



(12) 发明专利申请

(10) 申请公布号 CN 101939730 A

(43) 申请公布日 2011. 01. 05

(21) 申请号 200980104204. 7

(74) 专利代理机构 北京市中咨律师事务所

(22) 申请日 2009. 02. 09

11247

代理人 于静 杨晓光

(30) 优先权数据

12/030, 954 2008. 02. 14 US

(51) Int. Cl.

G06F 13/10(2006. 01)

(85) PCT申请进入国家阶段日

2010. 08. 05

(86) PCT申请的申请数据

PCT/EP2009/051450 2009. 02. 09

(87) PCT申请的公布数据

W02009/101053 EN 2009. 08. 20

(71) 申请人 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 G·希特曼三世 C·黄

U·恩尤库 D·莱迪

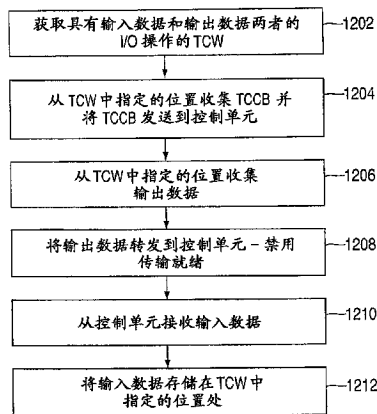
权利要求书 2 页 说明书 10 页 附图 13 页

(54) 发明名称

单个 I/O 操作内的双向数据传输

(57) 摘要

一种用于促进被配置为与控制单元通信的主计算机系统处的输入 / 输出 (I/O) 操作的 I/O 处理的制品、装置和方法。所述方法包括所述主计算机系统获取具有输入数据和输出数据的 I/O 操作的传输命令字 (TCW)。所述 TCW 指定了所述输出数据的位置以及用于存储所述输入数据的位置。所述主计算机系统将所述 I/O 操作转发到所述控制单元以便执行。所述主计算机系统响应于由所述 TCW 指定的所述输出数据的位置而收集所述输出数据,然后将所述输出数据转发到所述控制单元以在执行所述 I/O 操作中使用。所述主计算机系统从所述控制单元接收所述输入数据并将所述输入数据存储在与由所述 TCW 指定的位置处。



1. 一种促进被配置为与控制单元通信的主计算机系统处的输入 / 输出 I/O 操作的 I/O 处理的计算机程序产品, 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:

所述主计算机系统获取具有输入数据和输出数据的 I/O 操作的传输命令字 TCW, 所述 TCW 指定了所述输出数据的位置以及用于存储所述输入数据的位置;

所述主计算机系统将所述 I/O 操作转发到所述控制单元以便执行;

所述主计算机系统响应于由所述 TCW 指定的所述输出数据的位置而收集所述输出数据;

所述主计算机系统将所述输出数据转发到所述控制单元以在执行所述 I/O 操作中使用;

所述主计算机系统从所述控制单元接收所述输入数据; 以及

所述主计算机系统将所述输入数据存储在与由所述 TCW 指定的用于存储所述输入数据的位置处。

2. 如权利要求 1 中所述的计算机程序产品, 其中所述 TCW 还指定了所述输出数据的大小, 并且还响应于所述输出数据的大小而收集所述输出数据。

3. 如权利要求 1 中所述的计算机程序产品, 其中所述 I/O 操作包括一个或多个要由所述 I/O 操作执行的命令。

4. 如权利要求 1 中所述的计算机程序产品, 其中所述输出数据的位置以及用于存储所述输入数据的位置之一或两者是直接地址。

5. 如权利要求 1 中所述的计算机程序产品, 其中所述输出数据的位置以及用于存储所述输入数据的位置之一或两者是间接地址。

6. 如权利要求 1 中所述的计算机程序产品, 其中所述主计算机系统包括 I/O 处理系统, 并且由所述 I/O 处理系统执行所述方法。

7. 如权利要求 1 中所述的计算机程序产品, 其中所述主计算机系统包括通道子系统, 并且由所述通道子系统执行所述方法。

8. 如权利要求 1 中所述的计算机程序产品, 其中从主操作系统获取所述 TCW。

9. 一种用于在单个输入 / 输出 I/O 操作内提供双向数据传输的主计算机系统, 所述主计算机系统包括:

包括通道适配器的主计算机 I/O 子系统, 所述通道适配器适于与控制单元通信, 所述主计算机 I/O 子系统执行方法, 所述方法包括:

获取具有输入数据和输出数据的 I/O 操作的传输命令字 TCW, 所述 TCW 指定了所述输出数据的位置以及用于存储所述输入数据的位置;

响应于由所述 TCW 指定的所述输出数据的位置而收集所述输出数据;

将所述 I/O 操作和所述输出数据转发到所述控制单元以便执行;

从所述控制单元接收所述输入数据; 以及

将所述输入数据存储在与由所述 TCW 指定的用于存储所述输入数据的位置处。

10. 如权利要求 9 中所述的装置, 其中所述 TCW 还指定了所述输出数据的大小, 并且还响应于所述输出数据的大小而收集所述输出数据。

11. 如权利要求 9 中所述的装置,其中所述 I/O 操作包括一个或多个要由所述 I/O 操作执行的命令。

12. 如权利要求 9 中所述的装置,其中所述输出数据的位置以及用于存储所述输入数据的位置之一或两者是直接地址。

13. 如权利要求 9 中所述的装置,其中所述输出数据的位置以及用于存储所述输入数据的位置之一或两者是间接地址。

14. 如权利要求 9 中所述的装置,其中从主操作系统获取所述 TCW。

15. 一种用于促进被配置为与控制单元通信的主计算机系统处的输入 / 输出 I/O 操作的 I/O 处理的方法,所述方法包括:

获取具有输入数据和输出数据的 I/O 操作的传输命令字 TCW,所述 TCW 指定了所述输出数据的位置以及用于存储所述输入数据的位置;

响应于由所述 TCW 指定的所述输出数据的位置而收集所述输出数据;

将所述 I/O 操作和所述输出数据转发到所述控制单元以便执行;

从所述控制单元接收所述输入数据;以及

将所述输入数据存储在与由所述 TCW 指定的用于存储所述输入数据的位置处。

16. 如权利要求 15 中所述的方法,其中所述 TCW 还指定了所述输出数据的大小,并且还响应于所述输出数据的大小而收集所述输出数据。

17. 如权利要求 15 中所述的方法,其中所述 I/O 操作包括一个或多个要由所述 I/O 操作执行的命令。

18. 如权利要求 15 中所述的方法,其中所述输出数据的位置以及用于存储所述输入数据的位置之一或两者是直接地址。

19. 如权利要求 15 中所述的方法,其中所述输出数据的位置以及用于存储所述输入数据的位置之一或两者是间接地址。

20. 如权利要求 15 中所述的方法,其中所述主计算机系统包括 I/O 处理系统,并且由所述 I/O 处理系统执行所述方法。

21. 如权利要求 15 中所述的方法,其中所述主计算机系统包括通道子系统,并且由所述通道子系统执行所述方法。

22. 如权利要求 15 中所述的方法,其中从主操作系统获取所述 TCW。

单个 I/O 操作内的双向数据传输

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, 涉及提供包括输入数据和输出数据两者的 I/O 操作。

背景技术

[0002] 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。

[0003] 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。

[0004] 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。

[0005] 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。

[0006] 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利的。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。但是, 通过将两个或更多命令一起组合为单个 I/O 操作来改变命令序列可导致具有输入数据和输出数据两者的 I/O 操作。目前, I/O 操作可以支持单个数据区域, 所述单个数据区域可以用于同一 I/O 操作内的数据输入或数据输出, 但不能同时用于这两者。这限制了可以在单个 I/O 操作内组合在一起的命令的种类, 从而限制了可以通过组合命令获取的吞吐量的增加。相应地, 本领域需要在单个 I/O 操作内同时传输输入数据和输出数据的能力。

发明内容

[0007] 示例性实施例包括一种促进被配置为与控制单元通信的主计算机系统处的输入 / 输出 (I/O) 操作的 I/O 处理的计算机程序产品。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。所述方法包括所述主计算机系统获取具有输入数据和输出数据的 I/O 操作的传输命令字 (TCW)。所述 TCW 指

定了所述输出数据的位置以及用于存储所述输入数据的位置。所述主计算机系统将所述 I/O 操作转发到所述控制单元以便执行。所述主计算机系统响应于由所述 TCW 指定的所述输出数据的位置而收集所述输出数据,然后将所述输出数据转发到所述控制单元以在执行所述 I/O 操作中使用。所述主计算机系统从所述控制单元接收所述输入数据,以及将所述输入数据存储在与由所述 TCW 指定的位置处。

[0008] 另一示例性实施例包括一种用于在单个输入/输出 I/O 操作内提供双向数据传输的主计算机系统。所述主计算机系统包括主计算机 I/O 子系统,所述主计算机 I/O 子系统包括与控制单元通信的通道适配器。所述主计算机 I/O 子系统执行方法,所述方法包括所述主计算机系统获取具有输入数据和输出数据的 I/O 操作的传输命令字 TCW。所述 TCW 指定了所述输出数据的位置以及用于存储所述输入数据的位置。响应于由所述 TCW 指定的所述输出数据的位置而收集所述输出数据。将所述 I/O 操作和所述输出数据转发到所述控制单元以便执行。所述主计算机系统从所述控制单元接收所述输入数据,以及将其存储在由所述 TCW 指定的位置处。

[0009] 其他示例性实施例包括一种用于促进被配置为与控制单元通信的主计算机系统处的输入/输出 I/O 操作的 I/O 处理的方法。所述方法包括所述主计算机系统获取具有输入数据和输出数据的 I/O 操作的传输命令字 TCW。所述 TCW 指定了所述输出数据的位置以及用于存储所述输入数据的位置。所述主计算机系统将所述 I/O 操作转发到所述控制单元以便执行。所述主计算机系统响应于由所述 TCW 指定的所述输出数据的位置而收集所述输出数据,然后将所述 I/O 操作转发到所述控制单元以在执行所述 I/O 操作中使用。所述主计算机系统从所述控制单元接收所述输入数据,以及将其存储在由所述 TCW 指定的位置处。

[0010] 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/或方法对于本领域的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

[0011] 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:

[0012] 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;

[0013] 图 2A 示出了现有技术通道命令字的一个实例;

[0014] 图 2B 示出了现有技术通道命令字通道程序的一个实例;

[0015] 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;

[0016] 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;

[0017] 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以执行图 4 的 TCW 通道程序的链路协议的一个实施例;

[0018] 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;

[0019] 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；

[0020] 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一个实施例；

[0021] 图 9 示出了根据本发明的一个方面的 TCW 的一个实施例；

[0022] 图 10 示出了根据本发明的一个方面的 TCW 通道程序的一个实施例；

[0023] 图 11 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以执行图 10 的 TCW 通道程序的链路协议的一个实施例；

[0024] 图 12 示出了根据本发明的一个方面的用于单个 I/O 操作内的双向数据传输的过程的一个实施例；以及

[0025] 图 13 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。

[0026] 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

[0027] 根据本发明的一个方面,通过允许单个输入 / 输出 (I/O) 操作同时包括输入数据和输出数据来促进 I/O 处理。因此,每个 I/O 操作都可以用于传输输入流和输出流两者。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来促进 I/O 处理。例如,减少了诸如通道之类的 I/O 通信适配器和控制单元之间的交换和序列数。这通过将多个命令作为由控制单元执行的单个实体从 I/O 通信适配器发送到控制单元,并且通过由所述控制单元作为单个实体发送从所述命令产生的数据(如果有)来实现。作为单个实体发送到所述控制单元的多个设备命令字 (DCW) 可以包括读取命令和写入命令两者。

[0028] 将多个命令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。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。

[0029] 在一个示例性实施例中,所述 TCW 针对执行所述 I/O 操作所需的所有控制块提供指向通道的指针。在一个示例性实施例中,所述 TCW 包括指向输入数据地址和输出数据地址两者的指针。这允许在单个 I/O 操作内以两个方向(例如,从通道到控制单元以及从控制单元到通道)传输数据。

[0030] 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。

[0031] 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。

[0032] 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。

[0033] 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。

[0034] 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 可以以本领域公知的任何方式实现,包括光链路,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 的操作的执行的执行的路径管理功能。

[0035] 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,从而具有与一个或多个控制单元 110 连接的多个通道子系统(未示出)。

[0036] 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。

[0037] 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。

[0038] 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。

[0039] 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。

[0040] 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。

[0041] 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) 以及一个或多个读取 CCW 221 的另一区域 (例如应用区域)。每个读取 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。

[0042] 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。

[0043] 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 (图 2B), 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 和读取 CCW。它从查找记录 CCW 217 (图 2B) 查找记录命令 305 (图 3) 以及从查找记录数据 220 获取

数据。从读取 CCW 221 (图 2B) 取回读取命令 306 (图 3)。将它们都发送到控制单元 302。

[0044] 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 308。这可以在查找命令 305 和 / 或读取命令 306 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。

[0045] 控制单元 302 将所请求的数据 310 发送到通道 300。此外,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。

[0046] 以上 CCW 通道程序写入 4k 数据的处理需要打开和关闭两次交换以及七个序列。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。

[0047] 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。

[0048] 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。

[0049] 参考图 5,通道 500 打开与控制单元 502 的交换并将 TCCB 504 发送到控制单元 502。在一个实例中,在 FCP 命令 (称为 FCP_CMND 信息单元 (IU) 或传输命令 IU) 中将 TCCB 504 和序列主动性 (sequence initiative) 传输到控制单元 502。控制单元 502 执行 TCCB 504 的多个命令 (例如,定义范围命令、查找记录命令,将命令读取为设备控制字 (DCW)) 并经由例如 FCP 数据 IU 将数据 506 转发到通道 500。它还提供状态并关闭交换 508。作为一个实例,在 FCP 状态帧中发送最终状态,所述 FCP 状态帧例如在 FCP_RSP IU (也称为传输响应 IU) 有效负载的字节 10 或 11 中具有一个活动位。可使用 FCP_RSP_IU 有效负载传输 FICON 结束状态以及附加状态信息。

[0050] 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:

[0051] 1. 在 FCP_CMND IU 中传输 TCCB。

[0052] 2. 将数据 IU 和序列主动性传输到控制单元 502。(禁用 FCP 传输就绪)

[0053] 3. 在 FCP 状态帧中发送最终状态,所述 FCP 状态帧例如在 FCP_RSPIU 有效负载的字节 10 或 11 中具有一个活动位。使用 FCP_RES_INFO 字段或读出字段传输 FICON 结束状态以及附加状态信息。

[0054] 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 通道程序的七个序列相比 (参阅图

2B-3),对于 TCW 通道程序而言,存在三个通信序列(参阅图 4-5)。

[0055] 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 600、查找记录命令 601、读取命令 602、读取命令 604、读取命令 606、查找记录命令 607 以及读取命令 608)从通道 610 发送到控制单元 612。此外,在单独序列中将每个 4k 数据块(例如,数据 614-620)从控制单元 612 发送到通道 610。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)以及十四个通信序列。将其与图 7 的 TCW 通道程序的三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。

[0056] 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读取命令。响应于收到 TCCB 704,控制单元 702 执行所述命令并在单独序列中将 16k 数据 706 发送到通道 700。此外,控制单元 702 将状态提供给通道 700 并关闭交换 708。因此,所述 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。

[0057] 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。

[0058] 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。

[0059] 通道子系统 108 中的通道 124 包括支持与控制单元 110 通信的多个元件。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)。

[0060] 图 9 示出了传输控制字(TCW)900 的一个示例性实施例。通道 124 使用 TCW 900 建立 I/O 操作并且不将其发送到控制单元 110。图 9 中示出的 TCW 在单个 I/O 操作内提供输入数据和输出数据两者。

[0061] 在图 9 所示的示例性 TCW 900 中,等于“00b”的格式字段 902 指示后面紧跟着 TCW

900。TCW 900 还包括将来可能使用的保留位 904。

[0062] TCW 900 还包括标志字段 906。标志字段 906 的前五位保留以供将来使用并被设为 0。标志字段 906 的第六位是 TIDAL 读取标志。在一个示例性实施例中,当输入数据地址字段 918 包含 TIDAL 地址时,TIDAL 读取标志被设为 1。如果 TIDAL 读取标志被设为 0,则输入数据地址字段 918 包含数据地址。标志字段 906 的第七位是 TCCB TIDAL 标志。在一个示例性实施例中,当 TCCB 地址字段 922 包含 TIDAL 地址时,将 TCCBTIDAL 标志设为 1。如果将 TCCB TIDAL 标志设为 0,则 TCCB 地址字段 922 直接寻址 TCCB。TCCB TIDAL 标志允许操作系统软件或系统管理程序对功能进行分层以及为用户通道程序添加前缀。标志字段 906 的第八位是 TIDAL 写入标志。在一个示例性实施例中,当输出数据地址字段 916 包含 TIDAL 地址时,将 TIDAL 写入标志设为 1。如果将 TIDAL 写入标志设为 0,则输出数据地址字段 916 包含数据地址。

[0063] 标志字段 906 的第九位到第二十四位保留以供将来使用。

[0064] TCW 900 还包括间接表示 TCCB 的长度并可用于确定 TCCB 的实际长度的 TCCB 长度字段 910。

[0065] TCW 900 中的读 / 写位 912 用于指示是否作为执行 TCW 900 的结果而读和 / 或写数据。在一个示例性实施例中,读 / 写位 912 中的读取位被设为 1 以指示作为执行 TCW 900 的结果将输入数据从 I/O 设备 112 传输到主系统 101 中的系统存储 (例如,主存储器 102)。读 / 写位 912 中的写入位被设为 1 以指示作为执行 TCW 900 的结果将输出数据从主系统 101 中的系统存储 (例如,主存储器 102) 传输到 I/O 设备。

[0066] 输出数据地址字段 916 包括输出数据 (如果有) 的地址。如上所述,输出数据地址字段 916 的内容可以是输出数据的 TIDAL 地址 (例如,间接地址) 或输出数据的实际地址 (例如,直接地址)。输入数据地址字段 918 包括输入数据 (如果有) 的地址。如上所述,输入数据地址字段 918 的内容可以是输入数据的 TIDAL 地址或输入数据的实际地址。在一个示例性实施例中,输出数据地址字段 916 和输入数据地址字段 918 实现为六十四位地址。

[0067] TCW 900 还包括传输状态块地址字段 920。I/O 操作的传输响应 IU 中的完成状态的一部分 (例如,扩展状态部分) 存储在该地址中。TCW 900 中的 TCCB 地址字段 922 包括 TCCB 在系统存储中的地址。如上所述,TCCB 是要针对 TCW 900 执行的 DCW 所在的控制块。还如上所述,TCCB 地址字段 922 的内容可以是 TCCB 的 TIDAL 地址或 TCCB 的实际地址。在一个示例性实施例中,传输状态块地址字段 920 和 TCCB 地址字段 922 实现为六十四位地址。

[0068] TCW 900 中的输出计数字段 924 指示 TCW/TCCB 针对输出操作要传输的输出数据量。在一个示例性实施例中,输出计数字段 924 指定 TCW 所指定的输出存储区域 (输出数据地址 916) 中要传输的字节数。TCW 900 中的输入计数字段 926 指示 TCW/TCCB 针对输入操作要传输的输入数据量。在一个示例性实施例中,输入计数字段 926 指定 TCW 所指定的输入存储区域 (输入数据地址 918) 中要传输的字节数。将保留 TCW 900 中的若干附加字段:保留字段 928、保留字段 930 和保留字段 932。询问 TCW 地址字段 934 包含另一 TCW 的地址并且由通道 124 用于询问在取消子通道 I/O 指令的主动性下的操作的状态。

[0069] 图 9 中所示的 TCW 是如何配置命令字的一个实例。也可以存在其中包括附加字段和 / 或不包括图 9 中所示的字段的其他配置。

[0070] 图 10 示出了根据本发明的一个方面的当在单个 I/O 操作中包括输入和输出数据两者时的 TCW 通道程序的一个实施例。如图 10 所示,TCW 通道程序 1000 包括 TCW 1002 以及状态区域 1008,TCW 1002 指定了存储器中 TCCB 1004 的位置、存储器中用于存储输入数据 1006 或指向输入数据 1006 的位置的 TIDAL 1010(即,传输模式间接数据地址字(TIDAW)列表)的位置、存储器中输出数据区域 1014 或指向输出数据区域 1014 的 TIDAL 1012 的位置。

[0071] 参考图 11 描述了图 10 中所示的 TCW 通道程序 1000 的处理。现在参考图 11,通道 1100 打开与控制单元 1102 的交换,并将 TCCB 1104 和位于由 TCW 1002 指定的输出数据区域 1014 处的输出数据 1105 发送到控制单元 1102。通道 1100 根据 TCW 1002 中的输出计数 924 的值确定要发送的数据量。控制单元 1102 执行 TCCB 1104 的多个命令(例如,作为设备控制字(DCW)的定义范围命令、查找记录命令、写入命令以及读取命令)、从通道 1100 接收输出数据 1105,并例如通过 FCP_数据 IU 将依照所述 DCW 中的数据计数的输入数据 1106 转发到通道 1100。通道 1100 将输入数据 1106 存储在由 TCW 1002 指定的位置处。控制单元 1102 还提供状态并关闭交换 1108。通过这种方式,在单个 TCW 通道程序 1000(或 I/O 操作)中将数据输入到通道 1100 并输出到控制单元 1102。

[0072] 图 12 示出了根据本发明的一个方面的用于在单个 I/O 操作内进行双向数据传输的过程的一个实施例。在一个示例性实施例中,图 12 中所示的处理在与控制单元进行网络通信的主计算机系统处发生。所述主计算机系统可以包括执行所述过程的 I/O 处理系统。此外,所述 I/O 处理系统可以包括执行所述过程的通道子系统。在方块 1202,主计算机获取 TCW。在一个示例性实施例中,从主计算机上运行的操作系统获取(或接收)TCW。所述 TCW 同时包括输出数据地址 916 和输出计数字段 924 以及输入数据地址 918 和输入数据计数字段 926。在一个示例性实施例中,当读/写位 912 中的写入位设为 1 时,所述 TCW 包括输出数据,当读/写位 912 中的读取位设为 1 时,所述 TCW 包括输入数据。在方块 1204,取回由所述 TCW 指定的 TCCB 位置并将其转发到所述控制单元。所述 TCCB 包含通知所述控制单元执行何种 I/O 操作的 DCW。

[0073] 在方块 1206,从所述 TCW 所指定的位置收集输出数据(如果读/写位 912 中的写入位设为 1)。所收集的要包括在输出数据中的数据量基于输出数据计数字段 924 的值。如上所述,输出数据地址可以是输出数据的直接地址或输出数据的间接地址。间接地址是包含指向共同组成输出数据的多个存储位置的一个或多个地址的列表的地址(例如,TIDAL)。直接地址是包含输出数据的地址。在一个示例性实施例中,所述 TCW 内的标志字段 906 中的 TIDAL 写入标志在输出数据地址字段 916 包含 TIDAL 的地址时被设为 1,并且在输出数据地址字段 916 包含输出数据的地址时被设为 0。

[0074] 在方块 1208,将输出数据转发到所述控制单元。对于该实例,禁用 XFER_RDY。

[0075] 在方块 1210,作为执行 I/O 操作的结果而从所述控制单元接收输入数据。在方块 1212,将输入数据存储在由所述 TCW 指定的位置(输入数据地址 918)处。在一个示例性实施例中,当读/写位 912 中的读取位设为 1 时,所述 TCW 包括输入数据。如上所述,输入数据地址 918 可以是用于存储输入数据的直接地址,或备选地可以是到指向多个存储位置的地址列表的地址(例如,TIDAL 或间接地址),其中每个存储位置存储部分输入数据。在一个示例性实施例中,所述 TCW 内的标志字段 906 中的 TIDAL 读取标志在输入数据地址字段

918 包含 TIDAL 的地址时被设为 1, 并且在输入数据地址字段 918 包含输入数据的地址时被设为 0。

[0076] 示例性实施例的技术效果包括在单个 I/O 操作中包括输入数据和输出数据两者的能力。这提供了对 DCW 进行分组的灵活性, 并且可导致减少通道与控制单元之间所需的交换数。

[0077] 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 1302 上的如图 13 中所示的计算机程序产品 1300, 其中计算机程序代码逻辑 1304 包含在有形介质中体现为制品的指令。计算机可用介质 1302 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 1304 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 1304, 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 1304 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 1304 配置所述微处理器以创建专用逻辑电路。

[0078] 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。

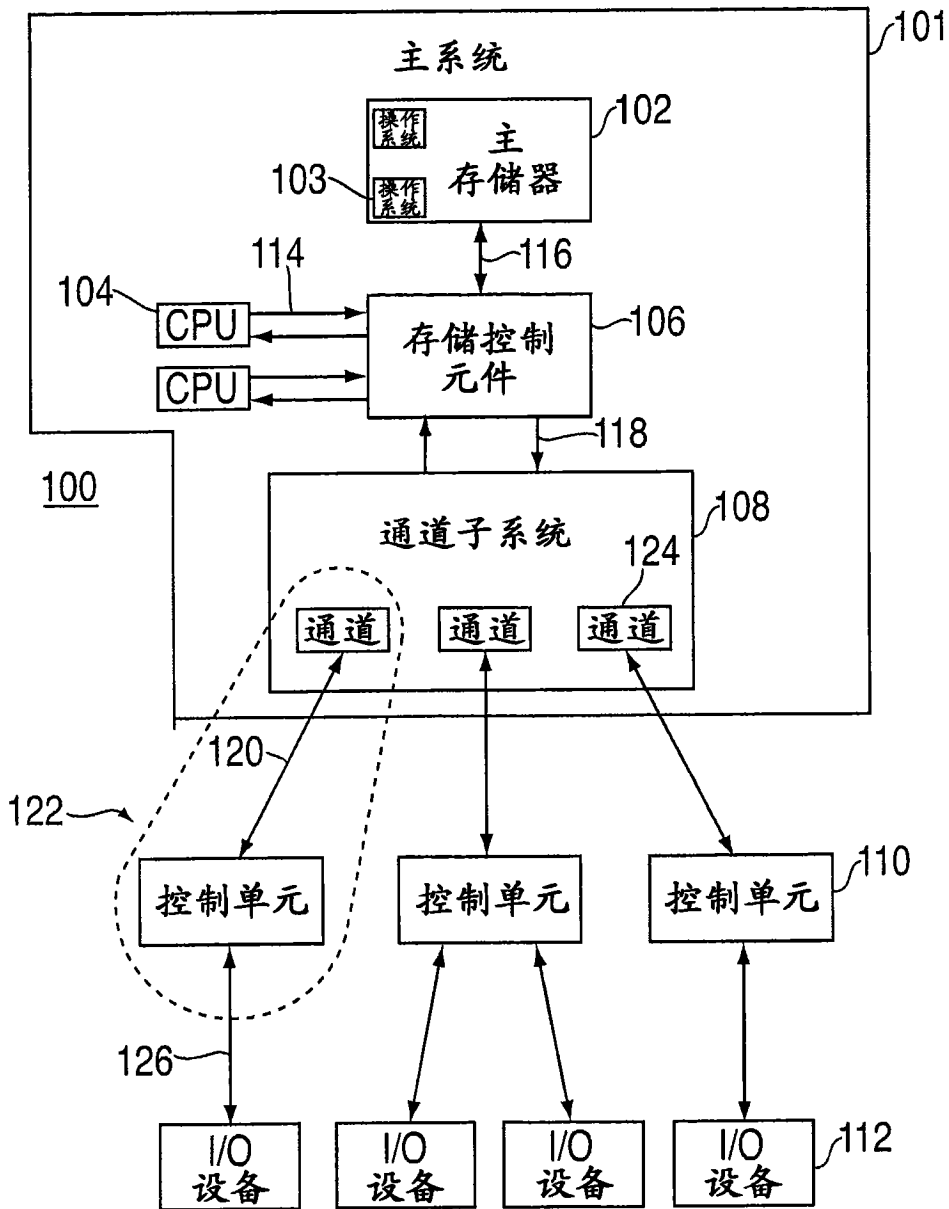


图 1

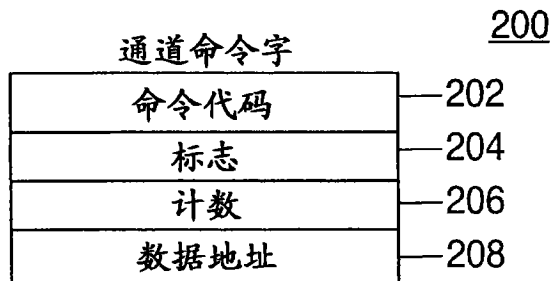


图 2A 现有技术

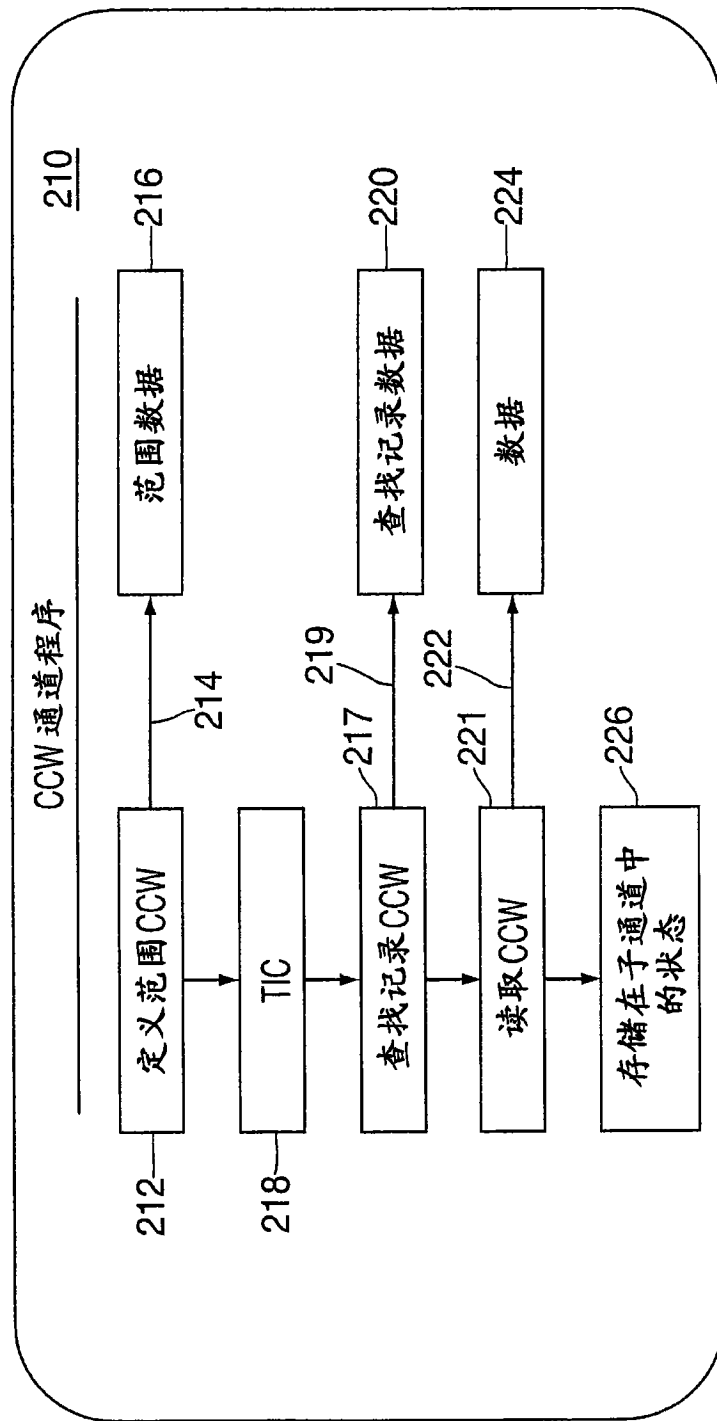


图 2B 现有技术

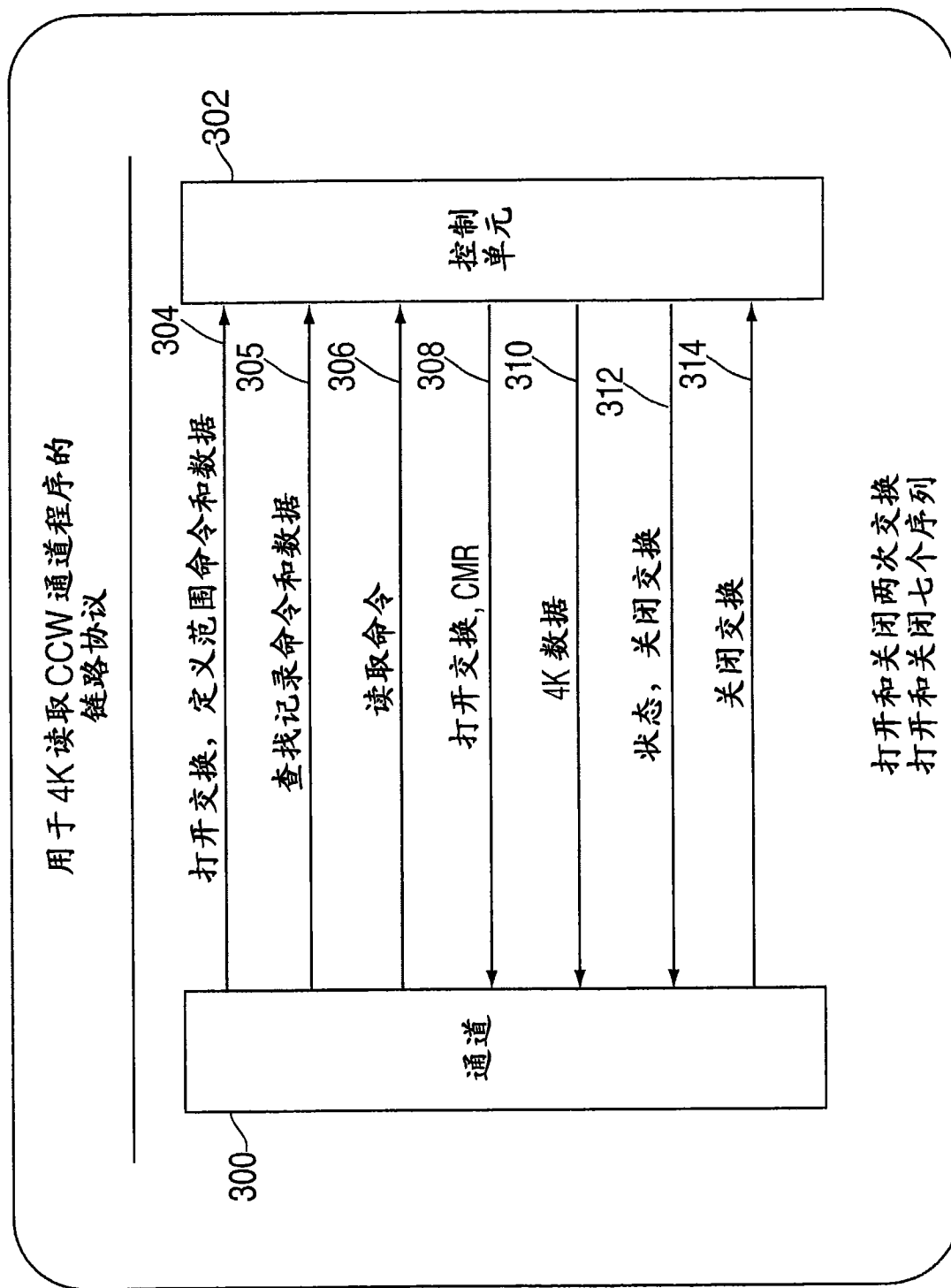


图3 现有技术

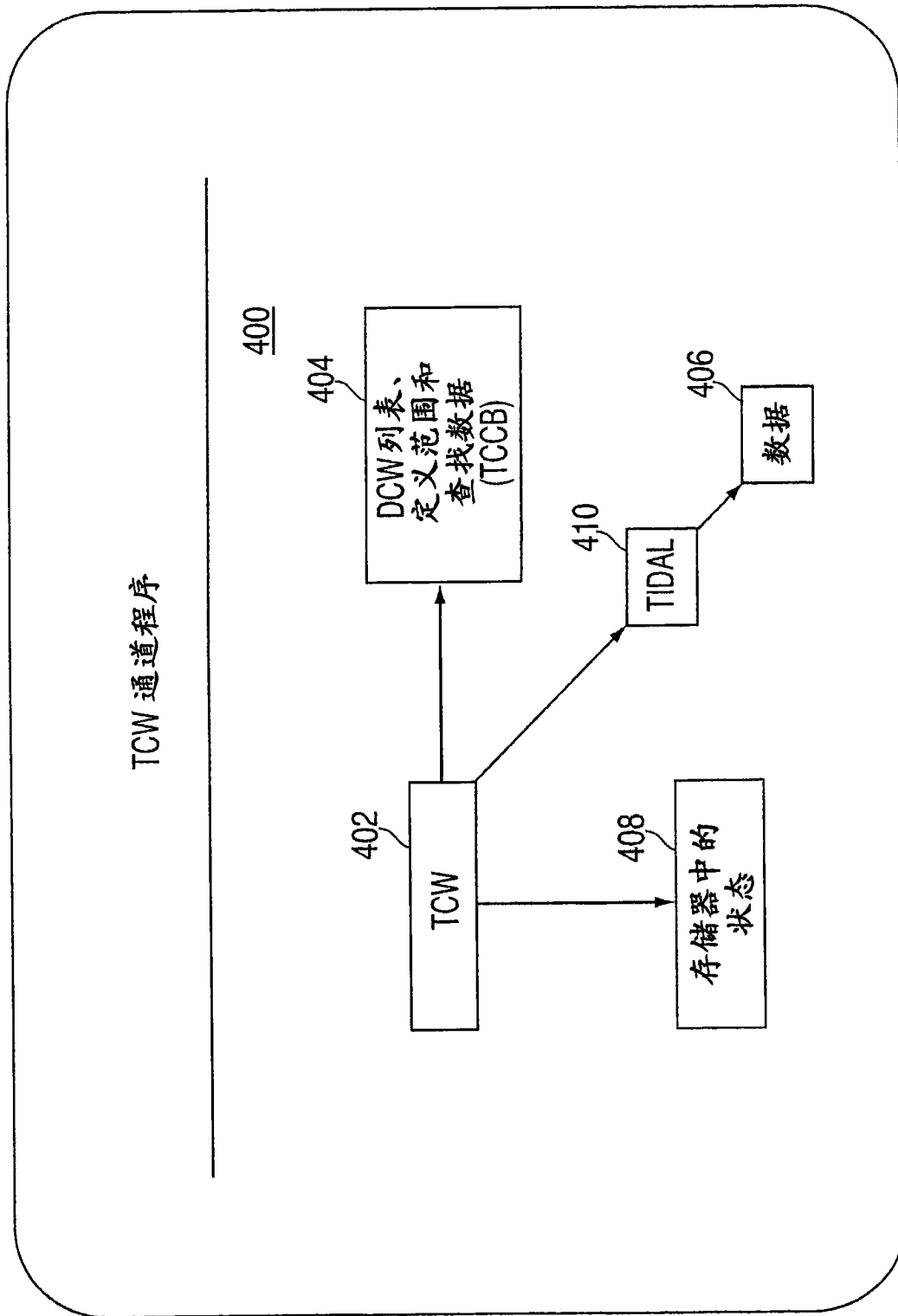


图 4

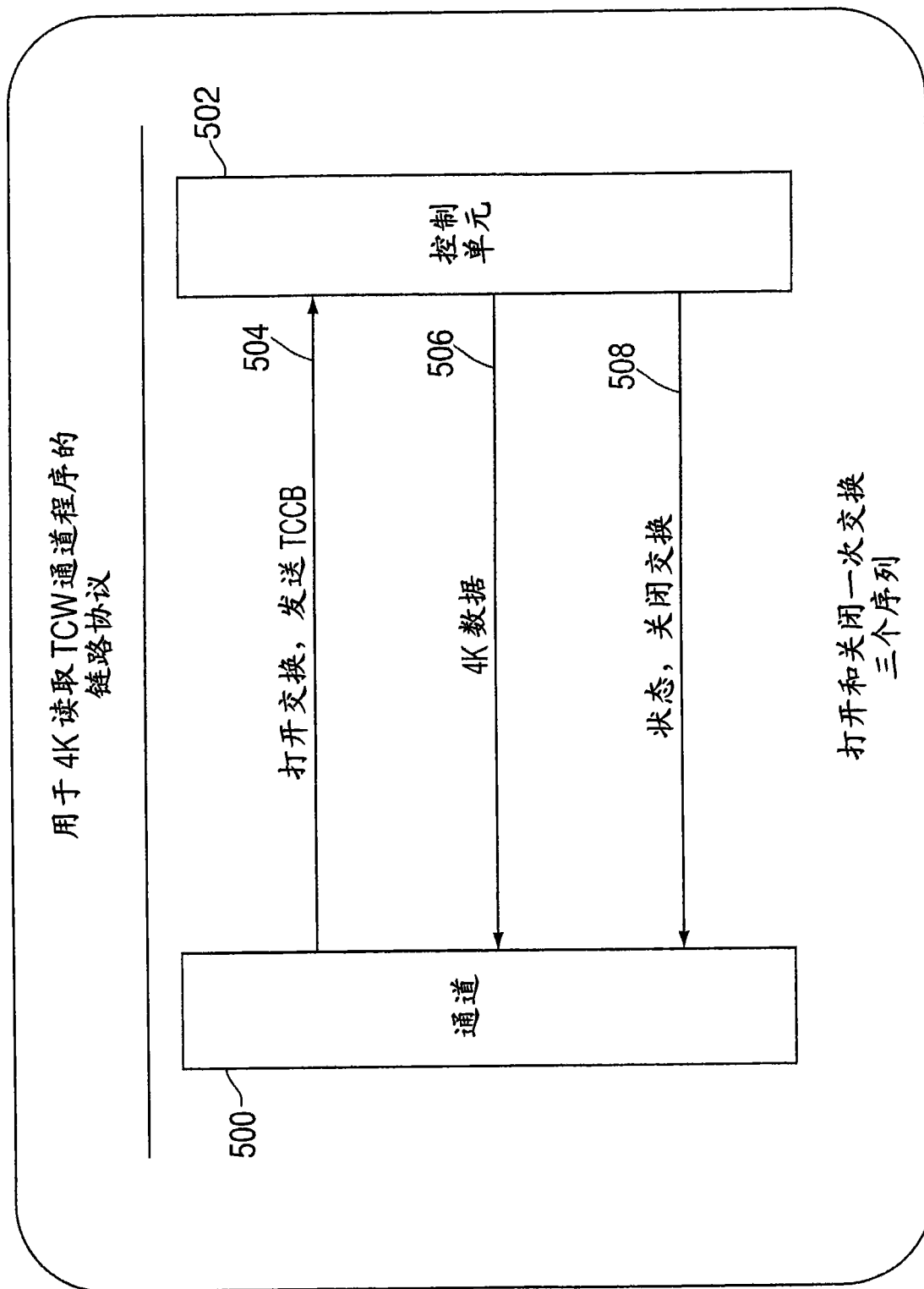


图 5

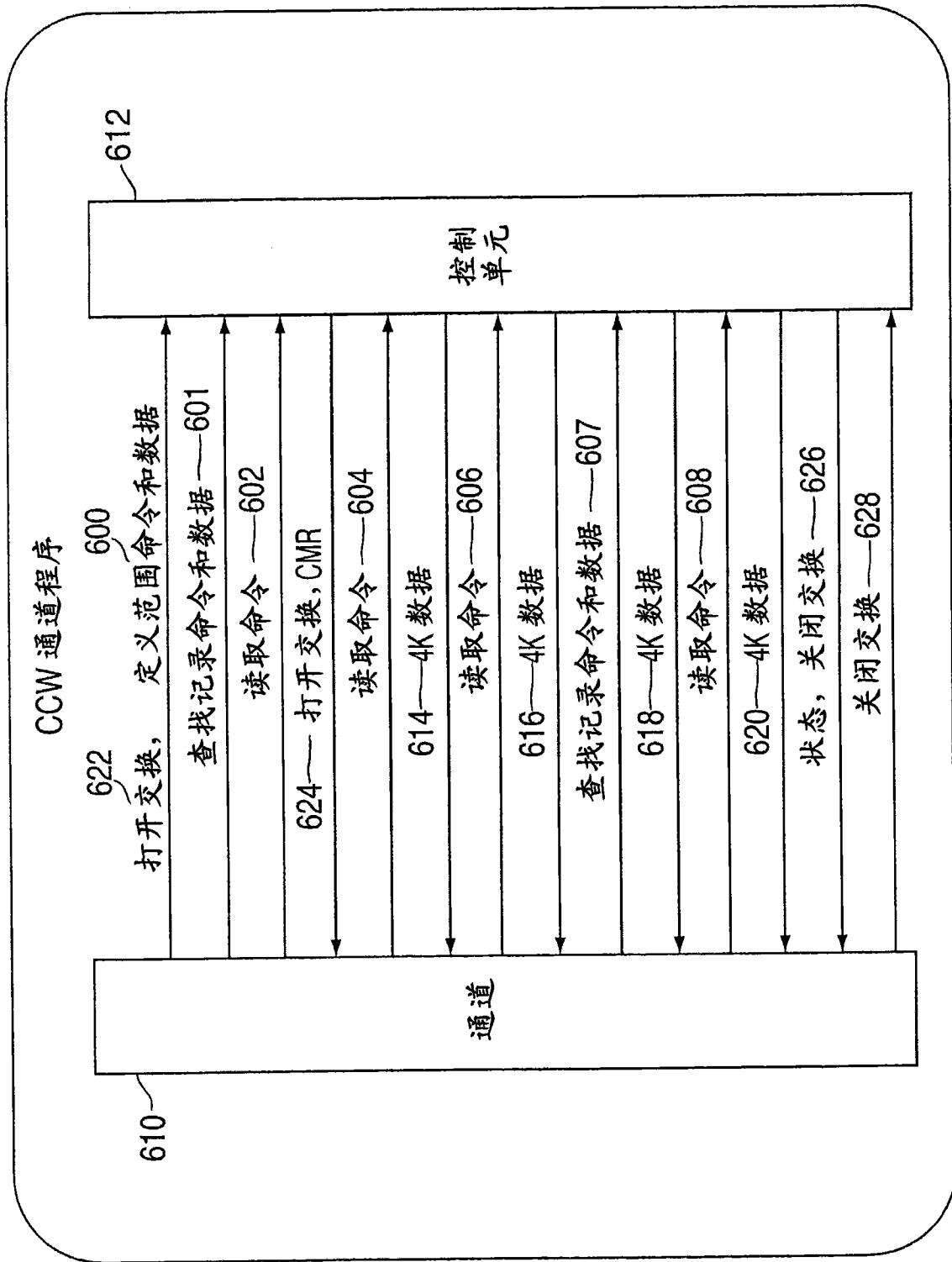


图 6 现有技术

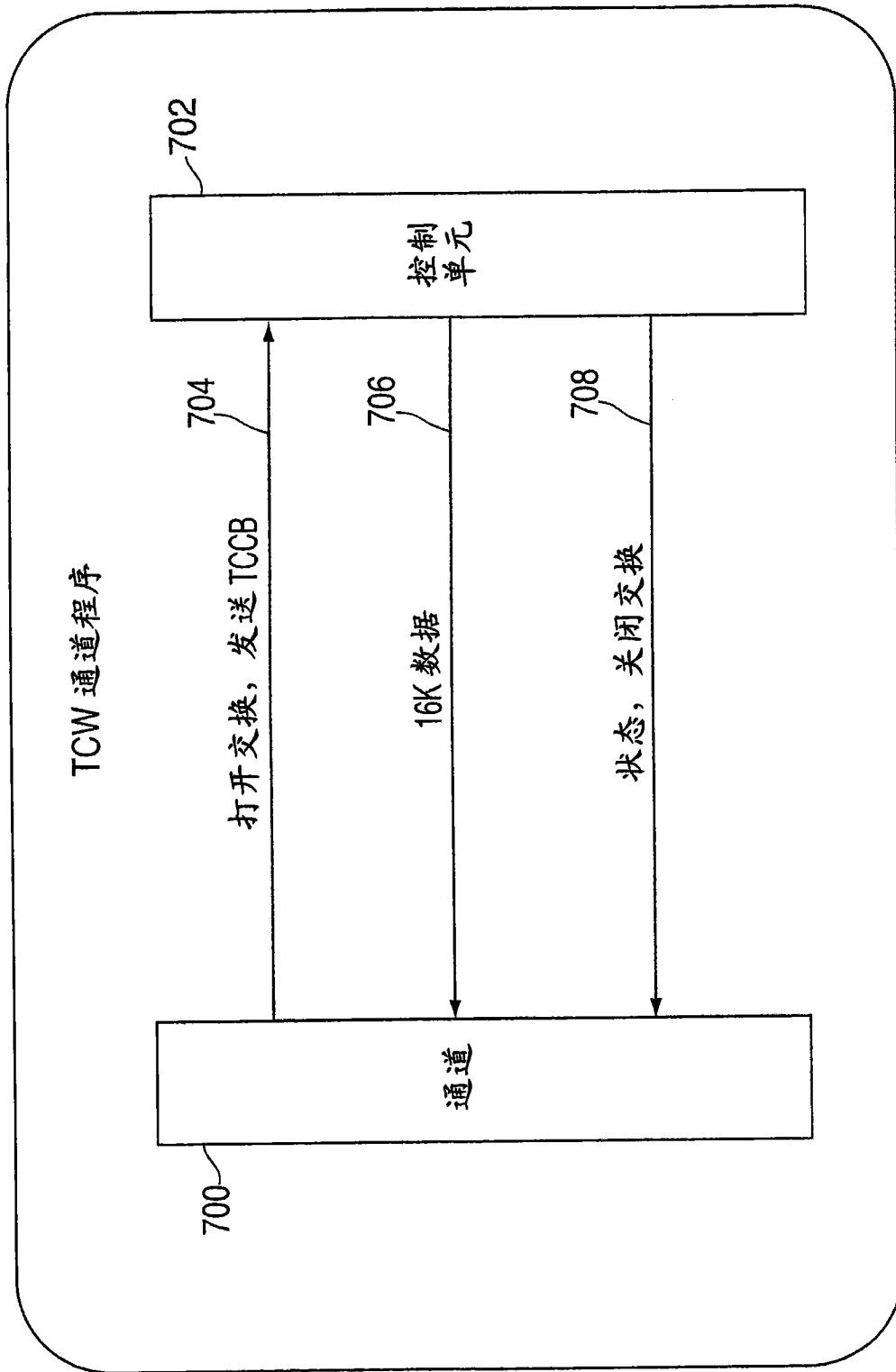


图 7

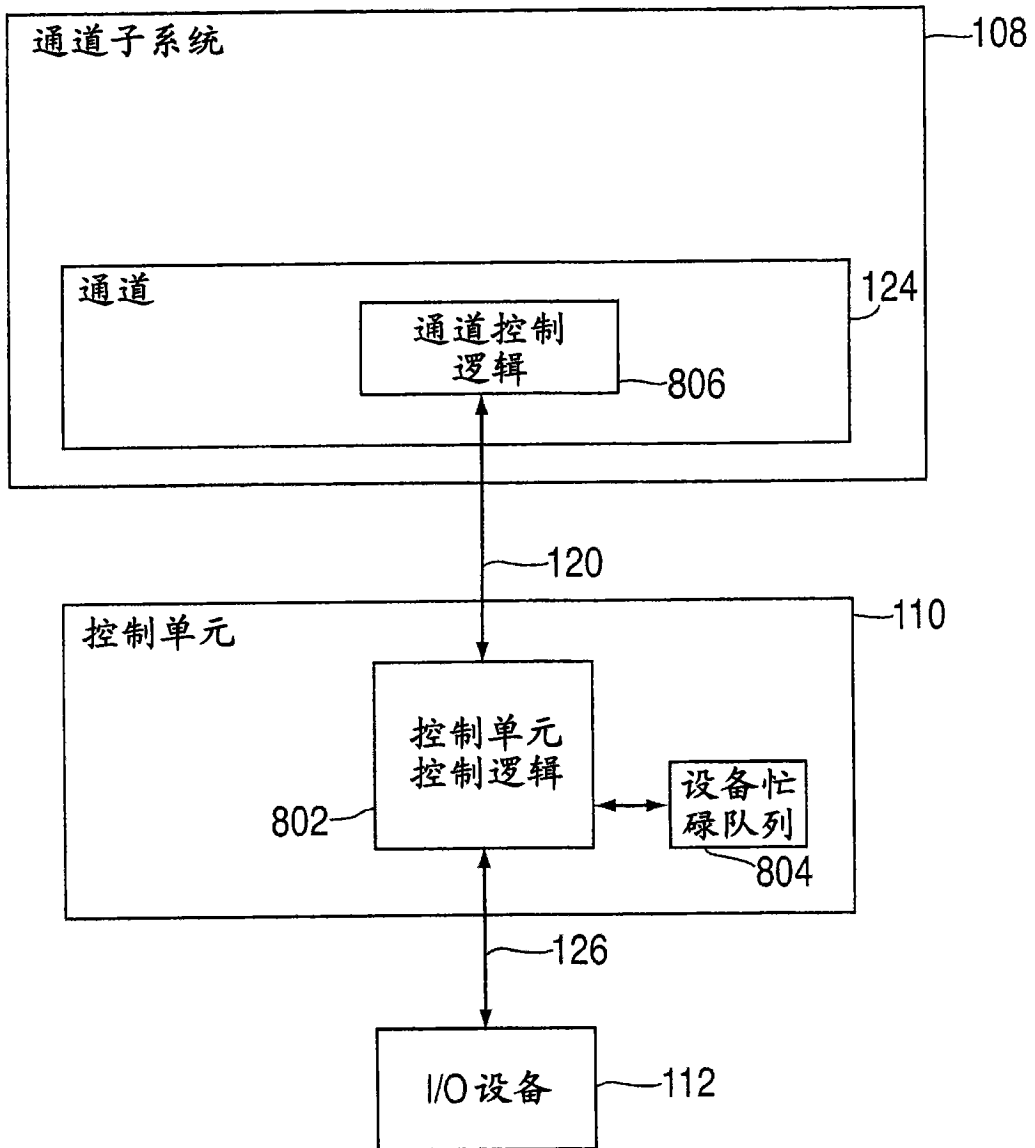


图 8

900

字	字节0	字节1	字节2	字节3	字节4	字节5	字节6	字节7
0	F000000 904	标志	906	保留	TCCBL (L1) ₉₁₀	R W	保留	914
1	916 输出数据地址							
2	918 输入数据地址							
3	920 传输状态块地址							
4	922 传输命令控制块地址							
5	输出计数				输入计数			
6	保留				保留			
7	保留				询问TCW地址			

902

图 9

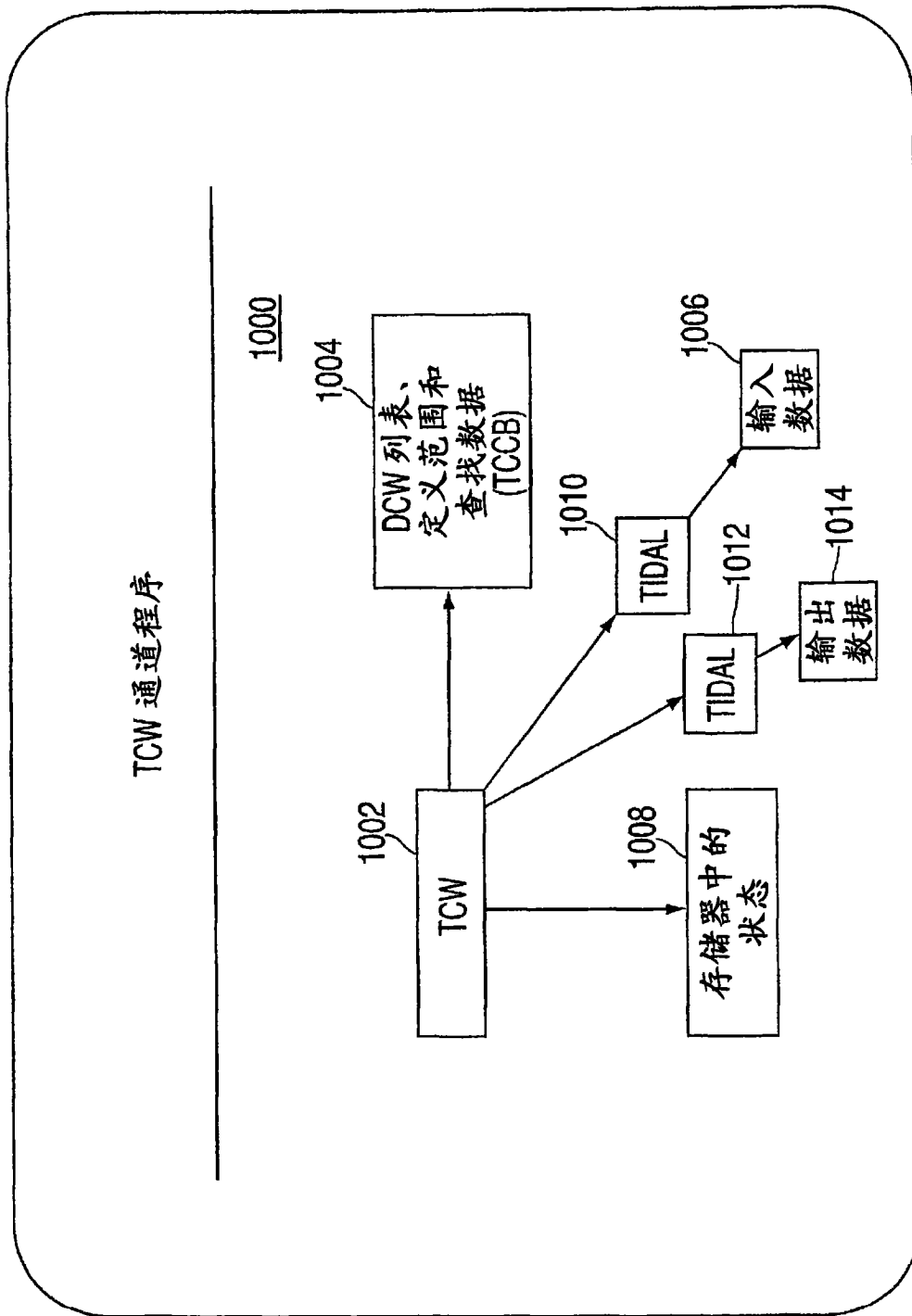


图 10

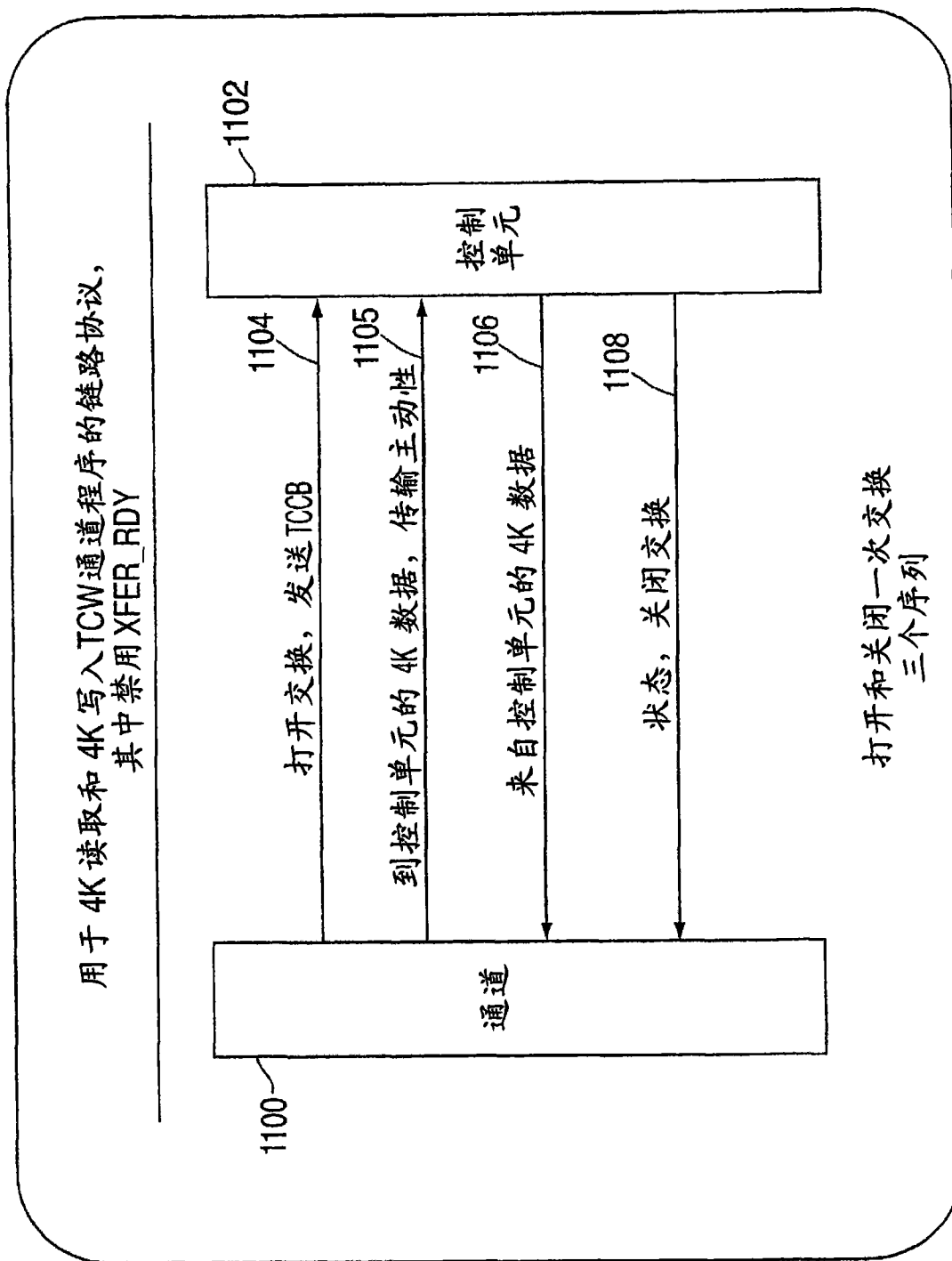


图 11

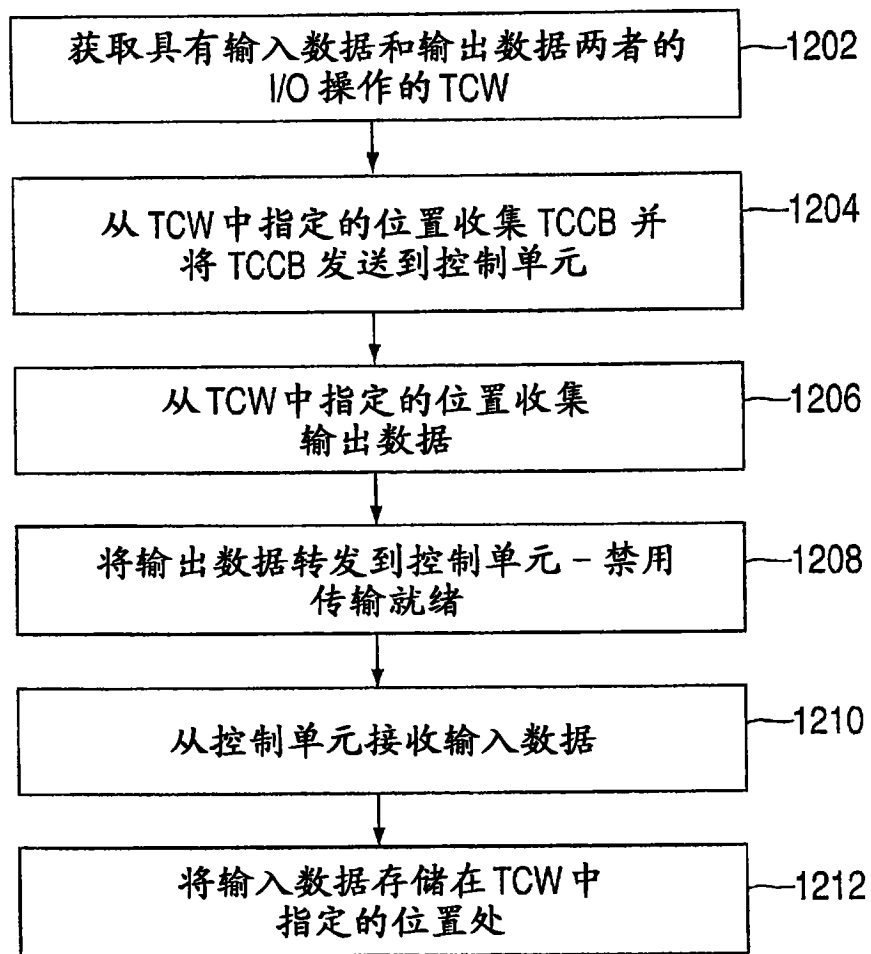


图 12

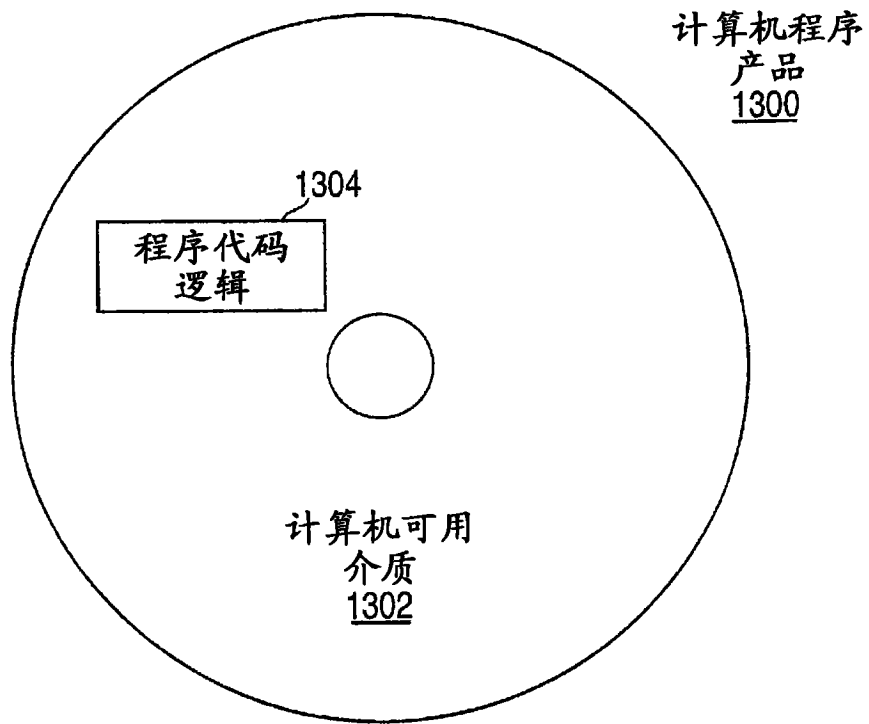


图 13