Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2704133 A1 2009/08/20

(21) 2 704 133

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2009/02/09
(87) Date publication PCT/PCT Publication Date: 2009/08/20
(85) Entrée phase nationale/National Entry: 2010/04/28
(86) N° demande PCT/PCT Application No.: EP 2009/051450
(87) N° publication PCT/PCT Publication No.: 2009/101053
(30) Priorité/Priority: 2008/02/14 (US12/030,954)

(51) Cl.Int./Int.Cl. *G06F 13/10* (2006.01)

(71) Demandeur/Applicant:
INTERNATIONAL BUSINESS MACHINES CORPORATION, US

(72) Inventeurs/Inventors:
CASPER, DANIEL, US;
FLANAGAN, JOHN, US;
KALOS, MATTHEW, US;
SITTMANN, GUSTAV, III, US;
HUANG, CATHERINE, US;
NJOKU, UGOCHUKWU, US;
RIEDY, DALE, US

(74) Agent: WANG, PETER

(54) Titre : TRANSFERT DE DONNEES BIDIRECTIONNEL EN UNE SEULE OPERATION D'E/S
(54) Title: BI-DIRECTIONAL DATA TRANSFER WITHIN A SINGLE I/O OPERATION

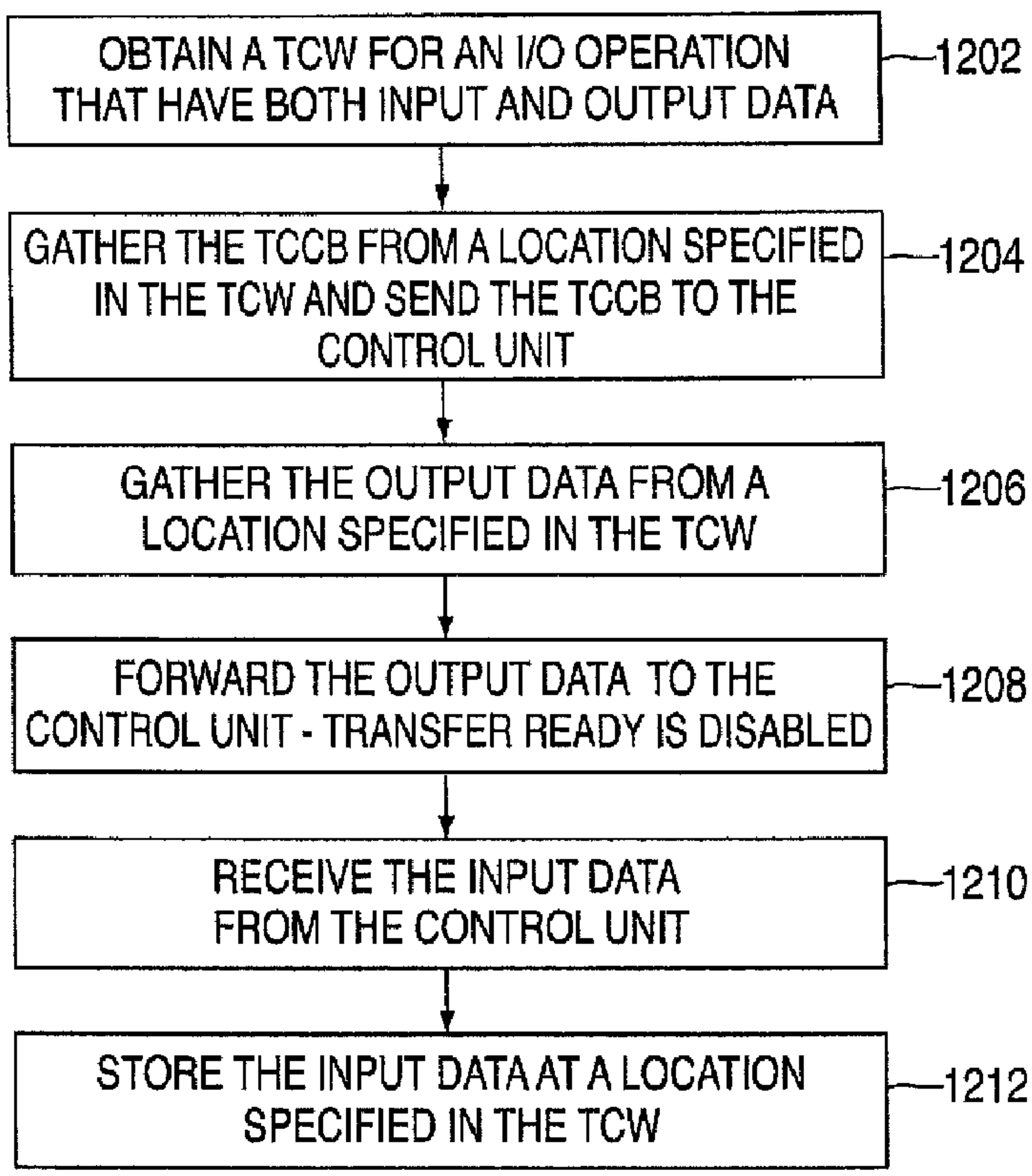

FIG. 12

(57) Abrégé/Abstract:
An article of manufacture, apparatus, and a method for facilitating input/output (I/O) pro-cessing for an I/O operation at a host computer system configured for communication with a control unit. The method includes the host computer system obtaining a

(57) Abrégé(suite)/Abstract(continued):
transport command word (TCW) for an I/O operation having both input and output data. The TCW specifies a location of the output data and a location for storing the input data. The host computer system forwards the I/O operation to the control unit for execution. The host computer system gathers the output data respon-sive to the location of the output data specified by the TCW, and then forwards the output data to the control unit for use in the execution of the I/O operation. The host computer system receives the input data from the control unit and stores the input data at the location specified by the TCW.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 August 2009 (20.08.2009)

PCT

(10) International Publication Number
WO 2009/101053 A1

(51) International Patent Classification:
G06F 13/10 (2006.01)

(21) International Application Number:
PCT/EP2009/051450

(22) International Filing Date:
9 February 2009 (09.02.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/030,954 14 February 2008 (14.02.2008) US

(71) Applicant *(for all designated States except US)*: **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, New York 10504 (US).

(71) Applicant *(for MG only)*: **IBM UNITED KINGDOM LIMITED** [GB/GB]; PO Box 41, North Harbour, Portsmouth Hampshire PO6 3AU (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **CASPER, Daniel** [US/US]; 13 Brett Place, Poughkeepsie, New York 12603 (US). **FLANAGAN, John** [US/US]; 15 Slate Hill Drive, Poughkeepsie, New York 12603 (US). **KALOS, Matthew** [US/US]; 5435 East Heatherwood Way, Tucson, Arizona 85718 (US). **SITTMANN III, Gustav** [US/US]; 422 South Park Avenue, Webster Groves, Missouri 63119 (US). **HUANG, Catherine** [US/US]; 6 Thornberry Way, Poughkeepsie, New York 12603 (US). **NJOKU, Ugochukwu** [US/US]; 426 Midland Avenue, Yonkers, New York 10704 (US). **RIEDY, Dale** [US/US]; 6 Oak Bend Road, Poughkeepsie, New York 12603 (US).

(74) Agent: **WILLIAMS, Julian, David**; IBM United Kingdom Limited, Intellectual Property Law, Hursley Park, Winchester Hampshire SO21 2JN (GB).

(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, BR, BW, BY, BZ, CA, CH, 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, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,

[Continued on next page]

(54) Title: BI-DIRECTIONAL DATA TRANSFER WITHIN A SINGLE I/O OPERATION

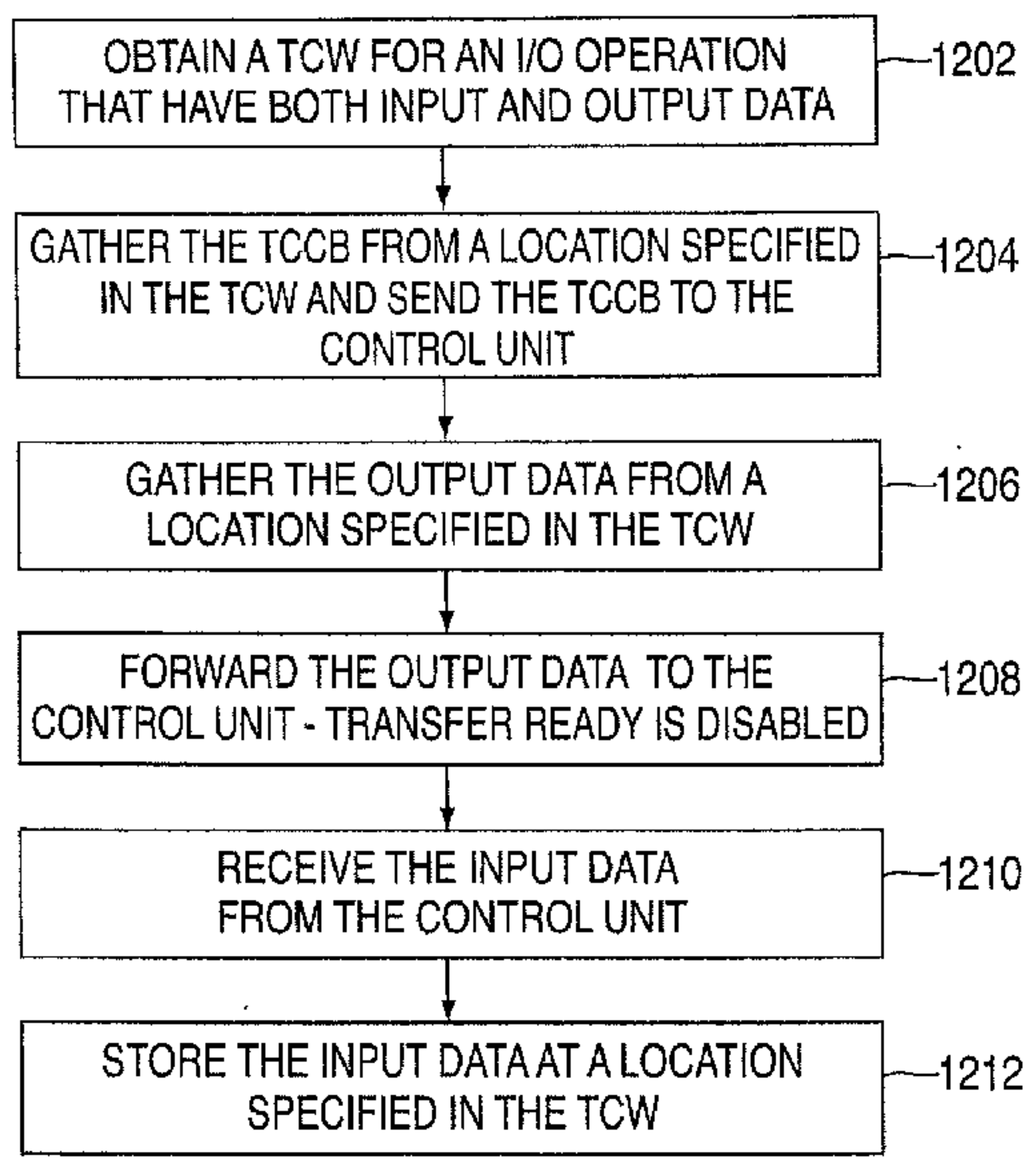

FIG. 12

(57) Abstract: An article of manufacture, apparatus, and a method for facilitating input/output (I/O) processing for an I/O operation at a host computer system configured for communication with a control unit. The method includes the host computer system obtaining a transport command word (TCW) for an I/O operation having both input and output data. The TCW specifies a location of the output data and a location for storing the input data. The host computer system forwards the I/O operation to the control unit for execution. The host computer system gathers the output data responsive to the location of the output data specified by the TCW, and then forwards the output data to the control unit for use in the execution of the I/O operation. The host computer system receives the input data from the control unit and stores the input data at the location specified by the TCW.

WO 2009/101053 A1

WO 2009/101053 A1

ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

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

BI-DIRECTIONAL DATA TRANSFER WITHIN A SINGLE I/O OPERATION

FIELD OF THE INVENTION

The present disclosure relates generally to input/output (I/O) processing, and in particular, to providing an I/O operation that includes both input and output data.

BACKGROUND OF THE INVENTION

Input/output (I/O) operations are used to transfer data between memory and I/O devices of an I/O processing system. Specifically, data is written from memory to one or more I/O devices, and data is read from one or more I/O devices to memory by executing I/O operations.

To facilitate processing of I/O operations, an I/O subsystem of the I/O processing system is employed. The I/O subsystem is coupled to main memory and the I/O devices of the I/O processing system and directs the flow of information between memory and the I/O devices. One example of an I/O subsystem is a channel subsystem. The channel subsystem uses channel paths as communications media. Each channel path includes a channel coupled to a control unit, the control unit being further coupled to one or more I/O devices.

The channel subsystem may employ channel command words (CCWs) to transfer data between the I/O devices and memory. A CCW specifies the command to be executed. For commands initiating certain I/O operations, the CCW designates the memory area associated with the operation, the action to be taken whenever a transfer to or from the area is completed, and other options.

During I/O processing, a list of CCWs is fetched from memory by a channel. The channel parses each command from the list of CCWs and forwards a number of the commands, each command in its own entity, to a control unit coupled to the channel. The control unit then processes the commands. The channel tracks the state of each command and controls when the next set of commands are to be sent to the control unit for processing. The channel ensures that each command is sent to the control unit in its own entity. Further, the channel infers certain information associated with processing the response from the control unit for each command.

Performing I/O processing on a per CCW basis may involve a large amount of processing overhead for the channel subsystem, as the channels parse CCWs, track state information, and react to responses from the control units. Therefore, it may be beneficial to shift much of the processing burden associated with interpreting and managing CCW and state information from the channel subsystem to the control units. Simplifying the role of channels in communicating between the control units and an operating system in the I/O processing system may increase communication throughput as less handshaking is performed. Simplifying the role of channels in communication may include grouping multiple commands into a single I/O operation. However, altering command sequences by grouping two or more commands together in a single I/O operation may result in the I/O operation having both input data and output data. Currently, an I/O operation can support a single data area that may be utilized for data input or data output, but not both within the same I/O operation. This limits the kinds of commands that can be grouped together in a single I/O operation and thus, limits the increase in throughput that can be gained by grouping commands. Accordingly, there is a need in the art to be able to transfer both input data and output data within a single I/O operation.

BRIEF SUMMARY OF THE INVENTION

An exemplary embodiment includes a computer program product for facilitating input/output (I/O) processing for an I/O operation at a host computer system configured for communication with a control unit. The computer program product includes a tangible storage medium readable by a processing circuit and storing instructions for execution by the processing circuit for performing a method. The method includes the host computer system obtaining a transport command word (TCW) for an I/O operation having both input and output data. The TCW specifies a location of the output data and a location for storing the input data. The host computer system forwards the I/O operation to the control unit for execution. The host computer system gathers the output data responsive to the location of the output data specified by the TCW, and then forwards the output data to the control unit for use in the execution of the I/O operation. The host computer system receives the input data from the control unit and stores the input data at the location specified by the TCW.

Another exemplary embodiment includes a host computer system for providing bi-directional data transfer within a single I/O operation. The host system includes a host computer I/O subsystem that includes a channel adapter that is in communication with a control unit. The host I/O subsystem performs a method that includes the host computer system obtaining a TCW for an I/O operation having both input and output data. The TCW specifies a location of the output data and a location for storing the input data. The output data is gathered in response to the location of the output data specified by the TCW. The I/O operation and the output data is forwarded to the control unit for execution. The host computer system receives the input data from the control unit and stores it at the location specified by the TCW.

A further exemplary embodiment includes a method for facilitating I/O processing for an I/O operation at a host computer system configured for communication with a control unit. The method includes the host computer system obtaining a TCW for an I/O operation having both input and output data. The TCW specifies a location of the output data and a location for storing the input data. The host computer system forwards the I/O operation to the control unit for execution. The host computer system gathers the output data responsive to the location of the output data specified by the TCW, and then forwards the output data to the control unit for use in the execution of the I/O operation. The host computer system receives the input data from the control unit and stores the input data at the location specified by the TCW.

Other articles of manufacture, apparatuses, and/or methods according to embodiments will be or become apparent to one with skill in the art upon review of the following drawings and detailed description. It is intended that all such additional articles of manufacture, apparatuses, and/or methods be included within this description, be within the scope of the present invention, and be protected by the accompanying claims.

BRIEF DESCRIPTION OF THE DRAWINGS

The subject matter which is regarded as the invention is particularly pointed out and distinctly claimed in the claims at the conclusion of the specification. The foregoing and other objects, features, and advantages of the invention are apparent from the following detailed description taken in conjunction with the accompanying drawings in which:

FIG. 1 depicts one embodiment of an I/O processing system incorporating and using one or more aspects of the present invention;

FIG. 2A depicts one example of a prior art channel command word;

FIG. 2B depicts one example of a prior art channel command word channel program;

FIG. 3 depicts one embodiment of a prior art link protocol used in communicating between a channel and control unit to execute the channel command word channel program of FIG. 2B;

FIG. 4 depicts one embodiment of a transport control word (TCW) channel program, in accordance with an aspect of the present invention;

FIG. 5 depicts one embodiment of a link protocol used to communicate between a channel and control unit to execute the TCW channel program of FIG. 4, in accordance with an aspect of the present invention;

FIG. 6 depicts one embodiment of a prior art link protocol used to communicate between a channel and control unit in order to execute four read commands of a channel command word channel program;

FIG. 7 depicts one embodiment of a link protocol used to communicate between a channel and control unit to process the four read commands of a TCW channel program, in accordance with an aspect of the present invention;

FIG. 8 depicts one embodiment of a control unit and a channel subsystem, in accordance with an aspect of the present invention;

FIG. 9 depicts one embodiment of a TCW in accordance with an aspect of the present invention;

FIG. 10 depicts one embodiment of a TCW channel program, in accordance with an aspect of the present invention;

FIG. 11 depicts one embodiment of a link protocol used to communicate between a channel and control unit to execute the TCW channel program of FIG. 10, in accordance with an aspect of the present invention;

FIG. 12 depicts one embodiment of a process for bi-directional data transfer within a single I/O operation, in accordance with an aspect of the present invention; and

FIG. 13 depicts one embodiment of an article of manufacture incorporating one or more aspects of the present invention.

The detailed description explains the preferred embodiments of the invention, together with advantages and features, by way of example with reference to the drawings.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with an aspect of the present invention, input/output (I/O) processing is facilitated by allowing a single I/O operation to include both input data and output data. Thus, each I/O operation can be utilized to transfer both an input stream and an output stream. This facilitates I/O processing by reducing communications between components of an I/O processing system used to perform the I/O processing. For instance, the number of exchanges and sequences between an I/O communications adapter, such as a channel, and a control unit is reduced. This is accomplished by sending a plurality of commands from the I/O communications adapter to the control unit as a single entity for execution by the control unit, and by the control unit sending the data resulting from the commands, if any, as a single entity. The plurality of device command words (DCWs) sent as a single entity to a control unit may include both read and write commands.

The plurality of commands are included in a block, referred to herein as a transport command control block (TCCB), an address of which is specified in a transport control word (TCW). The TCW is sent from an operating system (OS) or other application to the I/O communications adapter, which in turn forwards the TCCB in a command message to the control unit for processing. The control unit processes each of the commands absent a tracking of status relative to those individual commands by the I/O communications adapter. The plurality of commands is also referred to as a channel program, which is parsed and executed on the control unit rather than the I/O communications adapter.

In an exemplary embodiment, the TCW provides the pointers to the channel for all of the control blocks required to execute the I/O operation. In an exemplary embodiment, the TCW includes pointers to both an input data address and an output data address. This allows data to be transferred in both directions (e.g., from a channel to a control unit and from a control unit to a channel) within a single I/O operation.

One example of an I/O processing system incorporating and using one or more aspects of the present invention is described with reference to FIG. 1. I/O processing system 100 includes a host system 101, which further includes for instance, a main memory 102, one or more central processing units (CPUs) 104, a storage control element 106, and a channel subsystem 108. The host system 101 may be a large scale computing system, such as a mainframe or server. The I/O processing system 100 also includes one or more control units 110 and one or more I/O devices 112, each of which is described below.

Main memory 102 stores data and programs, which can be input from I/O devices 112. For example, the main memory 102 may include one or more operating systems (OSs) 103 that are executed by one or more of the CPUs 104. For example, one CPU 104 can execute a Linux® operating system 103 and a z/OS® operating system 103 as different virtual machine instances. The main memory 102 is directly addressable and provides for high-speed processing of data by the CPUs 104 and the channel subsystem 108.

CPU 104 is the controlling center of the I/O processing system 100. It contains sequencing and processing facilities for instruction execution, interruption action, timing functions, initial program loading, and other machine-related functions. CPU 104 is coupled to the storage control element 106 via a connection 114, such as a bidirectional or unidirectional bus.

Storage control element 106 is coupled to the main memory 102 via a connection 116, such as a bus; to CPUs 104 via connection 114; and to channel subsystem 108 via a connection

118. Storage control element 106 controls, for example, queuing and execution of requests made by one or more of the CPU 104 and the channel subsystem 108.

In an exemplary embodiment, channel subsystem 108 provides a communication interface between host system 101 and control units 110. Channel subsystem 108 is coupled to storage control element 106, as described above, and to each of the control units 110 via a connection 120, such as a serial link. Connection 120 may be implemented in any manner known in the art, including an optical link, employing single-mode or multi-mode waveguides in a Fibre Channel fabric (e.g., a fibre channel network). Channel subsystem 108 directs the flow of information between I/O devices 112 and main memory 102. It relieves the CPUs 104 of the task of communicating directly with the I/O devices 112 and permits data processing to proceed concurrently with I/O processing. The channel subsystem 108 uses one or more channel paths 122 as the communication links in managing the flow of information to or from I/O devices 112. As a part of the I/O processing, channel subsystem 108 also performs the path-management functions of testing for channel path availability, selecting an available channel path 122 and initiating execution of the operation with the I/O devices 112.

Each channel path 122 includes a channel 124 (channels 124 are located within the channel subsystem 108, in one example, as shown in FIG. 1), one or more control units 110 and one or more connections 120. In another example, it is also possible to have one or more dynamic switches (not depicted) as part of the channel path 122. A dynamic switch may be coupled to a channel 124 and a control unit 110 and provides the capability of physically interconnecting any two links that are attached to the switch. In another example, it is also possible to have multiple systems, and therefore multiple channel subsystems (not depicted) attached to one or more of the control units 110.

Also located within channel subsystem 108 are subchannels (not shown). One subchannel is provided for and dedicated to each I/O device 112 accessible to a program through the channel subsystem 108. A subchannel (e.g., a data structure, such as a table) provides the logical appearance of a device to the program. Each subchannel provides information concerning the associated I/O device 112 and its attachment to channel subsystem 108. The subchannel also provides information concerning I/O operations and other functions involving the associated I/O device 112. The subchannel is the means by which channel subsystem 108 provides information about associated I/O devices 112 to CPUs 104, which obtain this information by executing I/O instructions.

Channel subsystem 108 is coupled to one or more control units 110. Each control unit 110 provides logic to operate and control one or more I/O devices 112 and adapts, through the use of common facilities, the characteristics of each I/O device 112 to the link interface provided by the channel 124. The common facilities provide for the execution of I/O operations, indications concerning the status of the I/O device 112 and control unit 110, control of the timing of data transfers over the channel path 122 and certain levels of I/O device 112 control.

Each control unit 110 is attached via a connection 126 (e.g., a bus) to one or more I/O devices 112. I/O devices 112 receive information or store information in main memory 102 and/or other memory. Examples of I/O devices 112 include card readers and punches, magnetic tape units, direct access storage devices, displays, keyboards, printers, pointing devices, teleprocessing devices, communication controllers and sensor based equipment, to name a few.

One or more of the above components of the I/O processing system 100 are further described in “IBM® z/Architecture Principles of Operation,” Publication No. SA22-7832-05, 6th Edition, April 2007; U.S. Patent No. 5,461,721 entitled “System For Transferring Data Between I/O Devices And Main Or Expanded Storage Under Dynamic Control Of Independent Indirect Address Words (IDAWS),” Cormier et al., issued October 24, 1995; and U.S. Patent No. 5,526,484 entitled “Method And System For Pipelining The Processing Of Channel Command Words,” Casper et al., issued June 11, 1996, each of which is hereby incorporated herein by reference in its entirety. IBM is a registered trademark of International Business Machines Corporation, Armonk, New York, USA. Other names used herein may be registered trademarks, trademarks or product names of International Business Machines Corporation or other companies.

In one embodiment, to transfer data between I/O devices 112 and memory 102, channel command words (CCWs) are used. A CCW specifies the command to be executed, and includes other fields to control processing. One example of a CCW is described with reference to FIG. 2A. A CCW 200 includes, for example, a command code 202 specifying the command to be executed (e.g., read, read backward, control, sense and write); a plurality of flags 204 used to control the I/O operation; for commands that specify the transfer of data, a count field 206 that specifies the number of bytes in the storage area designated by the CCW to be transferred; and a data address 208 that points to a location in main memory that includes the data, when direct addressing is employed, or to a list (e.g., contiguous list) of modified indirect data address words (MIDAWs) to be processed, when modified indirect data addressing is employed. Modified indirect addressing is further described in U.S. Application Serial Number 11/464,613, entitled "Flexibly Controlling The Transfer Of Data Between Input/Output Devices And Memory," Brice et al., filed August 15, 2006, which is hereby incorporated herein by reference in its entirety.

One or more CCWs arranged for sequential execution form a channel program, also referred to herein as a CCW channel program. The CCW channel program is set up by, for example, an operating system, or other software. The software sets up the CCWs and obtains the addresses of memory assigned to the channel program. An example of a CCW channel program is described with reference to FIG. 2B. A CCW channel program 210 includes, for instance, a define extent CCW 212 that has a pointer 214 to a location in memory of define extent data 216 to be used with the define extent command. In this example, a transfer in channel (TIC) 218 follows the define extent command that refers the channel program to another area in memory (e.g., an application area) that includes one or more other CCWs, such as a locate record 217 that has a pointer 219 to locate record data 220, and one or more read CCWs 221. Each read CCW 220 has a pointer 222 to a data area 224. The data area includes an address to directly access the data or a list of data address words (e.g., MIDAWs or IDAWs) to indirectly access the data. Further, CCW channel program 210 includes a predetermined area in the channel subsystem defined by the device address called the subchannel for status 226 resulting from execution of the CCW channel program.

The processing of a CCW channel program is described with reference to FIG. 3, as well as with reference to FIG. 2B. In particular, FIG. 3 shows an example of the various exchanges and sequences that occur between a channel and a control unit when a CCW channel program is executing. The link protocol used for the communications is FICON (Fibre Connectivity), in this example. Information regarding FICON is described in "Fibre Channel Single Byte Command Code Sets-3 Mapping Protocol" (FC-SB-3), T11/Project

1357-D/Rev. 1.6, INCITS (March 2003), which is hereby incorporated herein by reference in its entirety.

Referring to FIG. 3, a channel 300 opens an exchange with a control unit 302 and sends a define extent command and data associated therewith 304 to control unit 302. The command is fetched from define extent CCW 212 (FIG. 2B) and the data is obtained from define extent data area 216. The channel 300 uses TIC 218 to locate the locate record CCW and the read CCW. It fetches the locate record command 305 (FIG. 3) from the locate record CCW 217 (FIG. 2B) and obtains the data from locate record data 220. The read command 306 (FIG. 3) is fetched from read CCW 221 (FIG. 2B). Each is sent to the control unit 302.

The control unit 302 opens an exchange 308 with the channel 300, in response to the open exchange of the channel 300. This can occur before or after locate command 305 and/or read command 306. Along with the open exchange, a response (CMR) is forwarded to the channel 300. The CMR provides an indication to the channel 300 that the control unit 302 is active and operating.

The control unit 302 sends the requested data 310 to the channel 300. Additionally, the control unit 302 provides the status to the channel 300 and closes the exchange 312. In response thereto, the channel 300 stores the data, examines the status and closes the exchange 314, which indicates to the control unit 302 that the status has been received.

The processing of the above CCW channel program to read 4k of data requires two exchanges to be opened and closed and seven sequences. The total number of exchanges and sequences between the channel and control unit is reduced through collapsing multiple commands of the channel program into a TCCB. The channel, e.g., channel 124 of FIG. 1, uses a TCW to identify the location of the TCCB, as well as locations for accessing and storing status and data associated with executing the channel program. The TCW is interpreted by the channel 124 and is not sent or seen by the control unit 110.

One example of a channel program to read 4k of data, as in FIG. 2B, but includes a TCCB, instead of separate individual CCWs, is described with reference to FIG. 4. As shown, a channel program 400, referred to herein as a TCW channel program, includes a TCW 402 specifying a location in memory of a TCCB 404, as well as a location in memory of a data area 406 or a TIDAL 410 (i.e., a list of transport mode indirect data address words (TIDAWs), similar to MIDAWs) that points to data area 406, and a status area 408.

The processing of a TCW channel program is described with reference to FIG. 5. The link protocol used for these communications is, for instance, Fibre Channel Protocol (FCP). In particular, three phases of the FCP link protocol are used, allowing host bus adapters to be used that support FCP to perform data transfers controlled by CCWs. FCP and its phases are described further in "Information Technology – Fibre Channel Protocol for SCSI, Third Version (FCP-3)," T10 Project 1560-D, Revision 4, September 13, 2005, which is hereby incorporated herein by reference in its entirety.

Referring to FIG. 5, a channel 500 opens an exchange with a control unit 502 and sends TCCB 504 to the control unit 502. In one example, the TCCB 504 and sequence initiative are transferred to the control unit 502 in a FCP command, referred to as FCP_CMND information unit (IU) or a transport command IU. The control unit 502 executes the multiple commands of the TCCB 504 (e.g., define extent command, locate record command, read command as device control words (DCWs)) and forwards data 506 to the channel 500 via, for instance, a FCP_Data IU. It also provides status and closes the exchange 508. As one example, final status is sent in a FCP status frame that has a bit active in, for instance, byte 10 or 11 of the payload of a FCP_RSP IU, also referred to as a transport response IU. The FCP_RSP_IU payload may be used to transport FICON ending status along with additional status information.

In a further example, to write 4k of customer data, the channel 500 uses the FCP link protocol phases, as follows:

1. Transfer a TCCB in the FCP_CMND IU.

2. Transfer the IU of data, and sequence initiative to the control unit 502. (FCP Transfer Ready Disabled)

3. Final status is sent in a FCP status frame that has a bit active in, for instance, byte 10 or 11 of the FCP_RSP IU Payload. The FCP_RES_INFO field or sense field is used to transport FICON ending status along with additional status information.

By executing the TCW channel program of FIG. 4, there is only one exchange opened and closed (see also FIG. 5), instead of two exchanges for the CCW channel program of FIG. 2B (see also FIG. 3). Further, for the TCW channel program, there are three communication sequences (see FIGs. 4-5), as compared to seven sequences for the CCW channel program (see FIGs. 2B-3).

The number of exchanges and sequences remain the same for a TCW channel program, even if additional commands are added to the program. Compare, for example, the communications of the CCW channel program of FIG. 6 with the communications of the TCW channel program of FIG. 7. In the CCW channel program of FIG. 6, each of the commands (e.g., define extent command 600, locate record command 601, read command 602, read command 604, read command 606, locate record command 607 and read command 608) are sent in separate sequences from channel 610 to control unit 612. Further, each 4k block of data (e.g., data 614-620) is sent in separate sequences from the control unit 612 to the channel 610. This CCW channel program requires two exchanges to be opened and closed (e.g., open exchanges 622, 624 and close exchanges 626, 628), and fourteen communications sequences. This is compared to the three sequences and one exchange for the TCW channel program of FIG. 7, which accomplishes the same task as the CCW channel program of FIG. 6.

As depicted in FIG. 7, a channel 700 opens an exchange with a control unit 702 and sends a TCCB 704 to the control unit 702. The TCCB 704 includes the define extent command, the two locate record commands, and the four read commands in DCWs, as described above. In response to receiving the TCCB 704, the control unit 702 executes the commands and sends, in a single sequence, the 16k of data 706 to the channel 700. Additionally, the control unit 702 provides status to the channel 700 and closes the exchange 708. Thus, the TCW channel program requires much fewer communications to transfer the same amount of data as the CCW channel program of FIG. 6.

Turning now to FIG. 8, one embodiment of the channel 124 in the channel subsystem 108 and the control unit 110 and the channel 124 of FIG. 1 that support TCW channel program execution are depicted in greater detail. The control unit 110 includes CU control logic 802 to parse and process command messages containing a TCCB, such as the TCCB 704 of FIG. 7, received from the channel 124 via the connection 120. The CU control logic 802 can extract DCWs and control data from the TCCB received at the control unit 110 to control a device, for instance, I/O device 112 via connection 126. The CU control logic 802 sends device commands and data to the I/O device 112, and receives status information and other feedback from the I/O device 112. For example, the I/O device 112 may be busy because of a previous reservation request targeting I/O device 112. To manage potential device reservation contention issues that can arise when the control unit 110 receives multiple requests to access the same I/O device 112, the CU control logic 802 keeps track of and stores device busy messages and associated data in a device busy queue 804. In an exemplary embodiment, an OS 103 of FIG. 1 reserves I/O device 112 to keep other OSs 103 from accessing the I/O device 112 while the reservation is active. Although device reservation is not required for all I/O operations, device reservation can be used to support operations that necessitate exclusive access for a fixed duration of time, e.g., disk formatting.

The control unit 110 may further include other buffer or memory elements (not depicted) to store multiple messages or status information associated with communications between the channel 124 and the I/O device 112. For example, a register located on the control unit 110 may include a maximum control unit exchange parameter that defines the maximum number of open control unit exchanges that the control unit 110 supports.

The channel 124 in the channel subsystem 108 includes multiple elements to support communication with the control unit 110. In an exemplary embodiment, the CHN control logic 806 controls communication between the channel subsystem 108 and the control unit 110. The CHN control logic 806 may directly interface to the CU control logic 802 via the connection 120 to send commands and receive responses, such as transport command and response IUs. Alternatively, messaging interfaces and/or buffers (not depicted) can be placed between the CHN control logic 806 and the CU control logic 802.

An exemplary embodiment of a transport control word (TCW) 900 is depicted in FIG. 9. The TCW 900 is utilized by the channel 124 to set up the I/O operation and is not sent to the control unit 110. The TCW depicted in FIG. 9 provides for both input and output data within a single I/O operation.

In an exemplary TCW 900 depicted in FIG. 9, a format field 902 equal to “00b” indicates that what follows is a TCW 900. The TCW 900 also includes reserved bits 904 for possible future use.

The TCW 900 also includes a flags field 906. The first five bits of the flags field 906 are reserved for future use and are set to zero. The sixth bit of the flags field 906 is a TIDAL read flag. In an exemplary embodiment, the TIDAL read flag is set to one when the input-data address field 918 contains an address of a TIDAL. If the TIDAL read flag is set to zero, then the input-data address field 918 contains a data address. The seventh bit of the flags field 906 is a TCCB TIDAL flag. In an exemplary embodiment, the TCCB TIDAL flag is set to one when the TCCB address field 922 contains an address of a TIDAL. If the TCCB TIDAL flag is set to zero, then the TCCB address field 922 directly addresses the TCCB. The TCCB TIDAL flag allows the operating system software or the hyper-visor to layer function and prefix user channel programs. The eighth bit of the flags field 906 is a TIDAL write flag. In an exemplary embodiment, the TIDAL write flag is set to one when the output-data address field 916 contains an address of a TIDAL. If the TIDAL write flag is set to zero, then the output-data address field 916 contains a data address.

The ninth through twenty-forth bits of the flags field 906 are reserved for future use.

The TCW 900 also includes a TCCB length field 910 which indirectly represents the length of the TCCB and may be utilized to determine the actual length of the TCCB.

The read/write bits 912 in the TCW 900 are utilized to indicate whether data is being read and/or written as a result of executing the TCW 900. In an exemplary embodiment, the read bit in the read/write 912 bits is set to one to indicate that input data is being transferred from an I/O device 112 to system storage (e.g., main memory 102) in the host system 101 as a result of executing the TCW 900. The write bit in the read/write bits 912 is set to one to indicate that output data is being transferred from system storage (e.g., main memory 102) in the host system 101 to an I/O device as a result of executing the TCW 900.

The output-data address field 916 includes the address for the output data (if any). As described previously, the contents of the output-data address field 916 may be an address of a TIDAL for output data (e.g., an indirect address) or the actual address of the output data (e.g., a direct address). The input-data address field 918 includes the address for the input data (if any). As described previously, the contents of the input-data address field 918 may be an address of a TIDAL for input data or the actual address of the input data. In an exemplary embodiment, the output-data address field 916 and the input data address field 918 are implemented as sixty-four bit addresses.

The TCW 900 also includes a transport-status-block address field 920. A portion (e.g., the extended status part) of a completion status in a transport response IU for an I/O operation is stored at this address. The TCCB address field 922 in the TCW 900 includes an address where the TCCB is located in system storage. As described previously, the TCCB is the control block where the DCWs to be executed for the TCW 900 reside. Also as described previously, the contents of the TCCB address field 922 may be an address of a TIDAL for the TCCB or the actual address of the TCCB. In an exemplary embodiment, the transport-status-block address field 920 and the TCCB address field 922 are implemented as sixty-four bit addresses.

The output count field 924 in the TCW 900 indicates the amount of output data to be transferred by the TCW/TCCB for an output operation. In an exemplary embodiment, the output count field 924 specifies the number of bytes in the output storage area designed by the TCW (the output-data address 916) to be transferred. The input count field 926 in the TCW 900 indicates the amount of input data to be transferred by the TCW/TCCB for an input operation. In an exemplary embodiment, the input count field 926 specifies the number of bytes in the output storage area designed by the TCW (the input-data address 918) to be transferred. Several additional fields in the TCW 900 are reserved: reserved field 928, reserved field 930 and reserved field 932. The interrogate-TCW address field 934 contains the address of another TCW and is used by the channel 124 to interrogate that state of an operation under the initiative of a cancel sub-channel I/O instruction.

The TCW depicted in FIG. 9 is one example of how a command word can be configured. Other configurations are possible where additional fields are included and/or fields depicted in FIG. 9 are not included.

FIG. 10 depicts one embodiment of a TCW channel program, in accordance with an aspect of the present invention when both input and output data is included in a single I/O operation. As shown in FIG. 10, a TCW channel program 1000 includes a TCW 1002 specifying a location in memory of a TCCB 1004, a location in memory for storing input data 1006 or a TIDAL 1010 (i.e., a list of transport mode indirect data address words (TIDAWs) that points to the location for the input data 1006, a location in memory of an output data area 1014 or a TIDAL 1012 that points to the output data area 1006, and a status area 1008.

The processing of the TCW channel program 1000 depicted in FIG. 10 is described with reference to FIG. 11. Referring to FIG. 11, a channel 1100 opens an exchange with a control unit 1102 and sends a TCCB 1104 and output data 1105 located at the output data area 1014 specified by the TCW 1002 to the control unit 1102. The channel 1100 determines how much data to send based on the value of the output count 924 in the TCW 1002. The control unit 1102 executes the multiple commands of the TCCB 1104 (e.g., define extent command, locate record command, write command and read command as device control words (DCWs)) receives the output data 1105 from the channel 1100 and forwards input data 1106 per the data count in the DCW to the channel 1100 via, for instance, a FCP_Data IU. The channel 1100 stores the input data 1106 at the location specified by the TCW 1002. The control unit 1102 also provides status and closes the exchange 1108. In this manner, data is input to the channel 110 and output to the control unit 1102 in a single TCW channel program 1000 (or I/O operation).

FIG. 12 depicts one embodiment of a process for bi-directional data transfer within a single I/O operation, in accordance with an aspect of the present invention. In an exemplary embodiment, the processing depicted in FIG. 12 occurs at a host computer system that is in network communication with a control unit. The host computer system may include an I/O processing system that executes the process. Additionally, the I/O processing system may include a channel subsystem that executes the process. At block 1202, a TCW is obtained by the host computer. In an exemplary embodiment, the TCW is obtained (or received) from an operating system running on the host computer. The TCW includes both an output data address 916 and output count field 924, and an input data address 918 and input data count field 926. In an exemplary embodiment, the TCW includes output data when the write bit in the read/write bits 912 is set to one and the TCW includes input data when the read bit in the read/write bits 912 is set to one. At block 1204 the TCCB location specified by the TCW

922 is fetched and forwarded to the control unit. The TCCB contains the DCWs that inform the control unit what I/O operations to execute.

At block 1206, the output data is gathered from the location specified by the TCW (if the write bit in the read/write bits 912 is set to one). The amount of data gathered to be included in the output data is based on the value of the output data count field 924. As described previously, the output data address may be a direct address of the output data or an indirect address of the output data. An indirect address refers to an address containing a list of one or more addresses (e.g., a TIDAL) that point to a plurality of storage locations that collectively make up the output data. A direct address refers to an address containing the output data. In an exemplary embodiment, the TIDAL write flag in the flags field 906 in the TCW is set to one when the output-data address field 916 contains an address of a TIDAL, and set to zero when the output-data address field 916 contains the address of the output data.

At block 1208, the output data is forwarded to the control unit. For this example XFER_RDY is disabled.

At block 1210, input data is received from the control unit as a result of executing the I/O operation. At block 1212, the input data is stored at the location specified by the TCW (the input-data address 918). In an exemplary embodiment, the TCW includes input data when the read bit in the read/write bits 912 is set to one. As described previously, the input data-address 918 may be a direct address for storing the input data, or alternatively it may be an address to a list of addresses (e.g., a TIDAL or indirect address) that point to a plurality of storage locations, each storing portions of the input data. In an exemplary embodiment, the TIDAL read flag in the flags field 906 in the TCW is set to one when the input-data address field 918 contains an address of a TIDAL, and set to zero when the input-data address field 918 contains the address of the input data.

Technical effects of exemplary embodiments include the ability to include both input and output data in a single I/O operation. This provides flexibility in grouping DCWs and may lead to a decrease in the number of exchanges required between a channel and a control unit.

As described above, embodiments can be embodied in the form of computer-implemented processes and apparatuses for practicing those processes. In exemplary embodiments, the invention is embodied in computer program code executed by one or more network elements. Embodiments include a computer program product 1300 as depicted in FIG. 13 on a computer usable medium 1302 with computer program code logic 1304 containing instructions embodied in tangible media as an article of manufacture. Exemplary articles of manufacture for computer usable medium 1302 may include floppy diskettes, CD-ROMs, hard drives, universal serial bus (USB) flash drives, or any other computer-readable storage medium, wherein, when the computer program code logic 1304 is loaded into and executed by a computer, the computer becomes an apparatus for practicing the invention. Embodiments include computer program code logic 1304, for example, whether stored in a storage medium, loaded into and/or executed by a computer, or transmitted over some transmission medium, such as over electrical wiring or cabling, through fiber optics, or via electromagnetic radiation, wherein, when the computer program code logic 1304 is loaded into and executed by a computer, the computer becomes an apparatus for practicing the invention. When implemented on a general-purpose microprocessor, the computer program code logic 1304 segments configure the microprocessor to create specific logic circuits.

While the invention has been described with reference to exemplary embodiments, it will be understood by those skilled in the art that various changes may be made and equivalents may be substituted for elements thereof without departing from the scope of the invention. In addition, many modifications may be made to adapt a particular situation or material to the teachings of the invention without departing from the essential scope thereof. Therefore, it is intended that the invention not be limited to the particular embodiment disclosed as the best mode contemplated for carrying out this invention, but that the invention will include all embodiments falling within the scope of the appended claims. Moreover, the use of the terms first, second, etc. do not denote any order or importance, but rather the terms first, second, etc. are used to distinguish one element from another. Furthermore, the use of the terms a, an, etc. do not denote a limitation of quantity, but rather denote the presence of at least one of the referenced item.

CLAIMS

1. A computer program product for facilitating input/output (I/O) processing for an I/O operation at a host computer system configured for communication with a control unit, the computer program product comprising:

a tangible storage medium readable by a processing circuit and storing instructions for execution by the processing circuit for performing a method comprising:

the host computer system obtaining a transport command word (TCW) for an I/O operation having both input and output data, the TCW specifying a location of the output data, and a location for storing the input data;

the host computer system forwarding the I/O operation to the control unit for execution;

the host computer system gathering the output data responsive to the location of the output data specified by the TCW;

the host computer system forwarding the output data to the control unit for use in the execution of the I/O operation;

the host computer system receiving the input data from the control unit; and

the host computer system storing the input data at the location specified by the TCW for storing the input data.

2. The computer program product of claim 1 wherein the TCW further specifies a size of the output data, and the gathering the output data is further responsive to the size of the output data.

3. The computer program product of claim 1 wherein the I/O operation includes one or more commands to be executed by the I/O operation.

4. The computer program product of claim 1 wherein one or both of the location of the output data and the location for storing the input data are direct addresses.

5. The computer program product of claim 1 wherein one or both of the location of the output data and the location for storing the input data are indirect addresses.

6. The computer program product of claim 1 wherein the host computer system includes an I/O processing system and the method is performed by the I/O processing system.

7. The computer program product of claim 1 wherein the host computer system includes a channel subsystem and the method is performed by the channel subsystem.

8. The computer program product of claim 1 wherein the TCW is obtained from a host operating system.

9. A host computer system for providing bi-directional data transfer within a single I/O operation, comprising:

a host computer I/O subsystem comprising a channel adapter, the channel adapter adapted for communication with a control unit, the host I/O subsystem performing a method comprising:

obtaining a TCW for an I/O operation having both input and output data, the TCW specifying a location of the output data and a location for storing the input data;

gathering the output data responsive to the location of the output data specified by the TCW;

forwarding the I/O operation and the output data and to the control unit for execution;

receiving the input data from the control unit; and

storing the input data at the location specified by the TCW for storing the input data.

10. The apparatus of claim 9 wherein the TCW further specifies a size of the output data, and the gathering the output data is further responsive to the size of the output data.

11. The apparatus of claim 9 wherein the I/O operation includes one or more commands to be executed by the I/O operation.

12. The apparatus of claim 9 wherein one or both of the location of the output data and the location for storing the input data are direct addresses.

13. The apparatus of claim 9 wherein one or both of the location of the output data and the location for storing the input data are indirect addresses.

14. The apparatus of claim 9 wherein the TCW is obtained from a host operating system.

15. A method for facilitating I/O processing for an I/O operation at a host computer system configured for communication with a control unit, the method comprising:

obtaining a TCW for an I/O operation having both input and output data, the TCW specifying a location of the output data and a location for storing the input data;

gathering the output data responsive to the location of the output data specified by the TCW;

forwarding the I/O operation and the output data to the control unit for execution;

receiving the input data from the control unit; and

storing the input data at the location specified by the TCW for storing the input data.

16. The method of claim 15 wherein the TCW further specifies a size of the output data, and the gathering the output data is further responsive to the size of the output data.

17. The method of claim 15 wherein the I/O operation includes one or more commands to be executed by the I/O operation.

18. The method of claim 15 wherein one or both of the location of the output data and the location for storing the input data are direct addresses.

19. The method of claim 15 wherein one or both of the location of the output data and the location for storing the input data are indirect addresses.

20. The method of claim 15 wherein the host computer system includes an I/O processing system and the method is performed by the I/O processing system.

21. The method of claim 15 wherein the host computer system includes a channel subsystem and the method is performed by the channel subsystem.

22. The method of claim 15 wherein the TCW is obtained from a host operating system.

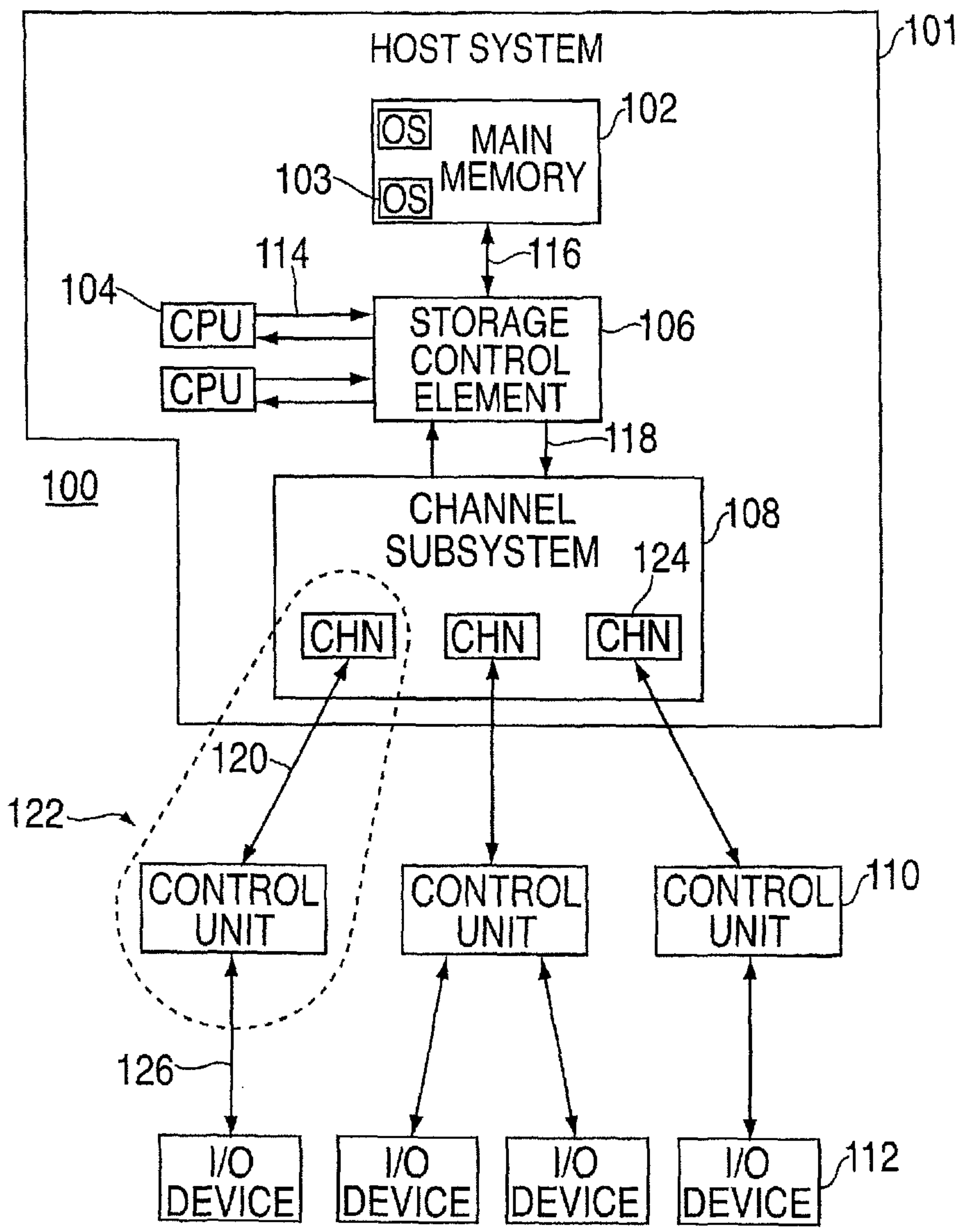
HOST SYSTEM
101
OS
MAIN MEMORY
102
103
OS
114
116
104
CPU
STORAGE CONTROL ELEMENT
106
CPU
118
100
CHANNEL SUBSYSTEM
108
124
CHN
CHN
CHN
120
122
CONTROL UNIT
CONTROL UNIT
CONTROL UNIT
110
126
I/O DEVICE
I/O DEVICE
I/O DEVICE
I/O DEVICE
112

FIG. 1

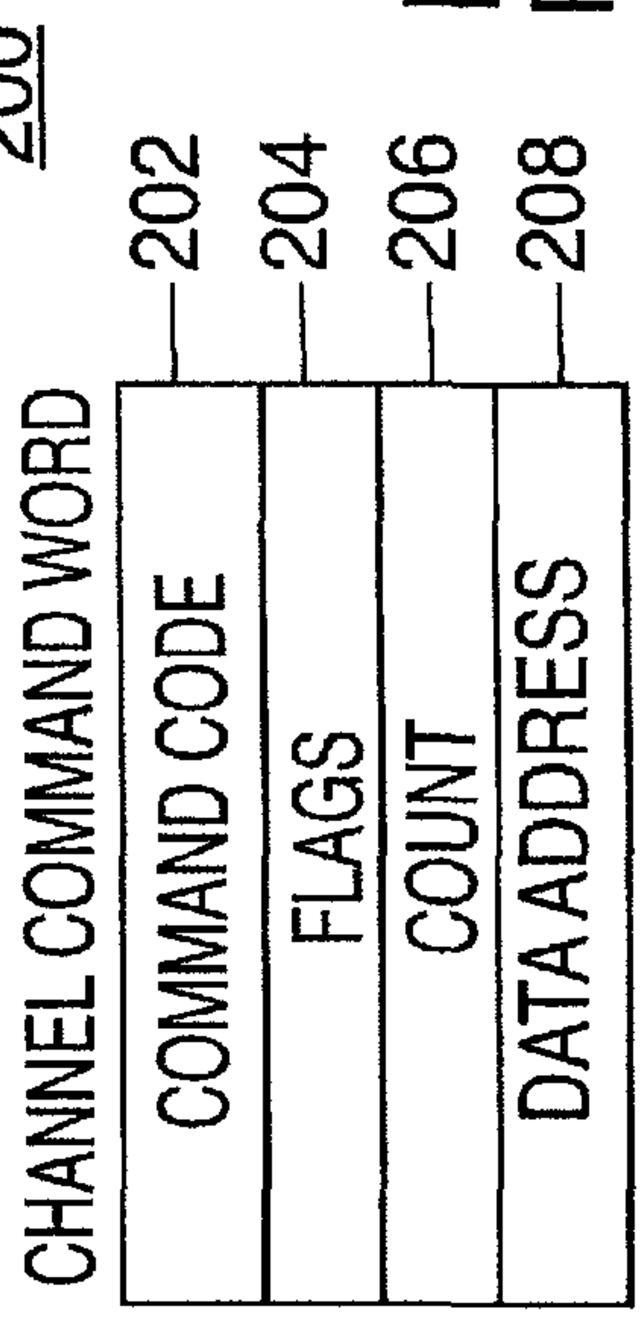

FIG. 2A
PRIOR ART

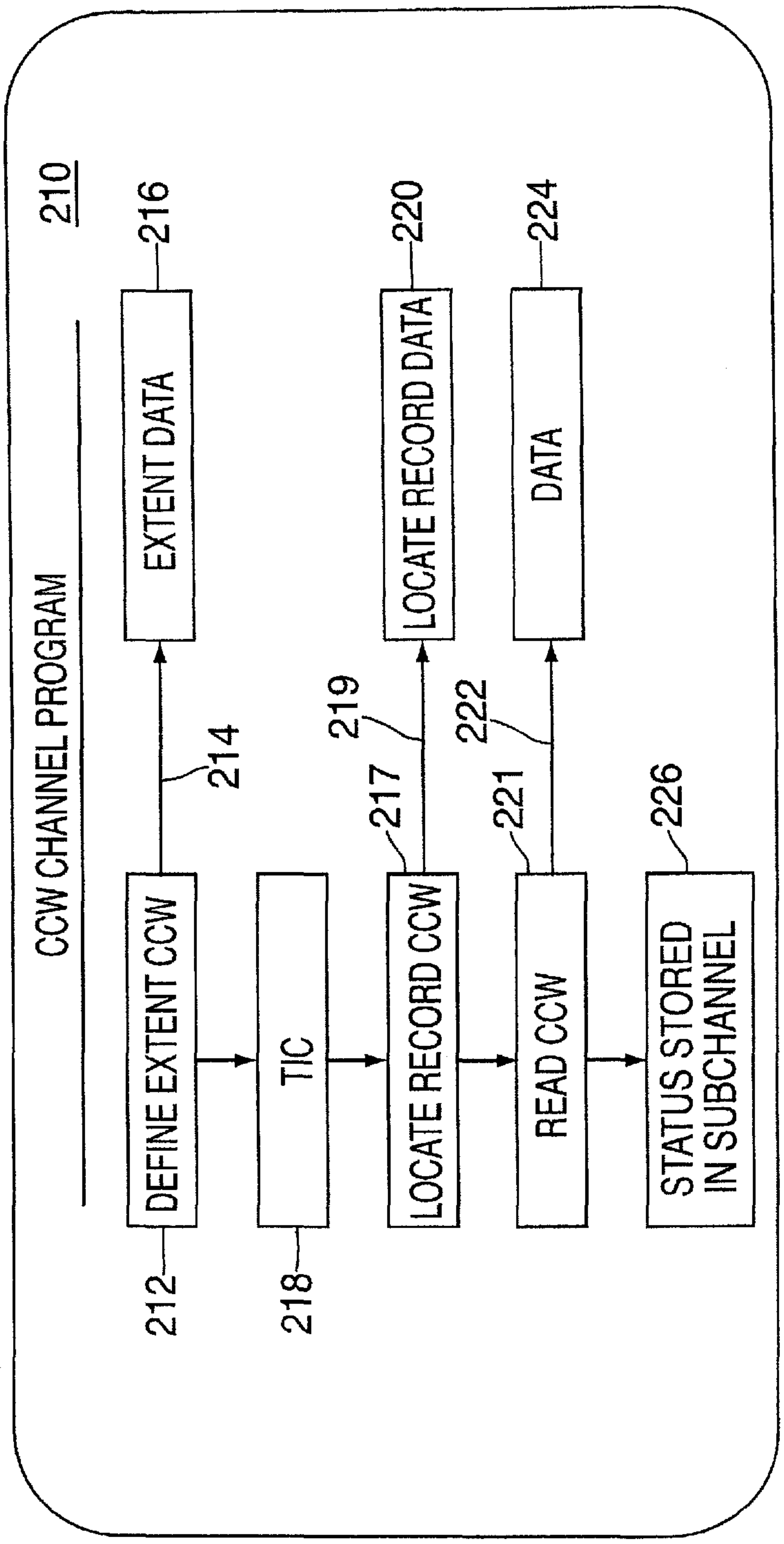

FIG. 2B
PRIOR ART

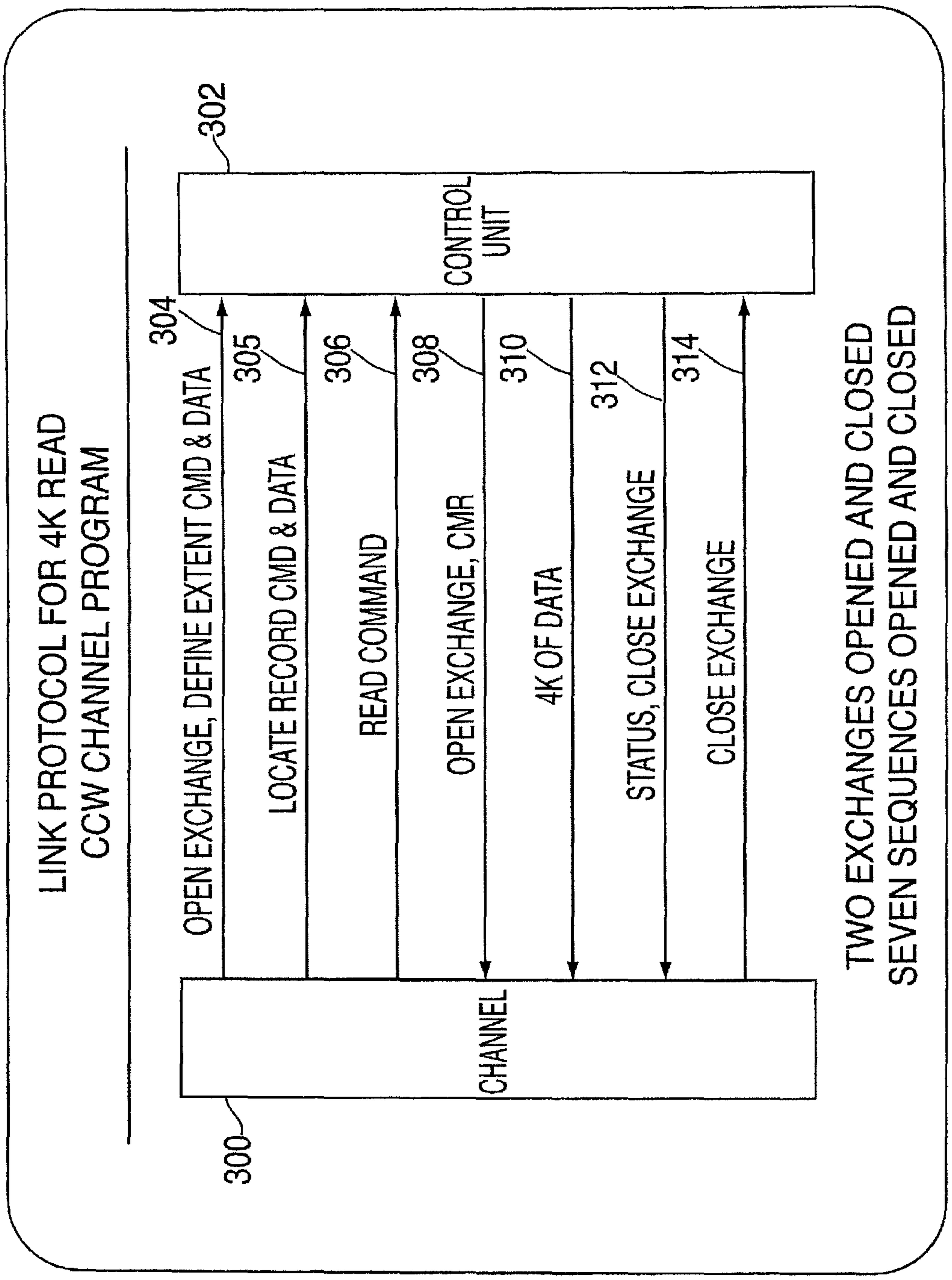

FIG. 3
PRIOR ART

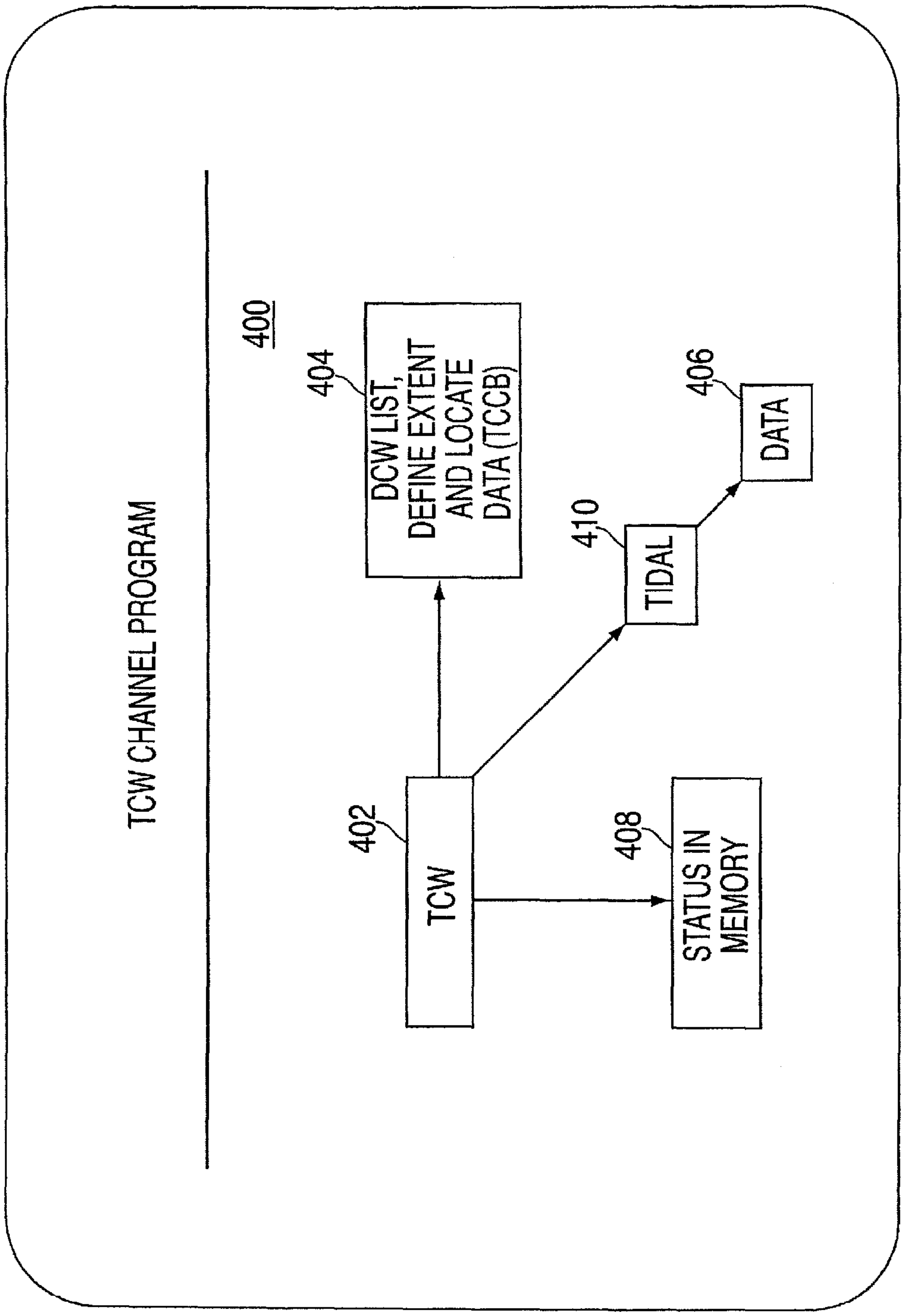
TCW CHANNEL PROGRAM
400
404
DCW LIST, DEFINE EXTENT AND LOCATE DATA (TCCB)
406
DATA
410
TIDAL
402
TCW
408
STATUS IN MEMORY

FIG. 4

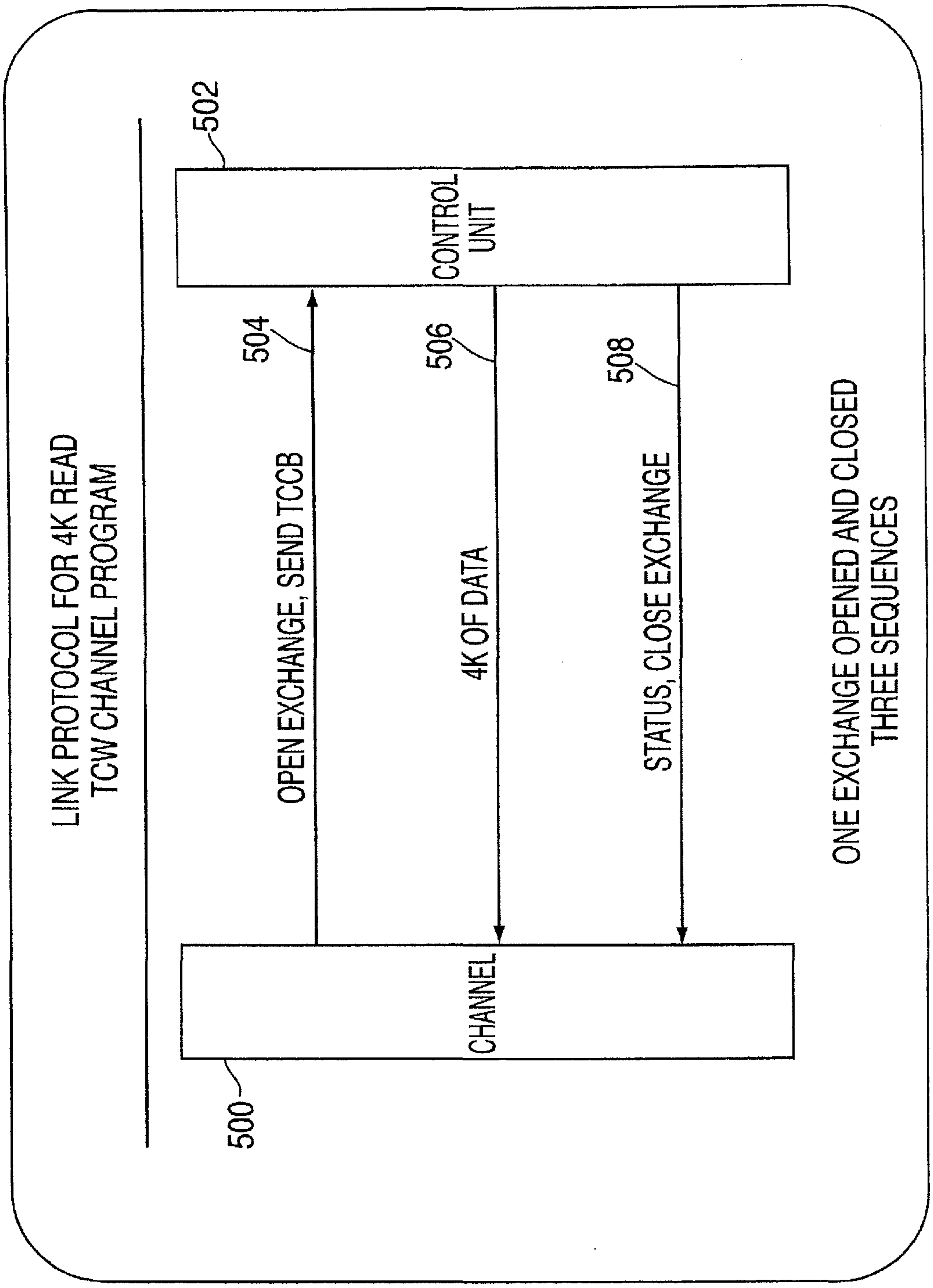
LINK PROTOCOL FOR 4K READ
TCW CHANNEL PROGRAM
502
CONTROL
UNIT
504
OPEN EXCHANGE, SEND TCCB
506
4K OF DATA
508
STATUS, CLOSE EXCHANGE
CHANNEL
500
ONE EXCHANGE OPENED AND CLOSED
THREE SEQUENCES

FIG. 5

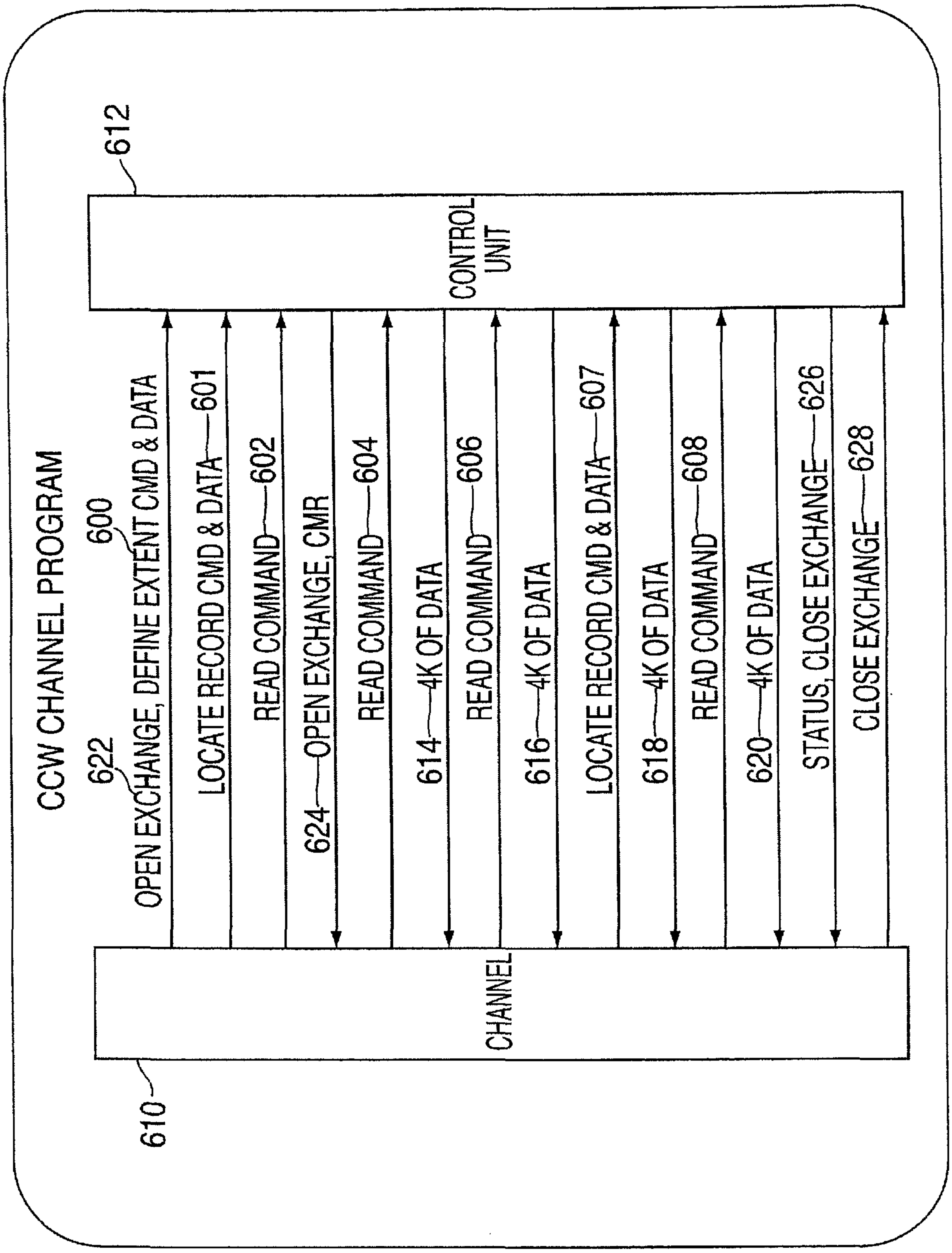

FIG. 6
PRIOR ART

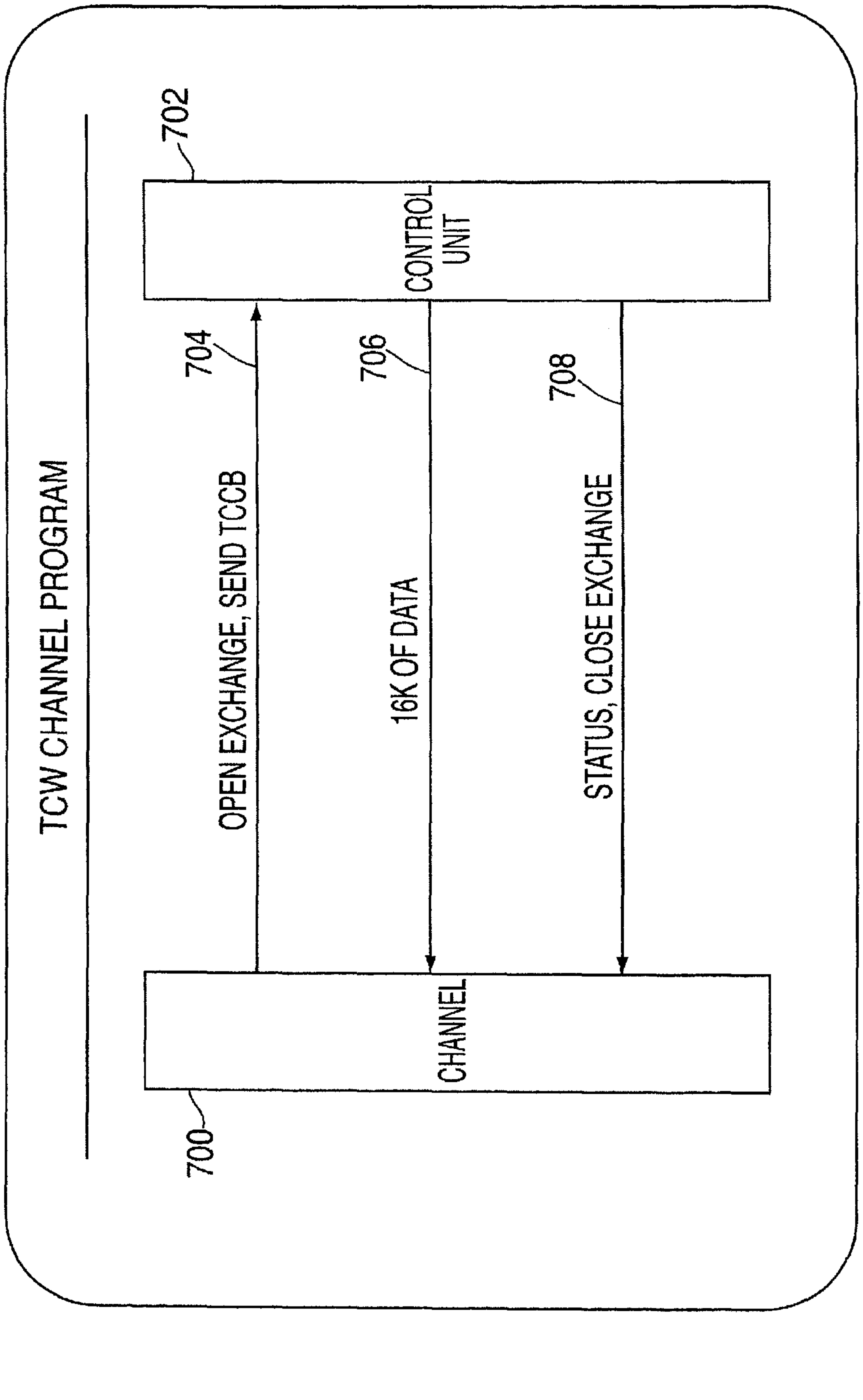
TCW CHANNEL PROGRAM
702
CONTROL UNIT
704
OPEN EXCHANGE, SEND TCCB
706
16K OF DATA
708
STATUS, CLOSE EXCHANGE
CHANNEL
700

FIG. 7

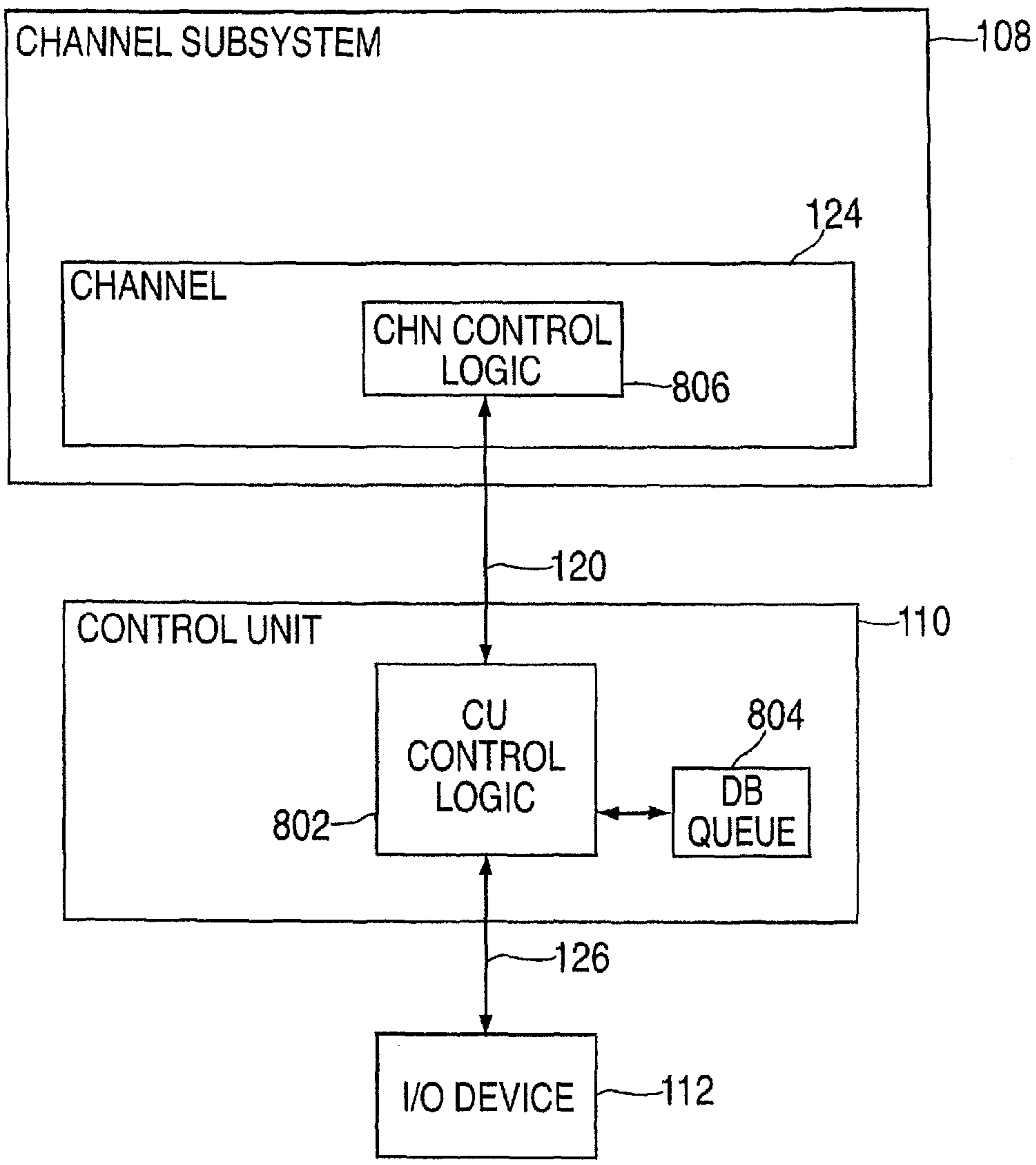
CHANNEL SUBSYSTEM
108
124
CHANNEL
CHN CONTROL LOGIC
806
120
CONTROL UNIT
110
CU CONTROL LOGIC
804
DB QUEUE
802
126
I/O DEVICE
112

FIG. 8

900

WORD	BYTE Ø (902)	BYTE 1	BYTE 2	BYTE 3	BYTE 4	BYTE 5	BYTE 6	BYTE 7
0	F 000000 904	FLAGS 906			Reserved 908	TCCBL (L1) 910 R W (912)	Reserved 914	
1	916 Output-Data Address							
2	918 Input-Data Address							
3	920 Transport-Status-Block Address							
4	922 Transport-Command-Control Block Address							
5	Output Count 924				Input Count 926			
6	reserved 928				reserved 930			
7	reserved 932				Interrogate-TCW Address 934			

FIG. 9

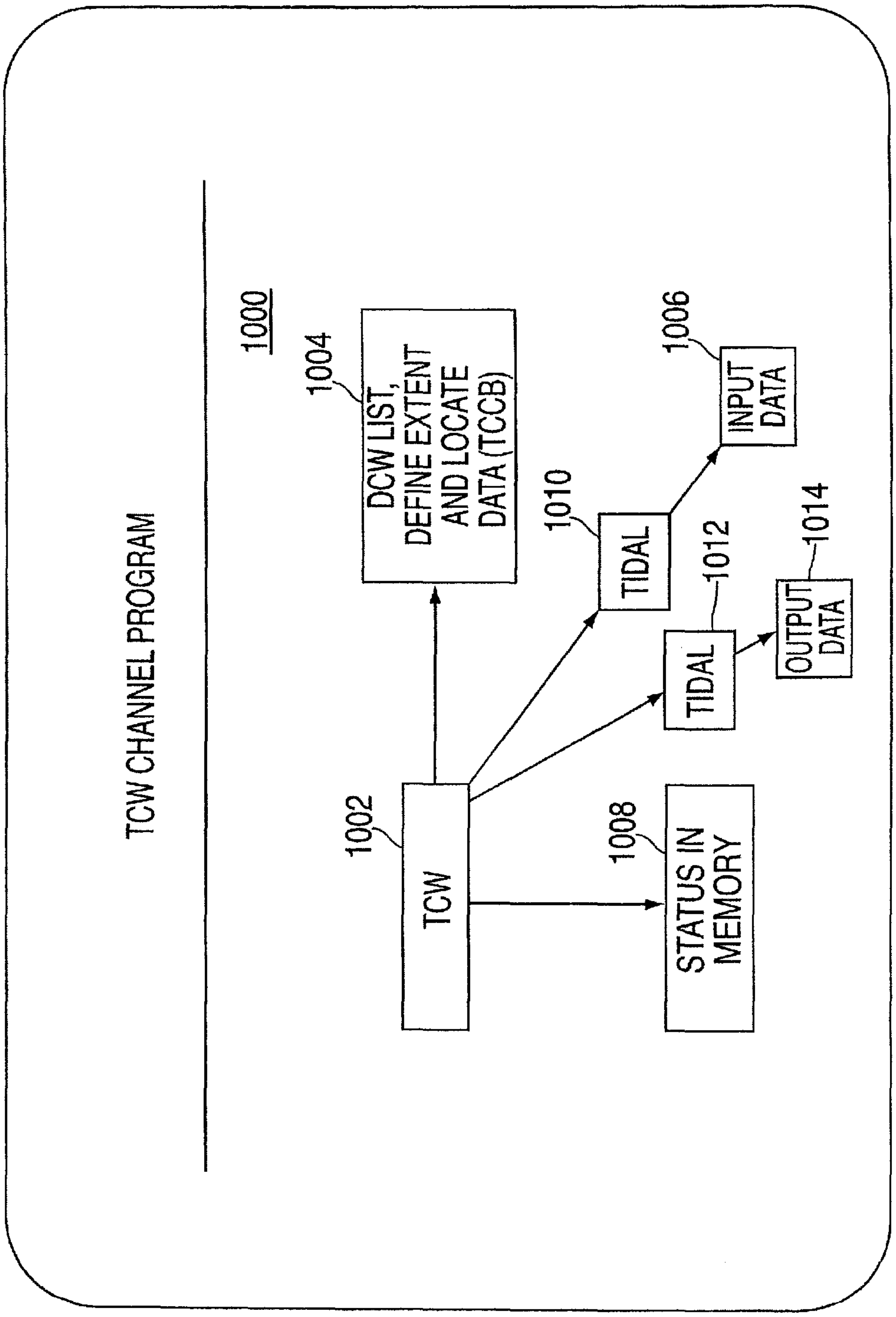
TCW CHANNEL PROGRAM
1000
1002
TCW
1004
DCW LIST, DEFINE EXTENT AND LOCATE DATA (TCCB)
1010
TIDAL
1006
INPUT DATA
1012
TIDAL
1014
OUTPUT DATA
1008
STATUS IN MEMORY

FIG. 10

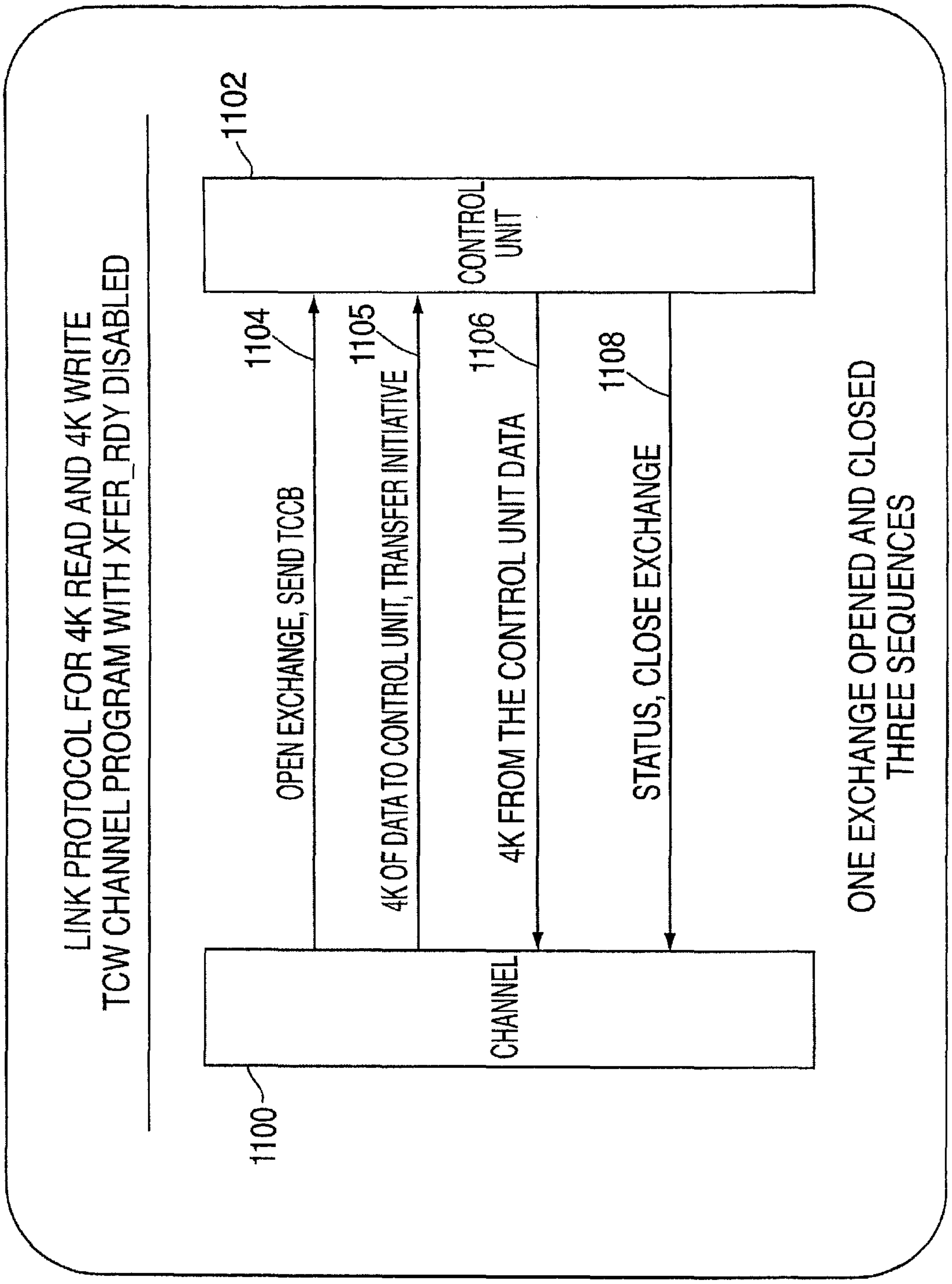
LINK PROTOCOL FOR 4K READ AND 4K WRITE
TCW CHANNEL PROGRAM WITH XFER_RDY DISABLED
CHANNEL
CONTROL UNIT
1100
1102
1104
1105
1106
1108
OPEN EXCHANGE, SEND TCCB
4K OF DATA TO CONTROL UNIT, TRANSFER INITIATIVE
4K FROM THE CONTROL UNIT DATA
STATUS, CLOSE EXCHANGE
ONE EXCHANGE OPENED AND CLOSED
THREE SEQUENCES

FIG. 11

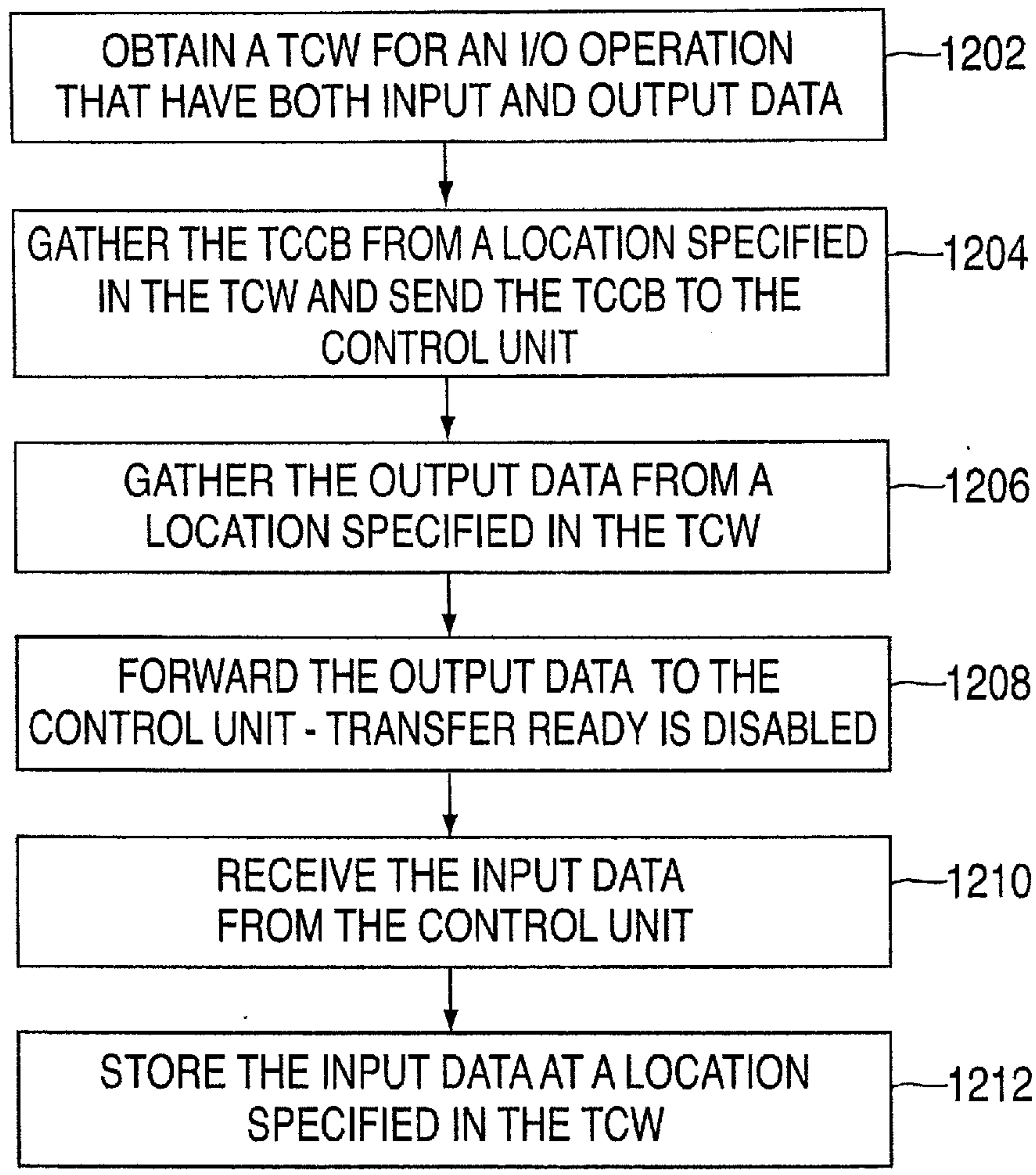
OBTAIN A TCW FOR AN I/O OPERATION THAT HAVE BOTH INPUT AND OUTPUT DATA
1202
GATHER THE TCCB FROM A LOCATION SPECIFIED IN THE TCW AND SEND THE TCCB TO THE CONTROL UNIT
1204
GATHER THE OUTPUT DATA FROM A LOCATION SPECIFIED IN THE TCW
1206
FORWARD THE OUTPUT DATA TO THE CONTROL UNIT - TRANSFER READY IS DISABLED
1208
RECEIVE THE INPUT DATA FROM THE CONTROL UNIT
1210
STORE THE INPUT DATA AT A LOCATION SPECIFIED IN THE TCW
1212

FIG. 12

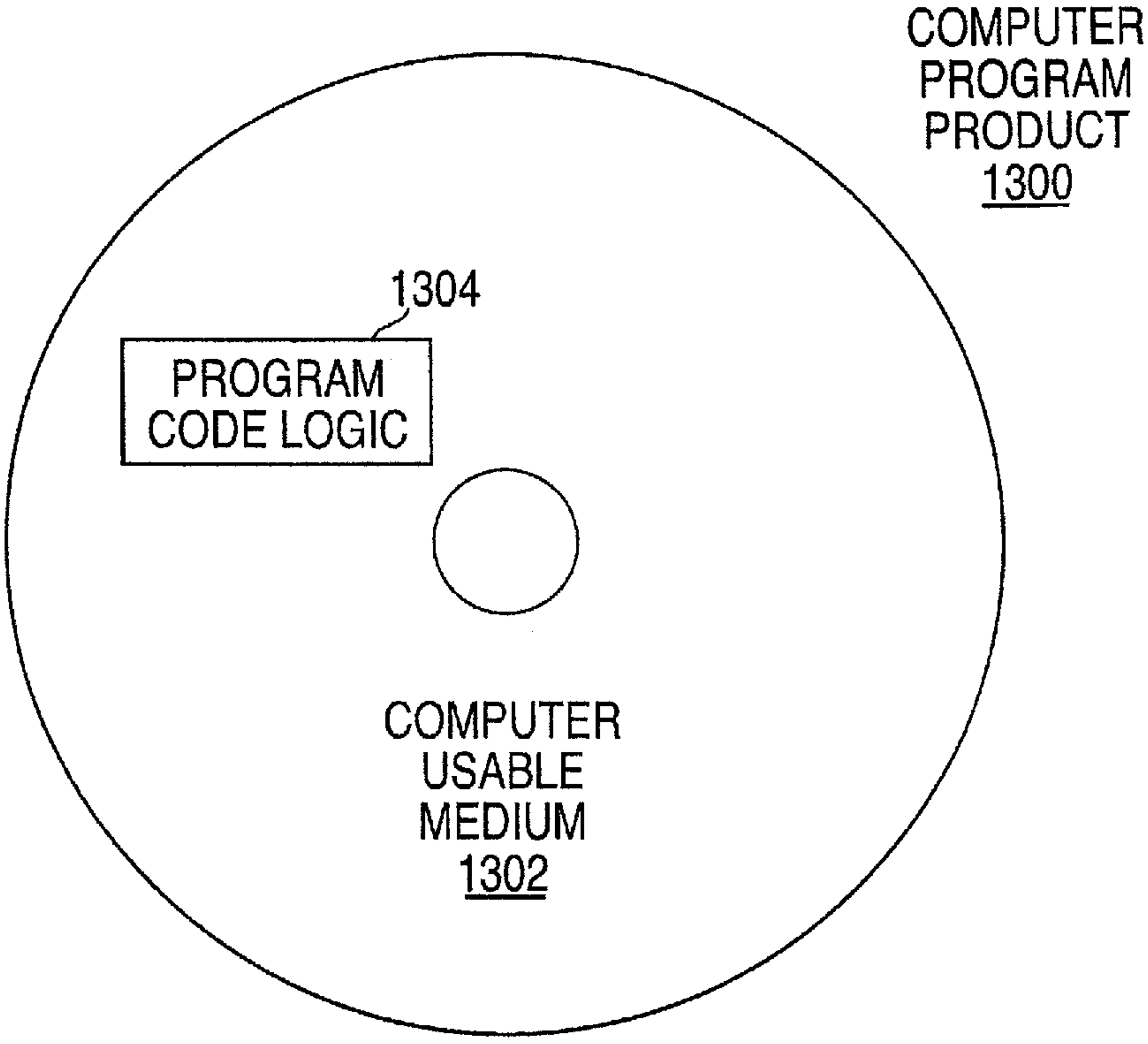
COMPUTER
PROGRAM
PRODUCT
1300
1304
PROGRAM
CODE LOGIC
COMPUTER
USABLE
MEDIUM
1302

FIG. 13

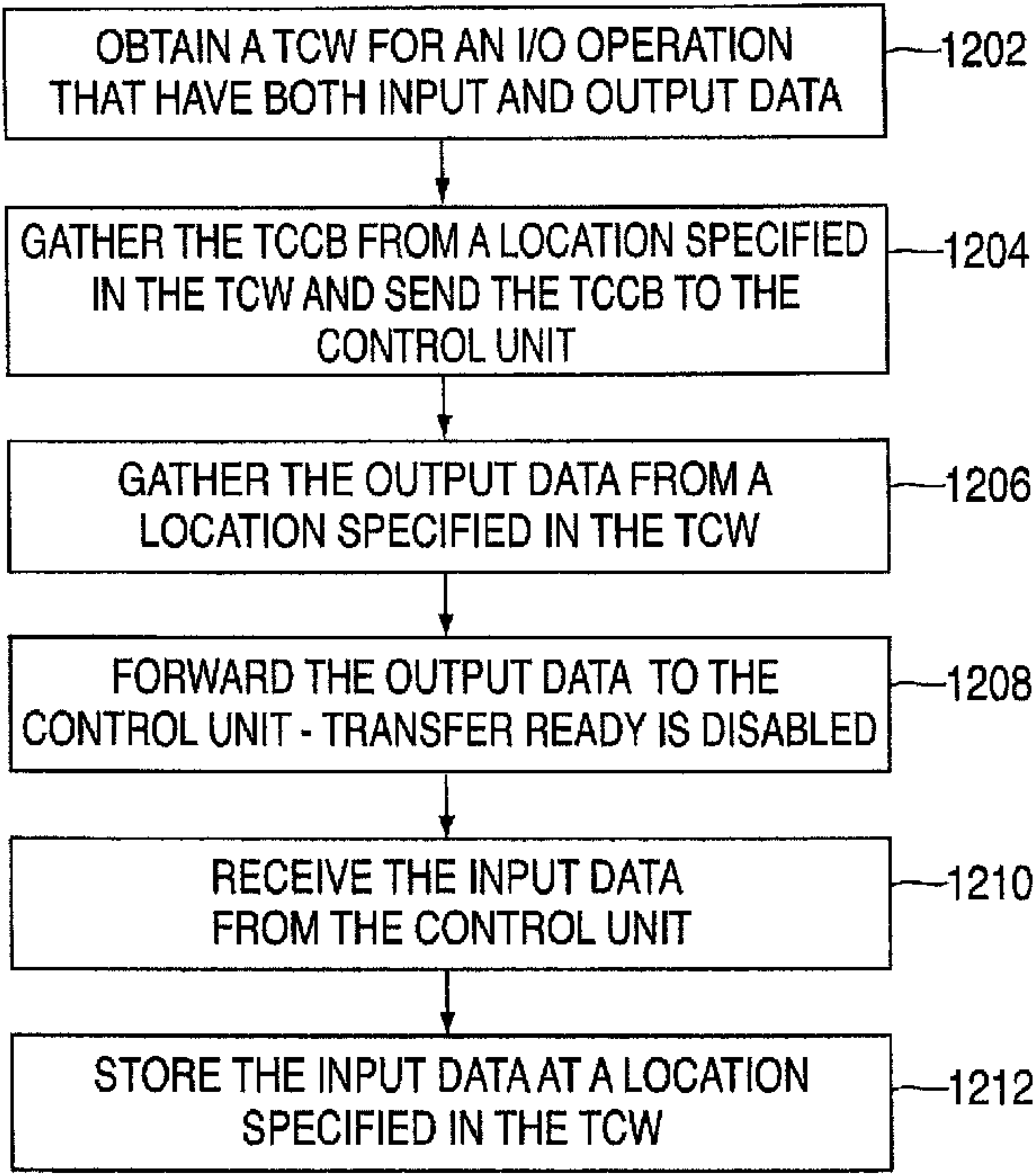
OBTAIN A TCW FOR AN I/O OPERATION THAT HAVE BOTH INPUT AND OUTPUT DATA
1202
GATHER THE TCCB FROM A LOCATION SPECIFIED IN THE TCW AND SEND THE TCCB TO THE CONTROL UNIT
1204
GATHER THE OUTPUT DATA FROM A LOCATION SPECIFIED IN THE TCW
1206
FORWARD THE OUTPUT DATA TO THE CONTROL UNIT - TRANSFER READY IS DISABLED
1208
RECEIVE THE INPUT DATA FROM THE CONTROL UNIT
1210
STORE THE INPUT DATA AT A LOCATION SPECIFIED IN THE TCW
1212

FIG. 12