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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 December 2016 (15.12.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/200390 A1

(51) International Patent Classification:
H04L 12/841 (2013.01) *H04L 12/937* (2013.01)

(21) International Application Number:
PCT/US2015/035333

(22) International Filing Date:
11 June 2015 (11.06.2015)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **HEWLETT PACKARD ENTERPRISE DEVELOPMENT LP** [US/US]; 11445 Compaq Center Drive West, Houston, TX 77070 (US).

(72) Inventors: **RAHUL, Kumar**; 153 Taylor Street, Littleton, Massachusetts 01460 (US). **MOHAN, Rupin T,**; 153 Taylor Street, Littleton, Massachusetts 01460 (US). **PUTTAGUNTA, Krishna**; 8000 Foothills Blvd, Roseville, California 95747 (US).

(74) Agent: **SURESH, Anup, A.**; Hewlett Packard Enterprise, 3404 E. Harmony Road, Mail Stop 79, Fort Collins, CO 80528 (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 the identity of the inventor (Rule 4.17(i))*

[Continued on next page]

(54) Title: NETWORK FRAME HOLD TIME PARAMETER

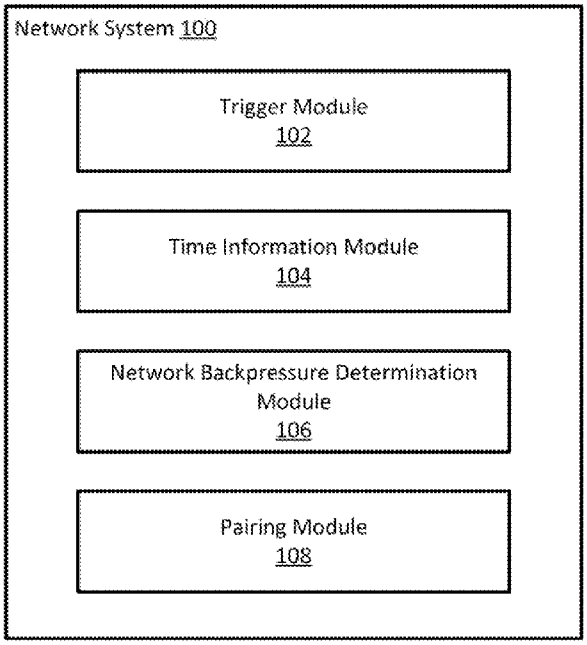

FIG. 1

(57) Abstract: According to an example, coordination of adjustments to network frame hold time parameters between network devices in a network is initiated based on a trigger condition. Time data may be obtained from a plurality of network devices in the network, where the time data describes a network frame hold time parameter of each network device in the plurality and a network frame processing time of each network device. A set of affected network devices affected by network back pressure in the network, and a set of non-affected network devices in the plurality not included in the set of affected network devices, may be determined. Based on the time data, a pairing may be determined between a time-available network device from the set of non-affected network devices and a time-needed network devices, from the set of affected network devices, to receive an allocation of time credit from the time-available network device.

WO 2016/200390 A1

Published:

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

NETWORK FRAME HOLD TIME PARAMETER

BACKGROUND

[1] Various network technologies, such as Fibre Channel, require that a network frame be dropped if the network frame is delayed more than a certain time duration in a network device (e.g., switch, router, etc.) while waiting to be forwarded to another network device. As a result, in storage area network (SAN) fabric implemented by Fibre Channel, network frames get dropped if they remain in the SAN for more than an allowed time threshold. Generally, the allowed time threshold is implemented as a configurable parameter to hold a network frame a certain duration before it is dropped, thereby ensuring that the network frame reaches its destination before an allowed delay limit.

BRIEF DESCRIPTION OF THE DRAWINGS

[2] Certain examples are described in the following detailed description in reference to the following drawings.

[3] FIGs. 1 and 2 are block diagrams illustrating example network systems that facilitate adjustment of a network frame time hold parameter according to the present disclosure.

[4] FIGs. 3 and 4 are flowcharts illustrating example methods performed by an example network device to facilitate adjustment of a network frame time hold parameter according to the present disclosure.

[5] FIG. 5 is a block diagram illustrating an example network environment that includes example network devices describes in the present disclosure.

[6] FIGs. 6 and 7 are block diagrams illustrating example network devices according to the present disclosure.

DETAILED DESCRIPTION

[7] Certain conventional network devices, such as those based on the Fibre Channel standard and used in storage area networks (SANs), network frames are dropped if they remain in the network more than an allowed threshold time limit specified by the Fibre Channel standard. To implement the threshold time

limit and ensure timely delivery, such conventional network devices (e.g., network switches, network routers, etc.) include a configurable parameter (hereafter, a network frame hold time) that determines how long a network frame (e.g., including a network packet), waiting to be forwarded to another network device (e.g., another network switch), will be held in buffer queue before it is dropped at the network device.

[8] As used herein, the network frame forward delay time value (hereafter, network frame forward delay) can be associated with a network port of a network device and may refer to the time required for that network device to forward a network frame arriving at the network port to the network frame's next network hop. Accordingly, the network frame forward delay associated with a network port may be calculated by subtracting the time the network frame arrived at a network device from the time the network frame is forwarded to the next network hop. The network frame forward delay can differ between different network ports of a network device. As also used herein, an average network frame forward delay can be associated with a network device and may be calculated by averaging together the network frame forward delay associated with each network port of the network device (i.e., for an *n*-port network device, the average is calculated by summing up all the network frame forward delays and dividing the sum by *n*). The average network frame forward delay can differ between conventional network devices in a network and, as a result, those conventional network devices having higher network frame forward delay than others may present a bottleneck for network traffic flowing through the network.

[9] Generally, when a network frame to be forwarded from a conventional network device to another network device (e.g., network switch) is received at the conventional network device, it is queued at the conventional network device until such time as the next network hop (i.e., the other network device) is ready to receive the network frame, or until the network frame is dropped from the network frame buffer queue based on the network frame hold time. In some contexts, the next network hop network device may not be ready to receive a network frame when the next network hop network device is experiencing higher work load than other network devices in its network, has reached its resource limits (e.g.,

hardware limits), or is slow to drain network frames. Typically, as the number of network frames on the network frame buffer queue increases for a conventional network device, so does the average network frame forward delay of that network device. Additionally, an increase in average network frame forward delay of a conventional network device, in turn, causes an increase in the number of network frames being dropped as a result of the network frame hold time limit being reached (by network frames on the network frame buffer queue). Consequently, a conventional network device can present a bottleneck and cause network back pressure in a network as its average frame forward delay increases and the number of network frames dropped increases.

[10] Various examples described herein provide systems and methods that permit a set of network devices in a network to dynamically adjust (e.g., tune) their respective network frame hold time parameter based on conditions in the network, such as the presence of network back pressure in the network (e.g., as a result of a network bottleneck). For some examples, the adjustment of network frame hold time parameters of various network devices in a network involves coordination between those network devices. Additionally, in some examples, the adjustment of network frame hold time parameters of various network devices is coordinated between those network devices by way of a network device designated as the principal network device in the network (e.g., designated by a manual selection or automatically elected by the network devices during handshake process between the network devices).

[11] Where a first network device experiencing an average network frame forward delay that is higher than a second network device, certain examples coordinate an adjustment of network frame hold time parameters between the two network devices such that the network frame hold time parameter of the second network device is reduced by a specific amount of time and the network frame hold time parameter of the first network device is increased by the specific amount of time. In this way, the second network device can allocate time credit from its network frame hold time parameter, the resulting allocation of time credit can be transferred from the second network device to the first network device, and the first network device can apply (e.g., add) the allocation of time credit to

its own network frame hold time parameter. In applying the allocation of time credit to its network frame hold time parameter, the first network device can attempt to alleviate network back pressure (e.g., network frame drops due to high network frame forward delay) being experienced by, and which may possibly be caused by (e.g., due to resource limitations), the first network device. Additionally, in this way, such examples can adjust network frame hold time parameters within a network while complying with overall timing requirements of the network as defined by a network standard the network implements (e.g., Fibre Channel standard, which requires that a network frame not be delayed in the network longer than 4 seconds).

[12] As used herein, a network device can include a network switch, network router, network gateway, network hub, or a network-enabled electronic device having a processor. Additionally, as used herein, a network frame can include a network packet that is transported over a communications network. With respect to various examples described herein, a network device in a network being affected by network back pressure in the network (and thus experience increased average network frame forward delay) may be referred to herein as an affected network device. A network device in a network not being affected by the same network back pressure may be referred to herein as a non-affected network device. Further, a network device in a network that is a candidate to provide an allocation of time credit to another network device in the network may be referred to herein as a time-available network device or a time-loaner network device. A network device in a network that is a candidate to receive an allocation of time credit from another network device in the network may be referred to herein as a time-needed network device or a time-receiver network device.

[13] According to various examples, a first network device in a given network monitors for when its average network frame forward is greater than (or equal to) a threshold value, where the threshold value may be a parameter for the first network device that is manually set by an individual (e.g., user-defined parameter set by a network administrator). Depending on the example, the threshold value may depend on network frame priority queues for the first network device, and may be adjusted (e.g., tuned) to improve overall performance or quality of service

(QoS) of the overall network. For various examples, the threshold value is set to be lesser than or equal to the current network frame hold time parameter of the first network device. Further, for various examples, the threshold value serves as a trigger for initiating the adjustment of the network frame hold time parameter of the first network device and of at least one other network device in the network. In particular, when the network frame forward delay of the first network device is greater than (or equal to) the threshold value, this can serve as an indication that the first network device is presenting a bottleneck in the network or the first network device is experiencing network back pressure.

[14] Accordingly, when the average network frame forward delay of the first network device is greater than (or equal to) the threshold value, the first network device can trigger a process by which the network frame hold time parameter of the first network device and of at least one other network device in the network are adjusted in effort to alleviate the network bottleneck or back pressure detected (by way of the threshold value being surpassed). The first network device, serving as the trigger network device, may cause the process for coordinating the adjustment of network frame hold time parameters to initiate at the first network device, or at another network device in the network, such as a one designated the principal network device in the network (e.g., a principal network switch in a Fibre Channel-based network). Hereinafter, a network device coordinating the adjustment of the network frame hold time parameters between network devices in a network may be referred to as the coordinating network device. As noted above, the trigger network device may cause the process for coordinating the adjustment to be initiated on itself (rather than another network device), in which case the trigger network device and the coordinating network device are same network device. A network device triggering the adjustment of the network frame hold time parameters between network devices in a network (e.g., when it average network frame forward delay surpasses a threshold value) is referred to as the trigger network device.

[15] Depending on the example, some or all network devices in the network may operate similarly to the first network device described above and include a

threshold value to determine when to trigger the adjustment of its network frame hold time parameter and that of another network device in the given network.

[16] At the beginning of the coordination process, the coordinating network device may obtain (e.g., request and then receive) from each network device, in a set of network devices in the network, a tuple containing the (time) value of their network frame hold time parameter and their current processing time. The current processing time may comprise the network device's current average network frame forward delay. The set of network devices from which tuples are obtained may include some or all of the network devices in a network associated with the coordinating network device. Based on the trigger network device, the coordinating network device determines direction of network back pressure from the trigger network device and the direction of the anti-back pressure from the network device. This determination may be determined based on the topology of the network and the direction of network traffic. Hereinafter, a network device affected by the determined network back pressure is referred to as an affected network device, while one that is not is referred to a non-affected network device. Based on the information regarding processing time (from the obtained tuples), the determined network back pressure, and the determined anti-back pressure, the coordinating network device determines pairings between non-affected network devices having available processing time (hereafter, time-available network devices) and affected network devices needing processing time (hereafter, time-needed network devices). These pairings may be one-to-one, where one non-affected network device is paired with one affected network devices, or may be one-to-many, where one non-affected network device is paired with many affected network devices. Additionally, these pairings may be many-to-one, where many non-affected network devices are paired with one affected network device. Based on the determined pairings, the coordinating network device may generate and provide data regarding the pairings to each of the time-available network devices included in at least one determined pairing. This data (hereafter, pairing data) sent to an individual time-available network device may describe all the pairings determined by the coordinating network device, or only those pairings that include the individual time-available network

device. Additionally, the pairing data may provide the values of the network frame hold time parameter of the time-available network device and the processing time (e.g., average network frame forward delay) of the time-available network device (as reported to the coordinating network device), and the values of the network frame hold time parameter and the processing time (e.g., average network frame forward delay) of the time-needed network device (as reported to the coordinating network device).

[17] Based on pairing data received from the coordinating network device, a time-available network device may determine (e.g., calculate) how much time credit it can allocate from its network frame hold time parameter to the time-needed network device to which it is paired according to the pairing data. This allocation time credit may be determined in view of the time-available network devices current processing time (e.g., current network frame forward delay) or the processing time of the time-available network device as described in the pairing data (and originally reported to the coordinating network device by the time-available network device). The time-available network device may then adjust (e.g., update) its own network frame hold time parameter based on allocation of time credit. The time-available network device may also request the time-needed network device update its network frame hold time parameter based on the allocation of time credit. Depending on the example, the time-available network device may adjust its own network frame hold time parameter (based on allocation of time credit) before it requests the time-needed network device update its network frame hold time parameter based on the allocation of time credit. After sending the request, the time-available network device may await acknowledgement from the time-needed network device that it has adjusted its network frame hold time parameter as requested.

[18] Based on the request received from the time-available network device, the time-needed may adjust (e.g., update) its network frame hold time parameter based on the allocation of time credit described in the request. Subsequent to this adjustment, the time-needed network device may send an acknowledgment of the time-available network device confirming the adjustment has been implemented.

[19] For some examples, after receiving acknowledgement from the time-needed network device confirming adjustment of its network frame hold time parameter as requested, the time-available network device may send an acknowledgement to the coordinating network device confirming completion of the adjustment for the time-available network device according to the pairing data. In the event the time-available network device does not receive an acknowledgement from the time-needed network device (e.g., before a pre-determined time limit expires), the time-available network device may not send an acknowledgement to the coordinating network device and, as a result, the coordinating network device may restart the adjustment process (e.g., start by re-obtaining from network tuples containing the value of their respective network frame hold time parameters and current processing times).

[20] The following provides a detailed description of examples illustrated by FIGs. 1-7.

[21] FIG. 1 is a block diagram illustrating an example network system 100 that facilitates adjustment of a network frame time hold parameter according to the present disclosure. For some examples, the network system 100 is included by a coordinating network device in a network, which in the context of a Fibre Channel network, may be the principal network device. As shown, the network system 100 includes a trigger module 102, a time information module 104, a network back pressure determination module 106, and a pairing module 108. Depending on the example, the network system 100 may be part of a network device that includes a processor, such as a networked computer system (e.g., desktop, laptop, server, etc.), network switch, network router, network gateway, or any device capable of receiving and sending a network packet. In various examples, the components or the arrangement of components in the network system 100 may differ from what is depicted in FIG. 1.

[22] As used herein, modules and other components of various examples may comprise, in whole or in part, machine-readable instructions or electronic circuitry. For instance, a module may comprise machine-readable instructions executable by a processor to perform one or more functions in accordance with various examples described herein. Likewise, in another instance, a module may

comprise electronic circuitry to perform one or more functions in accordance with various examples described herein. The elements of a module may be combined in a single package, maintained in several packages, or maintained separately.

[23] The trigger module 102 may initiate, based on a trigger condition, coordination of adjustments to network frame hold time parameters between at least two network devices in a network. The trigger condition may the detection of network congestion or a bottleneck in the network. Monitoring the network frame forward delay of various network devices in the network can assist in detecting such congestion or bottlenecks. Depending on the example, such monitoring may be conducted at the network system 100 or a separate network device in the network serving as a trigger network device. In particular examples, trigger condition comprises receiving, from a trigger network device in the network, a signal indicating that a network frame processing time of the trigger network device exceeds (or equals) a trigger threshold. As described herein, the trigger threshold may a parameter of the trigger network device that is manually set by an individual (e.g., a network administrator), and which may be set based on network frame priority queues of the trigger network device.

[24] The time information module 104 may facilitate obtaining time data from a plurality of network devices in the network. The plurality of network devices, which can comprise network switches, network routers, network gateways, and the like, may include some or all of the network devices associated with the network. The time information module 104 may facilitate obtaining the time data by requesting each network device report its respective time information (e.g., network frame processing time and network frame hold time parameter) to the network system 100. The time information may be received from the individual network devices as tuples containing requested time values. As described herein, a network frame hold time parameter associated with a given network device may be set to a value representing the amount of time a network frame to be forwarded to another network device is held in the buffer queue of the given network device before it is dropped. As also described herein, the network frame processing time associated with a given network device may comprise the average network frame forward delay of the given network device.

[25] The network back pressure determination module 106 may facilitate determining, based on the time data from the time information module 104, a set of affected network devices being affected by network back pressure in the network and a set of non-affected network devices not included in the set of affected network devices. The set of affected network devices can include a trigger network device that signaled start of the coordination efforts by the network system 100 (via the trigger module 102). Additionally, the set of affected network devices in the network may be those dropping network frames as a result if increased network frame processing time, which may be the result of network congestion or a network bottleneck in the network.

[26] Determining the set of affected network devices and the set of non-affected network devices may involve determining the direction of the network backpressure (e.g., from the trigger network device) caused by network congestion or a bottleneck present in the network. Determining the set of affected network devices and the set of non-affected network devices may also involve determining the direction opposite to the network backpressure (referred to herein as, "anti-back pressure"). Analysis of the topology of the network may be involved in determining either of these directions.

[27] The pairing module 108 may facilitate determining, based on the time data from the time information module 104, a pairing between a time-available network device in the set of non-affected network devices and a time-needed network device in the set of affected network devices, the time-needed network device being a candidate for receiving an allocation of time credit from the time-available network device to adjust its associated network frame hold time parameter based on the allocation. The time-available network device may be a network device in the network that the network system 100 identifies as having available network frame processing time based on the network frame processing time the time-available network device provided in the time data obtained by the time information module 104. According to various examples described herein, the time-available network device is to provide the allocation of time credit from its own network frame hold time hold parameter, thereby resulting in a reduction in the parameter value.

[28] For various examples described herein, including the network system 100, communication with network devices in the network is achieved by way of standard control information that that enables operation of the network. For instance, in the context of a Fibre Channel network, communication of signals, time data, or pairing data between the network system 100 and a network device in the network may be facilitated (e.g., transported within) through a link control frame, a link service frame, or an extended link service frame.

[29] FIG. 2 is a block diagram illustrating an example network system 200 that facilitates adjustment of a network frame time hold parameter according to the present disclosure. As shown, the network system 200 includes a trigger module 202, a time information module 204, a network back pressure determination module 206, a pairing module 208, and an adjustment coordination module 210. In various examples, the components or the arrangement of components in the network system 200 may differ from what is depicted in FIG. 2.

[30] For some examples, the trigger module 202, the time information module 204, the network back pressure determination module 206, and the pairing module 208 are respectively similar to the trigger module 102, the time information module 104, the network back pressure determination module 106, and the pairing module 108 of the network system 100 described above with respect to FIG. 1.

[31] The adjustment coordination module 210 may facilitate coordination of transfer of the allocation of time credit between the time-available network device and the time-needed network device. Coordinating the transfer of the allocation of time credit can comprise providing the time-available network device with pairing data describing the pairing between the time-available network device and the time-needed network device, describing the network frame hold time parameter of the time-needed network device, and describing the network frame processing time of the time-needed network device. The pairing data may also describe the network frame hold time parameter of the time-available network device and the network frame processing time of the time-available network device, which may the time-available network device rely on when determining

the allocation of time credit for the time-needed network device. For some examples, the network system 200 may provide the pairing data to at least all network devices in network that are included in at least one pairing described by the pairing data.

[32] Coordinating the transfer of the allocation of time credit may further comprise monitoring for an acknowledgment from the time-available network device at least indicating that the time-needed network device has adjusted its associated network frame hold time parameter. According to some examples, the adjustment coordination module 210 may cause the time information module 204, the network back pressure determination module 206, and the pairing module 208 to repeat their respective operations if the acknowledgement is not received from the time-available network device within a timeout period.

[33] FIG. 3 is a flowchart illustrating an example method 300 performed by an example network device to facilitate adjustment of a network frame time hold parameter according to the present disclosure. For some examples, the method 300 may be one performed by a time-available network device in a network. The method 300 may be implemented in the form of executable instructions stored on a machine-readable medium or in the form of electronic circuitry. Though the method 300 is described below from the perspective of a first network device in a network performing blocks 302, 304, 306 and 308, for some examples, less than all of the blocks be performed by the first network device.

[34] In FIG. 3, the method 300 may begin at block 302 by a first network device (e.g., time-available network device) in a network receiving pairing data from a second network device (e.g., coordinating network device) in the network. The pairing data may describe a pairing between the first network device and a third network device (e.g., time-needed network device) in the network, may describe a network frame hold time parameter of the third network device, and may describe the network frame processing time of the third network device. As described herein, the pairing may be one determined by the second network device based on time data the second network device obtained from various network devices in the network, including the first network device.

[35] The method 300 may continue to block 304 by the first network device determining an allocation of time credit for the third network device based on the pairing data and a network frame hold time parameter of the first network device. As described herein, for some examples, the network frame hold time parameter of the first network device may be adjusted down based on the allocation of time credit determined by the first network device.

[36] The method 300 may continue to block 306 by the first network device adjusting the network frame hold time parameter of the first network device by reducing the network frame hold time parameter of the first network device by the allocation of time credit. The method 300 may then continue to block 308 by the first network device requesting the third network device to adjust the network frame hold time parameter of the third network device based on the allocation of time credit. For some examples, the first network device adjusts its own network frame hold time parameter based on the allocation of time credit before it requests the third network device to adjust its network frame hold time parameter based on the allocation.

[37] FIG. 4 is a flowchart illustrating an example method 400 performed by an example network device to facilitate adjustment of a network frame time hold parameter according to the present disclosure. For some examples, the method 400 may be one performed by a time-available network device in a network. The method 400 may be implemented in the form of executable instructions stored on a machine-readable medium or in the form of electronic circuitry. Though the method 400 is described below from the perspective of a first network device in a network performing blocks 402, 404, 406, 408, 410, and 412, for some examples, less than all of the blocks be performed by the first network device.

[38] In FIG. 4, the method 400 may begin at block 402 and may continue to blocks 404, 406, and 408, which may be respectively similar to blocks 302, 304, 306, and 308 of the method 300 as described above with respect to FIG. 3. Additionally, blocks 402, 404, 406, and 408 of the method 400 may be exclusively performed by the first network device.

[39] The method 400 may continue to block 410 by the first network device monitoring for an acknowledgment from the third network device at least

indicating that the third network device has adjusted the network frame hold time parameter of the third network device based on the allocation of time credit. The method 400 may continue to block 412 by the first network device sending a second acknowledgment from the first network device to the second network device after receiving the acknowledgment at the first network device. Depending on the example, the second acknowledgment may at least indicate that the third network device has successfully adjusted its network frame hold time parameter based on the allocation of time credit. As described herein, where the first network device fails to send a second acknowledgment to the second network device within a certain timeout period, the second network device may cause the overall adjustment process to restart.

[40] FIG. 5 is a block diagram illustrating an example network environment 500 that includes example network devices describes in the present disclosure. In particular, the network environment includes a plurality of network devices that are part of a network experiencing network back pressure 512. As shown, the network environment 500 includes a host network device 502, a set of network devices affected by the network back pressure 512 (collectively referred hereafter as affected network devices 504-1 to 504-4), a set of network devices not affected by the network back pressure 512 (collectively referred hereafter as affected network devices 506-1 to 506-3), and a data storage array 508. The host network device 502 may include a networked computer system, and the data storage array 508 may include a plurality of data storage devices implementing a storage area network (SAN) device. As described herein, each of the network devices 504 and 506 may be a networked computer system, a network switch, a network router, a network router, or another device capable of receiving and sending a network packet. The network devices 504 and 506 may implement a storage area network (SAN) by which the host network device 502 may access (e.g., read data from, write data to, or modify data stored on) the data storage array 508. During such data access operations, the host device 502 and the data storage array 508 may exchange network packets through the network devices 504 and 506.

[41] During operation, the host network device 502 sends a network frame to the data storage array 508 via a network path 510. As shown in FIG. 5, the network path 510 passes through the affected network devices 504 and the non-affected network devices 506. After the host network device 502 sends the network frame, it may wait for acknowledgement before a timeout occurs. In the context of a SAN, the network frame may comprise a Small Computer Serial Interface (SCSI) command, and the host network device 502 may send an ABORT command if it does not receive an acknowledgement of the SCSI command before the timeout period. The occurrences of timeouts can have a negative impact on storage access performance of the SAN by the host network device 502. Based on the network back pressure, the network frame sent from the host network device 502 has the possibility of being dropped at the affected network device 504-1, 504-2, 504-3, or 504-4.

[42] According to some examples, the affected network device 504-4 can operate as a trigger network device, and signal (e.g., to a coordinating network device not shown) when the affected network device 504-4 detects that its network frame processing time (e.g., comprising its average network frame forward delay) has exceeded (or is equal to) a trigger threshold value. As described herein, when the network frame processing time exceeds the trigger threshold, this can be an indication that the affected network device 504-4 is experiencing network back pressure or is presenting a bottleneck in the network, which in turn can cause in an increase in dropped network frames at the affected network device 504-4 or the other affected network devices 504. As described herein, the signal from the affected network device 504-4 can trigger the adjustment of network frame hold time parameters of the affected network devices 504. Upon successful adjustment, one or more of the affected network devices 504-1, 504-2, 504-3, and 504-4 may have adjusted network frame hold time parameters that assist in alleviating the effects of a network bottleneck at the affected network device 504-4 or the network back pressure 512 caused by the network bottleneck.

[43] FIG. 6 is block diagram illustrating an example network device 600 according to the present disclosure. As described herein, the network device 600

may be any network device having a processor, such as a networked computer system (e.g., desktop, laptop, server, etc.), network switch, network router, or network gateway. For some examples, the network device 600 operates as a trigger network device in a network. As shown, the network device 600 includes a machine-readable medium 602, a processor 604, and a network interface 606. In various examples, the components or the arrangement of components of the network device 600 may differ from what is depicted in FIG. 6. For instance, the network device 600 can include more or less components than those depicted in FIG. 6.

[44] The machine-readable medium 602 may be any electronic, magnetic, optical, or other physical storage device that stores executable instructions. For example, the machine-readable medium 602 may be a Random Access Memory (RAM), an Electrically-Erasable Programmable Read-Only Memory (EEPROM), a storage drive, an optical disc, or the like. The machine-readable medium 602 can be encoded to store executable instructions that cause the processor 604 to perform operations in accordance with various examples described herein. In various examples, the machine-readable medium 602 is non-transitory. As shown in FIG. 6, the machine-readable medium 602 includes trigger detection instructions 610, trigger informing instructions 612, and time data providing instructions 614.

[45] The processor 604 may be one or more central processing units (CPUs), microprocessors, or other hardware devices suitable for retrieval and execution of one or more instructions stored in the machine-readable medium 602. The processor 604 may fetch, decode, and execute the instructions 610, 612 and 614 to enable the network device 600 to perform operations in accordance with various examples described herein. For some examples, the processor 604 includes one or more electronic circuits comprising a number of electronic components for performing the functionality of one or more of the instructions 610, 612, and 614.

[46] The network interface 606 may facilitate communication between the network device 600 and another network device in a network associated with the network device 600. The network interface 606 may include a set of hardware

interfaces, such as physical network ports, that permits exchange of data, over a communications network, with another network device.

[47] The trigger detection instructions 610 may cause the processor 604 to detect whether a network frame processing time of the network device exceeds (or equals) a trigger threshold. The trigger informing instructions 612 may cause the processor 604 to inform a second network device (e.g., the coordinating network device) when the network frame processing time of the network device exceeds the trigger threshold. The time data providing instructions 614 may cause the processor 604 to provide time data to the second network device, the time data including a network frame hold time parameter of the network device and a network frame processing time of the network device. Depending on the example, the time data providing instructions 614 may cause the processor 604 to provide time data in response to the second network device requesting such time information. For some examples, the time data may be provided as part of the informing the second network device when the network frame processing time of the network device exceeds the trigger threshold at the network device 600.

[48] FIG. 7 is block diagram illustrating an example network device 700 according to the present disclosure. As described herein, the network device 700 may be any network device having a processor, such as a networked computer system (e.g., desktop, laptop, server, etc.), network switch, network router, or network gateway. For some examples, the network device 700 operates as a trigger network device in a network. As shown, the network device 700 includes a machine-readable medium 702, a processor 704, and a network interface 706. In various examples, the components or the arrangement of components of the network device 700 may differ from what is depicted in FIG. 7. For instance, the network device 700 can include more or less components than those depicted in FIG. 7.

[49] As described herein, the machine-readable medium 702 may be any electronic, magnetic, optical, or other physical storage device that stores executable instructions. The machine-readable medium 702 can be encoded to store executable instructions that cause the processor 704 to perform operations in accordance with various examples described herein. Additionally, the

machine-readable medium 702 can be non-transitory. As shown in FIG. 7, the machine-readable medium 702 includes trigger detection instructions 710, trigger informing instructions 712, time data providing instructions 714, adjustment request receiving instructions 716, network frame hold time parameter adjustment instructions 718, and adjustment acknowledgment instructions 720.

[50] As described herein, the processor 704 may be one or more central processing units (CPUs), microprocessors, or other hardware devices suitable for retrieval and execution of one or more instructions stored in the machine-readable medium 702. The processor 704 may fetch, decode, and execute the instructions 710, 712, 714, 716, 718, and 720 to enable the network device 700 to perform operations in accordance with various examples described herein. For some examples, the processor 704 includes one or more electronic circuits comprising a number of electronic components for performing the functionality of one or more of the instructions 710, 712, 714, 716, 718, and 720.

[51] As also described herein, the network interface 706 may facilitate communication between the network device 700 and another network device in a network associated with the network device 700. The network interface 706 may include a set of hardware interfaces, such as physical network ports, that permits exchange of data, over a communications network, with another network device.

[52] For some examples, the trigger detection instructions 710, the trigger informing instructions 712, and the time data providing instructions 714 may are respectively similar to the trigger detection instructions 610, the trigger informing instructions 612, and the time data providing instructions 614 of the network device 600 described above with respect to FIG. 6.

[53] The adjustment request receiving instructions 716 may cause the processor 704 to receive a request from a third network device (e.g., time-available network device) to adjust the network frame hold time parameter based on an allocation of time credit. The network frame hold time parameter adjustment instructions 718 may cause the processor 704 to adjust the network frame hold time parameter based on the allocation of time credit. The adjustment acknowledgment instructions 720 may cause the processor 704 to send an acknowledgment to the third network device at least indicating that the network

device has adjusted the network frame processing time based on the allocation of time credit.

[54] In the foregoing description, numerous details are set forth to provide an understanding of the subject disclosed herein. However, various examples may be practiced without some or all of these details. Some examples may include modifications and variations from the details discussed above. It is intended that the appended claims cover such modifications and variations.

CLAIMS

1. A network system, comprising:

a trigger module to initiate, based on a trigger condition, coordination of adjustments to network frame hold time parameters between network devices in a network;

a time information module to obtain time data from a plurality of network devices in the network, the time data describing a network frame hold time parameter of each network device in the plurality and a network frame processing time of each network device in the plurality;

a network back pressure determination module to determine, based on the time data, a set of affected network devices in the plurality being affected by network back pressure in the network and a set of non-affected network devices in the plurality not included in the set of affected network devices; and

a pairing module to determine, based on the time data, a pairing between a time-available network device in the set of non-affected network devices and a time-needed network device in the set of affected network devices, the time-needed network device being a candidate for receiving an allocation of time credit from the time-available network device to adjust its associated network frame hold time parameter based on the allocation.

2. The network system of claim 1, comprising an adjustment coordination module to coordinate transfer of the allocation of time credit between the time-available network device and the time-needed network device.

3. The network system of claim 2, wherein coordinating the transfer of the allocation of time credit comprises providing the time-available network device with pairing data describing the pairing between the time-available network device and the time-needed network device, describing the network frame hold time parameter of the time-needed network device, and describing network frame processing time of the time-needed network device.

4. The network system of claim 2, wherein coordinating the transfer of the allocation of time credit comprises monitoring for an acknowledgment from the

time-available network device at least indicating that the time-needed network device has adjusted its associated network frame hold time parameter.

5. The network system of claim 4, wherein the adjustment coordination module is to cause the time information module, the network back pressure determination module, and the pairing module to repeat their respective operations if the acknowledgement is not received from the time-available network device within a timeout period.

6. The network system of claim 1, wherein the trigger condition comprises receiving, from a trigger network device in the network, a signal indicating that a network frame processing time of the trigger network device exceeds a trigger threshold.

7. The network system of claim 1, wherein the network is a Fibre Channel network, and the network system is a principal network device in the Fibre Channel network.

8. A method, comprising:

receiving, at a first network device in a network, pairing data from a second network device in the network, the pairing data describing a pairing between the first network device and a third network device in the network, describing a network frame hold time parameter of the third network device, and describing network frame processing time of the third network device;

determining, by the first network device, an allocation of time credit for the third network device based on the pairing data and on a network frame hold time parameter of the first network device;

adjusting, by the first network device, the network frame hold time parameter of the first network device by reducing the network frame hold time parameter of the first network device by the allocation of time credit; and

requesting, by the first network device, the third network device to adjust the network frame hold time parameter of the third network device based on the allocation of time credit.

9. The method of claim 8, comprising monitoring, at the first network device, for an acknowledgment from the third network device at least indicating that the third network device has adjusted the network frame hold time parameter of the third network device based on the allocation of time credit.

10. The method of claim 9, comprising after receiving the acknowledgment at the first network device, sending a second acknowledgment from the first network device to the second network device, the second acknowledgment at least indicating that the third network device has adjusted the network frame hold time parameter of the third network device based on the allocation of time credit.

11. The method of claim 8, wherein the network is a Fibre Channel network, and the second network device is a principal network device in the Fibre Channel network.

12. The method of claim 8, wherein the second network device and the third network device are the same network device.

13. The method of claim 8, comprising sending time data from the first network device to the second network device, the time data including a network frame hold time parameter of the first network device and a network frame processing time of the second network device.

14. A non-transitory machine-readable medium having instructions stored thereon, the instructions being executable by a processor of a network device, the instructions causing the processor to:

detect whether a network frame processing time of the network device exceeds a trigger threshold;

inform a second network device when the network frame processing time of the network device exceeds the trigger threshold; and

provide time data to the second network device, the time data including a network frame hold time parameter of the network device and a network frame processing time of the network device.

15. The non-transitory machine-readable medium of claim 14, the instructions causing the processor to:

receive a request from a third network device to adjust the network frame hold time parameter based on an allocation of time credit;

adjust the network frame hold time parameter based on the allocation of time credit; and

send an acknowledgment to the third network device at least indicating that the network device has adjusted the network frame processing time based on the allocation of time credit.

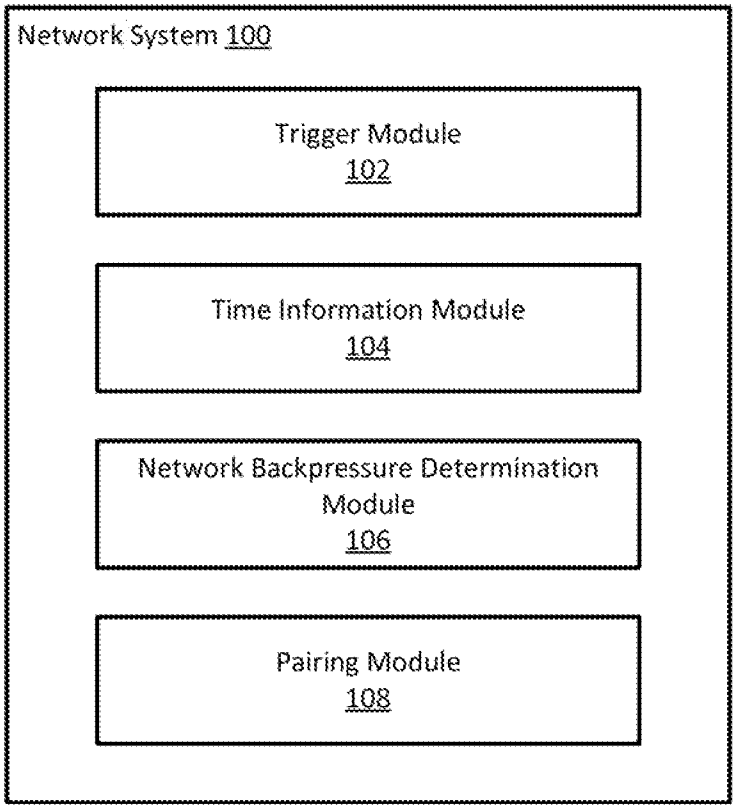
Network System 100
Trigger Module
102
Time Information Module
104
Network Backpressure Determination Module
106
Pairing Module
108

FIG. 1

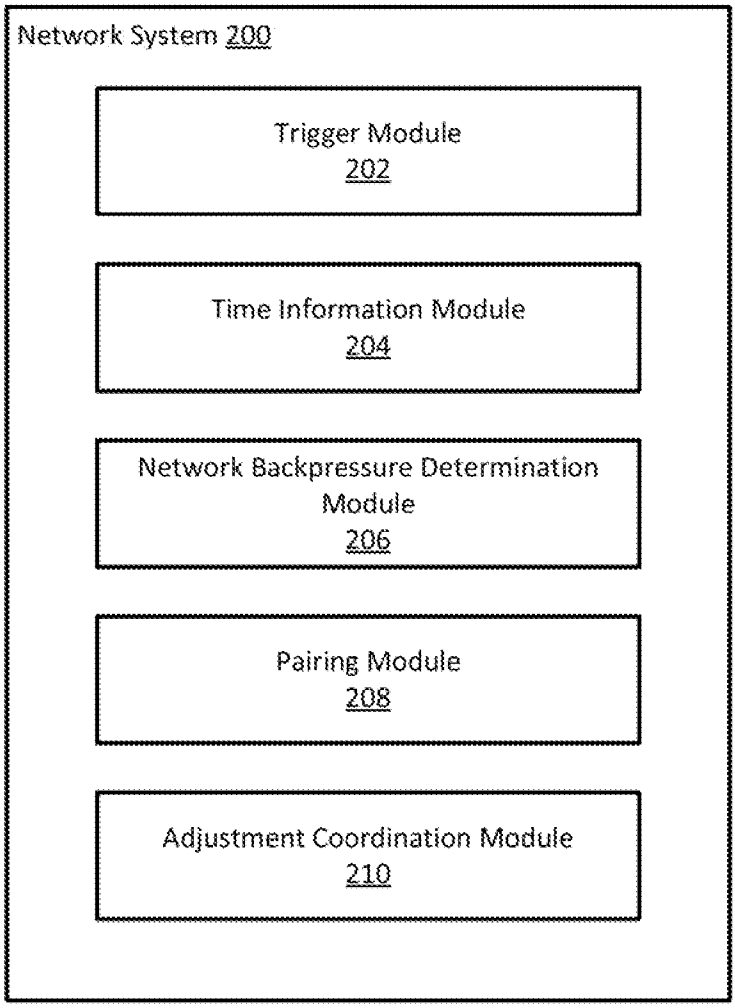
Network System 200
Trigger Module
202
Time Information Module
204
Network Backpressure Determination Module
206
Pairing Module
208
Adjustment Coordination Module
210

FIG. 2

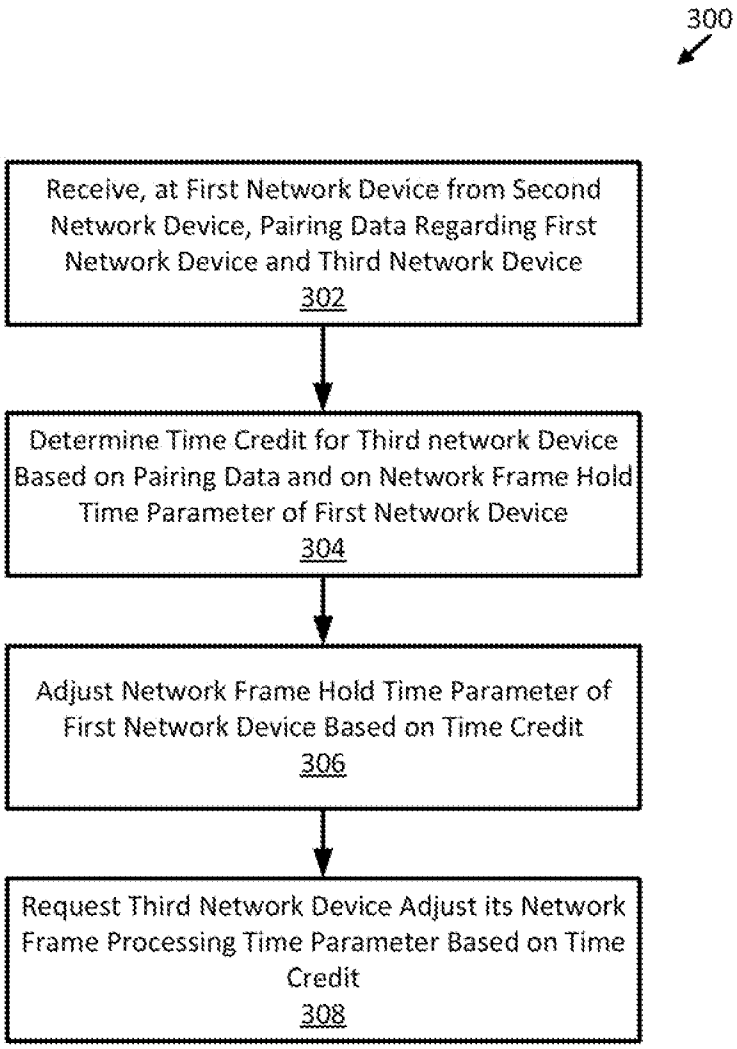
300
Receive, at First Network Device from Second Network Device, Pairing Data Regarding First Network Device and Third Network Device
302
Determine Time Credit for Third network Device Based on Pairing Data and on Network Frame Hold Time Parameter of First Network Device
304
Adjust Network Frame Hold Time Parameter of First Network Device Based on Time Credit
306
Request Third Network Device Adjust its Network Frame Processing Time Parameter Based on Time Credit
308

FIG. 3

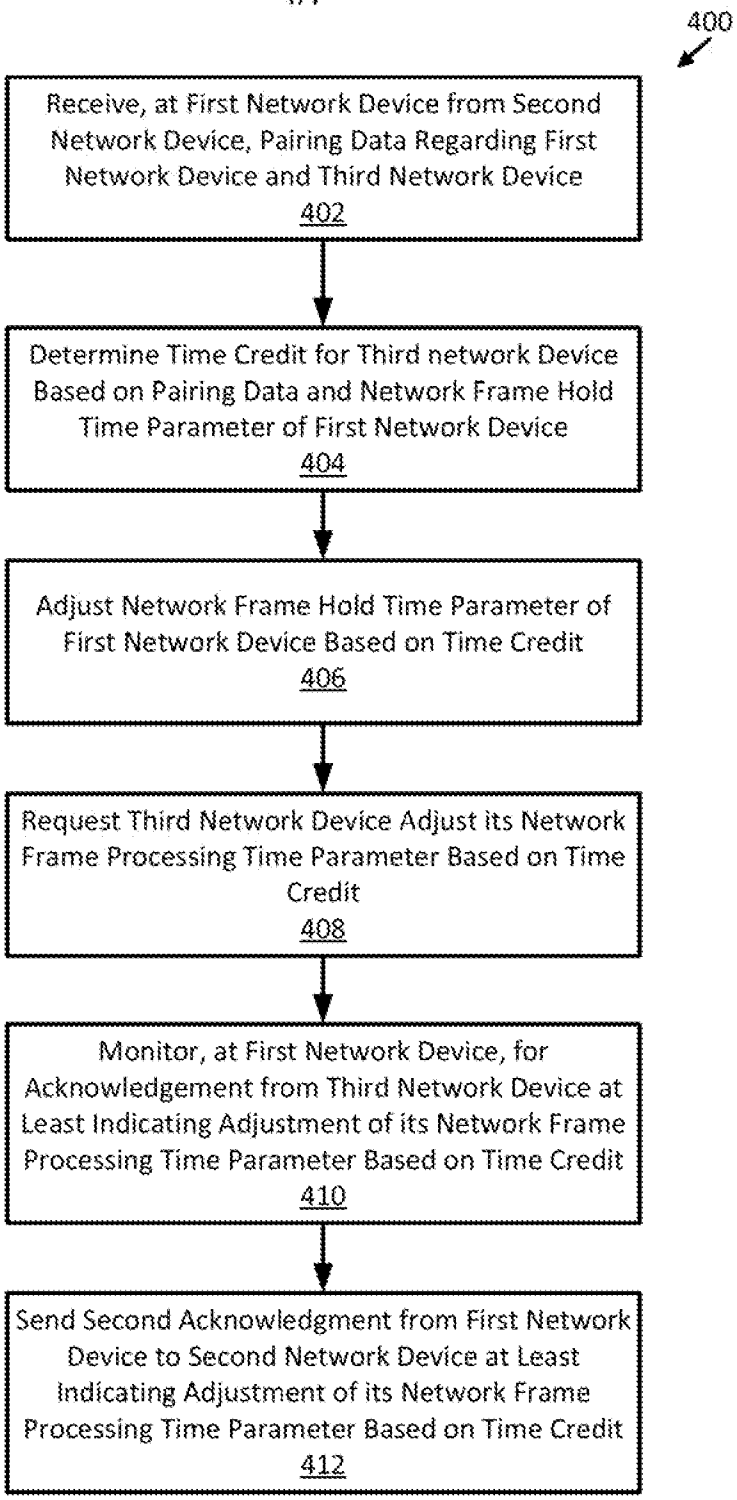
400
Receive, at First Network Device from Second Network Device, Pairing Data Regarding First Network Device and Third Network Device
402
Determine Time Credit for Third network Device Based on Pairing Data and Network Frame Hold Time Parameter of First Network Device
404
Adjust Network Frame Hold Time Parameter of First Network Device Based on Time Credit
406
Request Third Network Device Adjust its Network Frame Processing Time Parameter Based on Time Credit
408
Monitor, at First Network Device, for Acknowledgement from Third Network Device at Least Indicating Adjustment of its Network Frame Processing Time Parameter Based on Time Credit
410
Send Second Acknowledgment from First Network Device to Second Network Device at Least Indicating Adjustment of its Network Frame Processing Time Parameter Based on Time Credit
412

FIG. 4

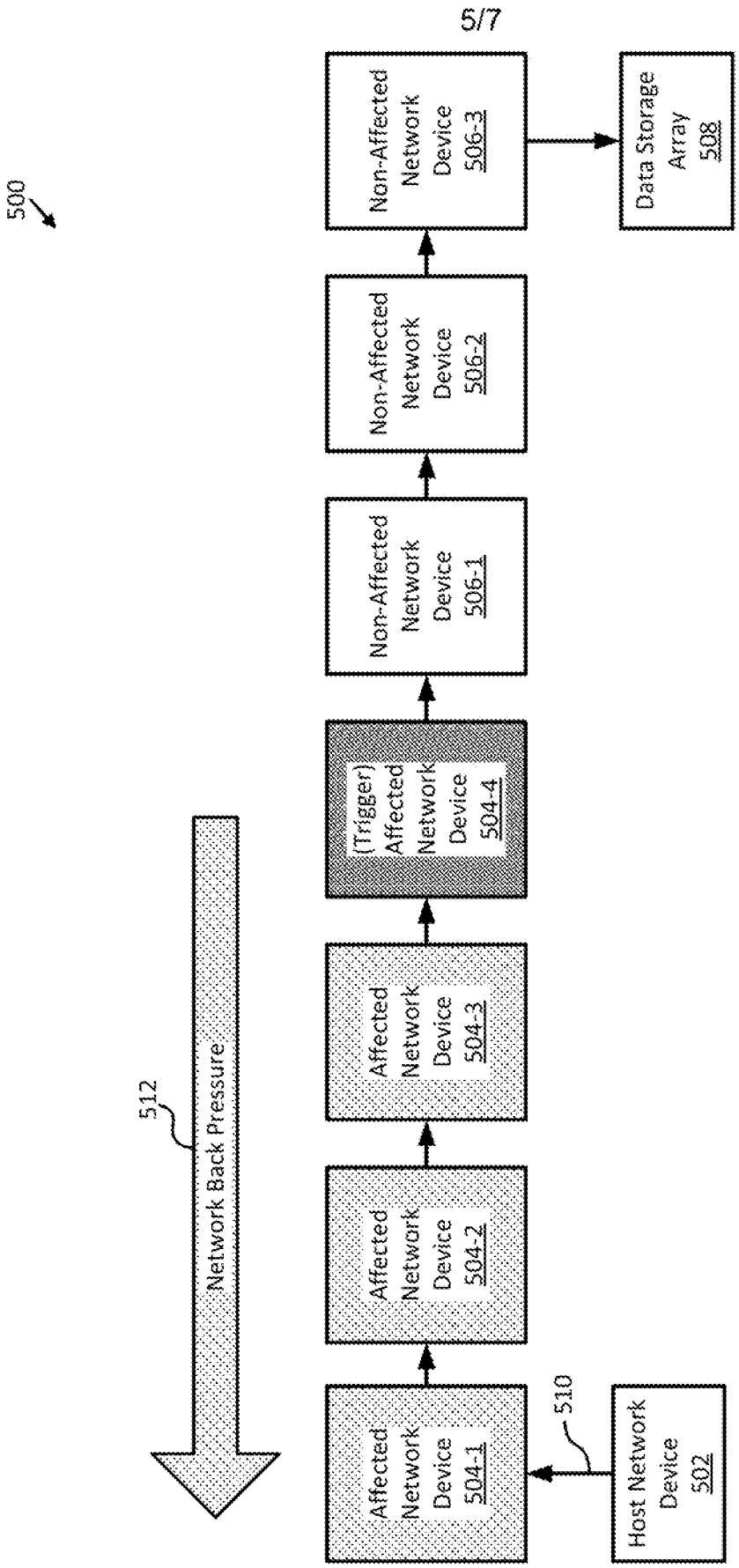
500
Host Network Device 502
510
Affected Network Device 504-1
Affected Network Device 504-2
Affected Network Device 504-3
(Trigger) Affected Network Device 504-4
Non-Affected Network Device 506-1
Non-Affected Network Device 506-2
Non-Affected Network Device 506-3
Data Storage Array 508
512
Network Back Pressure

FIG. 5

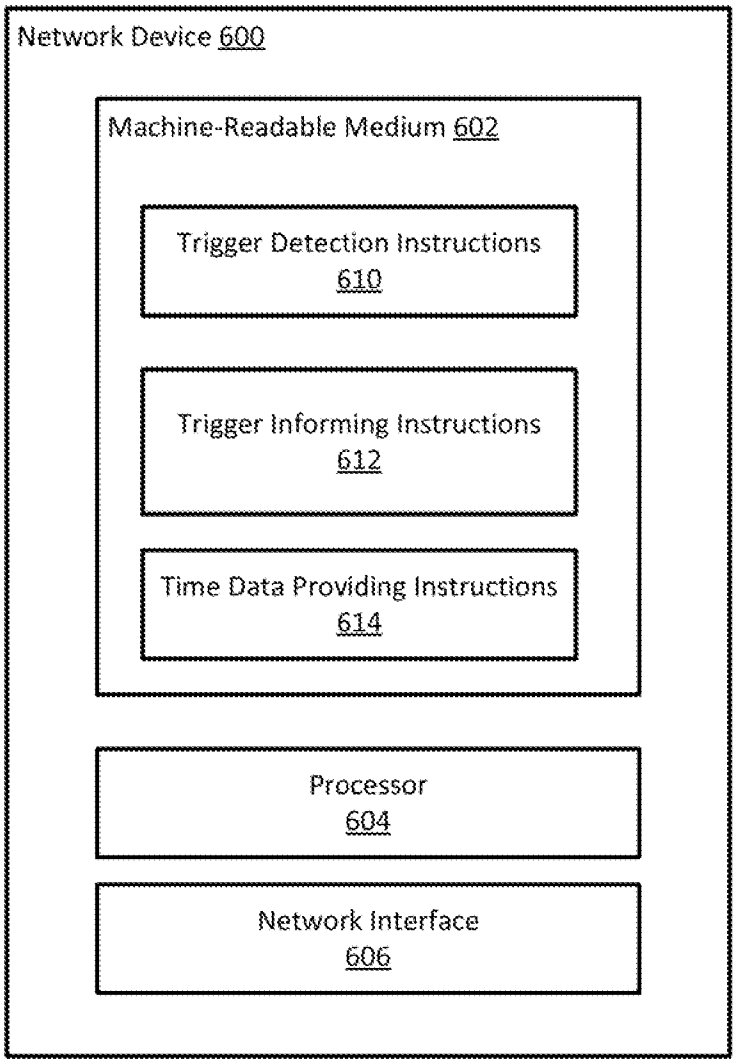
Network Device 600
Machine-Readable Medium 602
Trigger Detection Instructions 610
Trigger Informing Instructions 612
Time Data Providing Instructions 614
Processor 604
Network Interface 606

FIG. 6

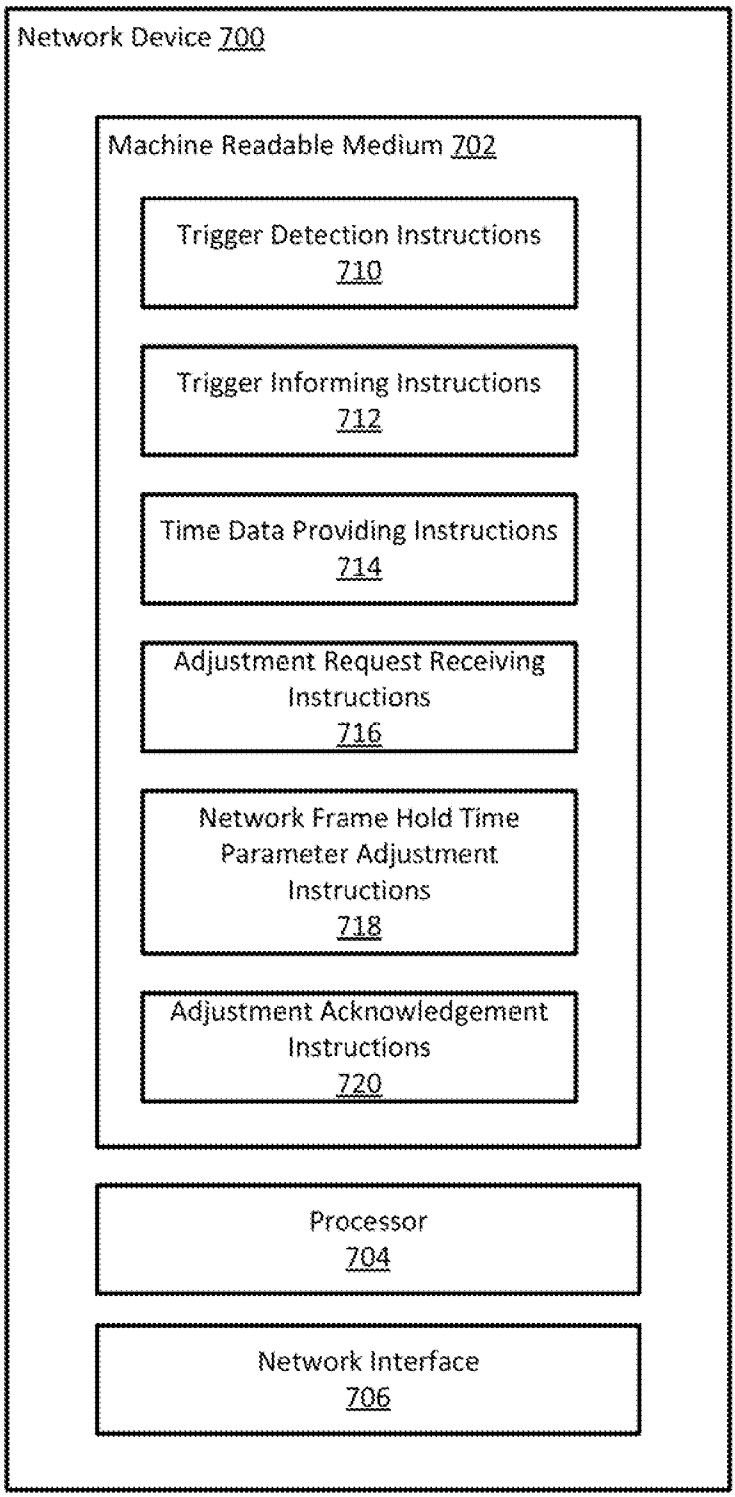
Network Device 700
Machine Readable Medium 702
Trigger Detection Instructions 710
Trigger Informing Instructions 712
Time Data Providing Instructions 714
Adjustment Request Receiving Instructions 716
Network Frame Hold Time Parameter Adjustment Instructions 718
Adjustment Acknowledgement Instructions 720
Processor 704
Network Interface 706

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2015/035333

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/841(2013.01)i, H04L 12/937(2013.01)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 12/841; G06F 15/16; H04L 1228; G06F 15/173; G06F 1/04; H04L 12/26; H04J 3/16; H04L 12/28; H04L 12/937

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: network, devices, frame, hold, time, parameter, paring, data, acknowledgement, back pressure

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2013-0286858 A1 (BROCADE COMMUNICATIONS SYSTEMS, INC.) 31 October 2013 See paragraphs [0025], [0035]-[0039], [0044], [0056], [0067], [0069]-[0070], [0079], [0090]-[0099], [0124]; and figures 4-5.	1-3,6-8,11-14
Y		4-5,9-10,15
Y	US 6449274 B1 (BRIAN D. HOLDEN et al.) 10 September 2002 See column 3, lines 46-48; column 10, lines 15-27; column 14, lines 46-55; and figure 1.	4-5,9-10,15
A	US 6647027 B1 (FRANK GASPARIK et al.) 11 November 2003 See column 6, line 22 - column 7, line 12; and figures 3-4D.	1-15
A	US 2005-0030978 A1 (FRANK R. DROPPS et al.) 10 February 2005 See paragraphs [0039]-[0041], [0174]-[0181]; and figure 6.	1-15
A	US 2005-0182856 A1 (CHARLES LEE MCKNETT) 18 August 2005 See paragraphs [0044]-[0059]; and figure 1.	1-15

☐ 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
14 March 2016 (14.03.2016)	**20 April 2016 (20.04.2016)**

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. +82-42-472-7140

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

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2015/035333

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2013-0286858 A1	31/10/2013	US 2012-063329 A1	15/03/2012
		US 8498213 B2	30/07/2013
		US 8767561 B2	01/07/2014
US 6449274 B1	10/09/2002	CA 2273208 A1	18/06/1998
		US 6188690 B1	13/02/2001
		US 6345050 B1	05/02/2002
		US 6396809 B1	28/05/2002
		US 6445705 B1	03/09/2002
		US 6724779 B1	20/04/2004
		WO 98-26539 A2	18/06/1998
		WO 98-26539 A3	08/10/1998
US 6647027 B1	11/11/2003	None	
US 2005-0030978 A1	10/02/2005	US 2005-0018604 A1	27/01/2005
		US 2005-0018621 A1	27/01/2005
		US 2005-0018649 A1	27/01/2005
		US 2005-0018650 A1	27/01/2005
		US 2005-0018663 A1	27/01/2005
		US 2005-0018672 A1	27/01/2005
		US 2005-0018673 A1	27/01/2005
		US 2005-0018674 A1	27/01/2005
		US 2005-0018675 A1	27/01/2005
		US 2005-0018676 A1	27/01/2005
		US 2005-0018680 A1	27/01/2005
		US 2005-0030954 A1	10/02/2005
		US 2005-0044267 A1	24/02/2005
		US 2008-0310306 A1	18/12/2008
		US 2009-0041029 A1	12/02/2009
		US 2009-0123150 A1	14/05/2009
		US 2009-0168772 A1	02/07/2009
		US 2009-0290584 A1	26/11/2009
		US 2009-0296715 A1	03/12/2009
		US 2009-0296716 A1	03/12/2009
		US 2009-0316592 A1	24/12/2009
		US 2010-0040074 A1	18/02/2010
		US 2010-0128607 A1	27/05/2010
		US 7406092 B2	29/07/2008
		US 7430175 B2	30/09/2008
		US 7466700 B2	16/12/2008
		US 7477655 B2	13/01/2009
		US 7525983 B2	28/04/2009
		US 7558281 B2	07/07/2009
		US 7573909 B2	11/08/2009
		US 7580354 B2	25/08/2009
		US 7583597 B2	01/09/2009
		US 7630384 B2	08/12/2009
		US 7646767 B2	12/01/2010

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2015/035333

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		US 7649903 B2	19/01/2010
		US 7684401 B2	23/03/2010
		US 7760752 B2	20/07/2010
		US 7792115 B2	07/09/2010
		US 7822061 B2	26/10/2010
		US 7894348 B2	22/02/2011
		US 8005105 B2	23/08/2011
		US 8072988 B2	06/12/2011
		US 8081650 B2	20/12/2011
		US 9118586 B2	25/08/2015
US 2005-0182856 A1	18/08/2005	AU 2004-311620 A1	21/07/2005
		CN 1918557 A	21/02/2007
		EP 1697853 A2	06/09/2006
		EP 1697853 A4	11/07/2007
		JP 2007-518159 A	05/07/2007
		WO 2005-065111 A2	21/07/2005
		WO 2005-065111 A3	19/01/2006