



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

(21) International Application Number:  
PCT/EP2009/051463

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
12/030,939 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, Tuc-

son, Arizona 85718 (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). **SITTMANN III, Gustav** [US/US]; 422 South Park Avenue, Webster Groves, Missouri 63119 (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, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[Continued on next page]

(54) Title: PROCESSING A VARIABLE LENGTH DEVICE COMMAND WORD AT A CONTROL UNIT IN AN I/O PROCESSING SYSTEM

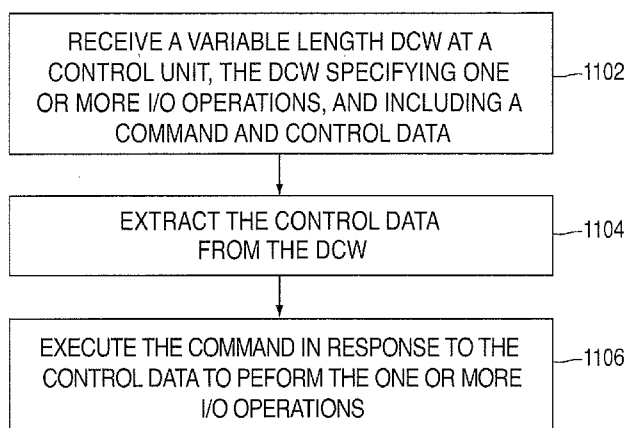


FIG. 11

(57) Abstract: A computer program product, apparatus and method for processing a variable length device command word (DCW) at a control unit configured for communication with an input/output (I/O) subsystem in an I/O processing system. 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 receiving a DCW at the control unit from the I/O subsystem. The DCW specifies one or more I/O operations and includes a command, a control data count, and control data having a varying length specified by the control data count. The control data is extracted in response to the control data count. The command is then executed in response to the extracted control data to perform the specified one or more I/O operations.



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).

— *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))*

**Published:**

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

**PROCESSING A VARIABLE LENGTH DEVICE COMMAND WORD AT A  
CONTROL UNIT IN AN I/O PROCESSING SYSTEM**

**FIELD OF THE INVENTION**

The present disclosure relates generally to input/output (I/O) processing, and in particular, to processing a variable length device command word at a control unit in an I/O processing system.

**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 I/O 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.

Some commands require both control data and customer data. Currently, a single command word (e.g., a CCW) cannot transfer both control data and customer data. A command that includes both control data and the customer data requires multiple transfers to specify both the control and customer data. This impacts the increase in throughput that can be gained by grouping commands. Performance may be improved by providing the ability for a single command word to transfer both control data and customer data.

## **BRIEF SUMMARY OF THE INVENTION**

An exemplary embodiment includes a computer program product for processing a variable length device command word (DCW) at a control unit configured for communication with an input/output (I/O) subsystem in an I/O processing system. 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 receiving a DCW at the control unit from the I/O subsystem. The DCW specifies one or more I/O operations and includes a command, a control data count, and control data having a varying length specified by the control data count. The control data is extracted in

response to the control data count. The command is then executed in response to the extracted control data to perform the specified one or more I/O operations.

Another exemplary embodiment includes an apparatus for processing a variable length DCW in an I/O processing system. The apparatus includes a control unit configured for communication with an I/O subsystem. The control unit performs a method that includes receiving a DCW at the control unit from the I/O subsystem. The DCW specifies one or more I/O operations and includes a command, a control data count, and control data having a varying length specified by the control data count. The control data is extracted in response to the control data count. The command is then executed in response to the extracted control data to perform the specified one or more I/O operations.

A further exemplary embodiment includes a method for processing a variable length DCW at a control unit configured for communication with an input/output (I/O) subsystem in an I/O processing system. The method comprises receiving a DCW at the control unit from the I/O subsystem. The DCW specifies one or more I/O operations and includes a command, a control data count, and control data having a varying length specified by the control data count. The control data is extracted in response to the control data count. The command is then executed in response to the extracted control data to perform the specified one or more I/O operations.

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, with XFER\_RDY not disabled, 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 write 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, with XFER\_RDY not disabled, to process the four write 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 TCCB in accordance with an aspect of the present invention;

FIG. 10 depicts one embodiment of a DCW in accordance with an aspect of the present invention;

FIG. 11 depicts one embodiment of a process for providing a variable length DCW in accordance with an aspect of the present invention; and

FIG. 12 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) is facilitated by allowing a single command word to include control data and also to transfer customer data. A one byte count field is added to the command word to define the byte count of the control data required by the command. Thus, the command word includes two count fields, one for the control data and the other for customer data. In an exemplary embodiment, the variable length control data directly follows the command word.

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 both control and customer data to the control unit as a single entity (e.g., as part of the same control word) for execution by the control unit.

The plurality of commands (e.g., device command words or “DCWs”) are included in a block, referred to herein as a transport command control block (TCCB), an address (indirect or direct) of which is specified in a transport control word (TCW). In an exemplary embodiment, 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.

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 CPU 104 and 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 as 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 is 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 control unit 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 instance, 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 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 instance, 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 write CCWs 221. Each write CCW 221 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 write CCW. It fetches the locate record command and data 306 (FIG. 3) from the locate record CCW 217 (FIG. 2B) and locate record data 220. The write command and data 308 (FIG. 3) is fetched from write CCW 221 and data area 224 (FIG. 2B). Each is sent to the control unit 302.

The control unit 302 opens an exchange 310 with the channel 300, in response to the open exchange of the channel 300. This can occur before or after locate command and data 306 and/or write command and data 308. 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 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 write 4k of data requires two exchanges to be opened and closed and six 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 and is not sent or seen by the control unit.

One example of a channel program to write 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. TIDAW zero 412, TIDAW one 414 and TIDAW two 416 (collectively TIDAWs 412 – 416) can reference different locations in the data area 406 for acquiring or storing data. The TIDAWs 412 – 416 can reference non-contiguous blocks of data or contiguous blocks of data. The TIDAWs 412 – 416 in TIDAL 410 may be located sequentially in memory or located non-contiguously relative to each other. While only three TIDAWs 412 – 416 are depicted in TIDAL 410, it will be understood that any number of TIDAWs can be included in the TIDAL 410.

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.

The embodiment of the link protocol depicted in FIG. 5 is utilized when XFER\_RDY is enabled. In the embodiment depicted in FIG. 5, the channel 500 cannot send the data 506 to the control unit 502 until it is requested by the control unit 502 via the XFER\_RDY IU 510. In an alternate exemplary embodiment, XFER\_RDY is disabled and the control unit does not transmit a XFER\_RDY IU 510 to the channel 500. Thus, the channel 500 does not have to wait for the control unit 502 to request the data 506 before sending the data 506. This alternate embodiment, where XFER\_RDY is disabled may be utilized when the channel 500 and the control unit 502 are located geographically far apart from each other (e.g., greater than twenty kilometers, greater than fifty kilometers) to improve performance. Unless otherwise specified, the discussion herein assumes that XFER\_RDY is enabled.

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 and sequence initiative to the control unit 502.
2. Wait for a XFER\_RDY IU indication that the control unit is ready to receive the data 502.
3. Transfer the IU of data, and sequence initiative to the control unit 502.
4. 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\_RSP\_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 four communication

sequences (see FIGs. 4-5), as compared to six 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 and data 600, locate record command and data 601, write command and data 602, write command and data 604, locate record command and data 606, write command and data 608, and write command and data 620) are sent in separate sequences from channel 610 to control unit 612. This CCW channel program requires two exchanges to be opened and closed (e.g., open exchanges 622, 624 and close exchanges 626, 628), and ten communications sequences. This is compared to the four 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 write commands in DCWs, as described above. The channel 700 transmits 16k of data 706 to the control unit 702 in a single sequence upon receipt of the XFER\_RDY IU 710. The channel 700 may insert a CRC every 4K of the 16k of data 706 in the sequence. The insertion of a CRC every 4K allows the control unit 702 to verify the 16K of data incrementally, rather than buffer the entire 16K for verification before completing the write commands in the TCCB 704. Additionally, the control unit 702 provides status to the channel 700 and closes the exchange 708. Thus, the TCW channel program of FIG. 7 requires much less communications to transfer the same amount of data as the CCW channel program of FIG. 6, while supporting incremental data verification via multiple CRC insertion in the output data stream from the channel 700.

Turning now to FIG. 8, one embodiment of channel 124 in the channel subsystem and the control unit 110 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 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.

FIG. 9 depicts one embodiment of a TCCB 900 in accordance with an aspect of the present invention. The TCCB 900 in FIG. 9 is located at the address indicated in by the TCW. This address may be a direct address or an indirect address, allowing the contents of the TCCB 900 to be in one storage location or to be spread among multiple non-contiguous storage locations. As described previously, the TCCB 900 is a control block built by software and then the channel 124 sends it to a control unit 110 (e.g., in a Transport Command IU) for execution. The TCCB 900 contains the commands to be executed by the control unit 110

and any control data required by the commands. The channel 124 does not look at the contents of the TCCB 900. The channel 124 packages the TCCB 900 and sends it to the control unit 110. This allows FCP transport protocols to be utilized instead of FICON.

The TCCB 900 includes a transport control area header (TCAH) 902 which, in an exemplary embodiment, includes information about the transport command area (TCA) 904 and operations within the TCA 904 (e.g., length, service code). The TCCB 900 depicted in FIG. 9 also includes a variable length TCA 904 which includes one or more DCWs 906 which may each include DCW control data 908. In an exemplary embodiment, the DCW control data 908 is variable length. In an exemplary embodiment, each DCW 906 includes a command flags (chaining), control data length, and read/write data length. DCW control data 908 is optional (depending on the DCW 906) and includes control parameters for the DCW 906. For example, DCW control data 908 may include define extent and/or prefix parameters. In an exemplary embodiment, the DCW control data is located at the end of the DCW 906 within the TCA 904 and is not pointed to by the DCW 906. Finally, the TCCB 900 includes a TCA trailer (TCAT) 910 that contains data such as the count of the bytes to be transferred in the TCCB 900.

FIG. 10 depicts one embodiment of a DCW 1000 in accordance with an aspect of the present invention. In an exemplary embodiment, the DCW 1000 is eight bytes in length plus the length of the control data 1002. The DCW includes a command field 1004, a flags field 1006, a reserved field 1008, a control data (CD) count field 1010, and a data byte count field 1012. The DCW command field 1004 is one byte in length and is the same as the CCW command byte utilized in a CCW (but may include additional command codes not utilized by a CCW). The flags field 1006 includes eight bits; in an exemplary embodiment, the second bit is a chain command to the next DCW 1000 in the TCA 904. When this flag bit is set to zero, it indicates that this is the last DCW 1000 of the DCW program in the TCA 904. The other bits in the flag field 1006 are reserved and set to zero.

The CD count field 1010 is a one byte field that contains the byte count of the control data 1002 used by this DCW 1000. The control data 1002 is tacked on to the end of the DCW 1000. In an exemplary embodiment, the space used between DCWs 1000 when control data 1002 is present is the size of the control data 1002 padded up to the next four byte boundary so that the DCWs 1000 always start on a four byte boundary. The data byte count field 1012



is a four byte DCW byte count field. This is the DCW count of customer data (e.g., the data 406 in FIG. 4). In an exemplary embodiment, this count does not include padding or CRC bytes to be transferred by this DCW command. This DCW byte count field 1012 has the same meaning as the count field in a CCW, except that it can be much larger. If the DCW command is a command immediate or a no-op command (i.e., contains no customer data), then the data byte count field 1012 is equal to zero.

In an exemplary embodiment, the control data 1002 is in the same TCA 904 at the end of the DCW 1000 with which the control data 1002 is associated. In an exemplary embodiment, the control data 1002 is padded out to the next four byte boundary in the TCA 904. The CD count field 1010 reflects the number of bytes of control data 1002. In an exemplary embodiment, the maximum length of the control data 1002 in a TCA 904 is two hundred and thirty two bytes assuming one DCW 1000 and the rest of the TCA 904 is taken up by the control data for the DCW 1000.

FIG. 11 depicts one embodiment of a process for providing a variable length DCW in accordance with an aspect of the present invention. At block 1102 a control unit receives a variable length DCW 1000. The DCW 1000 specifies multiple equivalent legacy I/O operations (e.g., define extent, locate record and then a read operation or a write operation, etc.) defined by command 1004, control data count 1010, and variable length control data 1002. The length of the control data is specified by the control data count. At block 1104, the control data 1002 is extracted from the DCW using the control data count to determine how many bytes at the end of the DCW 1000 are to be included in the control data. At block 1106, the command is executed by the control unit using the control data as input to the executing. The execution results in multiple equivalent legacy I/O operations being performed. If the DCW also includes customer data (e.g., the data byte count 1012 is non-zero) then the command is executed in response to the to the control data and to the customer data. In addition, the control unit may transmit a completion status of the command (e.g., completed successfully, error conditions, etc.) to the host computer system.

Thus, referring back to FIGs. 2B and 10, an exemplary embodiment of the present invention may be utilized to combine the define extent CCW 212, locate record CCW 217 and write CCW 221 into a single DCW. This DCW is a combined command that means: define extent, locate record, and write customer data. In this exemplary embodiment, the extent

and locate record parameter data is the control data that is part of the DCW whose length is defined by the control data count 1010. The customer data length is controlled by the data byte count 1012. The customer data transferred can only be a write (output) or a read (input), not both, in the same DCW. This combination of commands within a CCW could not have been performed in current implementations because the only way to get the control data (the define extent parameters and the locate record parameters) to the control unit was for the channel to transfer these parameters as data to the control unit.

Technical effects of exemplary embodiments include the ability to transfer customer data and control data within the same DCW. This may lead to improved performance by requiring fewer exchanges to perform the same number of functions.

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 1200 as depicted in FIG. 12 on a computer usable medium 1202 with computer program code logic 1204 containing instructions embodied in tangible media as an article of manufacture. Exemplary articles of manufacture for computer usable medium 1202 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 1204 is loaded into and executed by a computer, the computer becomes an apparatus for practicing the invention.

Embodiments include computer program code logic 1204, 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 1204 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 1204 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 processing a variable length device command word (DCW) at a control unit configured for communication with an input/output (I/O) subsystem in an I/O processing system, 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:

receiving a DCW at the control unit from the I/O subsystem, the DCW specifying one or more I/O operations and including a command, a control data count, and control data having a varying length specified by the control data count;

extracting the control data in response to the control data count; and

executing the command in response to the extracted control data to perform the specified one or more I/O operations.

2. The computer program product of claim 1, wherein the DCW further includes a customer data count specifying an amount of customer data associated with the DCW.

3. The computer program product of claim 2 wherein the customer data count has a value greater than zero and the method further comprises receiving customer data from the I/O subsystem.

4. The computer program product of claim 2 wherein the customer data count has a value greater than zero and the method further comprises transmitting the customer data to the I/O subsystem.

5. The computer program product of claim 1 wherein the method further comprises completing the I/O operations and transmitting a completion message to the I/O subsystem in response to the completing.

6. The computer program product of claim 1 wherein the DCW is received via a transport command control block (TCCB) that includes a transport command area (TCA) header, a TCA where the DCW is located, and a TCA trailer.

7. The computer program product of claim 6 wherein the TCA includes one or more other commands for execution by the control unit.
8. The computer program product of claim 1 wherein the I/O subsystem is a channel subsystem.
9. An apparatus for processing a variable length DCW in an I/O processing system, the apparatus comprising:
  - a control unit configured for communication with an I/O subsystem, the control unit performing a method comprising:
    - receiving a DCW specifying one or more I/O operations at the control unit from the I/O subsystem, the DCW including a command, a control data count, and control data having a varying length specified by the control data count;
    - extracting the control data in response to the control data count; and
    - executing the command in response to the extracted control data to perform the specified one or more I/O operations.
10. The apparatus of claim 9, wherein the DCW further includes a customer data count specifying an amount of customer data associated with the DCW.
11. The apparatus of claim 10 wherein the customer data count has a value greater than zero and the method further comprises receiving the customer data from the I/O subsystem.
12. The apparatus of claim 10 wherein the customer data count has a value greater than zero and the method further comprises transmitting the customer data to the I/O subsystem.
13. The apparatus of claim 9 wherein the method further comprises completing the I/O operations and transmitting a completion message to the I/O subsystem in response to completing.
14. The apparatus of claim 9 wherein the DCW is received via a TCCB that includes a TCA header, a TCA where the DCW is located, and a TCA trailer.

15. The apparatus of claim 14 wherein the TCA includes one or more other commands for execution by the control unit.

16. A method for processing a variable length DCW at a control unit configured for communication with an input/output (I/O) subsystem in an I/O processing system, the method comprising:

receiving a DCW specifying one or more I/O operations at the control unit from the host computer system, the DCW including a command, a control data count, and control data having a varying length specified by the control data count;

extracting the control data in response to the control data count; and

executing the command in response to the extracted control data to perform the specified one or more I/O operations.

17. The method of claim 16, wherein the DCW further includes a customer data count specifying an amount of customer data associated with the DCW.

18. The method of claim 17 wherein the customer data count has a value greater than zero and the method further comprises receiving the customer data from the I/O subsystem.

19. The method of claim 17 wherein the customer data count has a value greater than zero and the method further comprises transmitting the customer data to the I/O subsystem.

20. The method of claim 16 further comprising completing the I/O operations and transmitting a completion message to the I/O subsystem in response to the completing.

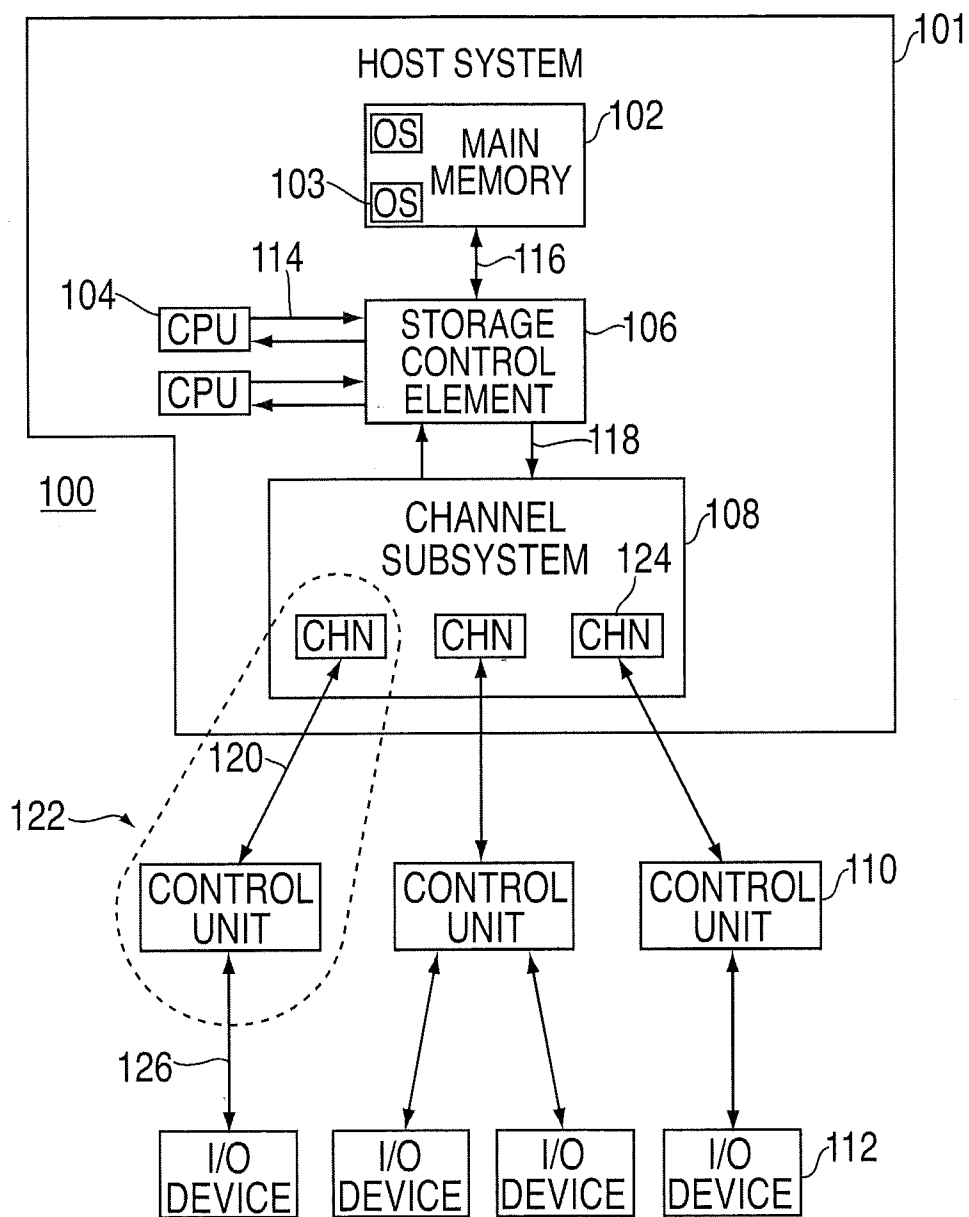


FIG. 1

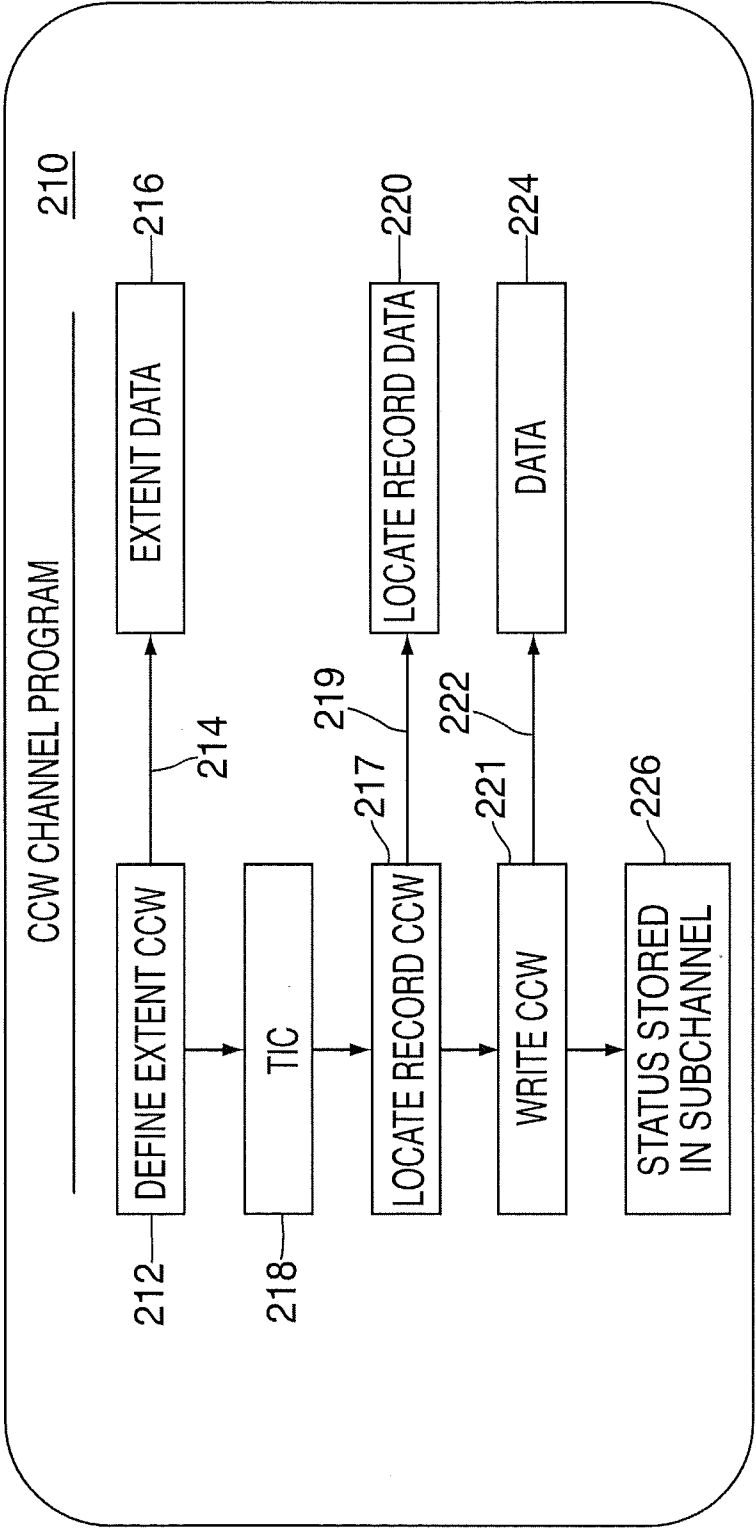
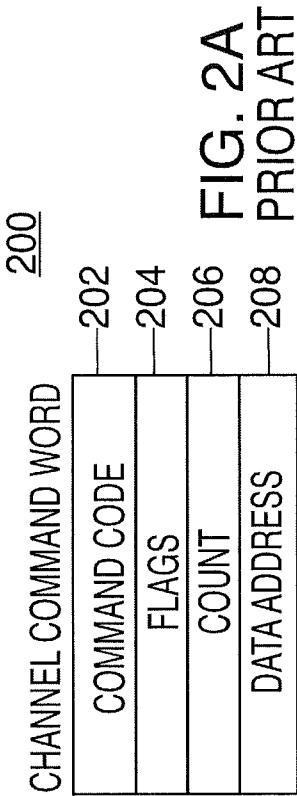


FIG. 2B  
PRIOR ART



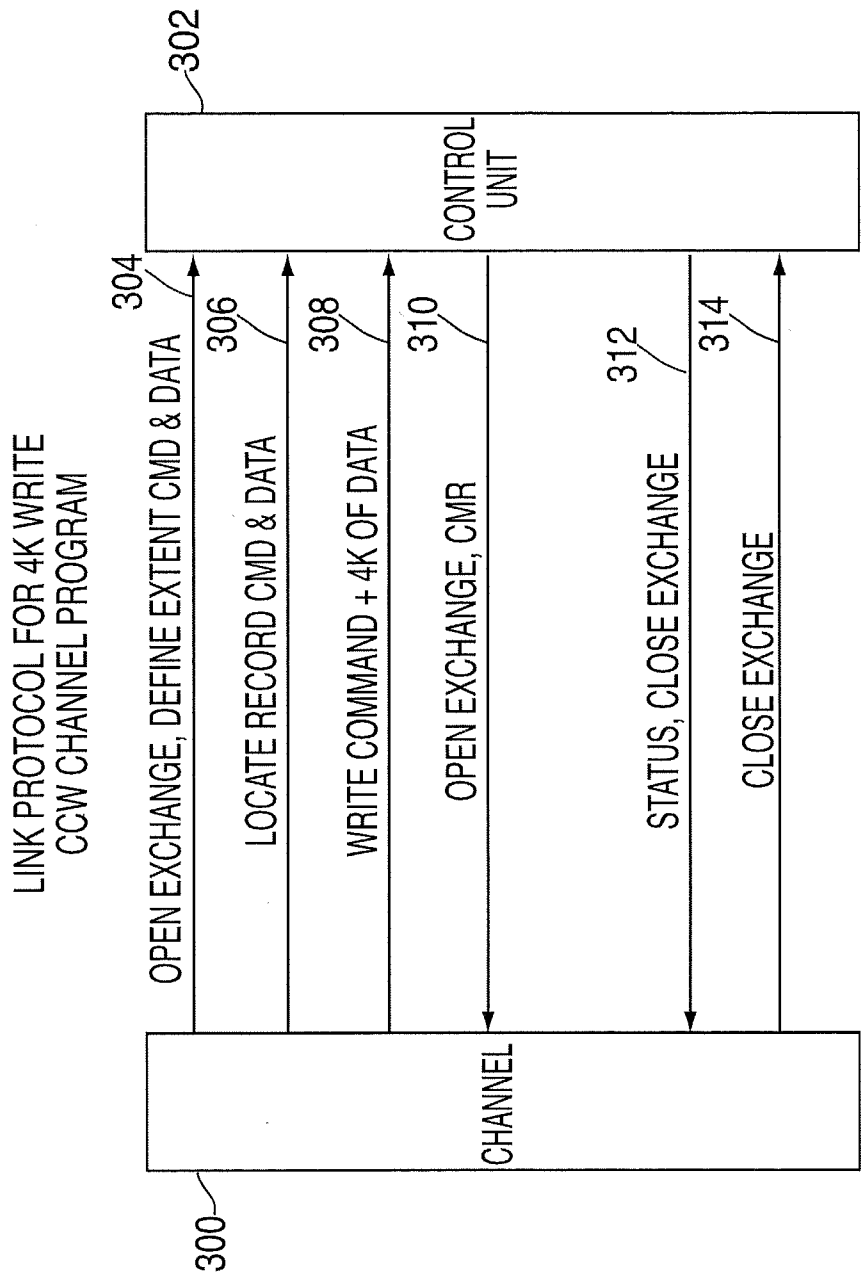


FIG. 3  
PRIOR ART

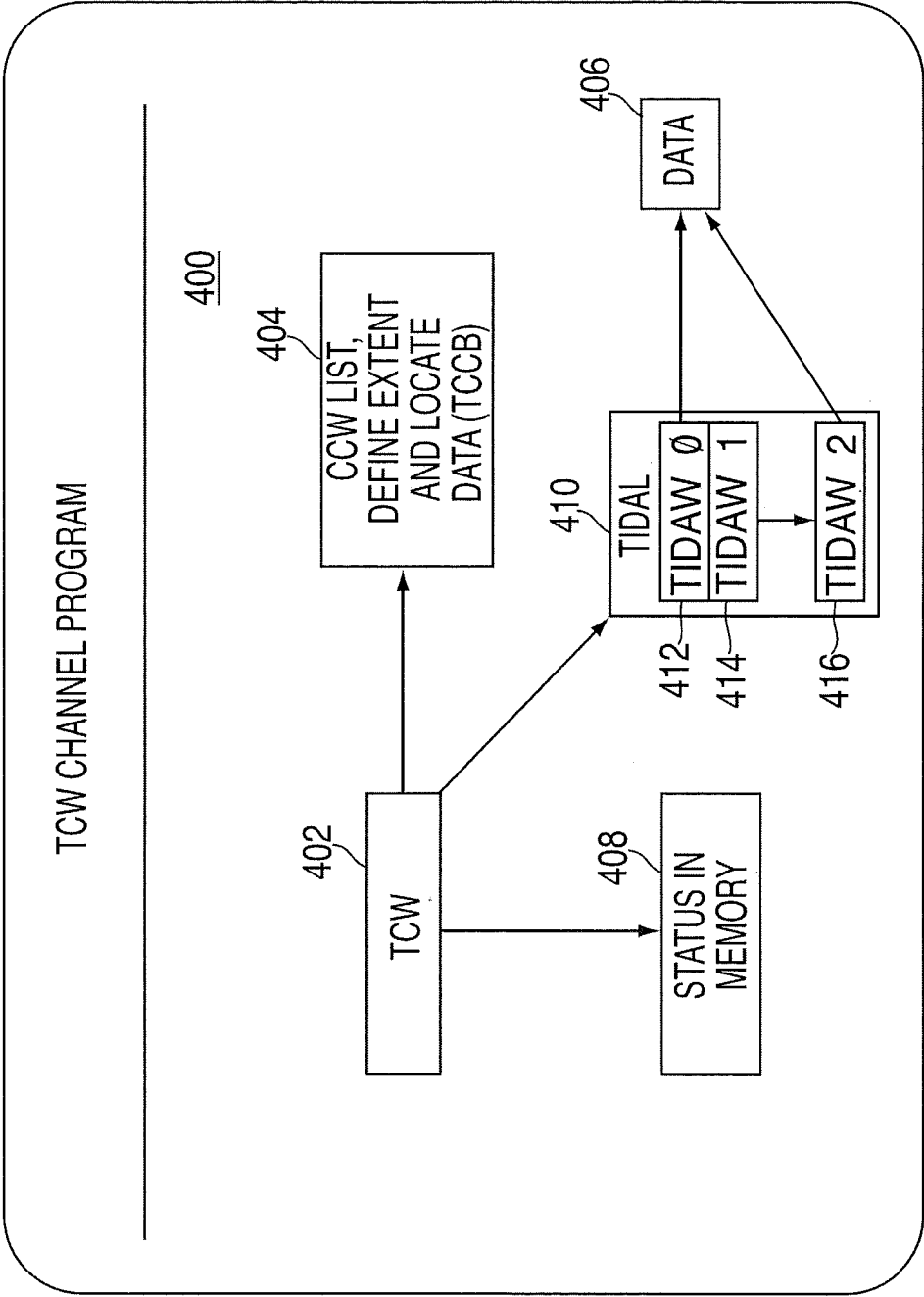


FIG. 4

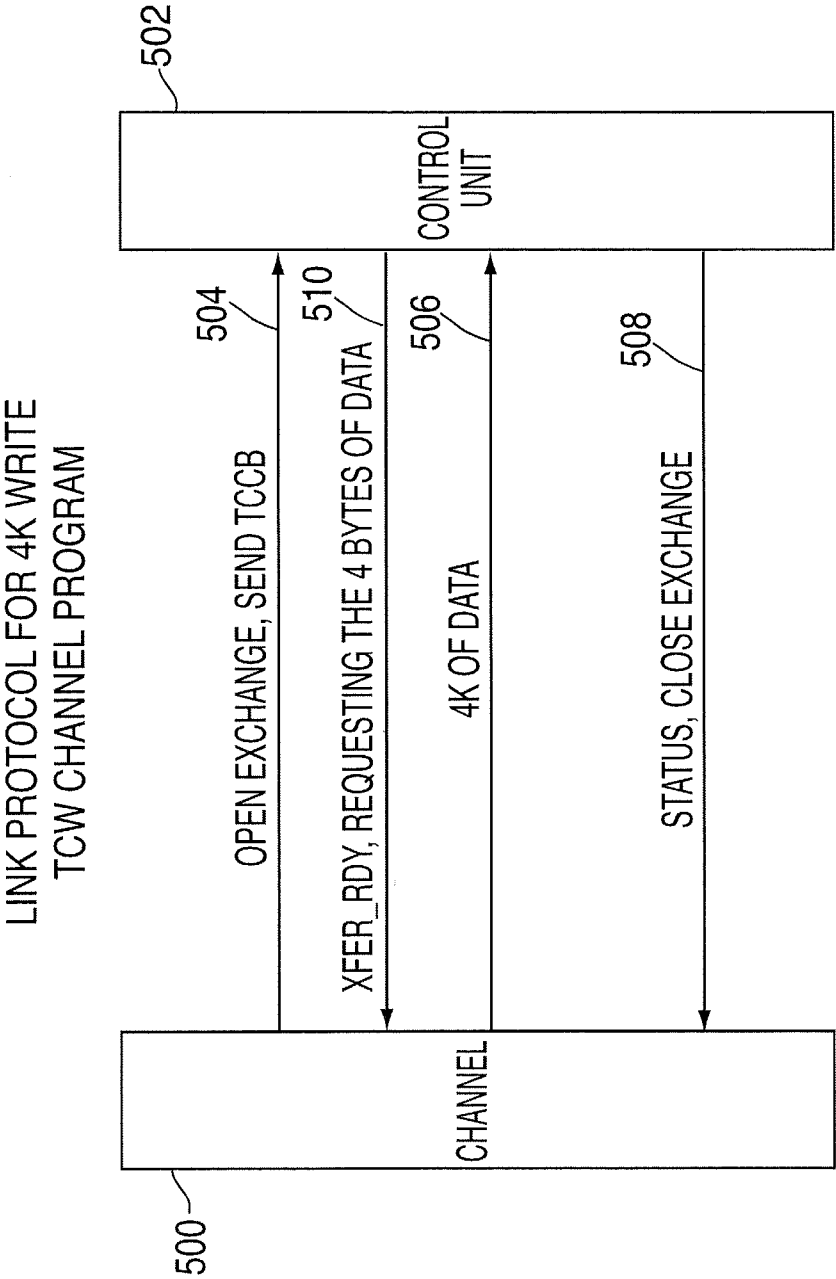


FIG. 5

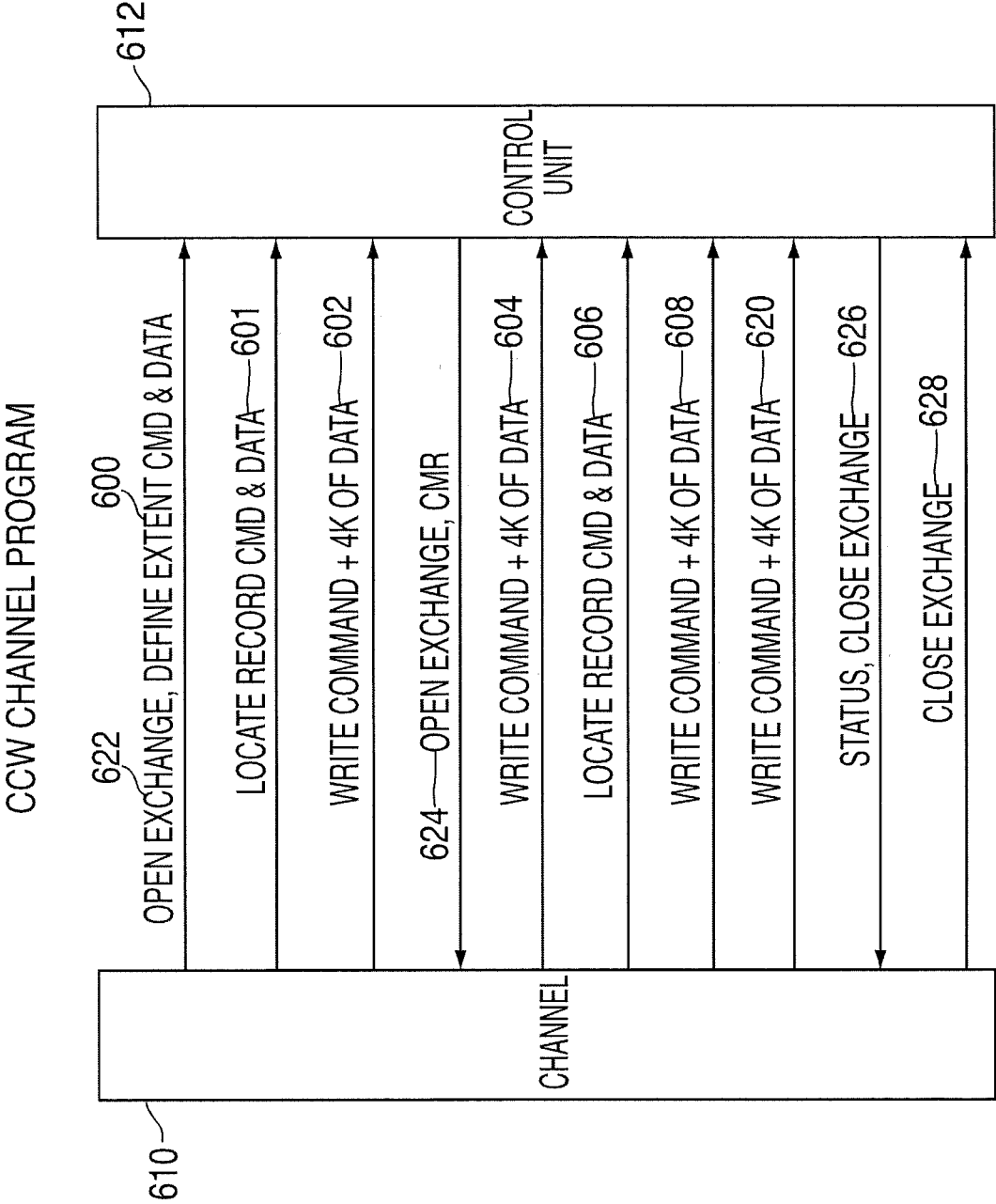


FIG. 6  
PRIOR ART

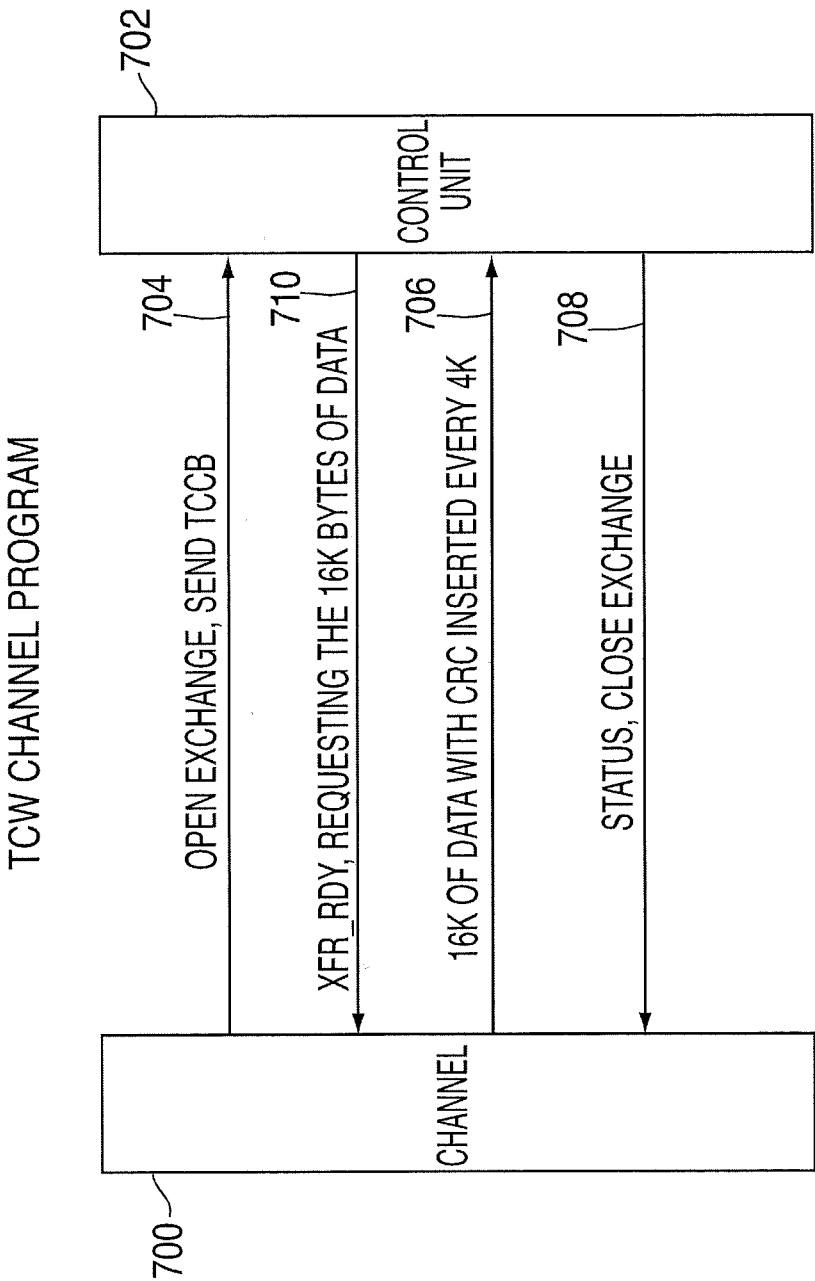


FIG. 7

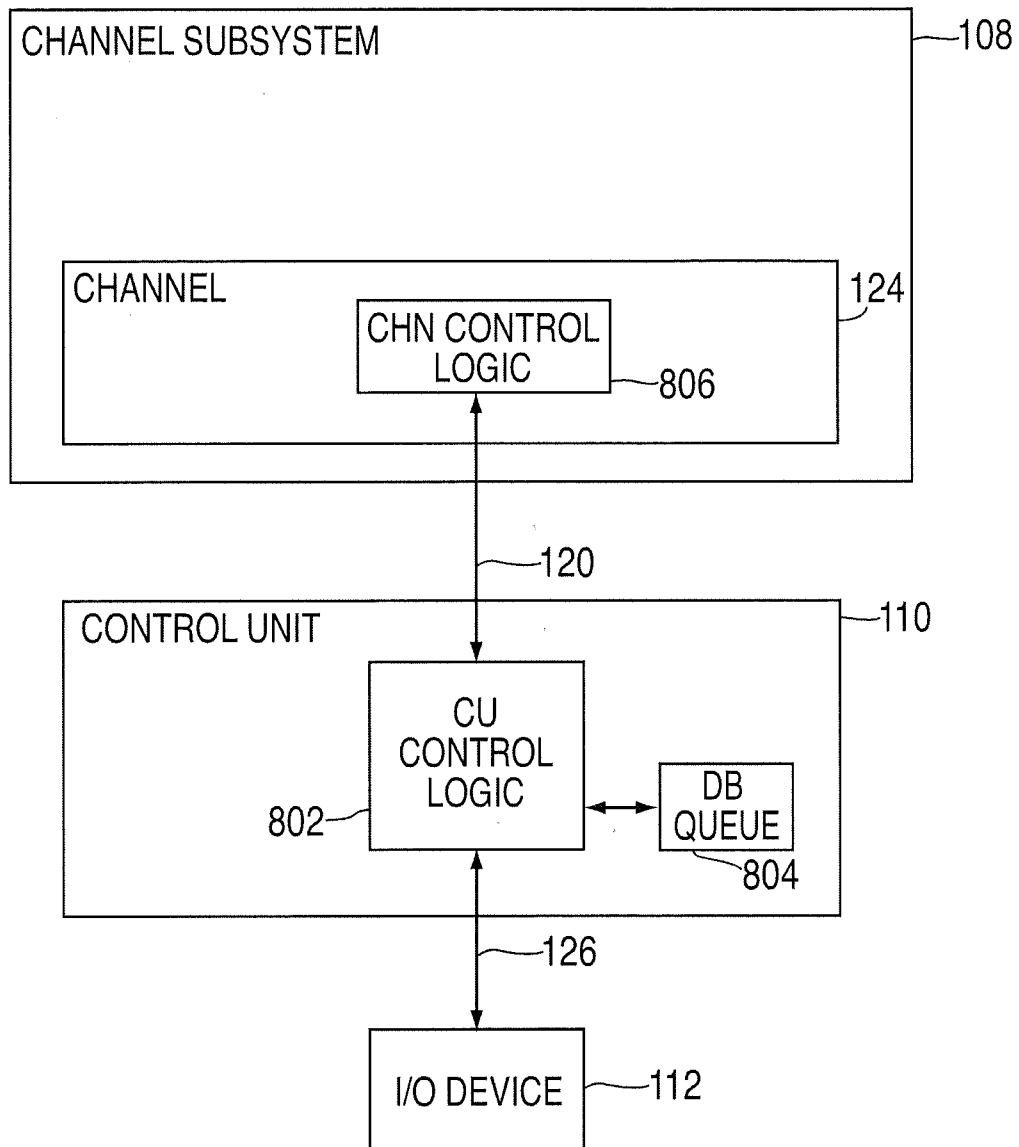


FIG. 8

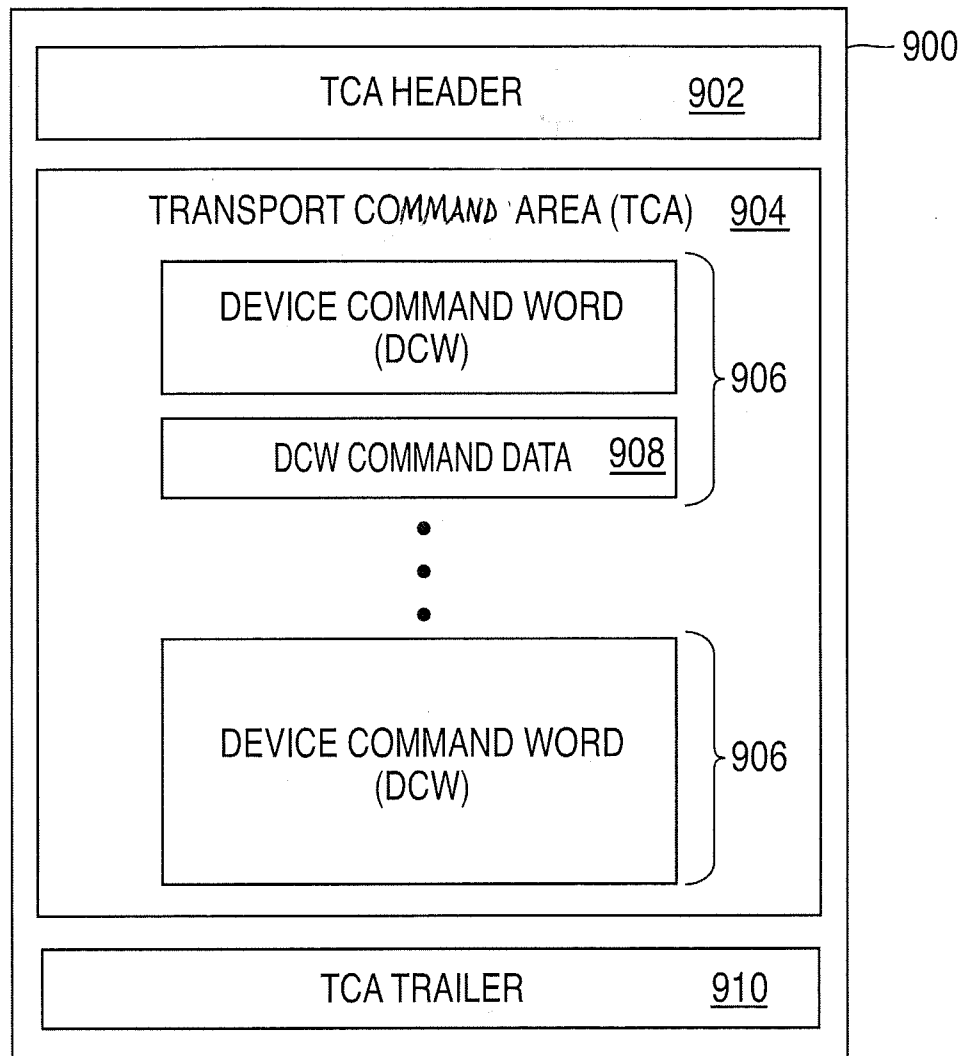


FIG. 9

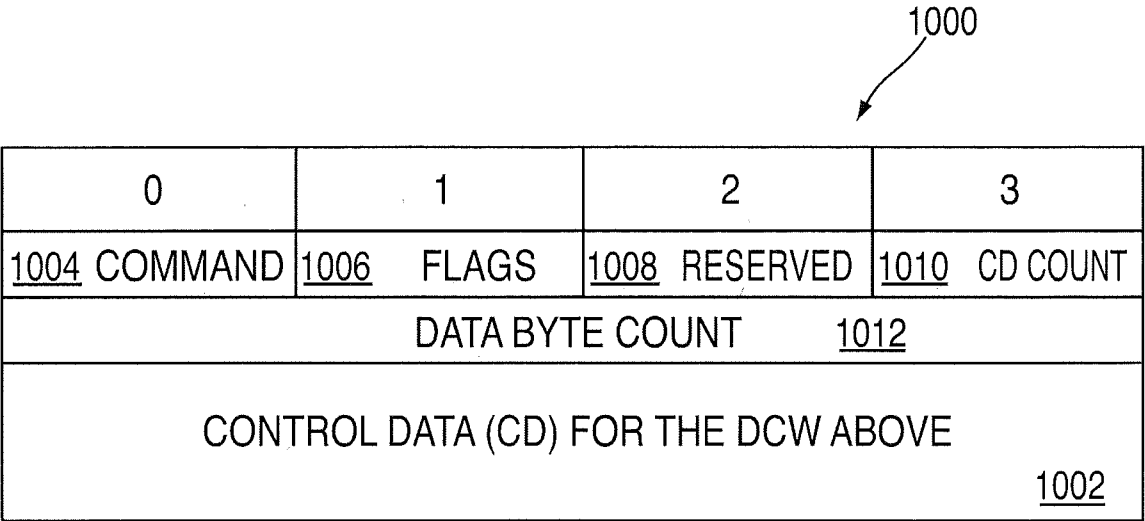


FIG. 10



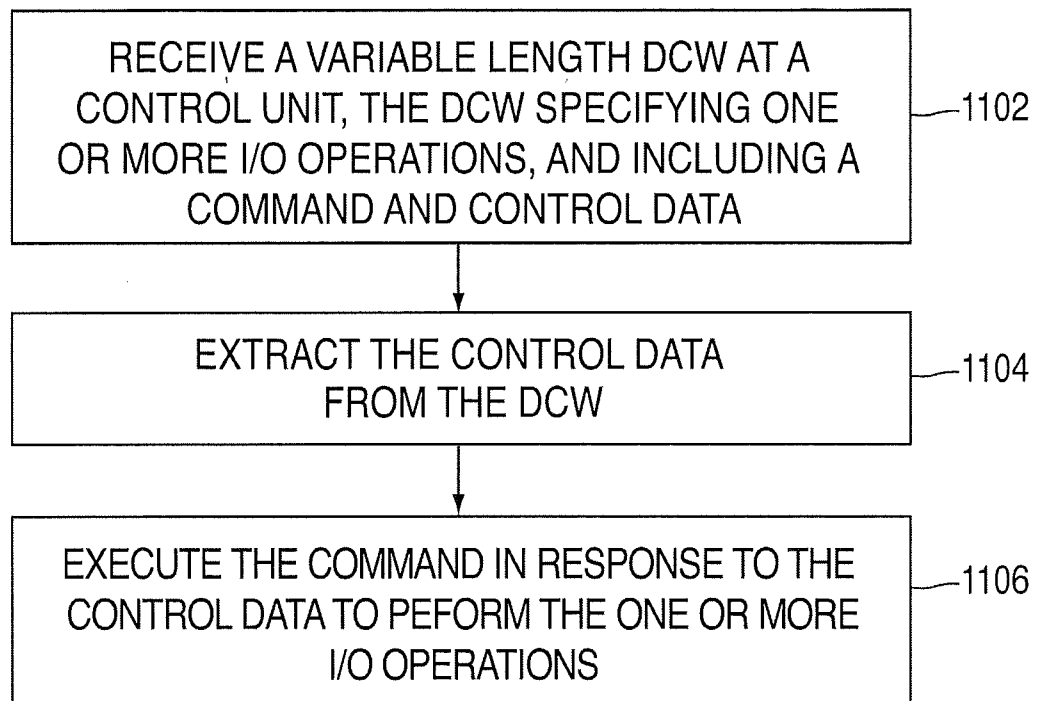


FIG. 11

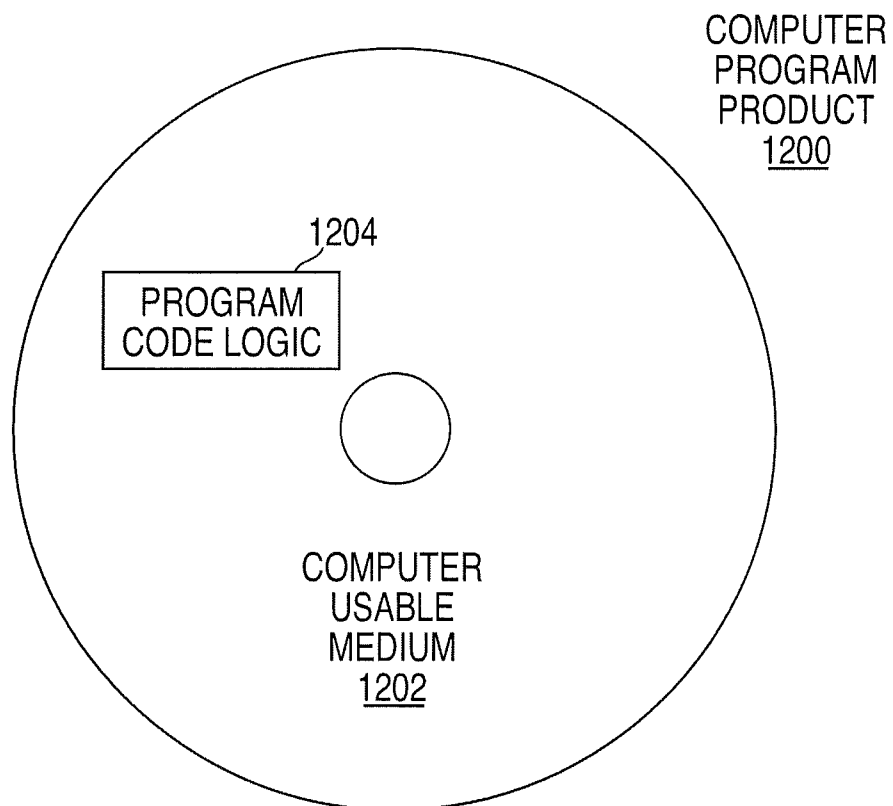


FIG. 12

## INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2009/051463

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. G06F13/10

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)  
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	WO 2006/102664 A (QUALCOMM INC [US]; CODRESCU LUCIAN [US]; PLONDKE ERICH [US]; AHMED MUH) 28 September 2006 (2006-09-28) abstract  page 2, paragraph 5 - page 5, paragraph 22 figures 1,4,5	1,5-9, 13-16,20  2-4, 10-12, 17-19
X A	US 5 500 942 A (EICKEMEYER RICHARD J [US] ET AL) 19 March 1996 (1996-03-19) the whole document	1,5-9, 13-16,20 2-4, 10-12, 17-19
A	US 2004/260851 A1 (TU CHUN-YEN [TW]) 23 December 2004 (2004-12-23) the whole document	1-20
-/--		

☒ 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 document 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.

\*&amp;\* document member of the same patent family

Date of the actual completion of the international search

14 July 2009

Date of mailing of the international search report

22/07/2009

Name and mailing address of the ISA/  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Rudolph, Stefan

# INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2009/051463

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>GB 1 264 096 A (INTERNATIONAL BUSINESS MACHINES CORP)  16 February 1972 (1972-02-16)  the whole document  -----</p>	1-20

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2009/051463

Patent document cited in search report		Publication date <sup>1</sup>	Patent family member(s)	Publication date
WO 2006102664	A	28-09-2006	CN 101176059 A US 2006218379 A1	07-05-2008 28-09-2006
US 5500942	A	19-03-1996	NONE	
US 2004260851	A1	23-12-2004	TW 271626 B	21-01-2007
GB 1264096	A	16-02-1972	BE 766468 A1 CA 935935 A1 CH 526166 A DE 2130299 A1 ES 392443 A1 FR 2095551 A5 JP 51018296 B NL 7108868 A SE 360759 B	16-09-1971 23-10-1973 31-07-1972 05-01-1972 01-10-1973 11-02-1972 09-06-1976 29-12-1971 01-10-1973