



(51) International Patent Classification:

H04L 29/06 (2006.01) **H04L 12/28** (2006.01)

(21) International Application Number:

PCT/US2017/038817

(22) International Filing Date:

22 June 2017 (22.06.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

15/218,681 25 July 2016 (25.07.2016) US

(71) Applicant: **INTEL CORPORATION** [US/US]; 2200 Mission College Boulevard, Santa Clara, California 95054 (US).(72) Inventors: **RAN, Adeo O.**; POB 88, 12220 Maayan Baruch (IL). **CHALUPSKY, David L.**; 36051 NW Hahn Rd, Banks, Oregon 97106 (US). **LILLIE, Kevan A.**; 21709 S 140th St, Chandler, Arizona 85286 (US). **MELLITZ, Richard I.**; 220 Golden Way, Prosperity, South Carolina 29127 (US). **LUSTED, Kent C.**; 840 SW 191st Ave, Aloha, Oregon 97006 (US).(74) Agent: **BURNETT, R. Alan**; Law Office of R. Alan Burnett, PS, 4108 131st Avenue SE, Bellevue, Washington 98006 (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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).

(54) Title: ETHERNET AUTO-NEGOTIATION WITH PARALLEL DETECT FOR 10G DAC OR OTHER NON-AUTO-NEGOTIATED MODES

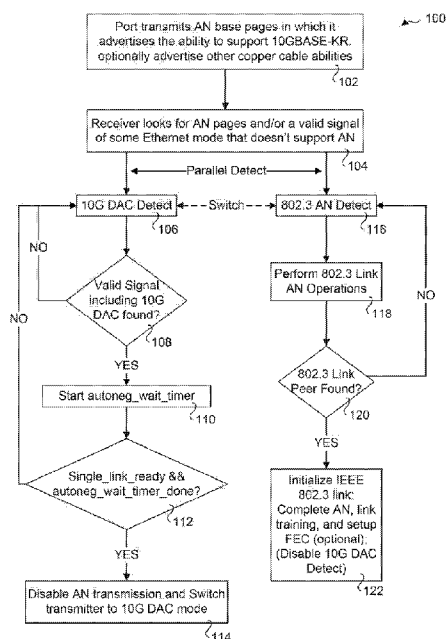


Fig. 1

(57) Abstract: Methods and apparatus for Ethernet auto-negotiation (AN) with parallel detect for 10G DAC or other non-auto-negotiated modes. AN base pages are transmitted from an Ethernet apparatus to advertise the ability to support at least one Institute of Electrical and Electronics Engineers (IEEE) 802.3 Ethernet specification supporting AN. A receiver and associated processing circuitry is configured to perform two detection modes in parallel, including a first detection mode that looks for a valid signal transmitted from an Ethernet link peer that does not support AN and a second detection mode looking for AN pages from an IEEE 802.3 Ethernet link peer that supports AN. If the link peer does not support AN, an Ethernet link is set up to use signaling in accordance with the Ethernet specification that does not support AN. If the link peer supports AN, an Ethernet link is set up using a corresponding IEEE 802.3 Ethernet link supporting AN. Supported non-AN Ethernet links include 10G DAC links.

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))*

ETHERNET AUTO-NEGOTIATION WITH PARALLEL DETECT FOR 10G DAC OR OTHER NON-AUTO-NEGOTIATED MODES

BACKGROUND INFORMATION

As computer processing has moved from local compute resources, such as a desktop or laptop computer system, to remote distributed compute resources, such as cloud-host services and the like, data transfer speeds across networks has become increasingly important. To this end, high-speed network links are critical for data center and high performance computing (HPC). One of the primary network technologies in data centers and HPC environments is high-speed Ethernet.

The original Ethernet standard was specified in Institute of Electrical and Electronics Engineers (IEEE) 802.3 in 1983. Subsequently, the vast majority of Ethernet standards have been developed and specified under an IEEE 802.3-based committee. The IEEE 802.3 Ethernet standard defines several data rates that can be used when connecting two systems. For the backplane and copper cables media, there is an auto-negotiation (AN) protocol defined in clause 73 of the standard (803.2-2012, Section Five), which enables selecting a port type, and other features such as forward error correction (FEC), based on advertised capabilities of both sides. Network management can be used to program what capabilities are advertised for a port, and thus control the data rate that will be selected by AN. For some rates, a link training period is included in the AN process, which enables improving received signal quality and operating over more challenging media (e.g. longer cables).

In contrast, the implementation of Ethernet operation over copper cables at 10 Gigabit per second (Gb/s) was defined outside of IEEE 802.3, in the Small Form Factor (SFF)-8418 committee, and labeled (Small Form Factor Pluggable) SFP+ direct attach or 10GSFP+Cu, also known as 10G Direct Attach Cable or 10G DAC. The form factor and electrical interface for 10G DAC are specified by a multi-source agreement, is widely-used in today's data centers and is envisioned to continue to see substantially usage in the future. In contrast to the 802.3 standard, 10G DAC does not support either auto-negotiation or FEC, nor is link training supported.

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 flowchart illustrating operations and logic for implemented parallel detection of 10G DAC signals and 10GBASE-KR (and/or other IEEE 802.3 Ethernet standards supporting

AN) signals, according to one embodiment;

Figure 2 is a table showing Technology Ability Field bits, including an extension to table 73-4 in clause 73 to further support 25GBASE-KR-S or 25GBASE-CR-S and 25GBASE-KR or 25GBASE-CR IEEE 802.3 Ethernet standards;

Figure 3 is diagram illustrating a link codeword Base Page in accordance with Figure 73-6 of IEEE 802.3-2012 clause 73;

Figures 4a and 4b show respective portions of an AN arbitration state diagram configured to support parallel detection of IEEE 802.3 Ethernet signals supporting AN and Ethernet signals in accordance with Ethernet standards that do not support AN, including 10G DAC, according to one embodiment; and

Figure 5 is a schematic diagram illustrating an architecture for a network node employing a network chip configured to implement prevention of false packet acceptance in accordance with the embodiments disclosed herein.

DETAILED DESCRIPTION

Embodiments of methods and apparatus for Ethernet auto-negotiation with parallel detect for 10G DAC or other non-auto-negotiated modes 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.*

Ethernet has defined port types for 40 Gb/s and 100 Gb/s data rates separately for the backplane and copper cable media (for 40G, 40GBASE-KR4 defined in clause 84 for backplane, and 40GBASE-CR4 defined in clause 85 for copper cables; for 100G, 100GBASE-KR4 defined in clause 93 for backplane, and 100GBASE-CR4 defined in clause 92 for copper cables). For these rates, capabilities are advertised in AN separately per medium. Also, the 802.3by task force is defining new port types for 25 Gb/s data rates (25GBASE-KR defined in clause 111 for backplane, 25GBASE-CR defined in clause 110 for copper cables). For the 25G rate, backplane and copper cable are advertised as a single capability. In addition, Ethernet has port types for backplane in lower data rates, 10 Gb/s (10GBASE-KR) and 1 Gb/s (1000BASE-KX), which can also be advertised in AN, but there is no designated port type for copper cable for these rates.

In backplane systems, auto-negotiation can be used to downgrade a link to a lower data rate if desired (assuming the partner supports the lower rate), for various reasons, e.g. power saving or due to poor signal or media quality. Thus, a port supporting 40GBASE-KR4 may be downgraded to run 10GBASE-KR (using only one out of four lanes), or a port supporting 25GBASE-KR link may be downgraded to run 10GBASE-KR or 1000BASE-KX (using lower signaling rates). Using AN provides an interoperable way of downgrading and upgrading a port and re-establishing the link with the partner. However, currently 10G DAC mode cannot be chosen for downgrade, since it does not include AN, and the capability cannot be advertised.

The 10GBASE-KR port defined for backplanes can also operate over copper cables, which would provide AN, link training and optional FEC over this medium, and also extend the downgrade/upgrade mode selection. However, the problem is that a port that uses 10GBASE-KR over copper cables cannot interoperate with a port that is configured as 10G DAC. Due to the large installed base of 10G DAC, this may slow adoption of 10GBASE-KR over copper cables, and thus prevent users from the additional capabilities offered by this feature. The current “solution” is to program both sides of the link to use 10G DAC instead of Ethernet AN. This requires controlling both sides of the link, which is more complex, and sometimes not feasible.

The existing AN protocol requires a device to send a designated signal (AN pages) on its transmitter, while on its receiver it looks for a partner that sends a similar signal. The AN page transmission continues until an AN-capable partner is detected.

AN also defines a “parallel detection” mode for interoperating with legacy devices that do not support AN (and thus do not transmit AN pages), but have a fixed signaling rate of 1.25 Gb/s or 3.125 Gb/s (1000BASE-KX or 10GBASE-KX4 respectively). A device that performs parallel

detection transmits AN pages but, in addition to looking for AN pages, also looks for a valid 1000BASE-KX or 10GBASE-KX4 signal. If any of these signals is detected, the device stops transmitting AN pages and instead switches directly to the detected mode. The partner (legacy device that does not use AN) will identify the new mode which matches its own mode and a link will be established.

Under aspects of the embodiments described herein, a similar “parallel detection” method is disclosed to detect 10G DAC signaling and thus enable linking and interoperating with 10G DAC devices. The scheme supports use of 10GBASE-KR with AN for copper cables instead of 10G DAC mode, which will enable link training, optional FEC, and automatic rate changes on copper cable media. In addition, the embodiments enables smooth upgrade of existing networks, which may include 10G DAC ports that cannot be configured as 10GBASE-KR.

Figure 1 shows a flowchart 100 illustrating operations and logic for implemented parallel detection of 10G DAC signals and 10GBASE-KR (and/or other IEEE 802.3 Ethernet standards supporting AN) signals, according to one embodiment. The process begins in a block 102, wherein a port transmits AN base pages in which it advertises 10GBASE-KR. In addition it may advertise other copper cable abilities (such as but not limited to 25GBASE-CR, 25GBASE-KR, and 40GBASE-KR).

Advertising is done by setting Technology Ability Field bits, as shown in the table 200 of Figure 2 (which is an augmented version of table 73-4 in clause 73), and Figure 3 (which is a reproduction of Figure 73-6 of clause 73). The base link codeword (Base Page) transmitted within a DME page conveys the encoding shown in Figure 3. The Auto-Negotiation function supports additional pages using the Next Page function. Encoding for the link codeword(s) used in the Next Page exchange are defined in 802.3-2012 clause 73.7.7. In a DME page, D0 shall be the first bit transmitted. D[4:0] contains the Selector Field. D[9:5] contains the Echoed Nonce field. D[12:10] contains capability bits to advertise capabilities not related to the PHY. C[1:0] is used to advertise pause capability. The remaining capability bit C[2] is reserved. D[15:13] contains the RF, Ack, and NP bits. These bits shall function as specified in 28.2.1.2. D[20:16] contains the Transmitted Nonce field. D[45:21] contains the Technology Ability Field. D[47:46] contains FEC capability (see clause 73.6.5).

Technology Ability Field (A[24:0]) is a 25-bit wide field containing information indicating supported technologies specific to the selector field value when used with the Auto-Negotiation for Backplane Ethernet. These bits are mapped to individual technologies such that abilities are advertised in parallel for a single selector field value. The current Technology Ability Field encoding for the IEEE 802.3 selector with Auto-Negotiation for Backplane Ethernet is described in Table 73-4.

As shown in Figure 2, two additional Technology Ability Field bits have been defined. Setting bit A9 advertises the ability to support either 25GBASE-KR-S or 25GBASE-CR-S, while setting bit A10 advertises the ability to support either 25GBASE-KR or 25GBASE-CR.

Returning to Figure 1, in a block 104, on the receiver, the device looks for AN pages and/or a valid signal of some supported Ethernet mode that does not support AN; this is marked as “link_control_[PD]←SCAN_FOR_CARRIER” 402 in the first portion of AN arbitration state diagram 400 in Figure 4a, wherein “PD” stands for “Parallel Detection.” The intent is that all modes that are supported for parallel detection would be enabled (SCAN_FOR_CARRIER). Under different embodiments of the parallel detection modes described herein, processing of received signals looking for AN pages or valid signals may be performed simultaneously, or may be performed by rapidly switching detection modes, as shown in the dashed double-headed arrow labeled “switch” in Figure 1.

One embodiment extends the PD modes defined in arbitration state diagram in Figure 73-11 of clause 73 to include a 10G DAC detection mode; in terms of clause 73, this would conceptually be a redefinition of “PD” as follows in TABLE 1:

PD	Represents all of the following that are present: 1000BASE-KD PMD; 10GBASE-KX4 (or 10GBASE-CX4) PMD; and 10G DAC PMD
----	--

TABLE 1

As illustrated in Figure 1, operations relating to detection of a 10G DAC and detection of an 802.3 link peer are respectively shown in the left-hand and right-hand portions of flowchart 100 below block 104. As further indicated, these operations are performed in parallel, either simultaneously or via rapid switching between detection modes.

Operations in a block 106 and a decision block 108 relating to 10G DAC detection are performed in loopwise manner, looking for a valid signal including a 10G DAC. As shown in the portion of arbitration state diagram 400 in Figure 4a, detection of a valid 10G DAC signal results in a link_state[10G_DAC]=OK state, which also coincides with a result of YES for decision block 108. In response, the arbitration state diagram 400 state proceeds to a LINK STATUS CHECK 404 in Figure 4a, which includes Start autoneg_wait_timer 404, which also is depicted in a block 110 in flowchart 100.

In a decision block 112, a determination is made to whether both single_link_ready state and autoneg_wait_time_done are true, indicating that a valid signal is present during the wait time defined for the auto-negotiation wait timer. If the answer to decision block 112 is YES, the logic proceeds to a block 114 in which transmission of the AN pages is disabled and the transmitter is switched to the 10G DAC mode.

If the answer to either decision block 108 or decision block 112 is NO, the logic loops back

to block 106. If the mode detection operations are performed simultaneously, the 10G DAC detection process is repeated in a continuous manner. If parallel detection is performed by switching between detection modes, once returning to block 106 the 10G DAC detection process can either be repeated or the logic can switch to performing 802.3 AN detection.

5 During 802.3 AN detection, the receiver is looking for AN pages that are transmitted by an IEEE 802.3 link peer supporting 10GBASE-KR or other 802.3 high-speed Ethernet links implemented over copper cables that support AN. The process includes operations in blocks 116 and 118, wherein 802.3 link AN operations are performed in a in accordance with arbitration state diagram 400. In a decision block 120, a determination is made to whether an IEEE 802.3
10 Ethernet link peer supporting AN is found. If the answer is YES, the logic proceeds to a block 122, wherein an IEEE 802.3 Ethernet link is initialized by completing AN operations, link training is performed, and optional FEC support is configured. In addition, if simultaneous detection is being performed, the 10G DAC detection mode is disabled. If the answer to decision block 120 is NO, the logic loops back to block 116, followed by either repeating the
15 802.3 AN detection operations or switching to the 10G DAC detection operations, depending on the parallel detection mode is performed simultaneously or switched.

It is noted that the criteria for deciding that a valid 10G DAC signal is found are implementation dependent. A receiver may, for example, use frequency lock indication, PCS/scrambler lock indication, signal quality or eye-opening measurement, or a combination of
20 the above. The receiver may have to be configured differently to enable detection of AN pages, KX, KX4, and 10G DAC modes. This can be solved by having the receiver switch periodically between these modes until a valid signal is detected. Variations on this scheme are possible and are implementation dependent.

In a similar way to the described for 10G DAC above, other modes can also be added. For
25 example, a device may also be able to parallel-detect 2.5G Ethernet and 5G Ethernet (which are currently being defined in 802.3cb), may detect other types of links that do not use auto-negotiation, or any other signaling method.

Figure 5 shows an architecture 500 for a compute node employing a multi-mode network
chip 502 configured to implement aspects of the embodiments disclosed herein. The compute
30 node is illustrative of various host platforms in which a multiple-mode network chip or NIC may be installed. Network chip 502 comprises multi-mode PHY circuitry 504 including an 802.3 PHY block 506, a 10G DAC PHY block 508, an auto-negotiation module 510, mode control logic 512, multiplexer circuitry 514, a transmitter (Tx) port 516 including transmitter circuitry 517, and a receiver (Rx) port 518 including receiver circuitry 519. The transmitter and
35 receiver circuitry may support a single lane (x1) or four lanes (x4) that may be reconfigured to

operate as a single lane when connected to a link partner that only supports a single lane Ethernet link.

802.3 PHY block 506 is illustrative of a PHY circuit block configured to support any of the 802.3 high-speed Ethernet technologies described herein, including 10GBASE-KR4, 25GBASE-KR4, and 40GBASE-KR4. 802.3 PHY block 506 includes a Physical Coding Sublayer (PCS) module 520, an FEC module 522 including an FEC decoder 523, a Physical Media Attachment (PMA) module 524, and a Physical Media Dependent (PMD) module 526. 10G DAC PHY block 508 is configured to implement the 10G DAC physical interface circuitry and logic in accordance with SFF-8418. In one embodiment 10G DAC PHY block 508 is implemented as an Application Specific Integrated Circuit including serializer and deserializer circuitry (ASIC/SerDes), as described in SFF-8418. As further illustrated, a PHY block that does not support auto-negotiation (i.e., a non-AN PHY block) may be implemented in place of 10G DAC PHY block 508, along with appropriate circuitry to implement the associated non-AN Ethernet link (not shown).

Network chip 502 further includes a DMA (Direct Memory Access) interface 528, a Peripheral Component Interconnect Express (PCIe) interface 530, a MAC module 532 and a Reconciliation Sublayer (RS) module 534. Compute node 500 also comprises a System on a Chip (SoC) 536 including a Central Processing Unit (CPU) 538 having one or more processor cores, coupled to a memory interface 540 and a PCIe interface 542 via an interconnect 544. Memory interface 540 is further depicted as being coupled to memory 546. Under a typical configuration, network chip 502, SoC 536 and memory 546 will be mounted on or otherwise operatively coupled to a circuit board 548 that includes wiring traces for coupling these components in communication, as depicted by single lines connecting DMA 528 to memory 546 and PCIe interface 530 to PCIe interface 542 at a PCIe port 550.

In one embodiment, MAC module 532, and RS module 534 are respectively configured to implement aspects of the MAC layer operations and reconciliation sub-layer operations as defined for one or more 802.3 high-speed Ethernet links, including 10GBASE-KR4, 25GBASE-KR4, and 40GBASE-KR4. MAC module 532 is also configured to support MAC layer operations for a 10G DAC link, and is configured to switch between one or more 802.3 Ethernet links and the 10G DAC link.

During link initialization, auto-negotiation module 510 is implemented for auto-negotiation of link speed and capabilities for an 802.3 Ethernet link. As described above, the AN format consists of a base-page, which is the first set of formatted information exchanged with the link partner, as depicted by a link partner 552 including a receiver port 554 and a transmitter port 556. As shown, compute node 500 and link partner 552 are linked in communication via an

Ethernet link 558 implemented using a copper cable.

Mode control logic 512 is configured to implement the operations and logic illustrated in flowchart 100 of Figure 1 described above. The includes controlling multiplexer circuitry 514, which is use to selectively couple the transmitter and receiver circuitry 517 and 519 to 802.3
5 PHY block 506 or 10G DAC PHY block 508. In one embodiment, multiplexer circuitry 514 is configured to simultaneously coupled signals received by receiver circuitry 517 to each of 802.3 PHY block 506 and 10G DAC PHY block 508.

In addition to implementing the circuit components and modules shown in Figure 5 in a network chip, such as a Network Interface Controller (NIC) chip, various components and
10 modules may be implemented as discreet components mounted to a circuit board or the like, or otherwise implemented using multi-component packaging. For example, 802.3 PHY block 506 and 10G DAC PHY block 508 might be implemented using separate chips. In addition to being implemented in a compute node, network chip 502, or a chip or multi-component module with similar functionality, may also be implemented in a network switch including multiple ports.

15 In addition to Further aspects of the subject matter described herein are set out in the following numbered clauses: [to be completed after the claims are approved]

In addition, embodiments of the present description may be implemented not only within a semiconductor chip but also within machine-readable media. For example, the designs described above may be stored upon and/or embedded within machine readable media associated with a
20 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. 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
25 other machine-readable media for semiconductor chip design may be used in a simulation environment to perform the methods of the teachings described above.

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
30 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
35 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.

As discussed above, various aspects of the embodiments herein may be facilitated by corresponding software and/or firmware components and applications, such as software and/or firmware executed by an embedded processor or the like. Thus, embodiments of this invention may be used as or to support a software program, software modules, firmware, and/or distributed software executed upon some form of processor, processing core or embedded logic a virtual

machine running on a processor or core or otherwise implemented or realized upon or within a computer-readable or machine-readable non-transitory storage medium. A computer-readable or machine-readable non-transitory storage medium includes any mechanism for storing or transmitting information in a form readable by a machine (*e.g.*, a computer). For example, a computer-readable or machine-readable non-transitory storage medium includes any mechanism that provides (*i.e.*, stores and/or transmits) information in a form accessible by a computer or computing machine (*e.g.*, computing device, electronic system, *etc.*), such as recordable/non-recordable media (*e.g.*, read only memory (ROM), random access memory (RAM), magnetic disk storage media, optical storage media, flash memory devices, *etc.*). The content may be directly executable (“object” or “executable” form), source code, or difference code (“delta” or “patch” code). A computer-readable or machine-readable non-transitory storage medium may also include a storage or database from which content can be downloaded. The computer-readable or machine-readable non-transitory storage medium may also include a device or product having content stored thereon at a time of sale or delivery. Thus, delivering a device with stored content, or offering content for download over a communication medium may be understood as providing an article of manufacture comprising a computer-readable or machine-readable non-transitory storage medium with such content described herein.

Various components referred to above as processes, servers, or tools described herein may be a means for performing the functions described. 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
5 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 performed by an Ethernet apparatus including a transmitter and receiver, comprising:

5 transmitting auto-negotiation (AN) base pages via the transmitter to an Ethernet link peer to advertise the Ethernet apparatus has the ability to support at least one Institute of Electrical and Electronics Engineers (IEEE) 802.3 Ethernet specification supporting AN;

detecting, at the receiver using a first detection mode, for a valid signal transmitted from the Ethernet link peer that does not support auto-negotiation (AN); and

10 detecting, at the receiver using a second detection mode, for AN pages transmitted from the Ethernet link peer using an IEEE 802.3 Ethernet link specification supporting AN.

2. The method of claim 1, wherein detecting for the valid signal and for the AN pages is performed in parallel by switching between the first and second detection modes.

15 3. The method of claim 1, wherein detecting for the valid signal and for the AN pages is performed in parallel by simultaneously processing signals received at the receiver that are transmitted from one of an Ethernet link peer that does not support AN and a IEEE 802.3 Ethernet link peer supporting AN.

20 4. The method of any of the preceding claims, wherein the Ethernet link that does not support AN is a 10 Gigabit per second Direct Attach Cable (10G DAC) link.

25 5. The method any of the preceding claims, wherein the Ethernet link that supports AN is one of a 10GBASE-KR link, a 25GBASE-KR link, and a 40GBASE-KR link.

6. The method any of the preceding claims, wherein the first detection mode comprises: detecting whether a signal received at the receiver is a valid signal transmitted from an Ethernet link peer that does not support AN;

30 in response to detecting the signal received at the receiver is a valid signal transmitted from an Ethernet link peer that does not support AN, starting an auto-negotiation wait timer; and detecting whether both a single_link_ready state and an autoneg_wait_timer_done are true.

7. The method of claim 6, wherein the Ethernet link that does not support AN is a 10 Gigabit per second Direct Attach Cable (10G DAC) link, further comprising:

in response to detecting that both the single_link_ready state and the autoneg_wait_timer_done are true,

5 disabling AN transmission and switching the transmitter to a 10G DAC mode.

8. The method of claim 6, further comprising:

in response to at least one of,

10 detecting the signal received at the receiver is not a valid signal transmitted from an Ethernet link peer that does not support AN; and

detecting that at least one of the single_link_ready state and the autoneg_wait_timer_done are not true,

switching from the first detection mode to the second detection mode.

15 9. The method of any of the preceding claims, wherein the second detection mode comprises:

detecting, via AN pages transmitted from an Ethernet link peer, that the Ethernet link peer supports one or more IEEE 802.3 Ethernet specifications supported by the Ethernet apparatus; and

20 completing auto-negotiation and link training with the Ethernet link peer to initialize link operations with the Ethernet link peer using one of the one or more IEEE 802.3 Ethernet link specifications supported by the Ethernet link peer and the Ethernet apparatus.

10. The method of claim 8 further comprising disabling the first detection mode.

25

11. An Ethernet apparatus comprising:

a receiver, configured to receive Ethernet signals in accordance with at least one Institute of Electrical and Electronics Engineers (IEEE) 802.3 Ethernet specification supporting auto-negotiation (AN) and Ethernet signals in accordance with at least one Ethernet specification that
30 does not support AN.

a transmitter, configured to transmit Ethernet signals in accordance with the at least one IEEE 802.3 Ethernet specification supporting AN and Ethernet signals in accordance with the at least one Ethernet specification that does not support AN;

a first block of physical layer (PHY) circuitry configured to perform PHY processing of Ethernet signals in accordance with the at least one IEEE 802.3 Ethernet specification supporting AN;

a second block of PHY circuitry configured to perform PHY processing of Ethernet

signals in accordance with the at least one Ethernet specification that does not support AN; and

an auto-negotiation module, comprising circuitry and logic configured to perform AN operations in accordance with the at least one IEEE 802.3 Ethernet specification supporting AN,

wherein the Ethernet apparatus is further configured to,

transmit AN base pages via the transmitter to advertise the Ethernet apparatus has

the ability to support at least one IEEE 802.3 Ethernet specification supporting AN;

detect, at the receiver using a first detection mode, for a valid signal transmitted from an Ethernet link peer that does not support auto-negotiation (AN); and

detect, at the receiver using a second detection mode, for AN pages transmitted from an IEEE 802.3 Ethernet link peer using an IEEE 802.3 Ethernet link technology

supporting AN.

12. The Ethernet apparatus of claim 11, wherein the first and second detection modes are performed in parallel by switching between the first and second detection modes.

13. The Ethernet apparatus of claim 11, wherein the first and second detection modes are performed in parallel by simultaneously processing received Ethernet signals using the first and second blocks of PHY circuitry.

14. The Ethernet apparatus of any of claims 11-13, wherein the Ethernet link peer that does not support AN is a 10 Gigabit per second Direct Attach Cable (10G DAC) Ethernet link peer.

15. The Ethernet apparatus of any of claims 11-14, wherein the Ethernet link peer that supports an IEEE 802.3 Ethernet specification supporting AN is one of a 10GBASE-KR, a 25GBASE-KR, and a 40GBASE-KR Ethernet link peer.

16. The Ethernet apparatus of any of claims 11-15, wherein the first detection mode comprises:

detecting whether a signal received at the receiver is a valid signal transmitted from an Ethernet link peer that does not support AN;

in response to detecting the signal received at the receiver is a valid signal transmitted from an Ethernet link peer that does not support AN, starting an auto-negotiation wait timer; and detecting whether both a single_link_ready state and an autoneg_wait_timer_done are true.

5

17. The Ethernet apparatus of claim 16, wherein the Ethernet link peer that does not support AN is a 10 Gigabit per second Direct Attach Cable (10G DAC) Ethernet link peer, and wherein the first detection mode further comprises:

in response to detecting that both the single_link_ready state and the autoneg_wait_timer_done are true, disabling transmission of the AN base pages and switching the transmitter to a 10G DAC mode.

18. The Ethernet apparatus of any of claims 11-17, wherein the first detection mode further comprises:

in response to at least one of, detecting the signal received at the receiver is not a valid signal transmitted from an Ethernet link peer that does not support AN; and detecting that at least one of the single_link_ready state and the autoneg_wait_timer_done are not true, switching from the first detection mode to the second detection mode.

19. The Ethernet apparatus of any of claims 11-18, wherein the second detection mode comprises:

detecting, via AN pages transmitted from an IEEE 802.3 Ethernet link peer, that the Ethernet link peer supports one or more 802.3 Ethernet link types supported by the Ethernet apparatus; and

completing auto-negotiation and link training with the IEEE 802.3 Ethernet link peer to initialize link operations with the IEEE 802.3 Ethernet link peer using one of the one or more 802.3 Ethernet link types supported by the IEEE 802.3 Ethernet link peer and the Ethernet apparatus.

20. A multi-mode Ethernet Network Interface Controller (NIC) comprising:
a receiver, configured to receive 10 Gigabit per second (10G) Ethernet signals;
a transmitter, configured to transmit 10G Ethernet signals;

a first physical layer (PHY) circuit block, selectively coupled to the receiver and transmitter and configured to process Institute of Electrical and Electronics Engineers (IEEE) 802.3 10GBASE-KR Ethernet signals;

a second PHY circuit block, selectively coupled to the receiver and transmitters and configured to process 10G Direct Attachment Cable (DAC) Ethernet signals,

wherein the multi-mode Ethernet NIC is further configured, when coupled to an Ethernet port to which an Ethernet link peer is communicatively coupled via a copper cable, to detect whether the Ethernet link peer is transmitting 10G DAC Ethernet signals or 10GBASE-KR Ethernet signals.

21. The multi-mode Ethernet NIC of claim 20, further configured to:

transmit auto-negotiation (AN) base pages via the transmitter to advertise the multi-mode Ethernet NIC has the ability to support 10GBASE-KR Ethernet signals supporting AN;

detect, at the receiver using a first detection mode, for a valid signal transmitted from an Ethernet link peer transmitting 10G DAC Ethernet signals; and

detect, at the receiver using a second detection mode, for AN pages transmitted from a 10GBASE-KR Ethernet link peer.

22. The multi-mode Ethernet NIC of claim 21, wherein the first and second detection modes are performed in parallel by switching between the first and second detection modes.

23. The multi-mode Ethernet NIC of claim 21 or 22, wherein the first detection mode comprises:

in response to detecting the Ethernet signal received at the receiver is a valid signal transmitted from a 10G DAC Ethernet link peer, starting an auto-negotiation wait timer; and detecting whether both a single_link_ready state and an autoneg_wait_timer_done are true.

24. The multi-mode Ethernet NIC of claim 23, wherein the first detection mode further comprises:

in response to detecting that both the single_link_ready state and the autoneg_wait_timer_done are true,

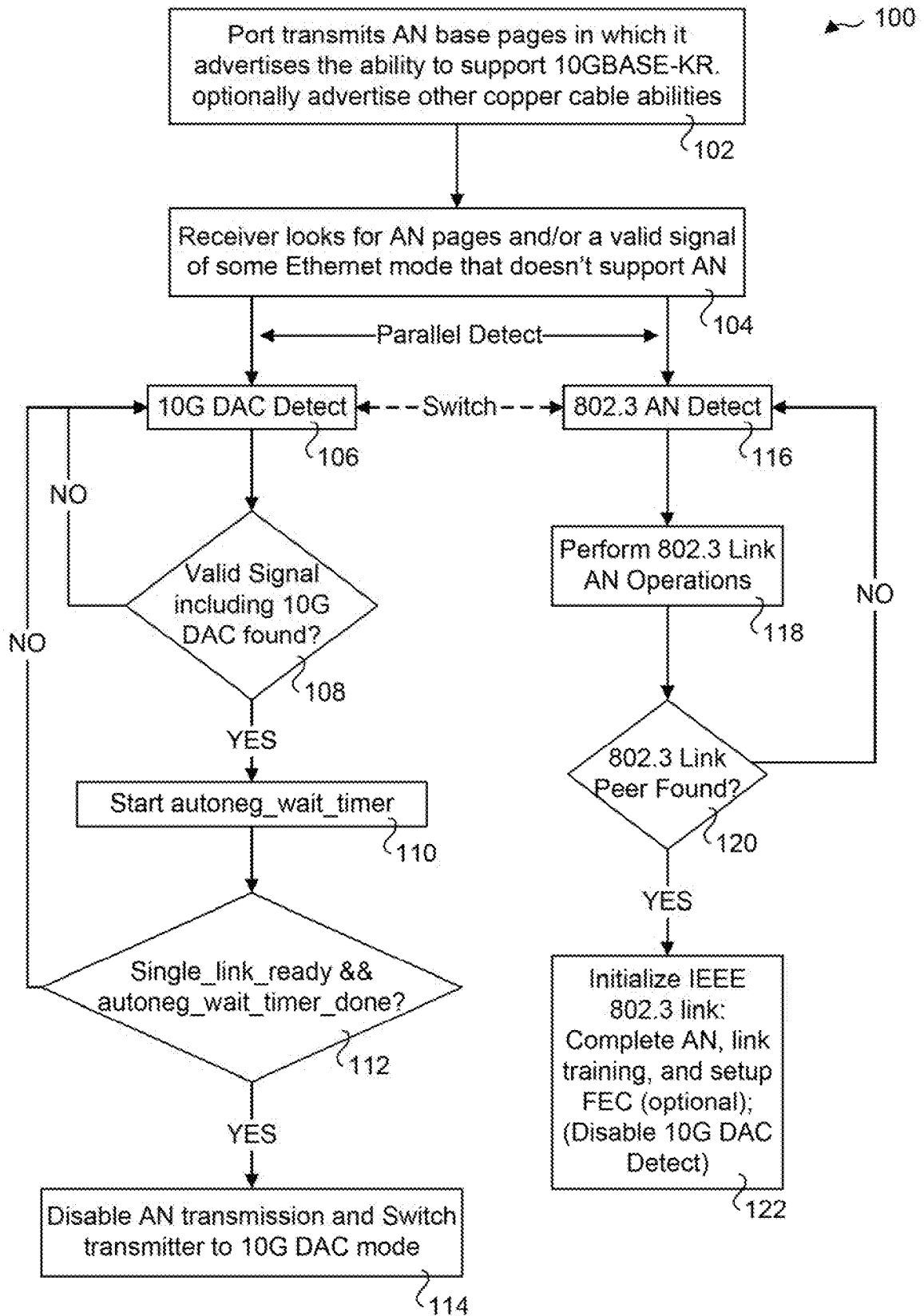
disabling transmission of the AN base pages and switching the transmitter to a 10G DAC mode.

25. The multi-mode Ethernet NIC of any of claims 21-24, wherein the second detection mode comprises:

detecting, via AN pages transmitted from a 10GBASE-KR Ethernet link peer, that the
5 Ethernet link peer supports one or more 802.3 Ethernet link types supported by the multi-mode Ethernet NIC; and

completing auto-negotiation and link training with the 10GBASE-KR Ethernet link peer to initialize link operations with the 10GBASE-KR Ethernet link peer using 10GBASE-KR Ethernet signaling.

10

**Fig. 1**

Bit	Technology
A0	1000BASE-KX
A1	10GBASE-KX4
A2	10GBASE-KR
A3	40GBASE-KR4
A4	40GBASE-CR4
A5	100GBASE-CR10
A6	100GBASE-KP4
A7	100GBASE-KR4
A8	100GBASE-CR4
A9	<u>25GBASE-KR-S or 25GBASE-CR-S</u>
A10	<u>25GBASE-KR or 25GBASE-CR</u>
<u>A9 Through A24</u> <u>A11 Through A22</u>	Reserved for Future Technology

100



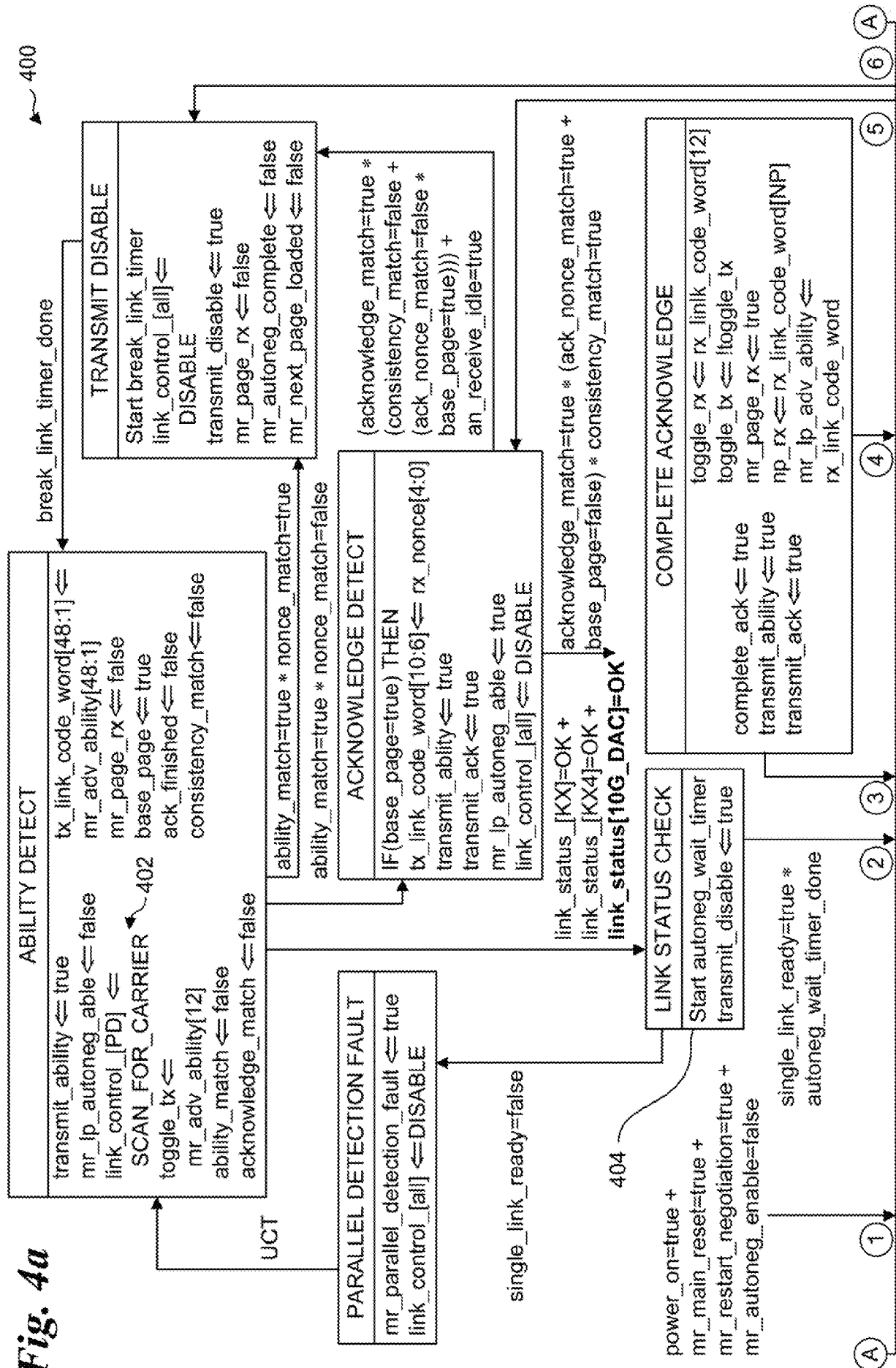
Fig. 2

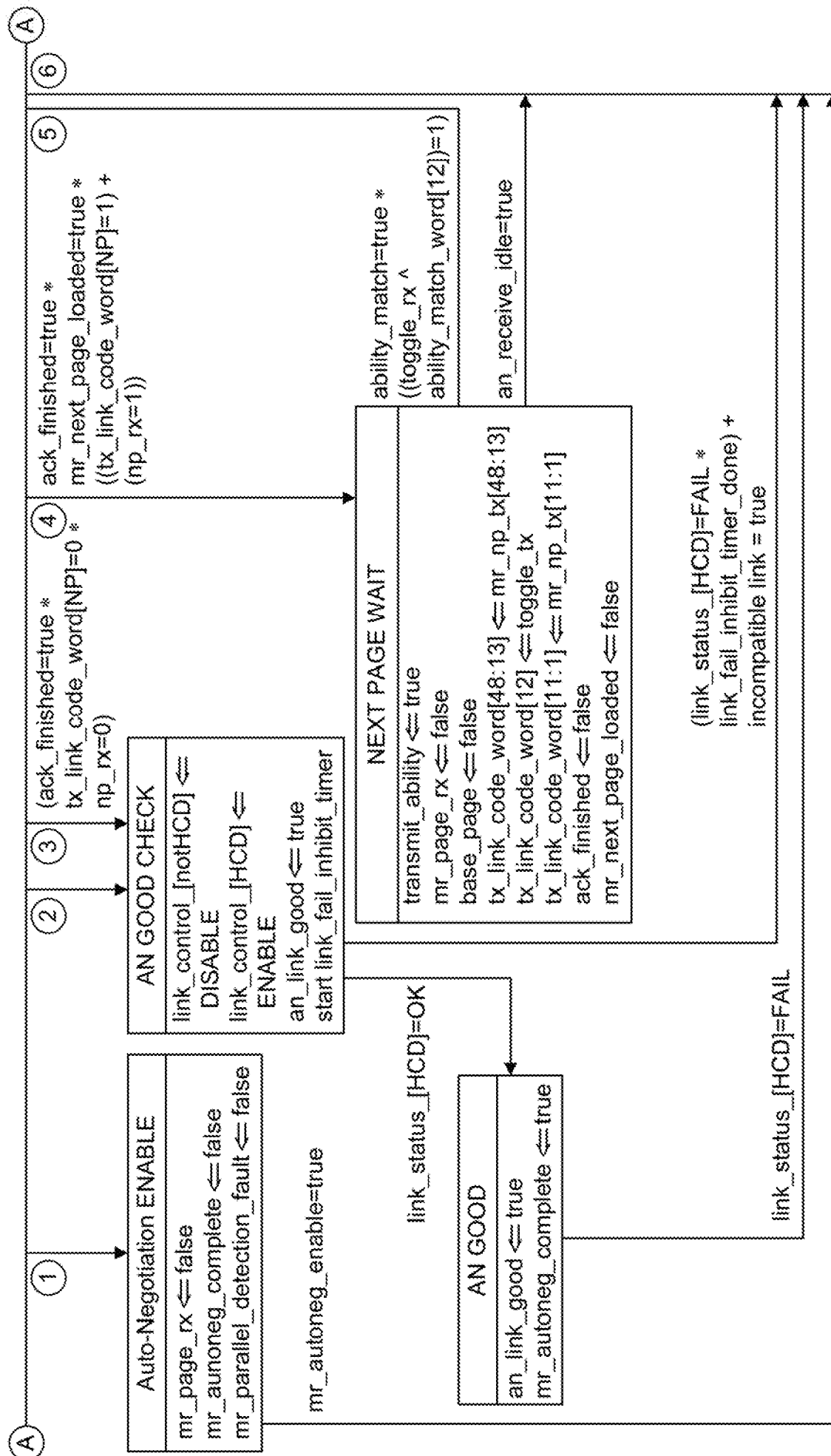
D 0	D 1	D 2	D 3	D 4	D 5	D 6	D 7	D 8	D 9	D 10	D 11	D 12	D 13	D 14	D 15
S 0	S 1	S 2	S 3	S 4	E 0	E 1	E 2	E 3	E 4	C 0	C 1	C 2	RF	Ack	NP

D 16	D 17	D 18	D 19	D 20	D 21	D 22	D 23	D 24	D 25	D 26	D 27	D 28	D 29	D 30	D 31	D 32	D 33	D 34	D 35	D 36	D 37	D 38	D 39	D 40	D 41	D 42	D 43	D 44	D 45	D 46	D 47	
T 0	T 1	T 2	T 3	T 4	A 0	A 1	A 2	A 3	A 4	A 5	A 6	A 7	A 8	A 9	F 0	F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8	F 9	F 10	F 11	F 12	F 13	F 14	F 15	F 16	F 17

Fig. 3 (prior art)

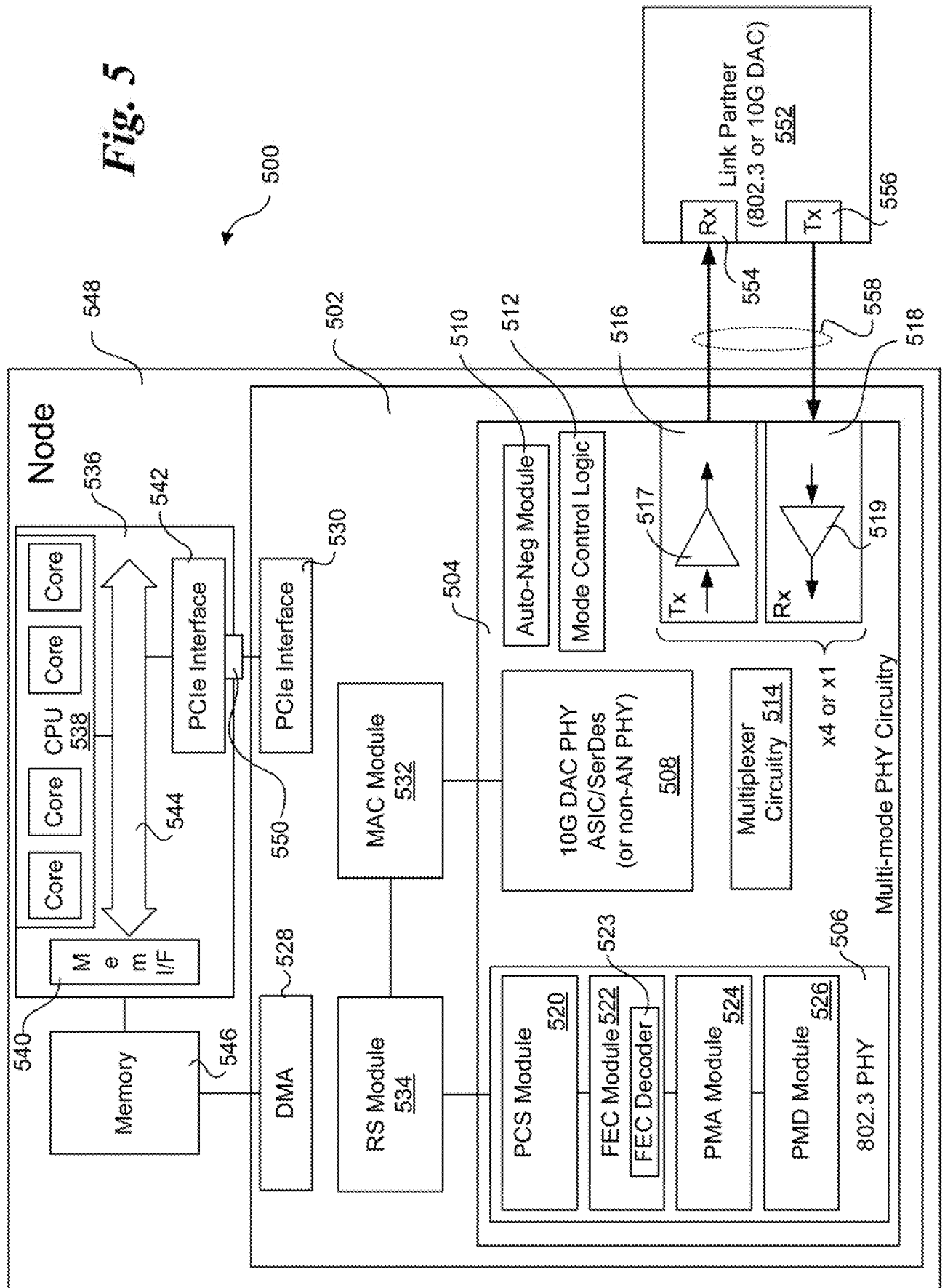
Fig. 4a





400

Fig. 4b



INTERNATIONAL SEARCH REPORTInternational application No.
PCT/US2017/038817**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 7-8, 10, 17
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Claims 7-8, 10 and 17 are not clear because they refer to claims 6 and 16 respectively, which do not comply with PCT Rule 6.4(a).
3. ☒ Claims Nos.: 5-6, 9, 15-16, 18-19, 25
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of any additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/038817**A. CLASSIFICATION OF SUBJECT MATTER****H04L 29/06(2006.01)i, H04L 12/28(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 29/06; H04L 12/24; H04J 14/00; H04B 10/08; H04B 10/20; H04B 1/38; H04L 12/28

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: auto-negotiation, base, page, 10g, ethernet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2016-0164736 A1 (KENT C. LUSTED et al.) 09 June 2016 See paragraphs [0015], [0028], [0038]-[0039]; claims 1, 12; and figures 1, 3.	1-3
A		4, 11-14, 20-24
Y	US 2010-0189168 A1 (BRADLEY J. BOOTH et al.) 29 July 2010 See paragraph [0054]; and figure 7.	1-3
A	US 2014-0169386 A1 (BROADCOM CORPORATION) 19 June 2014 See paragraphs [0026]-[0038]; and figures 1-4.	1-4, 11-14, 20-24
A	US 2002-0118412 A1 (PATRICK GIBSON et al.) 29 August 2002 See paragraphs [0021]-[0028]; and figures 4-6.	1-4, 11-14, 20-24
A	US 2009-0129777 A1 (VIKRAM SINGH) 21 May 2009 See paragraphs [0028]-[0038]; and figures 2-4.	1-4, 11-14, 20-24

☐ 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

12 October 2017 (12.10.2017)

Date of mailing of the international search report

12 October 2017 (12.10.2017)

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

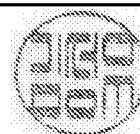


Facsimile No. +82-42-481-8578

Authorized officer

KIM, Seong Woo

Telephone No. +82-42-481-3348



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2017/038817

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2016-0164736 A1	09/06/2016	WO 2016-089520 A1	09/06/2016
US 2010-0189168 A1	29/07/2010	AU 2003-285044 A1	03/06/2004
		AU 2003-285044 A8	03/06/2004
		CN 100469046 C	11/03/2009
		CN 1711724 A	21/12/2005
		CN 1922828 A	28/02/2007
		CN 1922828 C	10/06/2007
		EP 1597874 A2	23/11/2005
		EP 1597874 B1	05/08/2009
		EP 1738533 A1	03/01/2007
		EP 1738533 B1	25/09/2013
		JP 2007-529961 A	25/10/2007
		JP 4443601 B2	31/03/2010
		TW 200423565 A	01/11/2004
		TW 200536715 A	16/11/2005
		TW 236237 A	11/07/2005
		TW 236237 B	11/07/2005
		TW 251540 B	21/03/2006
		TW I236237 B	11/07/2005
		TW I251540 B	21/03/2006
		US 2004-0091027 A1	13/05/2004
		US 2005-0111531 A1	26/05/2005
		US 7720135 B2	18/05/2010
		US 7885321 B2	08/02/2011
		WO 2004-045181 A2	27/05/2004
		WO 2004-045181 A3	23/09/2004
		WO 2005-091569 A1	29/09/2005
US 2014-0169386 A1	19/06/2014	US 2015-0350009 A1	03/12/2015
		US 9105377 B2	11/08/2015
US 2002-0118412 A1	29/08/2002	GB 2372669 A	28/08/2002
		GB 2372669 B	23/04/2003
		US 7023873 B2	04/04/2006
US 2009-0129777 A1	21/05/2009	US 8184556 B2	22/05/2012