just "controller") employs an autonegotiation or backchannel communication process in block 802 to determine whether the remote transmitter has a set of preset coefficient registers. A check is made in block 804 to determine whether any such registers remain untried. If so, in block 806, the controller selects the first remaining untried register by instructing the remote transmitter to program the transmit pre-equalizer in each lane with the preset coefficient values from that register. The remote transmitter uses the pre-equalizers to transmit data, and in block 808 the receiver demodulates the multiple receive signals, e.g., with a DFE, to estimate or measure error signals (the difference between the equalized receive signals and an idealized receive signal). For training patterns, the idealized receive signal is known, whereas for transmitted data, the idealized receive signal may be estimated. The energy or average amplitude of the difference is determined in block 810 to serve as an indicator of how well the selected preset coefficient values are performing.

Blocks 804-810 are repeated until the performance of each of the preset coefficients in each register have been evaluated. Then, in block 812, the best register is selected, with the remote transmitter being instructed to use those preset coefficients as a starting point for the next phase. The remote transmitter programs the pre-equalizer with the preset coefficients from the selected register and transmits a training pattern and/or data. In block 818, the receiver demodulates the receive signals, again obtaining error signals. The demodulated data streams are later combined to form a multi-lane data stream carrying user data to a destination application. In block 820 the controller uses the error signals to derive adaptation information for the local receive equalizers and the remote transmit equalizers, using any applicable adaptation algorithm that may be found in the open literature. The adaptation information may be applicable to the CTLE filter, any FFE filter that may be present, and the feedback filter of the DFE (referred to herein as "local" adaptation info), as well as to the pre-equalization filter of the remote node (referred to herein as "remote" adaptation info).

In block 822, the remote adaptation information is communicated to the remote transmitter, optionally in the form of modified alignment markers having backchannel information fields for conveying the remote adaptation info. In block 824, the backchannel information from the remote transmitter is obtained, optionally by extracting the contents of the backchannel information field from the received (modified) alignment markers, yielding adaptation information for the local pre-equalizer. The modified alignment markers are then replaced with reconstructed sets of grouped PCS alignment markers. In block 826, the local adaptation

15

information and the backchannel information are used to adjust the receive equalizers (CTLE, DFE, and FFE if any) and the local transmit pre-equalizer.

Numerous alternative forms, equivalents, and modifications will become apparent to those skilled in the art once the above disclosure is fully appreciated. For example, the foregoing description suggests specific locations and bit-widths for fields, yet any suitable location and width that can be standardized and agreed upon may be employed. For example, adaptation and status information fields may be present in each lane rather than limiting the adaption information to one lane and the status information to the other. It is intended that the claims be interpreted to embrace all such alternative forms, equivalents, and modifications that are encompassed in the scope of the appended claims.

What is claimed is:

1. A SerDes communications method that comprises, in a transceiver:

selecting one of multiple registers to specify initial pre-equalizer coefficient values, each of the multiple registers corresponding to a different channel model; updating the initial pre-equalizer coefficient values during a training phase; and using the updated pre-equalizer coefficient values to convey a transmit data stream.

2. The method of claim 1, wherein at least some of the different channel models are for chip-to-module (C2M) channels with different insertion loss and package loss characteristics.

3. The method of claim 1, wherein at least some of the different channel models presume different types of receiver equalization.

4. The method of claim 1, wherein said selecting includes using the initial pre-equalizer coefficient values to determine a performance characteristic for each of the multiple registers.

5. The method of claim 4, wherein the performance characteristic is an error signal energy.

6. The method of claim 4, wherein the performance characteristic is a bit error rate.

7. The method of claim 4, further comprising updating the selected register with the updated pre-equalizer coefficient values.

8. A chip-to-module communications link that comprises a port connector coupling a port transceiver to a pluggable module transceiver, the pluggable module transceiver including:

one or more transmit filters to each pre-equalize a corresponding serial symbol stream being transmitted to the port transceiver; and a controller having multiple registers, each of the multiple registers containing a set of initial coefficient values

16

corresponding to a different channel model, the controller using one of the registers to set initial coefficient values for the one or more transmit filters.

9. The communications link of claim 8, wherein at least some of the different channel models are for chip-to-module (C2M) channels with different insertion loss and package loss characteristics.

10. The communications link of claim 8, wherein at least some of the different channel models presume different types of receiver equalization.

11. A chip-to-module communications link that comprises a port connector coupling a port transceiver to a pluggable module transceiver, the pluggable module transceiver including:

one or more transmit filters to each pre-equalize a corresponding serial symbol stream being transmitted to the port transceiver; and

a controller having multiple registers, each of the multiple registers containing a set of initial coefficient values for which the port transceiver determines a performance characteristic,

the controller using one of the registers selected by the port transceiver to specify the initial coefficient values for the one or more transmit filters.

12. The communications link of claim 11, wherein the performance characteristic is an error signal energy.

13. The communications link of claim 11, wherein the performance characteristic is a bit error rate.

14. The communications link of claim 11, wherein the port transceiver generates updates for the coefficient values of the one or more transmit filters.

15. The communications link of claim 14, wherein the controller saves updated coefficient values in one of the registers.

16. A pluggable module transceiver including:

one or more transmit filters to each pre-equalize a corresponding serial symbol stream being transmitted to a port transceiver; and

a controller having multiple registers, each of the multiple registers containing a set of initial coefficient values corresponding to a different channel model, the controller using one of the registers to set initial coefficient values for the one or more transmit filters.

17. The pluggable module transceiver of claim 16, wherein the port transceiver determines a performance characteristic for each register.

18. The pluggable module transceiver of claim 17, wherein the performance characteristic is an error signal energy.

19. The pluggable module transceiver of claim 17, wherein the performance characteristic is a bit error rate.

* * * * *