(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 September 2016 (15.09.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/144953 A1

(51) International Patent Classification:
H04L 12/26 (2006.01) *H04L 1/20* (2006.01)

(21) International Application Number:
PCT/US2016/021366

(22) International Filing Date:
8 March 2016 (08.03.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/130,980 10 March 2015 (10.03.2015) US

(71) Applicant: **INTEL CORPORATION** [US/US]; 2200 Mission College Boulevard, Santa Clara, California 95054 (US).

(72) Inventor: **RAN, Adee O.**; POB 88, 12220 Maayan Baruch (IL).

(74) Agent: **BURNETT, R. Alan**; Law Office of R. Alan Burnett, PS, c/o CPA Global, P.O. Box 52050, Minneapolis, Minnesota 55402 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) Title: MONITORING ERRORS DURING IDLE TIME IN ETHERNET PCS

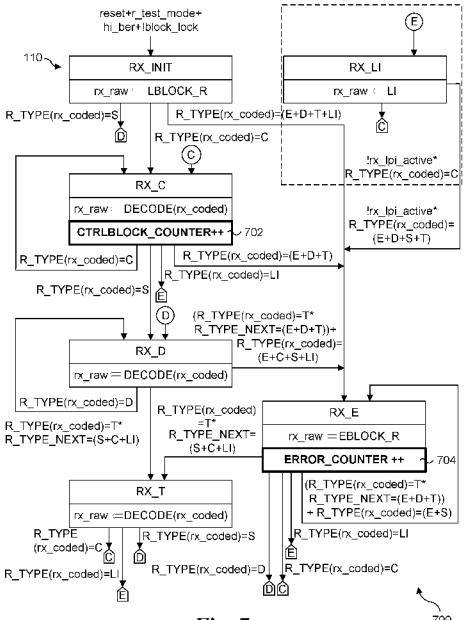


Fig. 7

(57) Abstract: Methods and apparatus for monitoring errors during idle times in high-speed links including a Physical Coding Sublayer (PCS). A denominator counter is implemented in a receive state machine of a PCS to count control blocks primarily consisting of idle control blocks or to count idle characters. An error counter is also implemented in the receive state machine to count errors while in an idle state. While operating the link in idle states, the counts of the denominator counter and error counter are used as inputs to various Bit Error Rate (BER) functions to estimate a BER of the PCS during idle times. The method and apparatus may be implemented in various high-speed link including Ethernet links.

WO 2016/144953 A1

MONITORING ERRORS DURING IDLE TIME IN ETHERNET PCS

BACKGROUND INFORMATION

25 Gigabit per second (Gb/s) Ethernet is a new standard for Ethernet being developed in the Institute of Electrical and Electronics Engineers (IEEE) P802.3by task force. Like other modern Ethernet standards, the Physical Layer (PHY) is specified to operate at a bit error ratio (BER) of less than 10^{-12}. The high data rate makes this BER requirement more challenging than in past generations, and for typical links it is almost certain that errors will occur occasionally (such that the BER is not practically equal to 0).

Errors can cause data loss if they occur within Ethernet frame boundaries. Ethernet frames include a 32-bit CRC (Cyclic Redundancy Check) field, so an Ethernet frame that was corrupted by errors is practically guaranteed to be detected by the MAC (Media Access Channel) and discarded. The MAC typically has counters for bad CRC frames, and these counters can be monitored to check the frame error rate. For example, a fully utilized 10 gigabit Ethernet link with the worst case BER is expected to encounter a CRC error approximately once in every 100 seconds. Many systems expect much lower error rates (such as less than once a week).

If the MAC CRC counters advance faster than expected, it can indicate a hardware problem. Network management systems can monitor the counters across multiple network nodes and trigger some maintenance action (such as board or cable replacement) on any indication of a high CRC error count. Thus, CRC error counting is a well-known and widely used feature of the Ethernet MAC.

A problem with MAC CRC error counting is that many Ethernet links are not fully utilized and may have long idle times between packets. Errors during idle times do not corrupt data so are considered safe, but they also do not cause CRC errors, so their rate cannot be monitored by the MAC counters.

Some Ethernet PHYs include forward error correction (FEC) codes, and the FEC decoding function can monitor the number of corrected errors. In a bad link, errors occur very often, and while the FEC decoder can recover the MAC data, it also counts the number of corrections. The Ethernet standard includes error counters for each kind of FEC. These counters provide error rate information with very fine resolution and operate in both data and idle times. However, FEC decoders increase the data delay (latency) in the receiver, and may increase power consumption, so some applications choose to disable them and some products do not even implement them. Therefore error monitoring may not always be available. Specifically, in the 25 Gb/s Ethernet project, although two kinds of FEC encoding are available, there is a strong desire to also enable operation without FEC encoding.

The 40 Gb/s and 100 Gb/s Ethernet varieties, which operate over multiple physical lanes,

use the PCS (Physical Coding Sublayer) defined in clause 82 of IEEE 802.3-2012. This PCS periodically inserts alignment markers into the data stream to enable assembling the multi-lane data in the receiver correctly. The alignment markers include a special field (bit interleaved parity, or BIP) that serves as an error monitoring mechanism at the PCS level. Any single bit error between two alignment markers would cause an incorrect BIP field and increment a BIP error counter. The BIP provides good error monitoring capability in both data and idle times. However, it is not available in the single-lane PCS chosen for the 25 Gb/s, which is based on IEEE 802.3-2012 clause 49.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing aspects and many of the attendant advantages of this invention will become more readily appreciated as the same becomes better understood by reference to the following detailed description, when taken in conjunction with the accompanying drawings, wherein like reference numerals refer to like parts throughout the various views unless otherwise specified:

Figure 1 is a schematic diagram depicting the relationship between the 10GBASE-R PCS and its associated sublayers;

Figure 2 is a schematic diagram illustrating the structure of an Ethernet Packet format;

Figure 3 is a schematic diagram illustrating the relationship of the reconciliation sublayer 102 and the 10 Gigabit Media Independent Interface (XGMII) to the ISO/IEC (IEEE) OSI reference model;

Figure 4 is a diagram illustrating the parts of an XGMII data stream;

Figure 5 is a schematic diagram illustrating mapping of data octets to lanes for the XGMII transmission and reception of bits in an XGMII data stream;

Figure 6 is a table depicting control codes user by the PCS sublayer; and

Figure 7 is a PCS receive state machine diagram in which a new control block counter and a new error counter are implemented to determine PCS BER during idle times, according to one embodiment.

DETAILED DESCRIPTION

Embodiments of methods and apparatus for monitoring errors during idle time in Ethernet Physical Coding Sublayer (PCS) are described herein. In the following description, numerous specific details are set forth to provide a thorough understanding of embodiments of the invention. One skilled in the relevant art will recognize, however, that the invention can be practiced without one or more of the specific details, or with other methods, components, materials, *etc.* In other instances, well-known structures, materials, or operations are not shown or described in detail to avoid obscuring aspects of the invention.

Reference throughout this specification to “one embodiment” or “an embodiment” means

that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, the appearances of the phrases "in one embodiment" or "in an embodiment" in various places throughout this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments.

For clarity, individual components in the Figures herein may also be referred to by their labels in the Figures, rather than by a particular reference number. Additionally, reference numbers referring to a particular type of component (as opposed to a particular component) may be shown with a reference number followed by "(typ)" meaning "typical." It will be understood that the configuration of these components will be typical of similar components that may exist but are not shown in the drawing Figures for simplicity and clarity or otherwise similar components that are not labeled with separate reference numbers. Conversely, "(typ)" is not to be construed as meaning the component, element, *etc.* is typically used for its disclosed function, implement, purpose, *etc.*

In accordance was aspect of the embodiments describe herein, techniques are disclosed for monitoring errors in high-speed Ethernet links during idle periods. The techniques are implemented at the Ethernet PCS sublayer. The Ethernet PCS sublayer (such as the 10GBASE-R PCS described in clause 49 of IEEE802.3-2012) is implemented in the Ethernet Physical layer (Ethernet PHY). Figure 1 depicts the relationship between the 10GBASE-R PCS and its associated sublayers.

In one embodiment, the techniques are implemented for the new 25 Gb/s standard for Ethernet being developed in the IEEE P802.3by task force. The proposed 25 Gb/s Ethernet employs the a similar PCS as defined by IEEE802.3-2012 clause 49, which defines the 10GBASE-R PCS, except the data rate is 25 Gb/s rather than 10 Gb/s. Figure 1 depicts the relationship between the 10GBASE-R PCS and its associated sublayers.

As illustrated, a 10GBASE-R PCS sublayer 100 interfaces with a reconciliation sublayer 102 that sits below the MAC layer 104 in the Data Link layer. 10GBASE-R PCS sublayer 100 also interfaces with a serial PMA (Physical Media Attachment) sublayer 106 in the 10GBASE-R PHY.

Figure 2 illustrates the structure of an Ethernet Packet format. The fields of the packet include a Preamble 200, a Start Frame Delimiter (SFD) 202, the addresses of the MAC frame's destination address 204 and source address 206, a length or type field 208 to indicate the length or protocol type of the following field that contains the MAC client data 210, a pad field 212 that contains padding if required, and a Frame Check Sequence (FCS) field 214 containing a cyclic

redundancy check value to detect errors in a received MAC frame. An optional extension field 216 is added, if required. Of these fields, all are of fixed size except for the MAC Client Data 210, pad field 212 and extension field 216, which may contain an integer number of octets between the minimum and maximum values that are determined by the specific implementation of the MAC.

Figure 3 illustrates the relationship of the reconciliation sublayer 102 and the 10 Gigabit Media Independent Interface (XGMII) 300 to the ISO/IEC (IEEE) OSI reference model. The purpose of the XGMII is to provide a simple, inexpensive, and easy-to-implement interconnection between the MAC sublayer 104 and PHY 302. A 10 Gigabit Attachment Unit Interface (XAUI) may optionally be used to extend the operational distance of the XGMII with reduced pin count (as defined in clause 47).

Reconciliation sublayer 102 adapts the bit serial protocols of the MAC to the parallel encodings of 10 Gb/s PHYs (when applied to 10GBASE-R). The 10 Gb/s PCS is specified to the XGMII, so if not implemented, a conforming implementation will behave functionally as if the reconciliation sublayer and XGMII were implemented. In one embodiment, the XGMII is proposed to be used for the 25 Gb/s Ethernet as defined in IEEE 802.3by. Optionally, a new 25G-MII (to be defined in clause 106) that is similar to XGMII is implemented (not shown)

Packets transmitted through the XGMII are transferred within the XGMII data stream. The data stream is a sequence of bytes, where each byte conveys either a data octet or control character. The parts of the data are shown in Figure 4, and include an inter-frame, followed by a preamble, and SFD, data, and the EFD. For the XGMII, transmission and reception of each bit and mapping of data octets to lanes is as shown in Figure 5.

The inter-frame <inter-frame> period on an XGMII transmit or receive path is an interval during which no frame data activity occurs. The <inter-frame> corresponding to the MAC interpacket gap begins with the Terminate control character, continues with Idle control characters and ends with the Idle control character prior to a Start control character. The length of the interpacket gap may be changed between the transmitting MAC and receiving MAC by one or more functions (e.g., reconciliation sub-layer (RS) lane alignment, PHY clock rate compensation, or 10GBASE-W data rate adaptation functions). The minimum interpacket gap at the XGMII of the receiving RS is five octets.

The signaling of link status information logically occurs in the <inter-frame> period, as described in IEEE802.3-2012 subclause 46.3.4. IEEE802.3-2012 subclause 46.3.3 describes frame processing when signaling of link status information is initiated or terminated.

Idle control characters (/I/) are transmitted when idle control characters are received from the XGMII. Idle characters may be added or deleted by the PCS to adapt between clock rates. /I/

insertion and deletion occurs in groups of 4. /I/s may be added following idle or ordered sets, but are not added while data is being received. When deleting /I/s, the first four characters after a /T/ are not be deleted.

To communicate Low-power idle (LPI) (a special idle-like sequence that may be sent to signal long idle periods), LPI control character /LI/ is sent continuously in place of /I/. LPI control characters are transmitted when LPI control characters are received from the XGMII. LPI characters may be added or deleted by the PCS to adapt between clock rates in a similar manner to idle control characters. /LI/ insertion and deletion occurs in groups of four. /LI/s may only be added following other LPI characters. The ability to send or receive LPI characters is an optional function.

Figure 6 is a table 600 depicting control codes user by the PCS sublayer. In addition to Idle /I/ and LPI /LI/, Table 600 includes a Start control character /S/, a Terminate control character /T/, an Error control character /E/, and a ordered set control character /Q/. The Start control character (/S/) indicates the start of a packet, while the terminate control character (/T/) indicates the end of a packet. The ordered_set control characters (/O/) indicate the start of an ordered_set. The Error /E/ is sent whenever an /E/ is received. It is also sent when invalid blocks are received. The /E/ allows physical sublayers such as the XGXS and PCS to propagate received errors.

The PCS maps XGMII signals into 66-bit blocks, and vice versa, using a 64B/66B coding scheme. The synchronization headers of the blocks allow establishment of block boundaries by the PCS Synchronization process. Blocks are unobservable and have no meaning outside the PCS. The PCS functions ENCODE and DECODE generate, manipulate, and interpret blocks as provided by the rules in IEEE802.3-2012 subclause 49.2.4.

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. The encodings defined by the transmission code ensure that sufficient transitions are present in the PHY bit stream to make clock recovery possible at the receiver. The encoding also preserves the likelihood of detecting any single or multiple bit errors that may occur during transmission and reception of information. In addition, the synchronization headers of the code enable the receiver to achieve block alignment on the incoming PHY bit stream. The 64B/66B transmission code specified for use in this standard has a high transition density and is a run-length-limited code.

The Ethernet PCS sublayer (such as the 10GBASE-R PCS) include a basic capability for monitoring errors, called BER monitor. This is a function of the PCS that detects errors only in predefined locations in the data stream (-3% of the total bits) in a short time window (125

microseconds). It is aimed at detecting high BER, which typically occurs when the medium is physically disconnected or when the remote link partner breaks the link. When triggered, it is visible to management. However, its BER threshold is 10^{-4}, much worse than the operational requirement, so it provides only an indication of extreme conditions, and does not provide link health assessment.

In accordance with aspects of the embodiments now disclosed, a PCS decoder is used to detect errors during idle times and define a new counter to count these events. This, in addition to CRC error counting at the PCS, provides monitoring of the error rate without the limit of not detecting errors during idle periods.

The PCS decoder has special rules for correctness of data blocks. As a result of these rules, a single error in a sequence of idle characters is practically guaranteed to be detected. For example, the PCS is configured to detect invalid sequences of characters and replace them with "error" characters.

The PCS decoding converts the serial bit stream into octets (bytes), where each octet can be either data or a control character. During idle times the transmitter repeatedly sends Idle (/I/) control characters. Exiting idle transmission is allowed only by sending "SPD" (start packet delimiter — another control character) followed by data characters. Therefore any error that corrupts an idle character into something other than SPD (idle error) can be detected as an invalid octet. The PCS decoding is specified to convert any block of 64 bits that contains an invalid octet into 8 "error" characters (/E/) so that corrupted data will not reach the MAC.

A single error or a burst of errors that occur within in an idle character (octet value 00) can change it to any other control character. Almost all other control characters are invalid and cause the current 64-bit block to be converted to 8 error characters /E/. The exception is the error character itself. Any such event can be recoded and counted.

IEEE802.3-2012 subclause 49.2.4 contains a full description of the PCS block structure, control character types and rules. Subclause 49.2.13.2.3 includes a definition of R_BLOCK TYPE which lists the possible received block types (idle characters are mostly included in blocks of type "C"). In further detail, the definition for R_BLOCK TYPE is,

R_BLOCK TYPE = {C, S, T, D, E, LI}

This function classifies each 66-bit rx_coded vector as belonging to one of the following types depending on its contents.

Values: C; The vector contains a sync header of 10 and one of the following:

a) A block type field of 0x1e and eight valid control characters other than /E/; and, if the EEE capability is supported, zero or four of the characters are /LI/;

b) A block type field of 0x2d or 0x4b, a valid O code, and four valid control characters;

c) A block type field of 0x55 and two valid O codes.

LI; For EEE capability, the LI type is supported where the vector contains a sync header of 10, a block type field of 0x1e, and eight control characters of 0x06 (/LI/).

S; The vector contains a sync header of 10 and one of the following:

a) A block type field of 0x33 and four valid control characters;

b) A block type field of 0x66 and a valid O code;

c) A block type field of 0x78.

T; The vector contains a sync header of 10, 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 sync header of 01.

E; The vector does not meet the criteria for any other value.

In one embodiment, a control block counter and an error counter are added to the receive state machine defined in clause 49, as shown in a receive state machine diagram 700 Figure 7. Receive state machine diagram 700 is similar to the receive state machine diagram of Figure 49-17 of IEEE802.3-2012 clause 49, with the addition of a control block counter (CTRLBLOCK_COUNTER) 702 and an error counter (ERROR_COUNTER) 704.

Receive state machine diagram 700 shows the state machine that detects legal character sequences. After a reset, the receive state enters RX_INIT (receiver initialization) state 706. Depending on the R_TYPE of the received LBLOCK_R, the state will advance to a state RX_C (for an R_BLOCK TYPE=C), a state RX_D (for an R_BLOCK TYPE=D), or a state RX_E (for an R_BLOCK TYPE=E). The state may advance to a state RX_T from either state of states RX_D and RX_E.

Under normal operation, idle characters (R_BLOCK TYPE C) keep the state machine in state RX_C. While in state RX_C, each control block (which includes an idle character or an ordered set of codes) is counted by control block counter 702. The only other valid block type (except for an additional idle) is S, which denotes start of a new frame, and causes transition to state RX_D. Any other code cause a transition to state RX_E and replacement of the current 8 characters with an "error" value EBLOCK_R. While in state RX_E, each error is counted by error counter 704.

As another option (not shown), the error counter may be implemented as part of the check performed on received 64-bit blocks, as described in the definition of R_BLOCK_TYPE above.

It should be noted that the PCS input is scrambled data, and it includes a descrambling function. One known feature of the descrambler is that it translates any single error on its input

to 3 errors on its output (also known as error multiplication). Since the PCS decodes the idle characters after descrambling, any error even on the scrambled data would cause 3 errors in the decoder. Accordingly, the error counter is expected to advance three times per single error event. An error counter with a value that is not divisible by 3 may indicate a burst of errors.

Based on the ratio of the error count to the received control block count, a BER may be calculated. Periodically, each of control block counter 702 and error counter 704 may be reset to prevent overflow.

In one embodiment, error counters are 32 bits wide and exposed to network management through the standard MDIO interface. The MDIO defines features useful for these counters, such as atomic reads and clear-on-read. Optionally, a rate of idle errors exceeding some threshold can be specified to break the link, since it signals that the link is unsafe.

Under some conventional approaches, estimation of the BER from error counters is performed, by monitoring the error counters over some period of time, and using the number of bits received during this time (roughly known from the data rate) as the denominator in the BER expression:

$$\text{estimated BER (general)} = \frac{\text{bit error count}}{\text{bit rate} \times \text{measurement time}}$$

For the proposed solution, counting the number of errors over a period does not provide sufficient information to estimate the BER, since this mechanism detects only errors in the idle characters between frames, and the number of idle characters depends on link utilization (or "sparseness" of frames) which is unknown. Therefore the denominator for the expression above is not available. In order to enable calculating the BER, another counter can be added (as discussed above), to estimate the denominator.

The denominator counter could count idle characters (each being 8 bits long), or idle blocks (blocks are 64 bits long, either blocks that contain at least one idle character, or blocks that contain only idle characters, or similar variations). Assuming idle periods are longer than non-idle periods (utilization of <50%), the number of idle characters will be approximately equal to 8 times the number of idle blocks, and most bit errors in the receiver will corrupt 3 idle characters residing in 2 idle blocks, due to error multiplication effect of the descrambler. This can be used to estimate the BER as follows

If counting at the character level:

$$\text{estimated BER (character counting)} = \frac{\text{value of character error counter}/3}{8 \times \text{number of idle characters received}}$$

If counting at the block level:

$$\text{estimated BER (block counting)} = \frac{\text{value of block error counter}/2}{64 \times \text{number of idle blocks received}}$$

The factor of 2 or 3 is often considered insignificant and estimation with either character error counting or block error counting in the numerator would practically yield the same result.

An idle character or block counter is expected to advance very quickly, especially when the link is under-utilized (for example, with a 25 Gb/s link that is 100% idle, a 32-bit counter of idle blocks will overflow after 11 seconds, while errors are expected to occur only once per 40 seconds if the BER is 10^{-12}). A counter should be able to accurately count to a high enough value, although the exact value is not important for calculation.

A possible solution is to use a long counter, e.g. 48 bits, but report only the most significant part, e.g., the 32 topmost bits. This can be implemented by various well-known mean in efficient ways. The number of idle characters/blocks received can then be approximated by multiplying the counter value by the appropriate power of 2, e.g. 2^{16}. A 48-bit idle blocks counter in an idle 25 Gb/s link can count about 200 hours without overflowing, and the number of blocks can be estimated from the 32 topmost bits with a resolution equivalent to less than 1 millisecond (negligible compared to a period of 40 seconds expected between errors when BER=10^{-12}). On an idle link, this enables estimating the BER with a resolution of 10^{-12} within a few minutes. With longer measurement times, BER can be estimated with a resolution better than 10^{-16}.

Figure 8 shows an architecture 800 for a network node employing a network chip 802 configured to BER determination during idle periods in accordance with aspects of the embodiments disclosed herein. Network chip 802 comprises PHY (Physical Layer) circuitry 804 including a Physical Coding Sublayer (PCS) module 806, a Physical Medium Attachment (PMA) module 807, a PMD module 808, a BER measurement module 809 including state machine logic 810 for implementing the state machine in Figure 7, a transmitter port 812 including transmitter circuitry 813 and a receiver port 814 including receiver circuitry 815. Network chip 802 further includes a DMA (Direct Memory Access) interface 816, an Input/Output (I/O) interface comprising a Peripheral Component Interconnect Express (PCIe) interface 818, a MAC (Media Access Channel) module 820 and a Reconciliation Sublayer (RS) module 822. Network node 800 also comprises a System on a Chip (SoC) 824 including a Central Processing Unit (CPU) 826 having one or more processor cores, coupled to a memory interface 828 and a PCIe interface 830 via an interconnect 832. Memory interface 828 is further depicted as being coupled to memory 834. Under a typical configuration, network chip 802, SoC 824 and memory 834 will be mounted on or otherwise operatively coupled to a circuit board 836 that includes wiring traces for coupling these components in communication, as depicted by single lines connecting DMA 816 to memory 834 and PCIe interface 818 to PCIe interface 830 at a PCIe port 838.

In one embodiment, MAC module 820 is configured to implement aspects of the MAC layer operations performed that are well-known in the art. Similar, RS module 822 is configured to implement reconciliation sub-layer operations.

During idle periods, BER measurement module 809 is implemented for determining BER. During both data and idle periods, data is exchanged between PHY transmitter and receiver ports 813 and 815 of node 800 and its link partner, as depicted by a link partner 844 including a receiver port 846 and a transmitter port 848. In one embodiment the configuration of node 800 and link partner 844 are similar, and are linked in communication via an Ethernet link 850.

In one embodiment, network chip 802 comprises a 25 Gb/s Ethernet Network Interface Controller (NIC) chip employing a 25GBASE-KR PHY or a 25GBASE-CR PHY. However, the circuitry and components of network chip 802 may also be implemented in other types of chips and components, including SoCs, multi-chip modules, and NIC chips including support for multiple network interfaces (*e.g.*, wired and wireless).

In general, aspects of the idle link error-detection embodiments disclosed herein may be implemented hardware (via, *e.g.*, embedded logic), or via a combination of hardware and software. For example, aspects of the operations performed by the embodiments may be implemented via embedded logic in a NIC, large-scale network interface, or the like.

In addition to implementation in 25 GB/s Ethernet links, aspects of the embodiments disclosed herein may be implemented in other high-speed links. These include, but are not limited to current and future Ethernet links, Peripheral Component Interconnect Express (PCIe) links, Universal Serial Bus (USB) links, Serial ATA (SATA) links, InfiniBand links, RapidIO links, and Intel® OmniPath links. The techniques disclosed herein are particularly advantages for high-speed links that do not employ forward error correction.

Further aspects of the subject matter described herein are set out in the following numbered clauses:

1. A method for monitoring errors during idle states in a high-speed link including a Physical Coding Sublayer (PCS), comprising:

implementing a denominator counter in a receive state machine of the PCS;

implementing an error counter in the receive state machine to count errors while in the idle state;

while the high-speed Ethernet link is operating in an idle state,

counting one of idle characters or idle blocks received over the high-speed link with the denominator counter

counting one of character errors or block errors that are detected in the idle characters or idle blocks received over the high-speed link with the error counter; and

estimating a Bit Error Rate (BER) when operating in the idle state as a function of outputs from the denominator counter and error counter.

2. The method of clause 1, wherein the high-speed link comprises an Ethernet link including an Ethernet PCS.

3. The method of clause 2, wherein the Ethernet link includes an Ethernet Physical Layer (PHY) including a PCS, and wherein the Ethernet PHY does not include a Forward Error Correction (FEC) sublayer.

4. The method of clause 2, wherein the Ethernet link comprises one of a 10 Gigabits per second (Gb/s), 25 Gb/s, or 40 Gb/s Ethernet link.

5. The method of clause 2, wherein the PCS has a block structure in accordance with IEEE802.3-2012 subclause 49.

6. The method of any of the preceding clauses, wherein the denominator counter is a control block counter that counts idle blocks, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/2}{64 \times number\ of\ idle\ blocks\ received.}$$

7. The method of any of the preceding clauses, wherein the denominator counter is an idle character counter that counts idle characters, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/3}{8 \times number\ of\ idle\ characters\ received.}$$

8. The method of any of the preceding clauses, further comprising:

employing an n-bit counter in the denominator counter;

reporting out m topmost bits of the n-bit counter; and

multiplying the m topmost bits by 2^{n-m} to determine one of a number of idle characters or idle blocks received.

9. The method of any of the preceding clauses, further comprising detecting that a BER during operation in an idle state exceeds a threshold, and in response thereto, one of resetting the link or disabling the link.

10. An apparatus configured to communicate with a link partner over a high-speed link, comprising:

Physical Layer (PHY) circuitry and logic, including,

a transmitter port including transmitter circuitry; and

a receiver port including receiver circuitry; and

Physical Coding Sublayer (PCS) circuitry and logic including a receive state machine having a denominator counter and an error counter,

wherein when the high-speed link is operating in an idle state the PCS circuitry and logic is configured to,

count one of idle characters or idle blocks received at the receiver port with the denominator counter;

count one of character errors or block errors that are detected in the idle characters or idle blocks received at the receiver port; and

estimate a Bit Error Rate (BER) when operating in the idle state as a function of outputs from the denominator counter and error counter.

11. The apparatus of clause 10, wherein the high-speed link comprises an Ethernet link.

12. The apparatus of clause 11, wherein the PHY Layer circuitry does not include circuitry for a Forward Error Correction (FEC) sublayer.

13. The apparatus of clause 11, wherein the Ethernet link comprises one of a 10 Gigabits per second (Gb/s), 25 Gb/s, or 40 Gb/s Ethernet link.

14. The apparatus of clause 11, wherein the PCS has a block structure in accordance with IEEE802.3-2012 subclause 49.

15. The apparatus of any of clauses 10-14, wherein the denominator counter is a control block counter that counts idle blocks, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/2}{64 \times number\ of\ idle\ blocks\ received.}$$

16. The apparatus of any of clauses 10-15, wherein the denominator counter is an idle character counter that counts idle characters, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/3}{8 \times number\ of\ idle\ characters\ received.}$$

17. The apparatus of any of clauses 10-16, wherein the PCS circuitry and logic includes an *n*-bit counter and is configured to:

employ the *n*-bit counter in the denominator counter;

output *m* topmost bits of the *n*-bit counter; and

multiply the *m* topmost bits by 2^{n-m} to determine one of a number of idle characters or idle blocks received.

18. The apparatus of any of clauses 10-17, wherein the PHY Layer circuitry and logic is further configured to detect that a BER while operating the high-speed link in an idle state exceeds a threshold, and in response thereto, one of reset the link or disable the link.

19. An apparatus configured to communicate with a link partner over a high-speed link, comprising:

Physical Layer (PHY) circuitry and logic, including,

a transmitter port including transmitter circuitry; and

a receiver port including receiver circuitry; and

Physical Coding Sublayer (PCS) circuitry and logic including a receive state machine having a denominator counter and an error counter;

a Media Access Control (MAC) module;

a Reconciliation Sublayer (RS) module; and

an Input/Output (I/O) interface;

wherein when the high-speed link is operating in an idle state the PCS circuitry and logic is configured to,

count one of idle characters or idle blocks received at the receiver port with the denominator counter;

count one of character errors or block errors that are detected in the idle characters or idle blocks received at the receiver port; and

estimate a Bit Error Rate (BER) when operating in the idle state as a function of outputs from the denominator counter and error counter.

20. The apparatus of clause 19, wherein the PHY Layer circuitry does not include circuitry for a Forward Error Correction (FEC) sublayer.

21. The apparatus of clause 19, wherein the high-speed link comprises an Ethernet link.

22. The apparatus of clause 21, wherein the Ethernet link comprises one of a 10 Gigabits per second (Gb/s), 25 Gb/s, or 40 Gb/s Ethernet link.

23. The apparatus of clause 21, wherein the PCS has a block structure in accordance with IEEE802.3-2012 subclause 49.

24. The apparatus of any of clauses 19-23, wherein the denominator counter is a control block counter that counts idle blocks, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/2}{64\ x\ number\ of\ idle\ blocks\ received.}$$

25. The apparatus of any of clauses 19-24, wherein the denominator counter is an idle character counter that counts idle characters, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/3}{8\ x\ number\ of\ idle\ characters\ received.}$$

Although some embodiments have been described in reference to particular implementations, other implementations are possible according to some embodiments. Additionally, the arrangement and/or order of elements or other features illustrated in the drawings and/or described herein need not be arranged in the particular way illustrated and

described. Many other arrangements are possible according to some embodiments.

In each system shown in a figure, the elements in some cases may each have a same reference number or a different reference number to suggest that the elements represented could be different and/or similar. However, an element may be flexible enough to have different implementations and work with some or all of the systems shown or described herein. The various elements shown in the figures may be the same or different. Which one is referred to as a first element and which is called a second element is arbitrary.

In the description and claims, the terms "coupled" and "connected," along with their derivatives, may be used. It should be understood that these terms are not intended as synonyms for each other. Rather, in particular embodiments, "connected" may be used to indicate that two or more elements are in direct physical or electrical contact with each other. "Coupled" may mean that two or more elements are in direct physical or electrical contact. However, "coupled" may also mean that two or more elements are not in direct contact with each other, but yet still co-operate or interact with each other.

An embodiment is an implementation or example of the inventions. Reference in the specification to "an embodiment," "one embodiment," "some embodiments," or "other embodiments" means that a particular feature, structure, or characteristic described in connection with the embodiments is included in at least some embodiments, but not necessarily all embodiments, of the inventions. The various appearances "an embodiment," "one embodiment," or "some embodiments" are not necessarily all referring to the same embodiments.

Not all components, features, structures, characteristics, *etc.* described and illustrated herein need be included in a particular embodiment or embodiments. If the specification states a component, feature, structure, or characteristic "may", "might", "can" or "could" be included, for example, that particular component, feature, structure, or characteristic is not required to be included. If the specification or claim refers to "a" or "an" element, that does not mean there is only one of the element. If the specification or claims refer to "an additional" element, that does not preclude there being more than one of the additional element.

An algorithm is here, and generally, considered to be a self-consistent sequence of acts or operations leading to a desired result. These include physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers or the like. It should be understood, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities.

In addition, embodiments of the present description may be implemented not only within a semiconductor chip such as a NIC, but also within non-transient machine-readable media. For example, the designs described above may be stored upon and/or embedded within non-transient machine readable media associated with a design tool used for designing semiconductor devices. Examples include a netlist formatted in the VHSIC Hardware Description Language (VHDL) language, Verilog language or SPICE language, or other Hardware Description Language. Some netlist examples include: a behavioral level netlist, a register transfer level (RTL) netlist, a gate level netlist and a transistor level netlist. Machine-readable media also include media having layout information such as a GDS-II file. Furthermore, netlist files or other machine-readable media for semiconductor chip design may be used in a simulation environment to perform the methods of the teachings described above.

The operations and functions performed by various components described herein may be implemented by software running on a processing element, via embedded hardware or the like, or any combination of hardware and software. Such components may be implemented as software modules, hardware modules, special-purpose hardware (*e.g.*, application specific hardware, ASICs, DSPs, *etc.*), embedded controllers, hardwired circuitry, hardware logic, *etc.* Software content (*e.g.*, data, instructions, configuration information, *etc.*) may be provided via an article of manufacture including computer-readable or machine-readable non-transitory storage medium, which provides content that represents instructions that can be executed. The content may result in a computer performing various functions/operations described herein.

As used herein, a list of items joined by the term "at least one of" can mean any combination of the listed terms. For example, the phrase "at least one of A, B or C" can mean A; B; C; A and B; A and C; B and C; or A, B and C.

The above description of illustrated embodiments of the invention, including what is described in the Abstract, is not intended to be exhaustive or to limit the invention to the precise forms disclosed. While specific embodiments of, and examples for, the invention are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the invention, as those skilled in the relevant art will recognize.

These modifications can be made to the invention in light of the above detailed description. The terms used in the following claims should not be construed to limit the invention to the specific embodiments disclosed in the specification and the drawings. Rather, the scope of the invention is to be determined entirely by the following claims, which are to be construed in accordance with established doctrines of claim interpretation.

CLAIMS

What is claimed is:

1. A method for monitoring errors during idle states in a high-speed link including a Physical Coding Sublayer (PCS), comprising:

implementing a denominator counter in a receive state machine of the PCS;

implementing an error counter in the receive state machine to count errors while in the idle state;

while the high-speed Ethernet link is operating in an idle state,

counting one of idle characters or idle blocks received over the high-speed link with the denominator counter

counting one of character errors or block errors that are detected in the idle characters or idle blocks received over the high-speed link with the error counter; and

estimating a Bit Error Rate (BER) when operating in the idle state as a function of outputs from the denominator counter and error counter.

2. The method of claim 1, wherein the high-speed link comprises an Ethernet link including an Ethernet PCS.

3. The method of claim 2, wherein the Ethernet link includes an Ethernet Physical Layer (PHY) including a PCS, and wherein the Ethernet PHY does not include a Forward Error Correction (FEC) sublayer.

4. The method of claim 2, wherein the Ethernet link comprises one of a 10 Gigabits per second (Gb/s), 25 Gb/s, or 40 Gb/s Ethernet link.

5. The method of claim 2, wherein the PCS has a block structure in accordance with IEEE802.3-2012 subclause 49.

6. The method of any of the preceding claims, wherein the denominator counter is a control block counter that counts idle blocks, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/2}{64 \times number\ of\ idle\ blocks\ received.}$$

7. The method of any of the preceding claims, wherein the denominator counter is an idle character counter that counts idle characters, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/3}{8 \times number\ of\ idle\ characters\ received.}$$

8. The method of any of the preceding claims, further comprising:

employing an *n*-bit counter in the denominator counter;

reporting out *m* topmost bits of the *n*-bit counter; and

multiplying the *m* topmost bits by 2^{n-m} to determine one of a number of idle characters or idle blocks received.

9. The method of any of the preceding claims, further comprising detecting that a BER during operation in an idle state exceeds a threshold, and in response thereto, one of resetting the link or disabling the link.

10. An apparatus configured to communicate with a link partner over a high-speed link, comprising:

Physical Layer (PHY) circuitry and logic, including,

a transmitter port including transmitter circuitry; and

a receiver port including receiver circuitry; and

Physical Coding Sublayer (PCS) circuitry and logic including a receive state machine having a denominator counter and an error counter,

wherein when the high-speed link is operating in an idle state the PCS circuitry and logic is configured to,

count one of idle characters or idle blocks received at the receiver port with the denominator counter;

count one of character errors or block errors that are detected in the idle characters or idle blocks received at the receiver port; and

estimate a Bit Error Rate (BER) when operating in the idle state as a function of outputs from the denominator counter and error counter.

11. The apparatus of claim 10, wherein the high-speed link comprises an Ethernet link.

12. The apparatus of claim 11, wherein the PHY Layer circuitry does not include circuitry for a Forward Error Correction (FEC) sublayer.

13. The apparatus of claim 11, wherein the Ethernet link comprises one of a 10 Gigabits per second (Gb/s), 25 Gb/s, or 40 Gb/s Ethernet link.

14. The apparatus of claim 11, wherein the PCS has a block structure in accordance with IEEE802.3-2012 subclause 49.

15. The apparatus of any of claims 10-14, wherein the denominator counter is a control block counter that counts idle blocks, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/2}{64 \times number\ of\ idle\ blocks\ received.}$$

16. The apparatus of any of claims 10-15, wherein the denominator counter is an idle character counter that counts idle characters, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/3}{8 \times number\ of\ idle\ characters\ received.}$$

17. The apparatus of any of claims 10-16, wherein the PCS circuitry and logic includes an *n*-bit counter and is configured to:

employ the *n*-bit counter in the denominator counter;

output *m* topmost bits of the *n*-bit counter; and

multiply the *m* topmost bits by 2^{n-m} to determine one of a number of idle characters or idle blocks received.

18. The apparatus of any of claims 10-17, wherein the PHY Layer circuitry and logic is further configured to detect that a BER while operating the high-speed link in an idle state exceeds a threshold, and in response thereto, one of reset the link or disable the link.

19. An apparatus configured to communicate with a link partner over a high-speed link, comprising:

Physical Layer (PHY) circuitry and logic, including,

a transmitter port including transmitter circuitry; and

a receiver port including receiver circuitry; and

Physical Coding Sublayer (PCS) circuitry and logic including a receive state machine having a denominator counter and an error counter;

a Media Access Control (MAC) module;

a Reconciliation Sublayer (RS) module; and

an Input/Output (I/O) interface;

wherein when the high-speed link is operating in an idle state the PCS circuitry and logic is configured to,

count one of idle characters or idle blocks received at the receiver port with the denominator counter;

count one of character errors or block errors that are detected in the idle characters or idle blocks received at the receiver port; and

estimate a Bit Error Rate (BER) when operating in the idle state as a function of outputs from the denominator counter and error counter.

20. The apparatus of claim 19, wherein the PHY Layer circuitry does not include circuitry for a Forward Error Correction (FEC) sublayer.

21. The apparatus of claim 19, wherein the high-speed link comprises an Ethernet link.

22. The apparatus of claim 21, wherein the Ethernet link comprises one of a 10 Gigabits per second (Gb/s), 25 Gb/s, or 40 Gb/s Ethernet link.

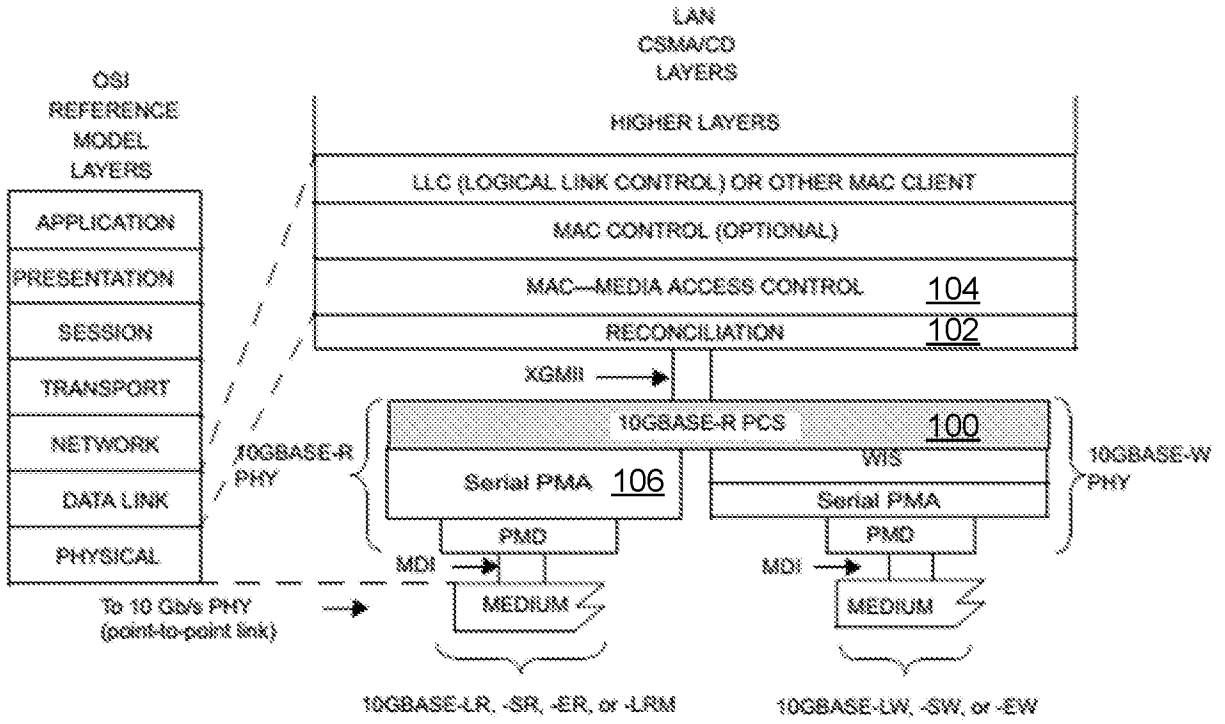
23. The apparatus of claim 21, wherein the PCS has a block structure in accordance with IEEE802.3-2012 subclause 49.

24. The apparatus of any of claims 19-23, wherein the denominator counter is a control block counter that counts idle blocks, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/2}{64\ x\ number\ of\ idle\ blocks\ received.}$$

25. The apparatus of any of claims 19-24, wherein the denominator counter is an idle character counter that counts idle characters, and the BER is estimated by the equation,

$$estimated\ BER = \frac{value\ of\ block\ error\ counter/3}{8\ x\ number\ of\ idle\ characters\ received.}$$



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

NOTE—The PMD sublayers are mutually independent.

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

Fig. 1 *(Prior Art)*

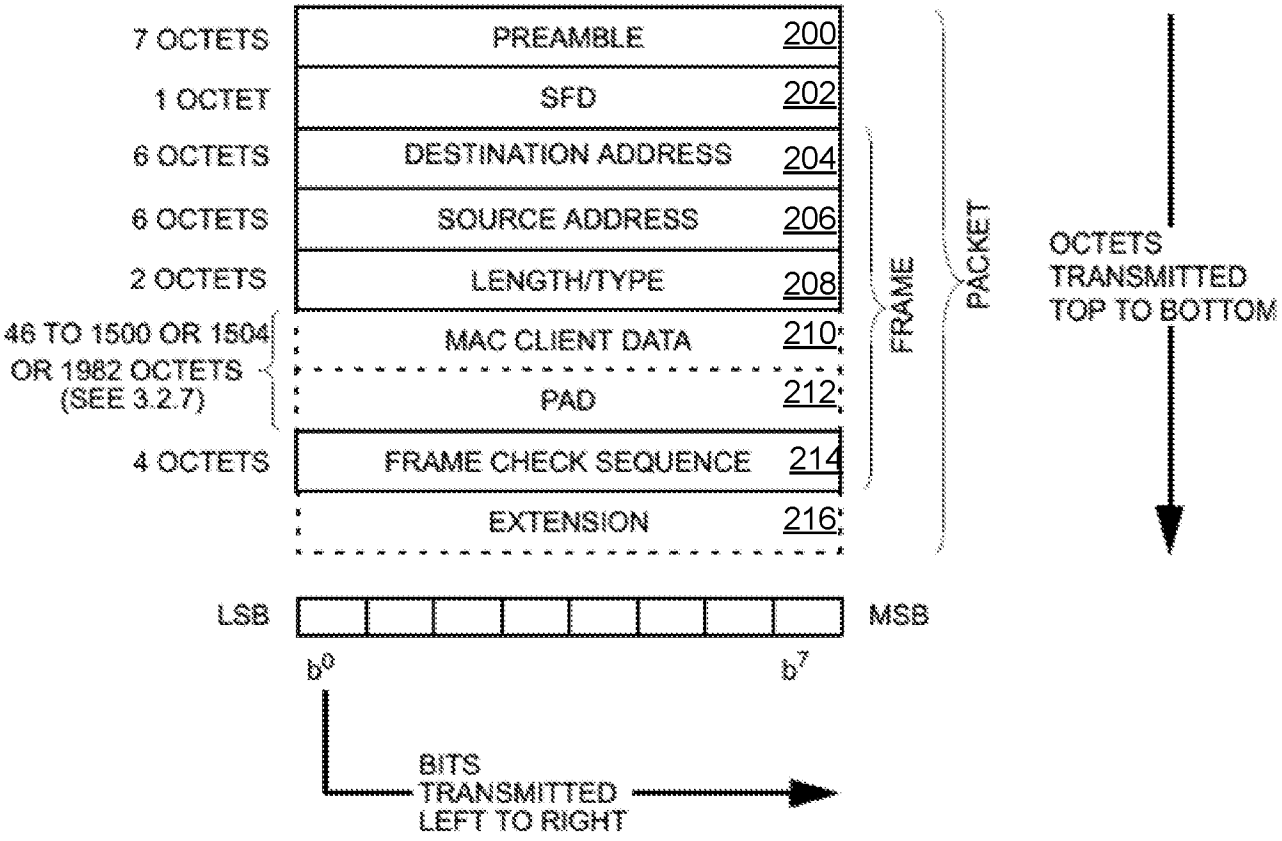
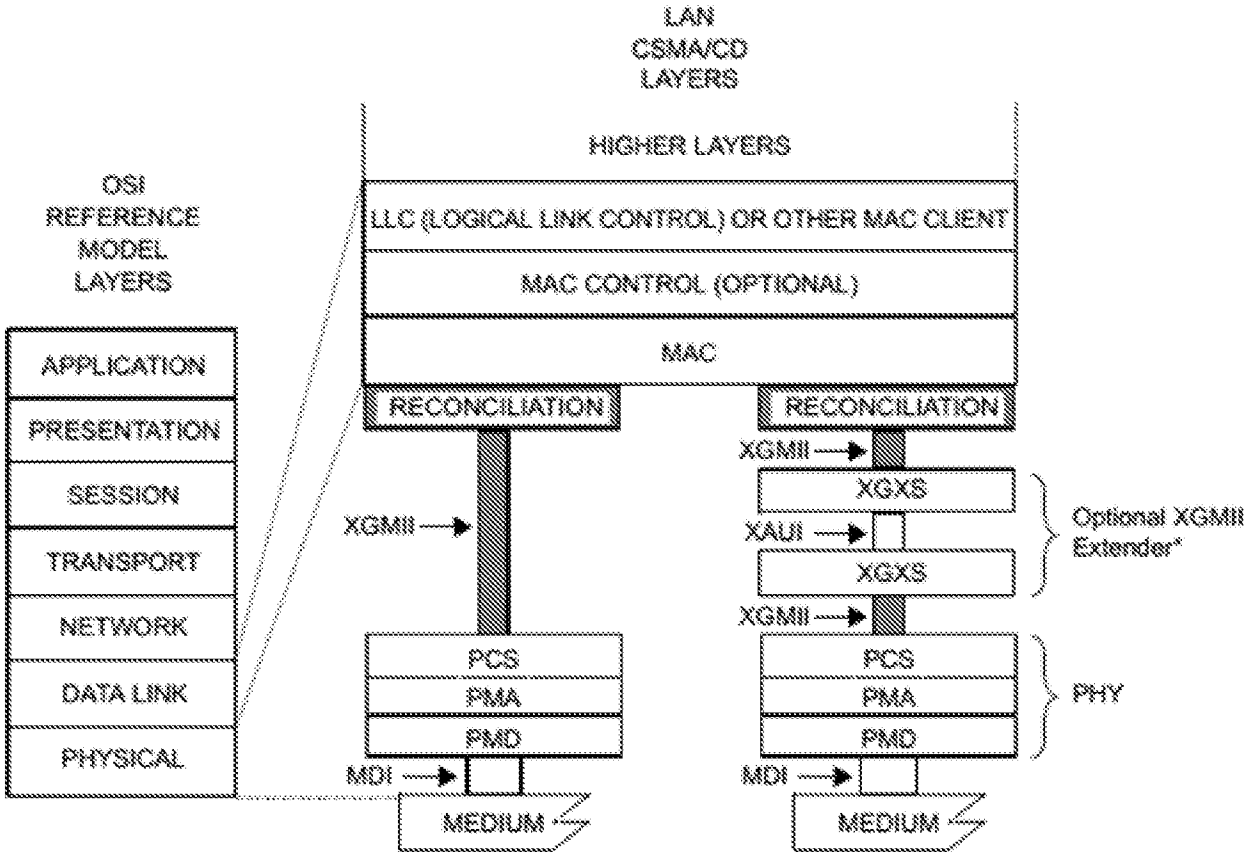


Fig. 2 *(Prior Art)*



MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE

PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XAUI = 10 GIGABIT ATTACHMENT UNIT INTERFACE
XGMII = 10 GIGABIT MEDIA INDEPENDENT INTERFACE
XGXS = XGMII EXTENDER SUBLAYER

*specified in Clause 47

Fig. 3 *(Prior Art)*

Fig. 4 (Prior Art)

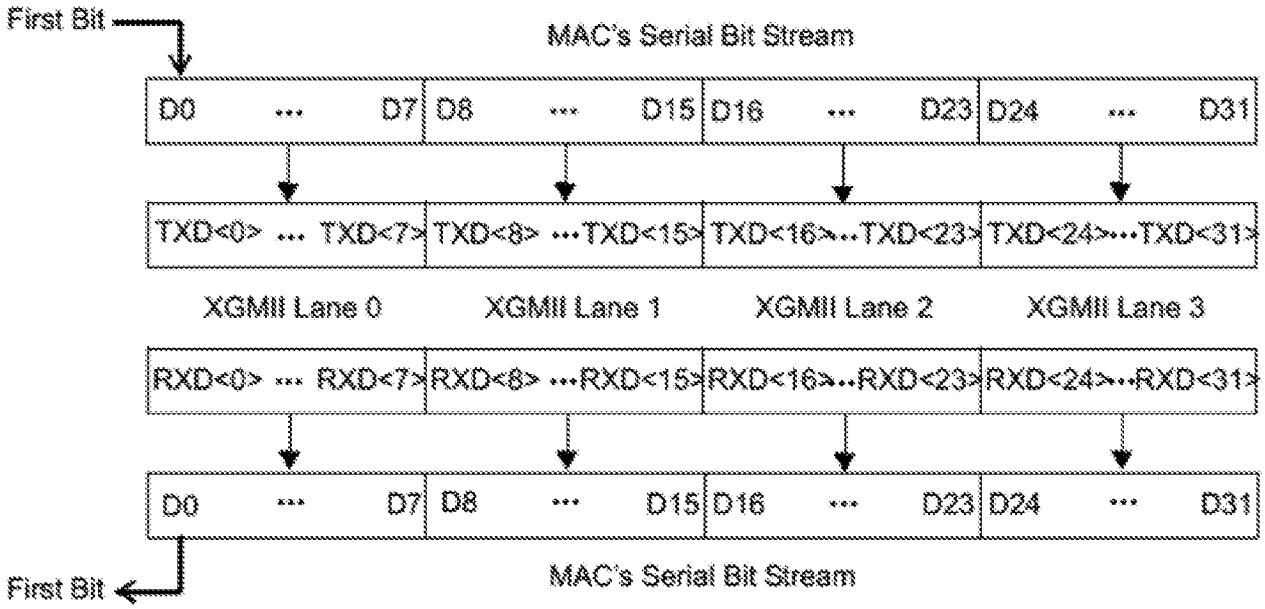


Fig. 5 (Prior Art)

Control Character	Notation	XGMII Control Code	10GBASE-R Control Code	10GBASE-R O Code	8B/10B Code[a]
Idle	/I/	0x07	0x00		K28.0 or K28.3 or K28.5
LPI	/LI/	0x06	0x06		K28.0 with D20.5 in One Row or K28.3 or K28.5 with D20.5 in One Row[b]
Start	/S/	0xfb	Encoded by Block Type Field		K27.7
Terminate	/T/	0xfd	Encoded by Block Type Field		K29.7
Error	/E/	0xfe	0x1e		K30.7
Sequence Ordered_Set	/Q/	0x9c	Encoded by Block Type Field Plus O Code	0x0	K28.4
Reserved0	/R/[c]	0x1c	0x2d		K28.0
Reserved1		0x3c	0x33		K28.1
Reserved2	/A/	0x7c	0x4b		K28.3
Reserved3	/K/	0xbc	0x55		K28.5
Reserved4		0xdc	0x66		K28.6
Reserved5		0xf7	0x78		K23.7
Signal Ordered_Set[d]	/Fsig/	0x5c	Encoded by Block Type Field Plus O Code	0xF	K28.2

600

Fig. 6 ***(Prior Art)***

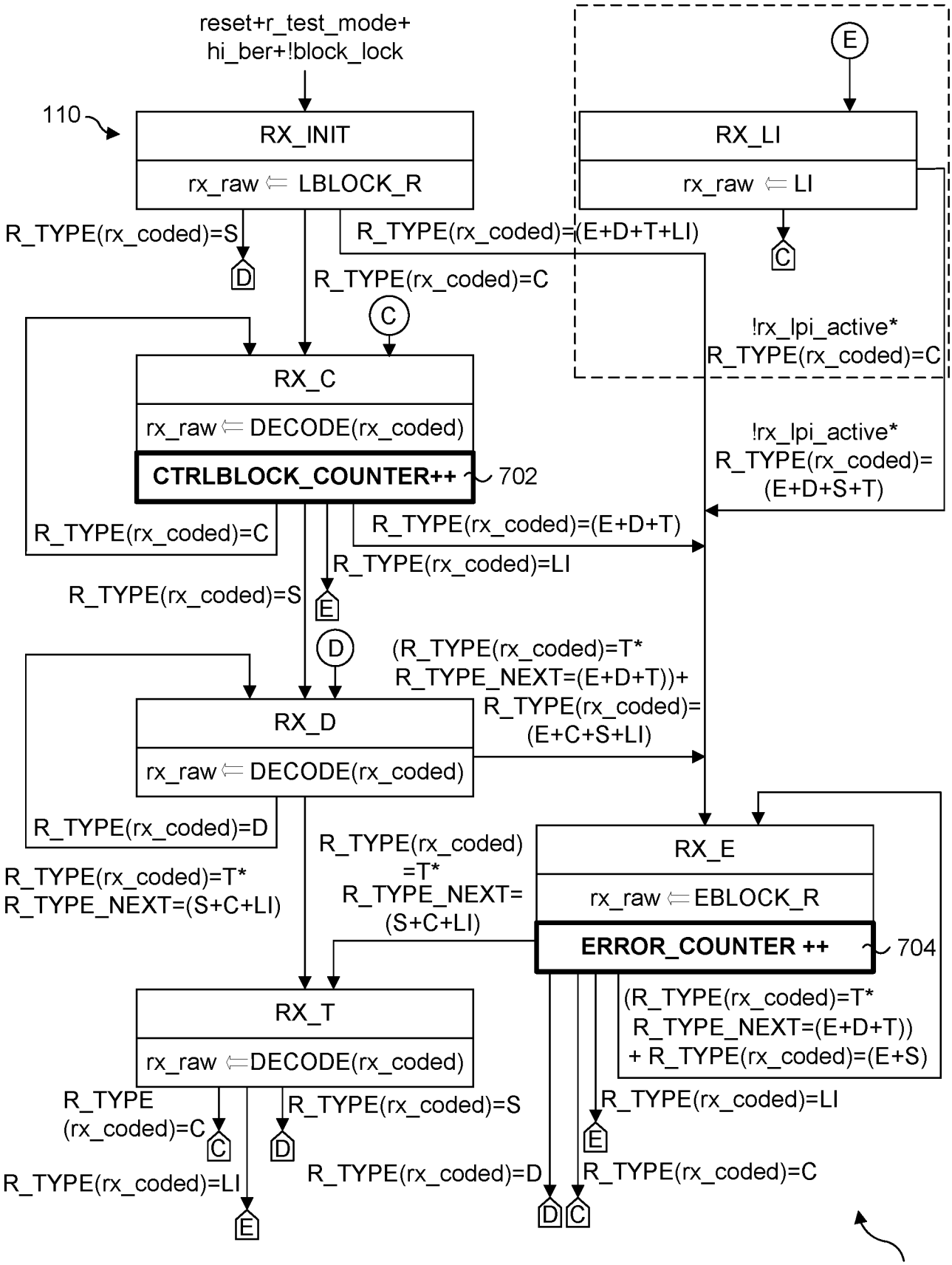
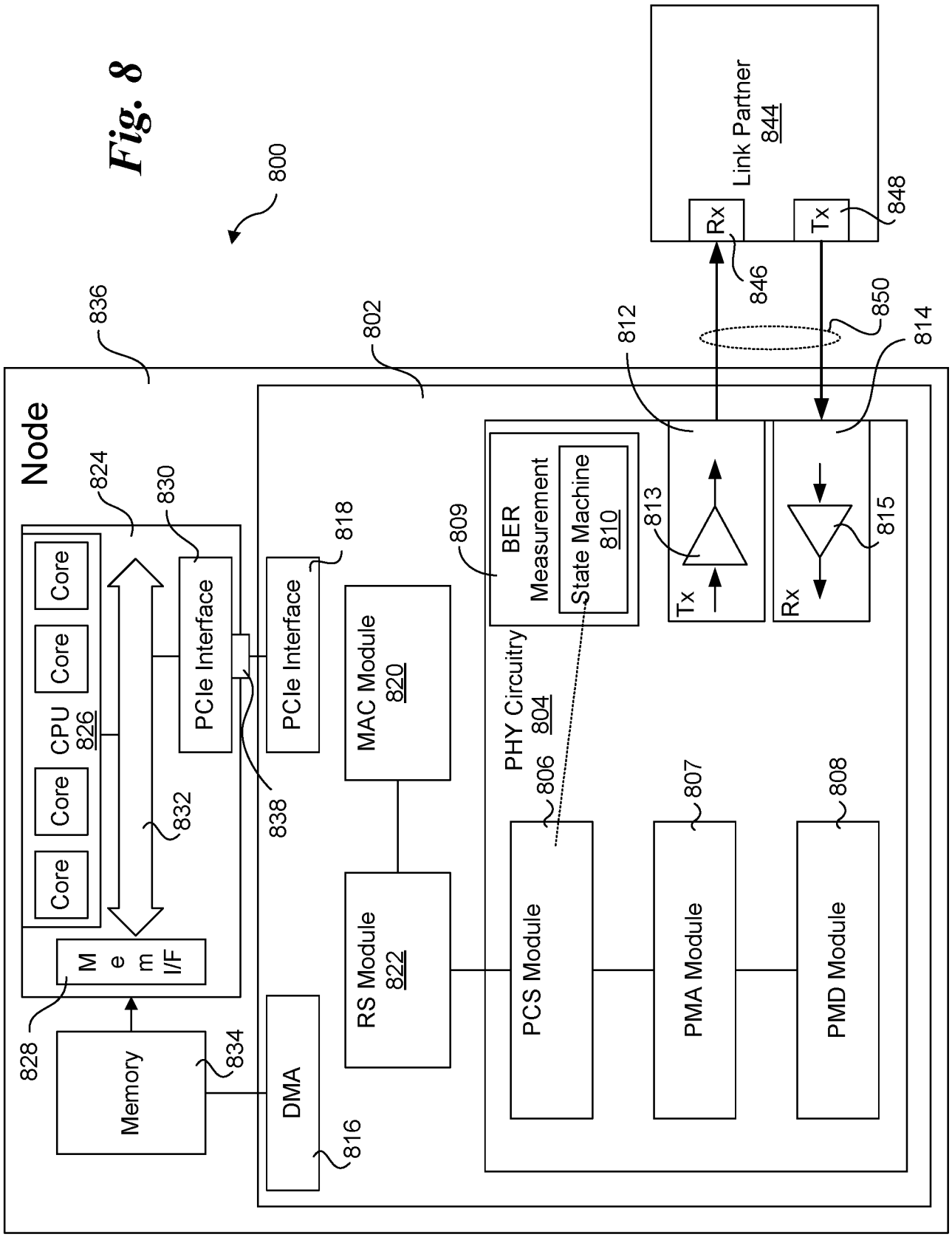

reset+r_test_mode+
hi_ber+!block_lock
110
RX_INIT
rx_raw ⇐ LBLOCK_R
R_TYPE(rx_coded)=S
D
R_TYPE(rx_coded)=(E+D+T+LI)
R_TYPE(rx_coded)=C
C
RX_C
rx_raw ⇐ DECODE(rx_coded)
CTRLBLOCK_COUNTER++
702
R_TYPE(rx_coded)=C
R_TYPE(rx_coded)=(E+D+T)
R_TYPE(rx_coded)=LI
R_TYPE(rx_coded)=S
E
D
(R_TYPE(rx_coded)=T*
R_TYPE_NEXT=(E+D+T))+
R_TYPE(rx_coded)=
(E+C+S+LI)
RX_D
rx_raw ⇐ DECODE(rx_coded)
R_TYPE(rx_coded)=D
R_TYPE(rx_coded)=T*
R_TYPE_NEXT=(S+C+LI)
R_TYPE(rx_coded)
=T*
R_TYPE_NEXT=
(S+C+LI)
RX_T
rx_raw ⇐DECODE(rx_coded)
R_TYPE
(rx_coded)=C
C
R_TYPE(rx_coded)=S
D
R_TYPE(rx_coded)=LI
E
R_TYPE(rx_coded)=D
D
R_TYPE(rx_coded)=LI
E
R_TYPE(rx_coded)=C
C
RX_E
rx_raw ⇐ EBLOCK_R
ERROR_COUNTER ++
704
(R_TYPE(rx_coded)=T*
R_TYPE_NEXT=(E+D+T))
+ R_TYPE(rx_coded)=(E+S)
E
RX_LI
rx_raw ⇐ LI
C
!rx_lpi_active*
R_TYPE(rx_coded)=C
!rx_lpi_active*
R_TYPE(rx_coded)=
(E+D+S+T)
700

Fig. 7

Fig. 8


800
Node
836
802
824
830
818
809
812
814
Core
Core
Core
Core
CPU 826
PCIe Interface
PCIe Interface
MAC Module 820
BER Measurement
State Machine 810
813
815
Tx
Rx
PHY Circuitry 804
806
807
808
832
838
RS Module 822
PCS Module
PMA Module
PMD Module
Mem I/F
828
Memory
834
DMA
816
Link Partner 844
Rx
Tx
846
848
850

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/021366

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/26(2006.01)i, H04L 1/20(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04L 12/26; H04J 1/16; H04L 1/00; G06F 15/173; H04L 12/28; H02H 3/05; H04L 1/20

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
eKOMPASS(KIPO internal) & keywords: Physical Coding Sublayer (PCS), idle state, denominator counter, error counter, Bit Error Rate (BER)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2008-0228941 A1 (PETRE POPESCU et al.) 18 September 2008 See paragraphs [0011], [0014], [0087], [0175] and figure 1.	1-25
A	US 2005-0005189 A1 (LIOR KHERMOSH et al.) 06 January 2005 See paragraphs [0047]-[0059] and figures 4a-4b.	1-25
A	US 6690650 B1 (RUDOLPH STENER) 10 February 2004 See column 6, line 45 - column 7, line 67 and figure 2.	1-25
A	US 2014-0258813 A1 (KENT C. LUSTED et al.) 11 September 2014 See paragraphs [0048]-[0049] and figures 4-5.	1-25
A	KR 10-2005-0062339 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 23 June 2005 See claim 1.	1-25

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 May 2016 (26.05.2016)	**11 July 2016 (11.07.2016)**

Name and mailing address of the ISA/KR
International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea
Facsimile No. +82-42-481-8578

Authorized officer
KIM, Seong Woo
Telephone No. +82-42-481-3348

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2016/021366

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2008-0228941 A1	18/09/2008	US 2005-0102419 A1	12/05/2005
		US 7949782 B2	24/05/2011
US 2005-0005189 A1	06/01/2005	AU 2002-304362 A1	10/11/2003
		CN 100576785 C	30/12/2009
		CN 1625859 A	08/06/2005
		JP 2005-524281A	11/08/2005
		KR 10-0835401 B1	04/06/2008
		KR 10-2005-0016369 A	21/02/2005
		US 7343540 B2	11/03/2008
		WO 03-092207 A1	06/11/2003
US 6690650 B1	10/02/2004	DE 69927424 D1	09/02/2006
		DE 69927424 T2	14/06/2006
		EP 0952700 A2	27/10/1999
		EP 0952700 A3	23/04/2003
		EP 0952700 B1	28/09/2005
		TW 429343 A	11/04/2001
US 2014-0258813 A1	11/09/2014	CN 104081797 A	01/10/2014
		JP 2015-508616 A	19/03/2015
		KR 10-2014-0097498 A	06/08/2014
		TW 201415829 A	16/04/2014
		US 9258082 B2	09/02/2016
		WO 2014-011224 A1	16/01/2014
KR 10-2005-0062339 A	23/06/2005	KR 10-0609695 B1	08/08/2006
		US 2005-0149822 A1	07/07/2005
		US 7600171 B2	06/10/2009