



(12) 发明专利

(10) 授权公告号 CN 101939731 B

(45) 授权公告日 2012. 10. 03

(21) 申请号 200980104698. 9

代理人 于静 杨晓光

(22) 申请日 2009. 02. 09

(51) Int. Cl.

(30) 优先权数据

G06F 13/10 (2006. 01)

12/030, 939 2008. 02. 14 US

(56) 对比文件

(85) PCT申请进入国家阶段日

WO 2006/102664 A2, 2006. 09. 28, 全文 .

2010. 08. 10

GB 1264096 A, 1972. 02. 16, 全文 .

(86) PCT申请的申请数据

US 2004/0260851 A1, 2004. 12. 23, 全文 .

PCT/EP2009/051463 2009. 02. 09

US 5500942 A, 1996. 03. 19, 全文 .

(87) PCT申请的公布数据

CN 1195821 A, 1998. 10. 14, 全文 .

W02009/101058 EN 2009. 08. 20

审查员 田越

(73) 专利权人 国际商业机器公司

地址 美国纽约

(72) 发明人 D·卡斯帕 J·弗拉纳根

M·凯洛斯 C·黄 U·恩尤库

D·莱迪 G·西格曼三世

(74) 专利代理机构 北京市中咨律师事务所

11247

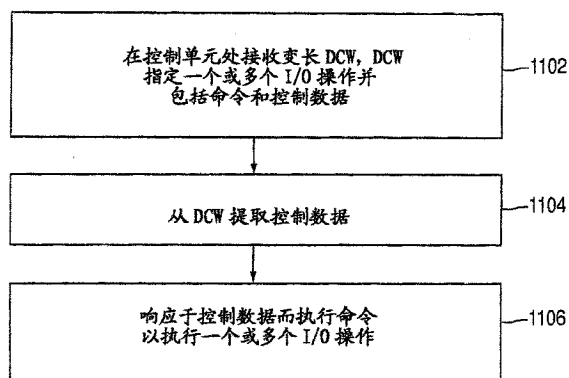
权利要求书 1 页 说明书 9 页 附图 10 页

(54) 发明名称

处理变长设备命令字的系统和方法

(57) 摘要

一种用于在输入 / 输出 (I/O) 处理系统内被配置为与 I/O 子系统通信的控制单元中处理变长设备命令字 (DCW) 的计算机程序产品、装置和方法。所述计算机程序产品包括有形存储介质, 其可由处理电路读取并存储由所述处理电路执行以便执行方法的指令。所述方法包括在所述控制单元处接收来自所述 I/O 子系统的设备命令字。所述设备命令字指定一个或多个 I/O 操作并包括命令、控制数据计数以及具有由所述控制数据计数指定的变化长度的控制数据。响应于所述控制数据计数而提取所述控制数据。然后响应于所提取的控制数据而执行所述命令以执行所指定的一个或多个 I/O 操作。



1. 一种用于在输入 / 输出 I/O 处理系统内被配置为与 I/O 子系统通信的控制单元中处理变长设备命令字 DCW 的系统,所述系统包括:

用于在所述控制单元处接收来自主计算机系统的指定一个或多个 I/O 操作的设备命令字的装置,所述设备命令字包括命令、控制数据计数以及具有由所述控制数据计数指定的变化长度的控制数据;

用于响应于所述控制数据计数而提取所述控制数据的装置;以及

用于响应于所提取的控制数据而执行所述命令以执行所指定的一个或多个 I/O 操作的装置;

其中所述设备命令字还包括指定与所述设备命令字关联的客户数据量的客户数据计数。

2. 如权利要求 1 中所述的系统,其中所述客户数据计数具有大于零的值,并且所述系统还包括用于从所述 I/O 子系统接收所述客户数据的装置。

3. 如权利要求 1 中所述的系统,其中所述客户数据计数具有大于零的值,并且所述系统还包括用于将所述客户数据传输到所述 I/O 子系统的装置。

4. 如权利要求 1 中所述的系统,还包括用于完成所述 I/O 操作并响应于所述完成而将完成消息传输到所述 I/O 子系统的装置。

5. 一种用于在输入 / 输出 I/O 处理系统内被配置为与 I/O 子系统通信的控制单元中处理变长设备命令字 DCW 的方法,所述方法包括:

在所述控制单元处接收来自主计算机系统的指定一个或多个 I/O 操作的设备命令字,所述设备命令字包括命令、控制数据计数以及具有由所述控制数据计数指定的变化长度的控制数据;

响应于所述控制数据计数而提取所述控制数据;以及

响应于所提取的控制数据而执行所述命令以执行所指定的一个或多个 I/O 操作;

其中所述设备命令字还包括指定与所述设备命令字关联的客户数据量的客户数据计数。

6. 如权利要求 5 中所述的方法,其中所述客户数据计数具有大于零的值,并且所述方法还包括从所述 I/O 子系统接收所述客户数据。

7. 如权利要求 5 中所述的方法,其中所述客户数据计数具有大于零的值,并且所述方法还包括将所述客户数据传输到所述 I/O 子系统。

8. 如权利要求 5 中所述的方法,还包括完成所述 I/O 操作并响应于所述完成而将完成消息传输到所述 I/O 子系统。

处理变长设备命令字的系统和方法

技术领域

[0001] 本发明一般地涉及输入 / 输出 (I/O) 处理, 具体地说, 涉及在 I/O 处理系统内的控制单元中处理变长设备命令字。

[0002] 背景技术

[0003] 输入 / 输出 (I/O) 操作用于在 I/O 处理系统的存储器和 I/O 设备之间传输数据。具体而言, 通过执行 I/O 操作, 将数据从存储器写入一个或多个 I/O 设备, 以及将数据从一个或多个 I/O 设备读取到存储器。

[0004] 为了便于 I/O 操作的处理, 采用 I/O 处理系统的 I/O 子系统。I/O 子系统与 I/O 处理系统的主存储器和 I/O 设备相连并在存储器和 I/O 设备之间引导信息流。I/O 子系统的实例是通道子系统。通道子系统使用通道路径作为通信介质。每个通道路径包括与控制单元相连的通道, 所述控制单元还与一个或多个 I/O 设备相连。

[0005] 所述通道子系统可以使用通道命令字 (CCW) 在 I/O 设备和存储器之间传输数据。CCW 指定要执行的 I/O 命令。对于发起特定 I/O 操作的命令, CCW 指定与所述操作关联的存储区、每当去往或来自所述存储区的传输操作完成时要采取的操作以及其他选项。

[0006] 在 I/O 处理期间, 通道从存储器取回 CCW 列表。所述通道分析 CCW 列表中的每个命令并将多个命令 (每个命令都位于其自己的实体内) 转发到与所述通道相连的控制单元。所述控制单元然后处理这些命令。所述通道跟踪每个命令的状态并控制何时将下一组命令发送到控制单元以进行处理。所述通道确保将每个命令通过其自己的实体发送到控制单元。进而, 所述通道针对每个命令推断与处理来自控制单元的响应关联的特定信息。

[0007] 在每个 CCW 的基础上执行 I/O 处理可能造成通道子系统的大量处理开销, 因为所述通道要分析 CCW, 跟踪状态信息以及对来自控制单元的响应做出反应。因此, 将与解释和管理 CCW 和状态信息关联的大量处理负担从通道子系统转移到控制单元将是有益的。简化通道在 I/O 处理系统内的控制单元与操作系统之间的通信中的作用可以因执行的握手减少而增加通信吞吐量。简化通道在通信中的作用可以包括将多个命令组合成单个 I/O 操作。

[0008] 某些命令需要控制数据和客户数据两者。当前, 单个命令字 (例如, CCW) 无法同时传输控制数据和客户数据。包括控制数据和客户数据两者的命令需要多个传输以指定控制数据和客户数据两者。这将影响可以通过组合命令获得的吞吐量的增加。可以通过为单个命令字提供传输控制数据和客户数据两者的能力来提高性能。

发明内容

[0009] 示例性实施例包括一种用于在输入 / 输出 (I/O) 处理系统内被配置为与 I/O 子系统通信的控制单元中处理变长设备命令字 (DCW) 的计算机程序产品。所述计算机程序产品包括有形存储介质, 其可由处理电路读取并存储由所述处理电路执行以便执行方法的指令。所述方法包括在所述控制单元处接收来自所述 I/O 子系统的 DCW。所述 DCW 指定一个或多个 I/O 操作并包括命令、控制数据计数以及具有由所述控制数据计数指定的变化长度

的控制数据。响应于所述控制数据计数而提取所述控制数据。然后响应于所提取的控制数据而执行所述命令以执行所指定的一个或多个 I/O 操作。

[0010] 另一示例性实施例包括一种用于在 I/O 处理系统中处理变长设备命令字 DCW 的装置。所述装置包括被配置为与 I/O 子系统通信的控制单元。所述控制单元执行方法,包括在所述控制单元处接收来自所述 I/O 子系统的 DCW。所述 DCW 指定一个或多个 I/O 操作并包括命令、控制数据计数以及具有由所述控制数据计数指定的变化长度的控制数据。响应于所述控制数据计数而提取所述控制数据。然后响应于所提取的控制数据而执行所述命令以执行所指定的一个或多个 I/O 操作。

[0011] 其他示例性实施例包括一种用于在输入 / 输出 (I/O) 处理系统内被配置为与 I/O 子系统通信的控制单元中处理变长设备命令字 (DCW) 的方法。所述方法包括在所述控制单元处接收来自所述 I/O 子系统的 DCW。所述 DCW 指定一个或多个 I/O 操作并包括命令、控制数据计数以及具有由所述控制数据计数指定的变化长度的控制数据。响应于所述控制数据计数而提取所述控制数据。然后响应于所提取的控制数据而执行所述命令以执行所指定的一个或多个 I/O 操作。

[0012] 通过参考以下附图和详细说明,根据实施例的其他制品、装置和 / 或方法对于本领域技术人员而言将是显而易见的或将变得显而易见。所有此类其他制品、装置和 / 或方法均旨在包括在本说明中、位于本发明的范围内,以及受所附权利要求的保护。

附图说明

[0013] 在说明书结尾处的权利要求中具体指出并明确要求保护被视为本发明的主题。从下面结合附图的详细说明,本发明的上述和其他目标、特性和优点是显而易见的,这些附图是:

[0014] 图 1 示出了结合和使用本发明的一个或多个方面的 I/O 处理系统的一个实施例;

[0015] 图 2A 示出了现有技术通道命令字的一个实例;

[0016] 图 2B 示出了现有技术通道命令字通道程序的一个实例;

[0017] 图 3 示出了用于通道和控制单元之间的通信以执行图 2B 的通道命令字通道程序的现有技术链路协议的一个实施例;

[0018] 图 4 示出了根据本发明的一个方面的传输控制字 (TCW) 通道程序的一个实施例;

[0019] 图 5 示出了根据本发明的一个方面的用于通道和控制单元之间的通信 (未禁用 XFER_RDY) 以执行图 4 的 TCW 通道程序的链路协议的一个实施例;

[0020] 图 6 示出了用于通道和控制单元之间的通信以执行通道命令字通道程序的四个写入命令的现有技术链路协议的一个实施例;

[0021] 图 7 示出了根据本发明的一个方面的用于通道和控制单元之间的通信 (未禁用 XFER_RDY) 以处理 TCW 通道程序的四个写入命令的链路协议的一个实施例;

[0022] 图 8 示出了根据本发明的一个方面的控制单元和通道子系统的实施例;

[0023] 图 9 示出了根据本发明的一个方面的 TCCB 的一个实施例;

[0024] 图 10 示出了根据本发明的一个方面的 DCW 的一个实施例;

[0025] 图 11 示出了根据本发明的一个方面的用于提供变长 DCW 的过程的一个实施例;以及

[0026] 图 12 示出了结合本发明的一个或多个方面的制品的一个实施例。

[0027] 详细描述通过实例的方式参考附图解释了本发明的优选实施例以及优点和特性。

具体实施方式

[0028] 根据本发明的一个方面,通过允许单个命令字包括控制数据并且还传输客户数据来促进输入/输出(I/O)。向命令字添加 1 字节计数字段以定义命令所需的控制数据的字节计数。因此,命令字包括两个计数字段,一个用于控制数据而另一个用于客户数据。在一个示例性实施例中,变长控制数据紧跟命令字。

[0029] 这通过减少用于执行 I/O 处理的 I/O 处理系统的组件之间的通信而促进了 I/O 处理。例如,减少了在诸如通道之类的 I/O 通信适配器和控制单元之间的交换和序列数。这通过作为由控制单元执行的单个实体(例如,作为同一控制字的一部分)将控制数据和客户数据发送到控制单元来实现。

[0030] 将多个命令(例如,设备命令字或“DCW”)包括在此处称为传输命令控制块(TCCB)的块中,在传输控制字(TCW)中指定所述块的地址(间接或直接地址)。在一个示例性实施例中,将 TCW 从操作系统(OS)或其他应用发送到 I/O 通信适配器,后者又在命令消息中将 TCCB 转发到控制单元进行处理。所述控制单元处理每个命令,但是不通过 I/O 通信适配器跟踪与这些命令相关的状态。所述多个命令也被称为通道程序,在控制单元而不是在 I/O 通信适配器上分析和执行所述通道程序。

[0031] 参考图 1 描述了结合和使用本发明的一个或多个方面的 I/O 处理系统的一个实例。I/O 处理系统 100 包括主系统 101,其进一步例如包括主存储器 102、一个或多个中央处理单元(CPU) 104、存储控制元件 106 以及通道子系统 108。主系统 101 可以是大规模计算系统,例如大型机或服务器。I/O 处理系统 100 还包括一个或多个控制单元 110 和一个或多个 I/O 设备 112,将在下面逐一进行描述。

[0032] 主存储器 102 存储数据和程序,所述数据和程序可以从 I/O 设备 112 输入。例如,主存储器 102 可以包括由一个或多个 CPU 104 执行的一个或多个操作系统(OS) 103。例如,一个 CPU 104 可以执行 **Linux®**操作系统 103 和 **z/OS®**操作系统 103 作为不同的虚拟机实例。主存储器 102 可以直接寻址并通过 CPU 104 和通道子系统 108 提供高速数据处理。

[0033] CPU 104 是 I/O 处理系统 100 的控制中心。它包含用于指令执行、中断操作、计时功能、初始程序加载以及其他机器相关功能的排序和处理工具。CPU 104 通过诸如双向或单向总线之类的连接 114 与存储控制元件 106 相连。

[0034] 存储控制元件 106 通过诸如总线之类的连接 116 与主存储器 102 相连;通过连接 114 与 CPU 104 相连;以及通过连接 118 与通信子系统 108 相连。存储控制元件 106 例如控制 CPU 104 和通信子系统 108 所发出的请求的排队和执行。

[0035] 在一个示例性实施例中,通道子系统 108 提供主系统 101 和控制单元 110 之间的通信接口。如上所述,通道子系统 108 与存储控制元件 106 相连,并且通过诸如串行链路之类的连接 120 与每个控制单元 110 相连。连接 120 可以实现为光链路,在光纤通道结构(例如,光纤通道网络)中采用单模或多模光导。通道子系统 108 在 I/O 设备 112 和主存储器 102 之间引导信息流。它免除了 CPU 104 直接与 I/O 设备 112 通信的任务并且允许数据处

理与 I/O 处理同时进行。通道子系统 108 使用一个或多个通道路径 122 作为通信链路来管理去往或来自 I/O 设备 112 的信息流。作为 I/O 处理的一部分,通道子系统 108 还执行测试通道路径可用性、选择可用通道路径 122 以及启动与 I/O 设备 112 的操作的执行的途径管理功能。

[0036] 每个通道路径 122 包括通道 124(在一个实例中,通道 124 位于通道子系统 108 内,如图 1 所示)、一个或多个控制单元 110 以及一个或多个连接 120。在另一实例中,也可以具有一个或多个动态开关(未示出)作为通道路径 122 的一部分。动态开关与通道 124 和控制单元 110 相连并提供将任意两个与开关连接的链路进行物理互连的能力。在另一实例中,还可以具有多个系统,从而具有多个与控制单元 110 连接的通道子系统(未示出)。

[0037] 通道子系统 108 中还包括子通道(未示出)。通过通道子系统 108 为程序可访问的每个 I/O 设备 112 提供一个子通道并且该子通道专用于该程序可访问的每个 I/O 设备 112。子通道(例如,诸如表之类的数据结构)为程序提供设备的逻辑形式。每个子通道提供有关关联的 I/O 设备 112 以及所述 I/O 设备与通道子系统 108 的连接的信息。所述子通道还提供有关 I/O 操作以及涉及关联的 I/O 设备 112 的其他功能的信息。子通道是通道子系统 108 将有关关联的 I/O 设备 112 的信息提供给 CPU 104 时所用的手段,CPU 104 通过执行 I/O 指令获取此信息。

[0038] 通道子系统 108 与一个或多个控制单元 110 相连。每个控制单元 110 提供操作和控制一个或多个 I/O 设备 112 的逻辑并通过使用公共工具使每个 I/O 设备 112 的特性适合通道 124 提供的链路接口。所述公共工具提供 I/O 操作的执行、有关 I/O 设备 112 和控制单元 110 的状态指示、在通道路径 122 上数据传输的计时,以及特定级别的 I/O 设备 112 控制。

[0039] 每个控制单元 110 通过连接 126(例如,总线)与一个或多个 I/O 设备 112 相连。I/O 设备 112 接收信息或将信息存储在主存储器 102 和 / 或其他存储器中。I/O 设备 112 的实例包括读卡器和穿孔器、磁带单元、直接 存取存储设备、显示器、键盘、打印机、指点设备、远程处理设备、通信控制器和基于传感器的设备,等等。

[0040] 在“ **IBM®** z/Architecture Principles of Operation(**IBM®** z/Architecture 操作原理,出版号 SA22-7832-05,第六版,2007 年 4 月)” ;Cormier 等人的标题为“System For Transferring Data Between I/O Devices And Main Or Expanded Storage Under Dynamic Control Of Independent Indirect Address Words(IDAWS)(用于在独立间接地址字(IDAWS)的动态控制下在 I/O 设备与主 / 扩展存储装置间传输数据的系统,1995 年 10 月 24 日公告)” 的美国专利第 5,461,721 号 ;以及 Casper 等人的标题为“Method And System For Pipelining The Processing Of Channel Command Words(用于通道命令字流水线处理的方法和系统,1996 年 6 月 11 日公告)” 的美国专利第 5,526,484 号(所有这些文献的全部内容都在此引入作为参考)中进一步描述了 I/O 处理系统 100 的上述组件中的一个或多个。IBM 是位于美国纽约阿蒙克的国际商业机器公司的注册商标。此处所用的其他名称可能是国际商业机器公司或其他公司的注册商标、商标或产品名。

[0041] 在一个实施例中,为了在 I/O 设备 112 与存储器 102 之间传输数据,将使用通道命令字 (CCW)。CCW 指定要执行的命令,并包括其他控制处理的字段。将参考图 2A 描述 CCW 的一个实例。CCW 200 例如包括指定要执行的命令(例如,读取、回读、控制、读出和写入)

的命令代码 202 ;多个用于控制 I/O 操作的标志 204 ;对于指定数据传输的命令,包括指定 CCW 200 所指定的存储区中要传输的字节数的计数字段 206 ;以及当采用直接寻址时,指向主存储器中包括数据的位置,或者当采用修改后的间接数据寻址时,指向要处理的修改后的间接数据地址字 (MIDAW) 列表 (例如,连续列表) 的数据地址 208。在 Brice 等人的标题为 “Flexibly Controlling The Transfer Of Data Between Input/Output Devices And Memory (灵活控制输入 / 输出设备和存储器间的数据传输, 2006 年 8 月 15 日申请)” 的美国申请第 11/464,613 号 (其全部内容在此引入作为参考) 中进一步描述 了修改后的间接寻址。

[0042] 排列为顺序执行的一个或多个 CCW 形成一个通道程序,所述通道程序在此也称为 CCW 通道程序。CCW 通道程序例如由操作系统或其他软件建立。所述软件建立 CCW 并获取分配给通道程序的存储器地址。将参考图 2B 描述 CCW 通道程序的一个实例。CCW 通道程序 210 例如包括定义范围 CCW 212,其具有指向要与定义范围命令一起使用的定义范围数据 216 的存储器中的某个位置的指针 214。在该实例中,通道转移 (TIC) 218 在定义范围命令之后,通道转移 (TIC) 218 使通道程序引用存储器中包括一个或多个其他 CCW (如具有指向查找记录数据 220 的指针 219 的查找记录 217) 以及一个或多个写入 CCW 221 的另一区域 (例如应用区域)。每个写入 CCW 221 具有指向数据区域 224 的指针 222。所述数据区域包括直接访问数据的地址或间接访问数据的数据地址字列表 (例如, MIDAW 或 IDAW)。进而, CCW 通道程序 210 包括通道子系统中由被称为状态 226 的子通道 (从执行 CCW 通道程序产生) 的设备地址所定义的预定区域。

[0043] 参考图 3 以及参考图 2B 描述 CCW 通道程序的处理。具体而言,图 3 示出了当执行 CCW 通道程序时在通道与控制单元之间发生的各种交换和序列的一个实例。在该实例中,用于通信的链路协议为 FICON (光纤连通性)。在 “Fibre Channel Single Byte Command Code Sets-3 Mapping Protocol (FC-SB-3) (光纤通道单字节命令代码集 -3 映射协议 (FC-SB-3), T11/ 项目 1357-D/ 修订版 1.6, INCITS, 2003 年 3 月,其全部内容在此引入作为参考)” 中说明了有关 FICON 的信息。

[0044] 现在参考图 3,通道 300 打开与控制单元 302 的交换并将与交换关联的定义范围命令和数据 304 发送到控制单元 302。从定义范围 CCW 212 (图 2B) 取回所述命令,并从定义范围数据区域 216 获取所述数据。通道 300 使用 TIC 218 定位查找记录 CCW 和写入 CCW。它从查找记录 CCW 217 (图 2B) 和查找记录数据 220 取回查找记录命令和数据 306 (图 3)。从写入 CCW 221 和数据区域 224 (图 2B) 取回写入命令和数据 308 (图 3)。将它们都发送到控制单元 302。

[0045] 响应于通道 300 的打开交换,控制单元 302 打开与通道 300 的交换 310。这可以在查找命令和数据 306 和 / 或写入命令和数据 308 之前或之后发生。伴随着打开交换,将响应 (CMR) 转发到通道 300。所述 CMR 向通道 300 提供控制单元 302 处于活动状态并正在运行的指示。

[0046] 控制单元 302 将状态提供给通道 300 并关闭交换 312。响应于此操作,通道 300 存储所述数据、检查状态并关闭交换 314,以此向控制单元 302 指示已收到所述状态。

[0047] 以上 CCW 通道程序写入 4k 数据的处理需要打开和关闭两个交换以及六个序列。通过将通道程序的多个命令折叠成 TCCB,将减少通道与控制单元之间的交换和序列的总数。

通道（例如，图 1 的通道 124）使用 TCW 标识 TCCB 的位置以及用于访问和存储与执行通道程序关联的状态和数据的位置。所述 TCW 由通道解释，并且控制单元不发送或查看所述 TCW。

[0048] 参考图 4 描述了如图 2B 中所示但包括 TCCB 而不是各个单独 CCW 的用于写入 4k 数据的通道程序的一个实例。如图所示，通道程序 400（在此称为 TCW 通道程序）包括 TCW 402，TCW 402 指定存储器中 TCCB 404 的位置以及存储器中数据区域 406 或指向数据区域 406 的 TIDAL 410（例如，传输模式间接数据地址字（TIDAW）列表，类似于 MIDAW）以及状态区域 408 的位置。TIDAW 0412、TIDAW 1414 和 TIDAW 2416（统称为 TIDAW 412-416）可以引用数据区域 406 中的不同位置以获取或存储数据。TIDAW 412-416 可以引用非连续数据块或连续数据块。TIDAL 410 中的 TIDAW 412-416 可以按顺序位于存储器中或相对于彼此不连续地位于存储器中。虽然 TIDAL 410 中仅示出三个 TIDAW 412-416，但将理解的是，TIDAL 410 中可以包括任何数量的 TIDAW。

[0049] 参考图 5 描述了 TCW 通道程序的处理。用于这些通信的链路协议例如为光纤通道协议（FCP）。具体而言，使用 FCP 链路协议的三个阶段，从而允许使用支持 FCP 的主机总线适配器以执行由 CCW 控制的数据传输。在“Information Technology-Fibre Channel Protocol for SCSI, Third Version (FCP-3)（信息技术-用于 SCSI 的光纤通道协议第三版（FCP-3），T10 项目 1560-D，修订版 4，2005 年 9 月 13 日，其全部内容在此引入作为参考）”中进一步说明了 FCP 及其各个阶段。

[0050] 当启用 XFER_RDY 时，使用图 5 中示出的链路协议的实施例。在图 5 中示出的实施例中，直到控制单元 502 经由 XFER_RDY IU 510 请求数据 506 时，通道 500 才能将数据 506 发送到控制单元 502。在备选示例性实施例中，禁用 XFER_RDY 并且所述控制单元不将 XFER_RDY IU 510 传输到通道 500。因此，通道 500 在发送数据 506 之前不必等待控制单元 502 请求数据 506。当通道 500 和控制单元 502 在地理上彼此相距很远（例如，大于二十千米、大于五十千米）时，可以使用禁用了 XFER_RDY 的该备选实施例以提高性能。除非另行指定，否则此处的讨论都假设启用 XFER_RDY。

[0051] 在进一步的实例中，为了写入 4k 客户数据，通道 500 使用 FCP 链路协议的各个阶段，如下所示：

[0052] 1. 在 FCP_CMND IU 中将 TCCB 和序列主动性（sequence initiative）传输到控制单元 502。

[0053] 2. 等待控制单元准备好接收数据 506 的 XFER_RDY IU 指示。

[0054] 3. 将数据 IU 和序列主动性传输到控制单元 502。

[0055] 4. 在 FCP 状态帧中发送最终状态，所述 FCP 状态帧例如在 FCP_RSP IU 有效负载的字节 10 或 11 中具有一个活动位。使用 FCP_RSP_INFO 字段或读出字段传输 FICON 结束状态以及附加状态信息。

[0056] 通过执行图 4 的 TCW 通道程序，仅打开和关闭一次交换（参阅图 5），而不是图 2B 的 CCW 通道程序的两次交换（参阅图 3）。此外，与 CCW 通道程序的六个序列相比（参阅图 2B-3），对于 TCW 通道程序而言，存在四个通信序列（参阅图 4-5）。

[0057] 对于 TCW 通道程序，即使向此程序添加附加命令，交换和序列数也保持不变。例如，将图 6 的 CCW 通道程序的通信与图 7 的 TCW 通道程序的通信进行比较。在图 6 的 CCW 通道程序中，在单独序列中将每个命令（例如，定义范围命令和数据 600、查找记录命令和

数据 601、写入命令和数据 602、写入命令和数据 604、查找记录命令和数据 606、写入命令和数据 608 以及写入命令和数据 620) 从通道 610 发送到控制单元 612。此 CCW 通道程序需要打开和关闭两次交换 (例如, 打开交换 622、624 和关闭交换 626、628) 以及十个通信序列。将其与图 7 的 TCW 通道程序的四个序列和一次交换相比, 图 7 的 TCW 通道程序与图 6 的 CCW 通道程序完成相同的任务。

[0058] 如图 7 所示, 通道 700 打开与控制单元 702 的交换并将 TCCB 704 发送到控制单元 702。如上所述, TCCB 704 在 DCW 中包括定义范围命令、两个查找记录命令以及四个写入命令。收到 XFER_RDY IU 710 之后, 通道 700 在单个序列中将 16k 数据 706 传输到控制单元 702。通道 700 可以在序列中的 16k 数据 706 中每 4K 插入一个 CRC。每 4K 插入一个 CRC 允许控制单元 702 在完成 TCCB 704 中的写入命令之前增量地检验 16K 数据, 而不是缓冲整个 16K 以便进行检验。此外, 控制单元 702 将状态提供给通道 700 并关闭交换 708。因此, 图 7 的 TCW 通道程序需要少得多的通信即可传输与图 6 的 CCW 通道程序相同的数据量, 同时通过在来自通道 700 的输出数据流中插入多个 CRC 而支持增量数据检验。

[0059] 现在转到图 8, 其中更详细地示出了图 1 中支持 TCW 通道程序执行的通道子系统 108 中的通道 124 和控制单元 110 的一个实施例。控制单元 110 包括控制单元控制逻辑 802, 控制单元控制逻辑 802 用于分析和处理经由连接 120 从通道 124 接收的包含 TCCB (如图 7 的 TCCB 704) 的命令消息。控制单元控制逻辑 802 可以从在控制单元 110 处接收的 TCCB 中提取 DCW 和控制数据, 以便经由连接 126 控制诸如 I/O 设备 112 之类的设备。控制单元控制逻辑 802 将设备命令和数据发送到 I/O 设备 112 并从 I/O 设备 112 接收状态信息和其他反馈。例如, I/O 设备 112 可能因为针对 I/O 设备 112 的先前保留请求而忙碌。为了管理当控制单元 110 接收到多个访问同一 I/O 设备 112 的请求时可能引起的潜在设备保留竞争问题, 控制单元控制逻辑 802 跟踪设备忙碌消息和关联的数据并将其存储在设备忙碌队列 804 中。在一个示例性实施例中, 图 1 的 OS 103 保留 I/O 设备 112 以便在保留处于活动状态时阻止其他 OS 103 访问 I/O 设备 112。虽然并非所有 I/O 操作都需要设备保留, 但是可以使用设备保留支持需要在固定持续时间内进行独占访问的操作, 例如磁盘格式化。

[0060] 控制单元 110 还可以包括其他缓冲或存储元件 (未示出) 以存储与通道 124 和 I/O 设备 112 之间的通信关联的多个消息或状态信息。例如, 位于控制单元 110 中的寄存器可以包括最大控制单元交换参数, 其定义了控制单元 110 所支持的打开控制单元交换的最大数量。

[0061] 通道子系统 108 中的通道 124 包括支持与控制单元 110 通信的元件。在一个示例性实施例中, 通道控制逻辑 806 控制通道子系统 108 和控制单元 110 之间的通信。通道控制逻辑 806 可以经由连接 120 直接与控制单元控制逻辑 802 通过接口连接以发送命令和接收响应, 例如传输命令和响应 IU。备选地, 可以在通道控制逻辑 806 和控制单元控制逻辑 802 之间布置消息传送接口和 / 或缓冲器 (未示出)。

[0062] 图 9 示出了根据本发明的一个方面的 TCCB 900 的一个实施例。图 9 中的 TCCB 900 位于由 TCW 指示的地址处。该地址可以是直接地址, 也可以是间接地址, 从而允许 TCCB 900 的内容位于一个存储位置中或分布在多个非连续存储位置间。如上所述, TCCB 900 是由软件生成的控制块, 然后通道 124 将其发送到控制单元 110 (例如, 在传输命令 IU 中) 以便执行。TCCB 900 包含要由控制单元 110 执行的命令以及所述命令所需的任何控制数据。通

道 124 不会查看 TCCB 900 的内容。通道 124 将 TCCB 900 打包并将其发送到控制单元 110。这允许使用 FCP 传输协议而非 FICON。

[0063] TCCB 900 包括传输控制区域头 (TCAH) 902, 在一个示例性实施例中, 传输控制区域头 (TCAH) 902 包括有关传输命令区域 (TCA) 904 和 TCA 904 内的操作的信息 (例如, 长度、服务代码)。图 9 中示出的 TCCB 900 还包括变长 TCA 904, 变长 TCA 904 包括一个或多个 DCW 906, 每个 DCW 906 都可包括 DCW 控制数据 908。在一个示例性实施例中, DCW 控制数据 908 具有可变长度。在一个示例性实施例中, 每个 DCW 906 包括命令标志 (链)、控制数据长度以及读 / 写数据长度。DCW 控制数据 908 是可选的 (取决于 DCW 906) 并且包括用于 DCW 906 的控制参数。例如, DCW 控制数据 908 可以包括定义范围和 / 或前缀参数。在一个示例性实施例中, DCW 控制数据位于 TCA 904 内的 DCW 906 的结尾并且 DCW 906 不指向 DCW 控制数据。最后, TCCB 900 包括 TCA 尾 (TCAT) 910, TCA 尾 (TCAT) 910 包含诸如要在 TCCB 900 中传输的字节计数之类的数据。

[0064] 图 10 示出了根据本发明的一个方面的 DCW 1000 的一个实施例。在一个示例性实施例中, DCW 1000 长度为 8 个字节再加上控制数据 1002 的长度。DCW 包括命令字段 1004、标志字段 1006、保留字段 1008、控制数据 (CD) 计数字段 1010 以及数据字节计数字段 1012。DCW 命令字段 1004 长度为 1 个字节并且与 CCW 中使用的 CCW 命令字节相同 (但可以包括 CCW 未使用的附加命令代码)。标志字段 1006 包括 8 个位; 在一个示例性实施例中, 第二位是指向 TCA 904 中的下一 DCW 1000 的链命令。当此标志位被设为 0 时, 指示这是 TCA 904 中的 DCW 程序的最后一个 DCW 1000。标志字段 1006 中的其他位被保留并被设为 0。

[0065] CD 计数字段 1010 是包含由此 DCW 1000 使用的控制数据 1002 的字节计数的 1 字节字段。控制数据 1002 被添加到 DCW 1000 的结尾。在一个示例性实施例中, 当存在控制数据 1002 时, 在 DCW 1000 之间使用的空间是填充到下一 4 字节边界的控制数据 1002 的大小, 以便 DCW 1000 始终在 4 字节边界上开始。数据字节计数字段 1012 是 4 字节的 DCW 字节计数字段。这是客户数据 (例如, 图 4 中的数据 406) 的 DCW 计数。在一个示例性实施例中, 此计数不包括要由此 DCW 命令传输的填充字节或 CRC 字节。此 DCW 字节计数字段 1012 与 CCW 中的计数字段具有相同的含义, 只是它可以更大。如果 DCW 命令为立即命令或非操作命令 (例如, 不包含任何客户数据), 则数据字节计数字段 1012 等于 0。

[0066] 在一个示例性实施例中, 控制数据 1002 位于同一 TCA 904 中与控制数据 1002 关联的 DCW 1000 的结尾。在一个示例性实施例中, 控制数据 1002 被填充到 TCA 904 中的下一 4 字节边界之外。控制数据计数字段 1010 反映控制数据 1002 的字节数。在一个示例性实施例中, TCA 904 中的控制数据 1002 的最大长度为 232 个字节, 假设具有一个 DCW 1000 并且 TCA 904 的其余部分被此 DCW 1000 的控制数据占据。

[0067] 图 11 示出了根据本发明的一个方面用于提供变长 DCW 的过程的一个实施例。在方块 1102, 控制单元接收变长 DCW 1000。所述 DCW 1000 指定由命令 1004 定义的多个等效的遗留 I/O 操作 (例如, 定义范围、查找记录以及读取操作或写入操作等)、控制数据计数 1010 以及变长控制数据 1002。控制数据的长度由控制数据计数指定。在方块 1104, 使用所述控制数据计数从所述 DCW 提取控制数据 1002 以确定要将 DCW 1000 结尾处的多少个字节包括在控制数据中。在方块 1106, 所述控制单元使用所述控制数据作为要执行的输入来执行所述命令。所述执行导致执行多个等效的遗留 I/O 操作。如果所述 DCW 还包括客户数据

(例如,数据字节计数 1012 为非零),则响应于所述控制数据和所述客户数据而执行所述命令。此外,所述控制单元可以将所述命令的完成状态(例如,已成功完成、错误条件等)传输到主计算机系统。

[0068] 因此,再次参考图 2B 和 10,可以使用本发明的一个示例性实施例将定义范围 CCW 212、查找记录 CCW 217 和写入 CCW 221 组合成单个 DCW。此 DCW 是组合式命令,意味着:定义范围、查找记录和写入客户数据。在此示例性实施例中,范围和查找记录参数数据是作为 DCW 一部分的控制数据,其长度由控制数据计数 1010 定义。客户数据长度由数据字节计数 1012 控制。在同一 DCW 中,所传输的客户数据只能是写入(输出)或读取(输入),而不能同时是这两者。CCW 内的命令的这种组合无法在当前实施方式中执行,因为使控制数据(定义范围参数和查找记录参数)到达控制单元的唯一方式是使通道将这些参数作为数据传输到控制单元。

[0069] 示例性实施例的技术效果包括在同一 DCW 内传输客户数据和控制数据的能力。这可导致通过需要较少的交换而执行相同数量的功能来提高性能。

[0070] 如上所述,各实施例可以体现在计算机实现的过程以及用于实现这些过程的装置的形式中。在一个示例性实施例中,本发明体现在由一个或多个网络元件执行的计算机程序代码中。各实施例包括计算机可用介质 1202 上的如图 12 中所示的计算机程序产品 1200,其中计算机程序代码逻辑 1204 包含在有形介质中体现为制品的指令。计算机可用介质 1202 的示例性制品可包括软盘、CD-ROM、硬盘、通用串行总线(USB)闪存或其他任何计算机可读存储介质,其中,当计算机程序代码逻辑 1204 加载到计算机上并由所述计算机执行时,所述计算机变为用于实现本发明的装置。各实施例包括计算机程序代码逻辑 1204,例如,无论是存储在存储介质中,加载到计算机上和/或由所述计算机执行,还是通过某些传输介质传输,例如通过电连线或线缆,通过光纤或通过电磁辐射,其中,当计算机程序代码逻辑 1204 加载到计算机上并由所述计算机执行时,所述计算机变为用于实现本发明的装置。当在通用微处理器上实现时,计算机程序代码逻辑 1204 配置所述微处理器以创建专用逻辑电路。

[0071] 尽管参考示例性实施例对本发明进行了说明,但是本领域的技术人员将理解,可以在不偏离本发明范围的情况下做出各种变化并且使用等价的元件替换其中的元件。此外,可以做出许多修改以便在不偏离本发明的基本范围的情况下使特定的情况或材料符合本发明的教导。因此,本发明并非旨在限于被构想为执行本发明的最佳模式的所披露的特定实施例,本发明还包括所有处于所附权利要求的范围之内的实施例。而且,使用术语“第一”、“第二”等不表示任何顺序性或重要性,术语“第一”、“第二”等用于对元件进行区分。另外,使用术语“一”、“一个”等不表示数量限制,而是表示存在至少一个所引用的项。

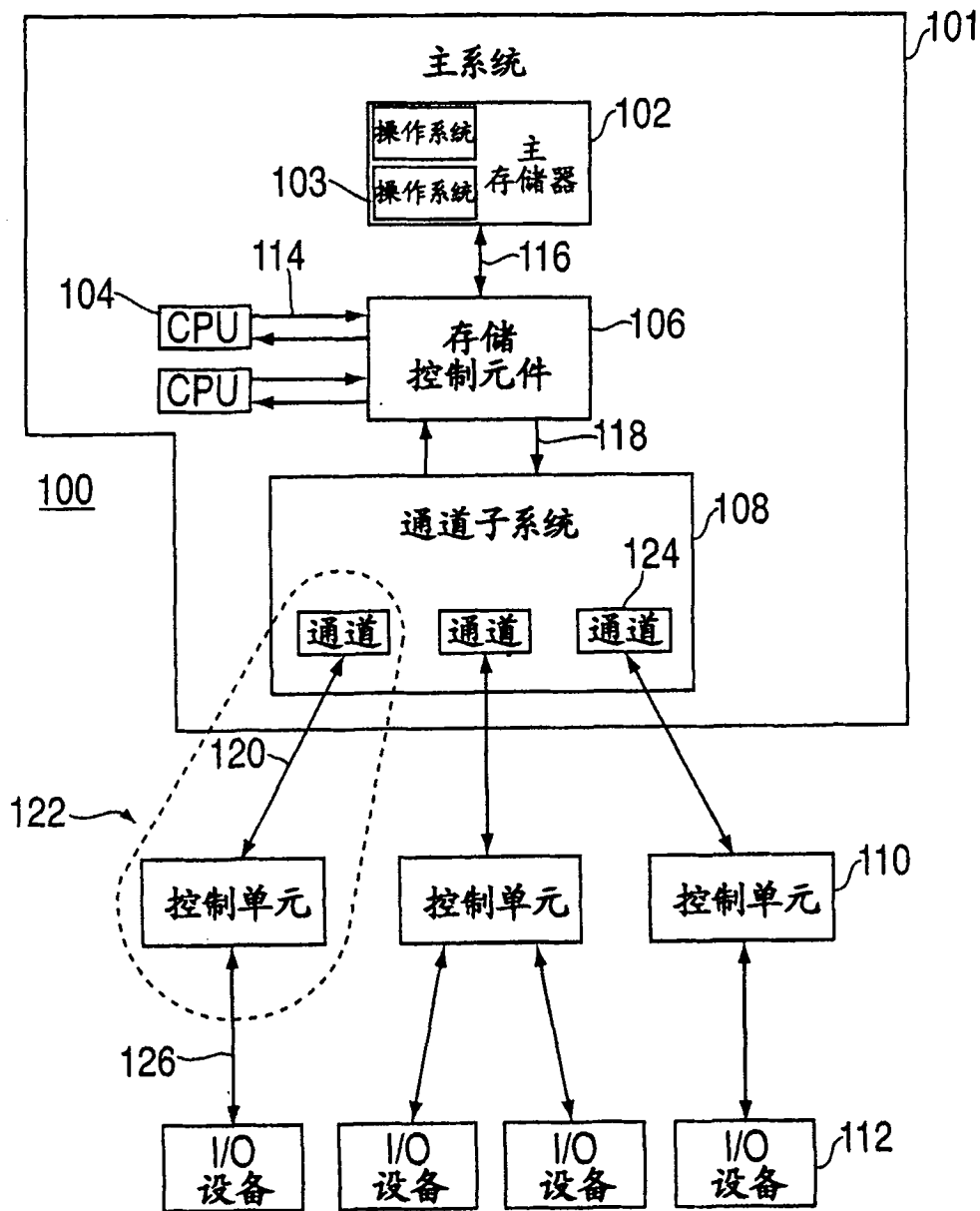


图 1

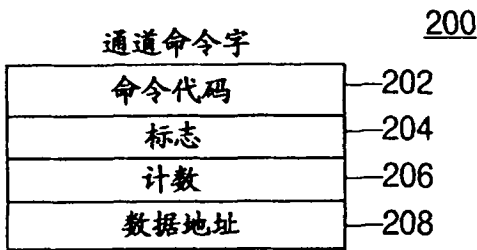


图 2A 现有技术

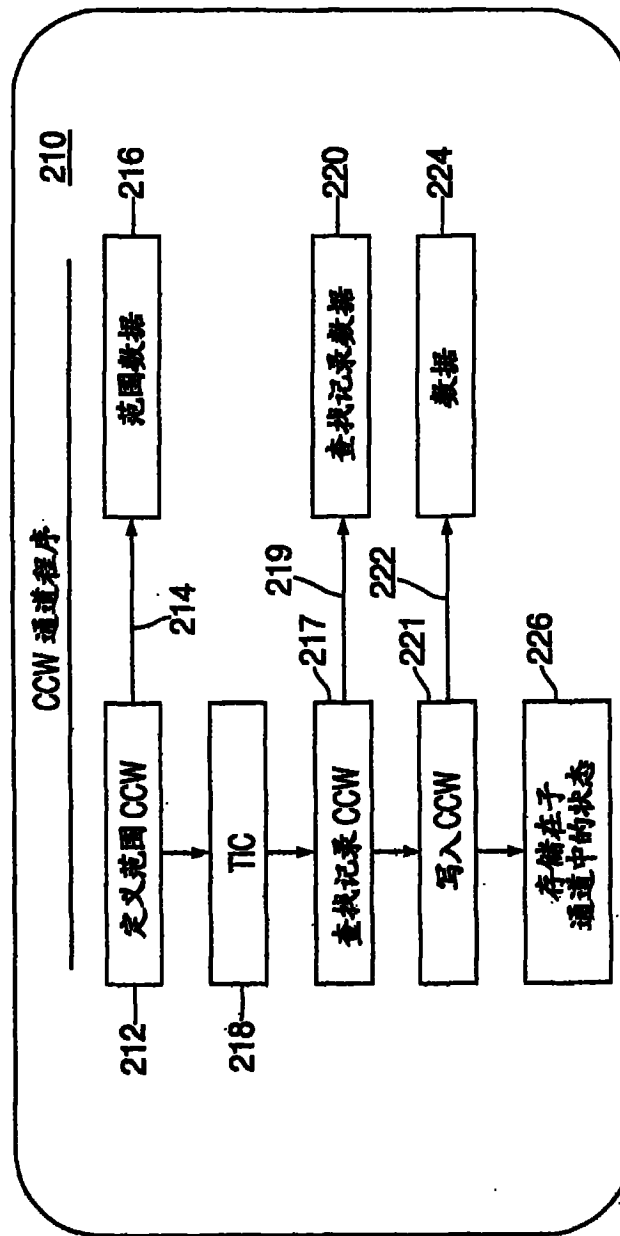


图 2B 现有技术

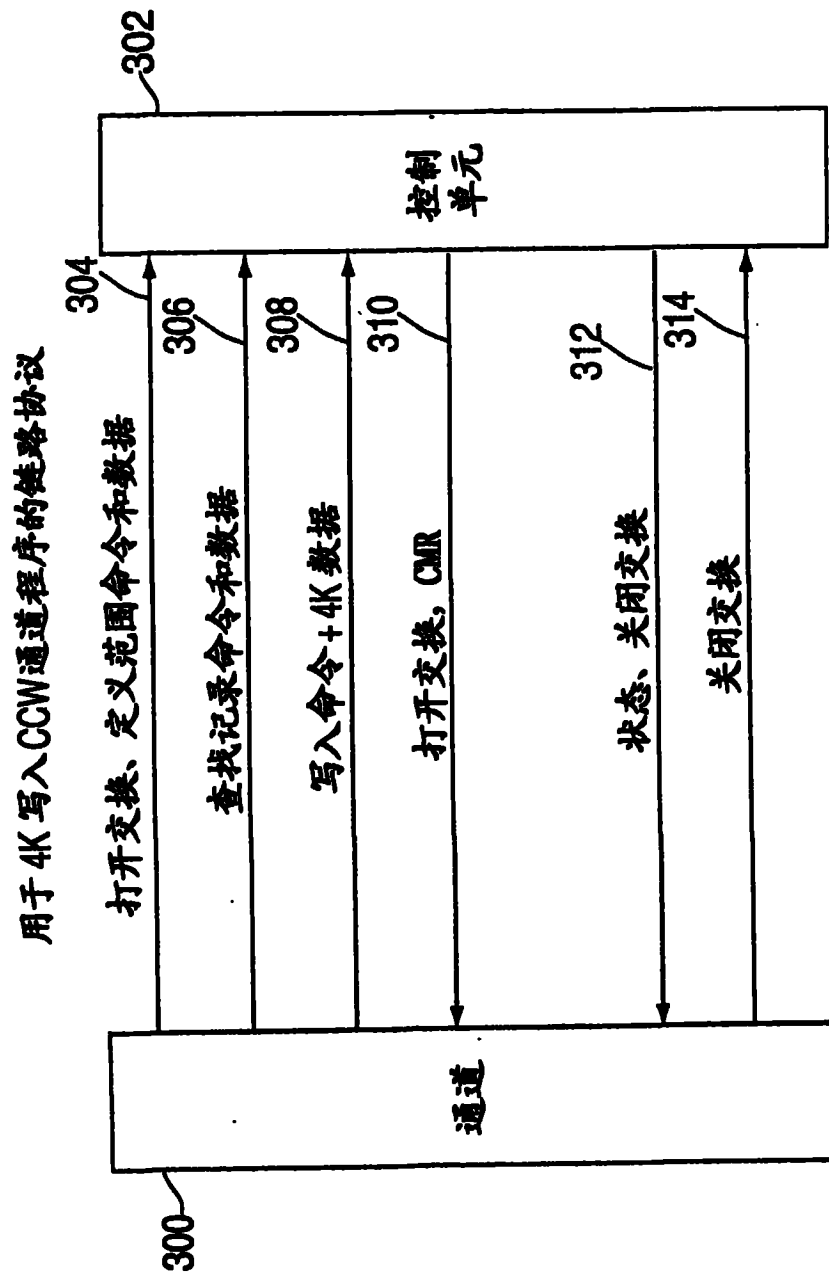


图3 现有技术

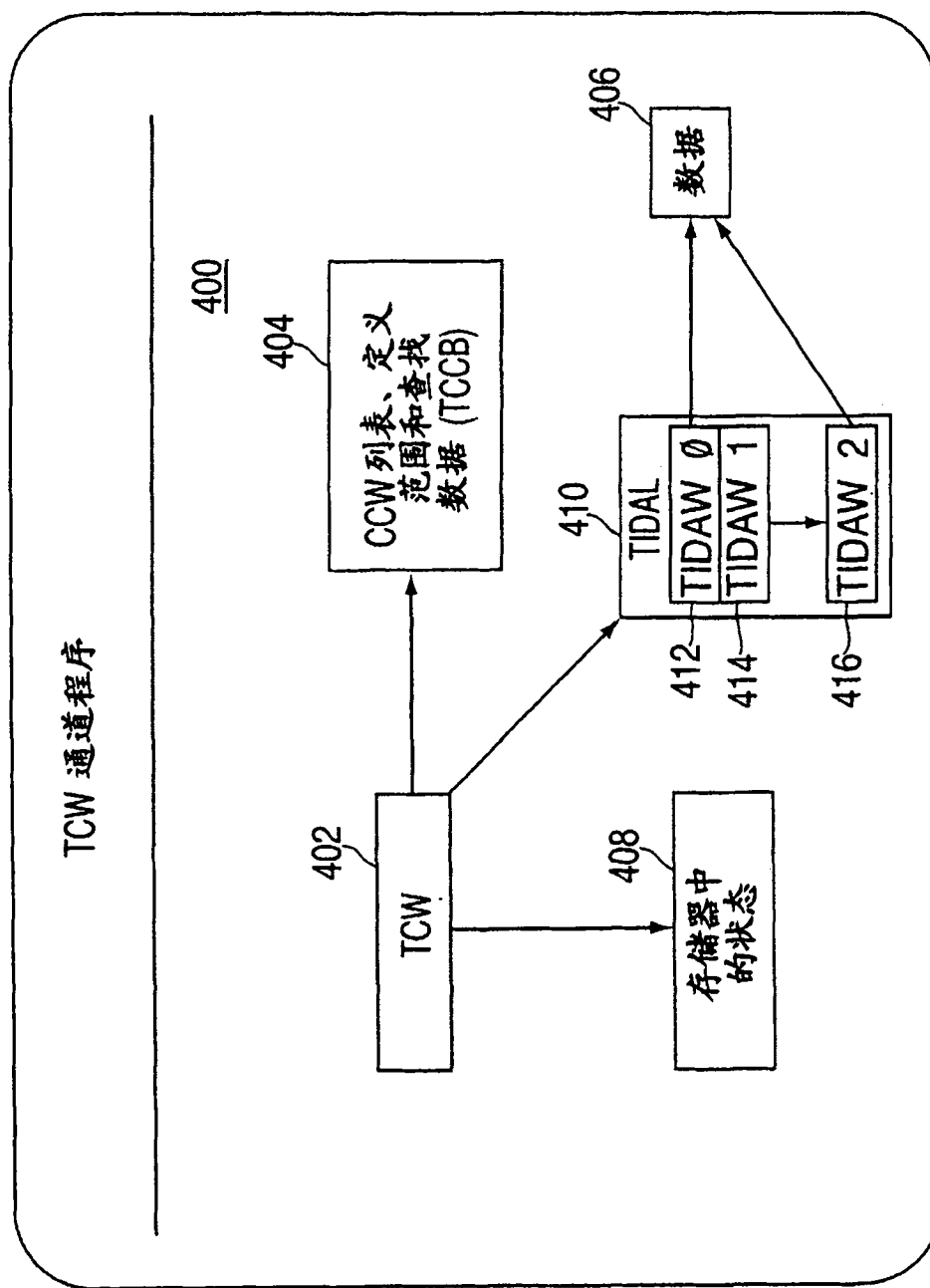


图 4

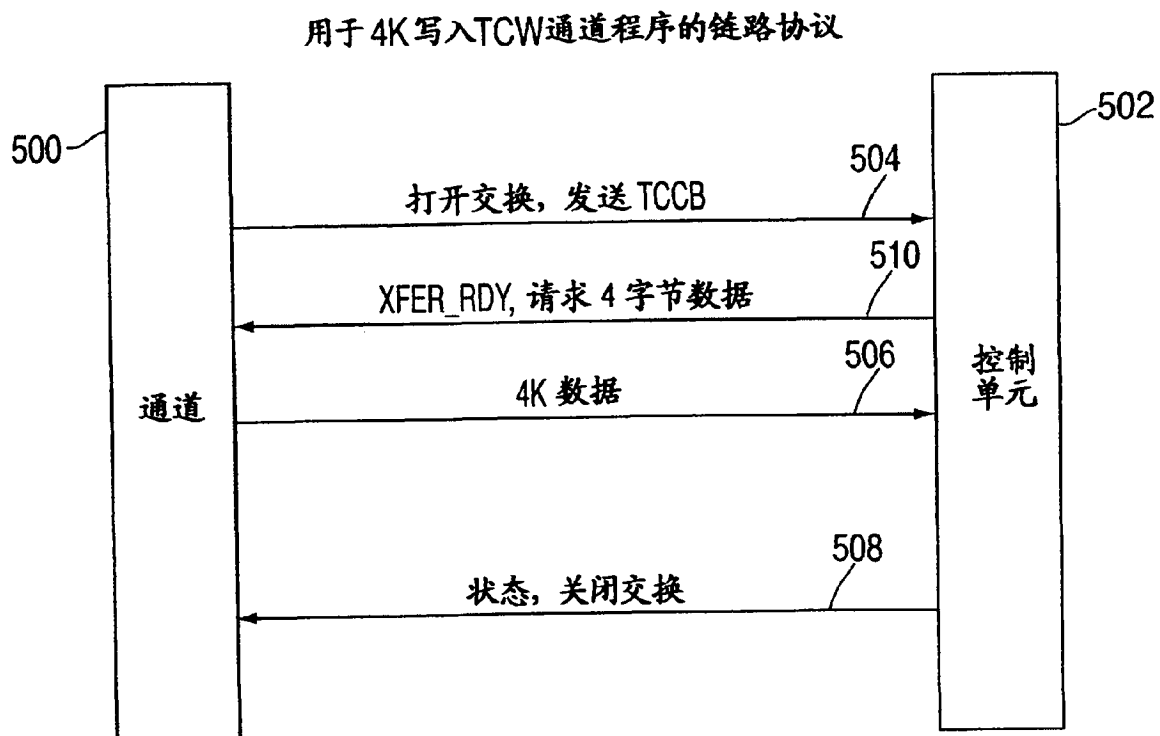


图 5

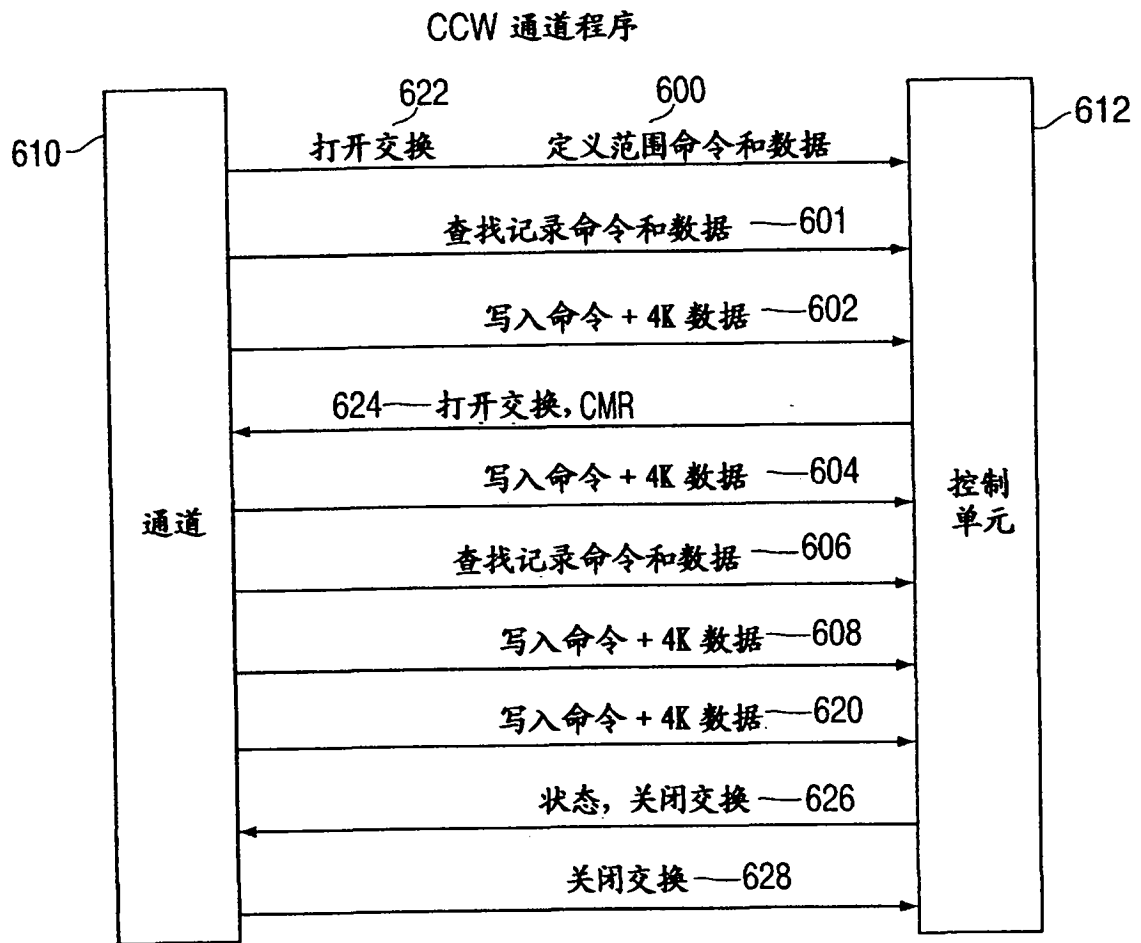


图 6 现有技术

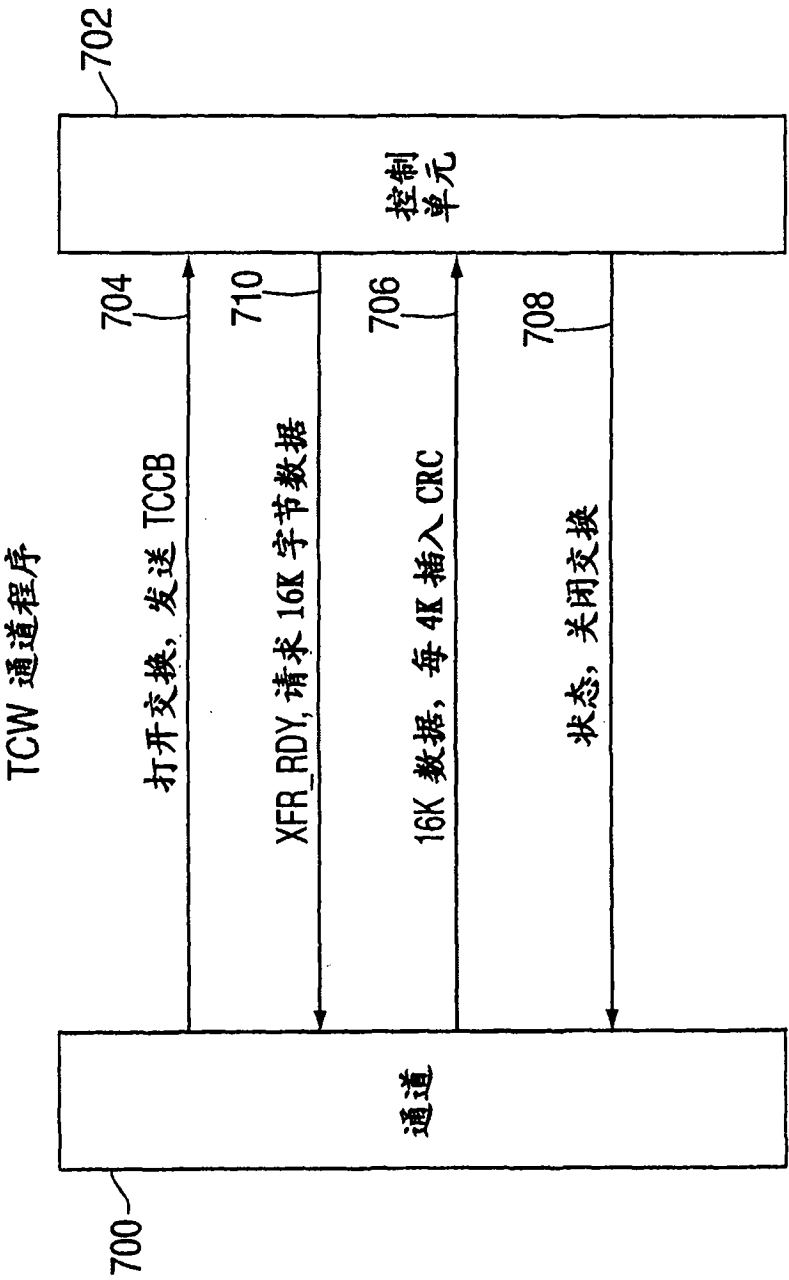


图 7

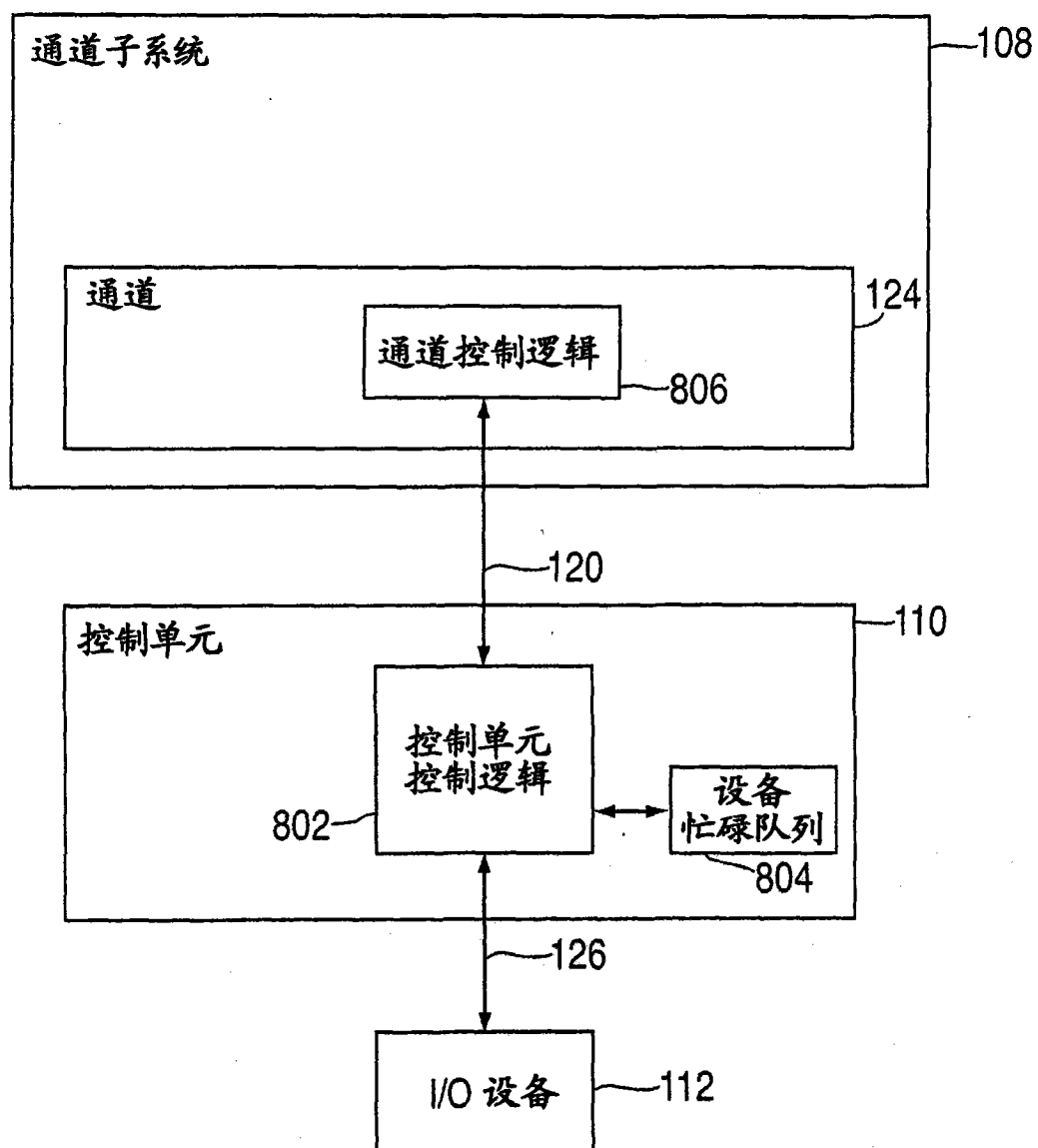


图 8

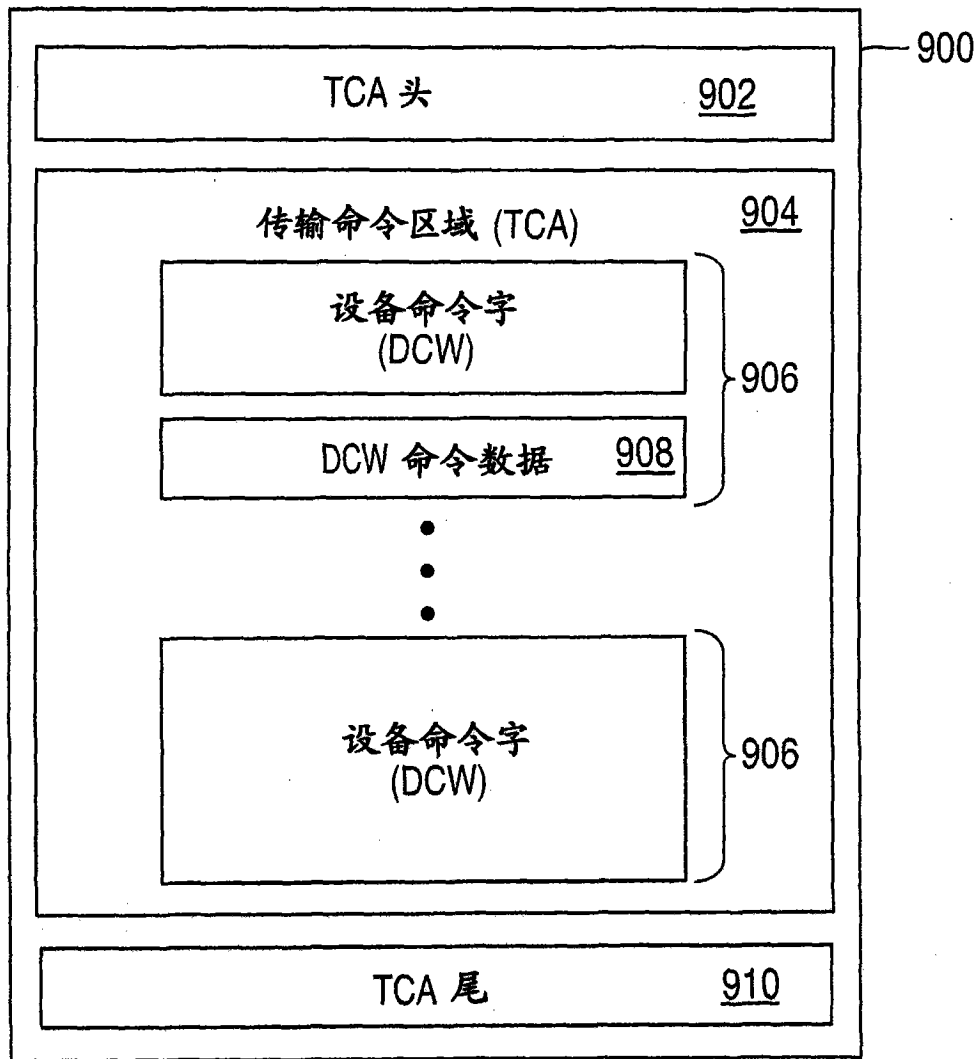


图 9

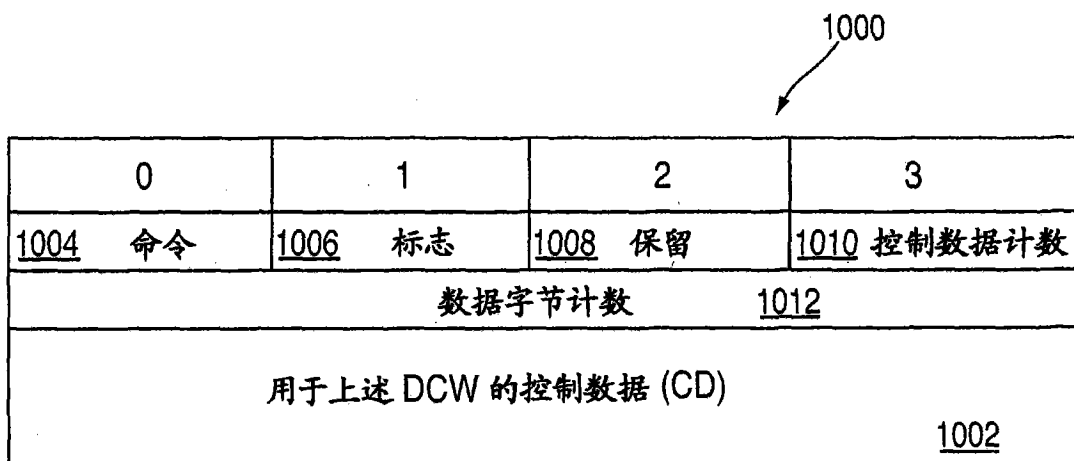


图 10

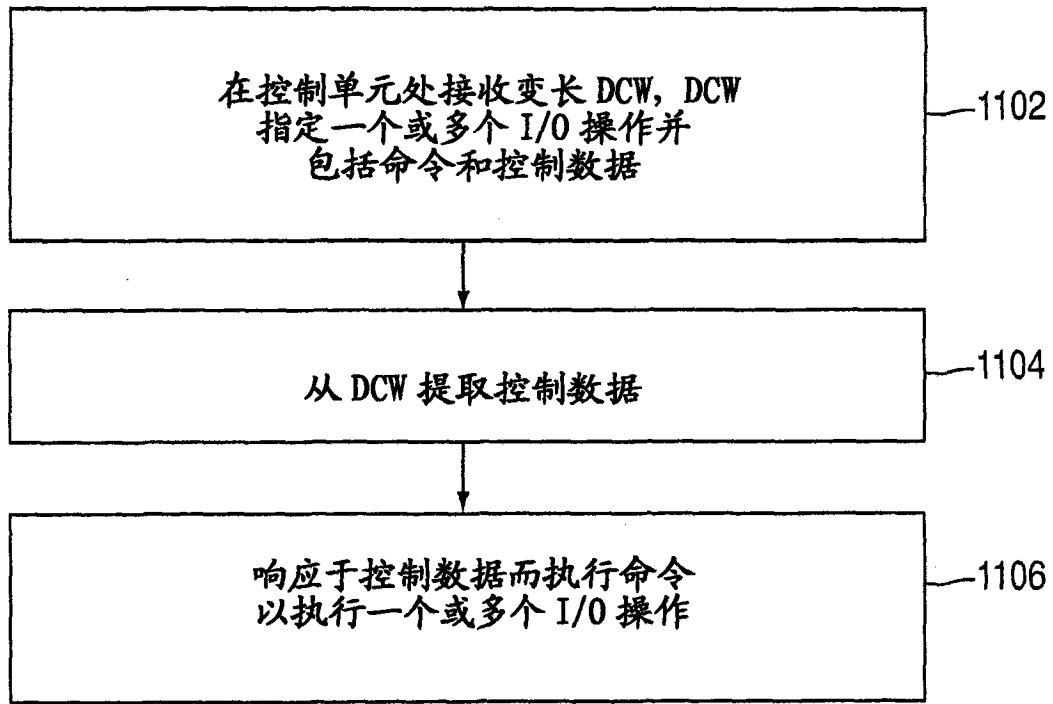


图 11

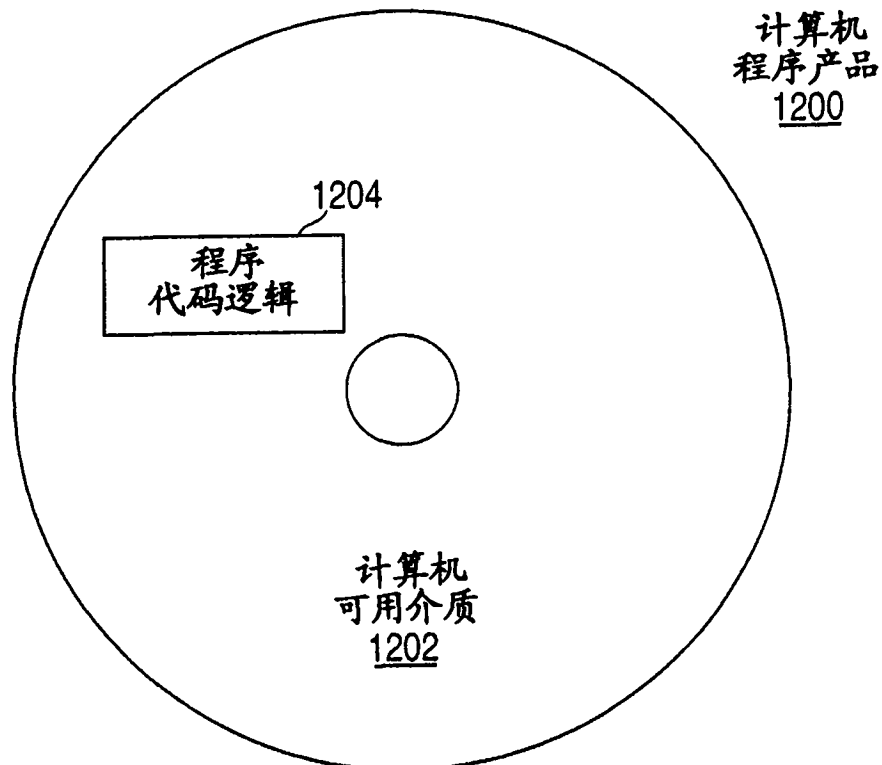


图 12