



(12) 发明专利申请

(10) 申请公布号 CN 101971154 A

(43) 申请公布日 2011. 02. 09

(21) 申请号 200980108974. 9

(22) 申请日 2009. 02. 09

(30) 优先权数据

12/030, 989 2008. 02. 14 US

(85) PCT申请进入国家阶段日

2010. 09. 14

(86) PCT申请的申请数据

PCT/EP2009/051461 2009. 02. 09

(87) PCT申请的公布数据

W02009/101056 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 丹尼尔·卡斯珀 斯科特·卡尔森

约翰·弗拉纳甘 罗杰·哈索恩

凯瑟琳·黄 马修·卡洛斯

路易斯·里奇 戴尔·雷迪

古斯塔夫·希特曼三世

乔库·乌戈丘克乌

哈里·尤登弗兰德

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 黄小临

(51) Int. Cl.

G06F 13/12(2006. 01)

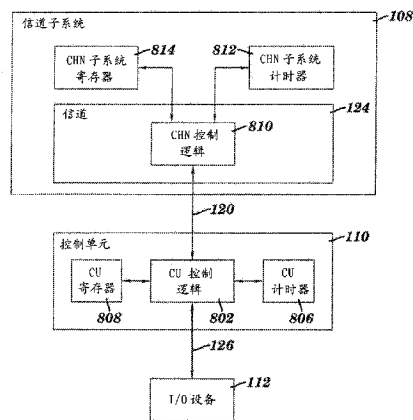
权利要求书 3 页 说明书 14 页 附图 14 页

(54) 发明名称

用于确定输入 / 输出操作的状态的计算机命令和响应

(57) 摘要

本发明提供用于确定输入 / 输出 (I/O) 操作的状态的计算机命令和响应。确定 I/O 处理系统中的 I/O 操作的状态。在信道子系统处接收来自 I/O 操作系统的用于执行所述 I/O 操作的请求, 并将所述请求转发到控制用于执行所述 I/O 操作的 I/O 设备的控制单元。在预定时间量经过而没有接收到来自所述控制单元的、所述 I/O 操作已完成的指示后, 在所述信道子系统处接收来自所述 I/O 操作系统的用于确定所述 I/O 操作的状态的询问请求。从所述信道子系统向所述控制单元发送询问命令。接收来自所述控制单元的响应, 所述响应指示指示执行所述 I/O 操作的所述 I/O 设备的状态、控制执行所述 I/O 操作的所述 I/O 设备的控制单元的状态、以及正被执行的所述 I/O 操作的状态。



1. 一种计算机程序产品,用于确定输入/输出(I/O)处理系统中的I/O操作的状态,该计算机程序产品包括:

实体存储介质,可由处理电路读取,并存储了由所述处理电路执行来进行一种方法的指令,该方法包括:

在信道子系统处接收来自I/O操作系统的用于执行所述I/O操作的请求;

将所述请求从所述信道子系统转发到控制用于执行所述I/O操作的I/O设备的控制单元;

在预定时间量经过而没有接收到来自所述控制单元的、所述I/O操作已完成的指示后,在所述信道子系统处接收来自所述I/O操作系统的用于确定所述I/O操作的状态的询问请求;

响应于在所述信道子系统处接收到所述询问请求,从所述信道子系统向所述控制单元发送询问命令以确定所述I/O操作的状态;以及

在所述信道子系统处接收来自所述控制单元的、指示所述I/O操作的状态的响应,其中所述响应包括指示执行所述I/O操作的所述I/O设备的状态、控制执行所述I/O操作的所述I/O设备的控制单元的状态、以及正被执行的所述I/O操作的状态的信息。

2. 根据权利要求1所述的计算机程序产品,其中,所述I/O操作与用于完成的时间段有关,并且所述预定时间量短于所述用于完成的时间段。

3. 根据权利要求1所述的计算机程序产品,其中,所述询问命令包括指示为什么和何时发出了所述询问命令的信息。

4. 根据权利要求1所述的计算机程序产品,其中,包括在所述响应中的、指示所述控制单元的状态的信息指示所述控制单元是否忙碌、正在执行恢复操作或者正响应于其他询问请求。

5. 根据权利要求1所述的计算机程序产品,其中,包括在所述响应中的、指示所述I/O设备的状态的信息指示所述I/O设备是否被路径组预定、所述控制单元是否处于长时间忙碌状态、或所述I/O设备是否正在执行恢复处理。

6. 根据权利要求1所述的计算机程序产品,其中,包括在所述响应中的、指示所述I/O操作的状态的信息指示操作是否出现、正在执行、等待另一I/O操作的完成、或等待依赖于设备的操作的执行。

7. 根据权利要求1所述的计算机程序产品,其中,包括在来自所述控制单元的所述响应中的信息指示所述I/O设备是否被另一I/O处理系统预定。

8. 一种被适配用于与控制单元通信的装置,该控制单元控制用于执行输入/输出(I/O)处理系统中的I/O操作的I/O设备,所述装置包括:

信道子系统,用于与I/O操作系统和所述控制单元通信,所述信道子系统执行一种方法,该方法包括:

接收来自所述I/O操作系统的用于执行所述I/O操作的请求;

将所述请求转发到控制用于执行所述I/O操作的所述I/O设备的控制单元;

在预定时间量经过而没有接收到来自所述控制单元的、所述I/O操作已完成的指示后,接收来自所述I/O操作系统的用于确定所述I/O操作状态的询问请求;

响应于接收到所述询问请求,向所述控制单元发送询问命令以确定所述I/O操作的状

态 ; 以及

接收来自所述控制单元的、指示所述 I/O 操作的状态的响应, 其中, 所述响应包括指示执行所述 I/O 操作的所述 I/O 设备的状态、控制执行所述 I/O 操作的所述 I/O 设备的控制单元的状态、和正被执行的所述 I/O 操作的状态的信息。

9. 根据权利要求 8 所述的装置, 其中, 所述 I/O 操作与用于完成的时间段相关, 并且所述预定时间量短于所述用于完成的时间段。

10. 根据权利要求 8 所述的装置, 其中, 所述询问命令包括指示为什么和何时发出了所述询问命令的信息。

11. 根据权利要求 8 所述的装置, 其中, 包括在所述响应中的、指示所述控制单元的状态的信息指示所述控制单元是否忙碌、正在执行恢复操作或者正响应于其他询问请求。

12. 根据权利要求 8 所述的装置, 其中, 包括在所述响应中的、指示所述 I/O 设备的状态的信息指示所述 I/O 设备是否被路径组预定、所述控制单元是否处于长时间忙碌状态、或所述 I/O 设备是否正在执行恢复处理。

13. 根据权利要求 8 所述的装置, 其中, 包括在所述响应中的、指示所述 I/O 操作的状态的信息指示操作是否出现、正在执行、等待另一 I/O 操作的完成、或等待依赖于设备的操作的执行。

14. 根据权利要求 8 所述的装置, 其中, 包括在来自所述控制单元的所述响应中的信息指示所述 I/O 设备是否被另一 I/O 处理系统预定。

15. 一种用于确定输入 / 输出 (I/O) 处理系统中的 I/O 操作的状态的方法, 包括:  
在信道子系统处接收来自 I/O 操作系统的用于执行所述 I/O 操作的请求;  
将所述请求从所述信道子系统转发到控制用于执行所述 I/O 操作的 I/O 设备的控制单元;

在预定时间量经过而没有接收到来自所述控制单元的、所述 I/O 操作已完成的指示后, 在所述信道子系统处接收来自所述 I/O 操作系统的用于确定所述 I/O 操作状态的询问请求;

响应于在所述信道子系统处接收到所述询问请求, 从所述信道子系统向所述控制单元发送询问命令以确定所述 I/O 操作的状态; 以及

在所述信道子系统处接收来自所述控制单元的、指示所述 I/O 操作的状态的响应, 其中, 所述响应包括指示执行所述 I/O 操作的所述 I/O 设备的状态、控制执行所述 I/O 操作的所述 I/O 设备的控制单元的状态、和正被执行的所述 I/O 操作的状态的信息。

16. 根据权利要求 15 所述的方法, 其中, 所述 I/O 操作与用于完成的时间段相关, 而所述预定时间量短于所述用于完成的时间段。

17. 根据权利要求 15 所述的方法, 其中, 所述询问命令包括指示为什么和何时发出了所述询问命令的信息。

18. 根据权利要求 15 所述的方法, 其中, 包括在所述响应中的、指示所述控制单元的状态的信息指示所述控制单元是否忙碌、正在执行恢复操作或者正响应于其他询问请求。

19. 根据权利要求 15 所述的方法, 其中, 包括在所述响应中的、指示所述 I/O 设备的状态的信息指示所述 I/O 设备是否被路径组或另一 I/O 处理系统预定、所述控制单元是否处于长时间忙碌状态、或所述 I/O 设备是否正在执行恢复处理。

20. 根据权利要求 15 所述的方法,其中,包括在所述响应中的、指示所述 I/O 操作的状态的信息指示操作是否出现、正在执行、等待另一 I/O 操作的完成、或等待依赖于设备的操作的执行。

## 用于确定输入 / 输出操作的状态的计算机命令和响应

### 技术领域

[0001] 本公开一般涉及输入 / 输出处理,具体地,涉及确定 I/O 操作的状态。

### 背景技术

[0002] 输入 / 输出 (I/O) 操作被用于在 I/O 处理系统的存储器和 I/O 设备间传送数据。具体地,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,和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,采用了 I/O 处理系统的 I/O 子系统。I/O 子系统与主存储器和 I/O 处理系统的 I/O 设备相耦接并引导在存储器和 I/O 设备间的信息流。I/O 子系统的例子是信道子系统。信道子系统利用信道路径作为通信介质。每个信道路径包括与控制单元相耦接的信道,该控制单元又和一个或多个 I/O 设备相耦接。

[0004] 操作系统可以采用信道命令字 (CCW),通过把信道命令字传送给信道子系统,从而在 I/O 设备和存储器间传输数据。CCW 规定了将要执行的命令。对于启动某些 I/O 操作的命令,CCW 指定与操作相关的存储器区域、每当到或从该区域的传输结束时将采取的动作和其他选择。

[0005] 在 I/O 处理期间,通过信道从存储器取回 CCW 的列表。该信道解析来自 CCW 的列表的每个命令,并转发一些命令到与该信道相耦接的控制单元,这些命令的每个命令在该信道自身的实体 (entity) 中。然后,控制单元处理这些命令。该信道跟踪每个命令的状态并控制下一组命令何时将被发送到控制单元进行处理。该信道确保在该信道自身的实体中每个命令被发送到控制单元。此外,该信道获取与处理来自控制单元的对每个命令的响应有关的某些信息。

[0006] 取决于使用的链路协议,关于对花费比期望的或分配的时间更长的时间来完成的 I/O 操作采取什么动作,操作系统做出有根据的决定可能有困难。因此,需要提供具有确定 I/O 操作的状态和确定需花费比期望的或分配的时间更长的时间来执行的 I/O 操作采取的动作的方式的操作系统。

### 发明内容

[0007] 本发明的实施例包括用于确定输入 / 输出 (I/O) 处理系统中的 I/O 操作的状态的计算机程序产品。该计算机程序产品包括:实体存储介质,可由处理电路读取,并存储了由所述处理电路执行来进行一种方法的指令。该方法包括:在信道子系统处接收来自 I/O 操作系统的用于执行所述 I/O 操作的请求;并将所述请求从所述信道子系统转发到控制用于执行所述 I/O 操作的 I/O 设备的控制单元。所述方法还包括:在预定时间量经过而没有接收到来自所述控制单元的、所述 I/O 操作已完成的指示后,在所述信道子系统处接收来自所述 I/O 操作系统的用于确定所述 I/O 操作的状态的询问请求;响应于在所述信道子系统处接收到所述询问请求,从所述信道子系统向所述控制单元发送询问命令以确定所述 I/O 操作的状态;以及在所述信道子系统处接收来自所述控制单元的、指示所述 I/O 操作的状

态的响应。所述响应包括指示执行所述 I/O 操作的所述 I/O 设备的状态、控制执行所述 I/O 操作的所述 I/O 设备的控制单元的状态、以及正被执行的所述 I/O 操作的状态的信息。

[0008] 另外的实施例包括被适配用于与控制单元通信的装置,该控制单元控制在 I/O 处理系统中执行 I/O 操作的输入 / 输出 (I/O) 设备。该装置包括信道子系统,用于与 I/O 操作系统和所述控制单元通信。所述信道子系统执行一种方法,该方法包括:接收来自所述 I/O 操作系统的用于执行所述 I/O 操作的请求;并将所述请求转发到控制用于执行所述 I/O 操作的所述 I/O 设备的控制单元。所述方法还包括:在预定时间量经过而没有接收到来自所述控制单元的、所述 I/O 操作已完成的指示后,接收来自所述 I/O 操作系统的用于确定所述 I/O 操作状态的询问请求;响应于接收到所述询问请求,向所述控制单元发送询问命令以确定所述 I/O 操作的状态;以及接收来自所述控制单元的、指示所述 I/O 操作的状态的响应。所述响应包括指示执行所述 I/O 操作的所述 I/O 设备的状态、控制执行所述 I/O 操作的所述 I/O 设备的控制单元的状态、和正被执行的所述 I/O 操作的状态的信息。

[0009] 另外的实施例包括一种用于确定输入 / 输出 I/O 处理系统中的 I/O 操作的状态的方法。该方法包括:在信道子系统处接收来自 I/O 操作系统的用于执行所述 I/O 操作的请求;将所述请求从所述信道子系统转发到控制用于执行所述 I/O 操作的 I/O 设备的控制单元。所述方法还包括:在预定时间量经过而没有接收到来自所述控制单元的、所述 I/O 操作已完成的指示后,在所述信道子系统处接收来自所述 I/O 操作系统的用于确定所述 I/O 操作状态的询问请求;响应于在所述信道子系统处接收到所述询问请求,从所述信道子系统向所述控制单元发送询问命令以确定所述 I/O 操作的状态;以及在所述信道子系统处接收来自所述控制单元的、指示所述 I/O 操作的状态的响应。所述响应包括指示执行所述 I/O 操作的所述 I/O 设备的状态、控制执行所述 I/O 操作的所述 I/O 设备的控制单元的状态、和正被执行的所述 I/O 操作的状态的信息。

[0010] 在回顾下述附图和详细描述时,根据各实施例的其他系统、方法和 / 或计算机程序产品对于在本领域的普通技术人员将是或者变得显而易见的。意要所有这样的另外的系统、方法和 / 或各件产品被包括在本说明书中、在本发明的范围内、并且由所附的权利要求保护。

## 附图说明

[0011] 在本说明书的结论中的权利要求中,具体指出和清楚地要求保护作为本发明的主题。通过结合附图的下述具体说明,本发明的前述和其他的目的、特征及有益效果是显而易见的,在附图中:

[0012] 图 1 描述包含并使用本发明的一个或多个方面的 I/O 处理系统一个实施例;

[0013] 图 2A 描述信道命令字的一个例子;

[0014] 图 2B 描述信道命令字信道程序的一个例子;

[0015] 图 3 描述用于信道和控制单元间的通信以执行图 2B 信道命令字信道程序的链路协议的一个例子;

[0016] 图 4 描述根据本发明的一个方面的传输控制字信道程序的一个实施例;

[0017] 图 5 描述根据本发明的一个方面的用于在信道和控制单元之间通信以执行图 4 的信道命令字信道程序的链路协议的一个实施例;

- [0018] 图 6 描述用于在信道和控制单元之间通信以便执行信道命令字信道程序的 4 个读取命令的链路协议的一个实施例；
- [0019] 图 7 描述根据本发明的一个方面的用于在信道和控制单元之间通信以处理传输控制字信道程序的 4 个读取命令的链路协议的一个实施例；
- [0020] 图 8 描述根据本发明的一个方面的控制单元和信道的一个实施例；
- [0021] 图 9 描述根据本发明的一个方面的包括询问 TCW 地址字段的传输控制字 (TCW) 的一个实施例；
- [0022] 图 10 描述根据本发明的一个方面的询问 DCW 的一个实施例；
- [0023] 图 11 描述根据本发明的一个方面的传送响应 IU 的一个实施例；
- [0024] 图 12A 描述根据本发明的一个方面的由 I/O 操作系统执行的用于确定何时请求来自控制单元的 I/O 操作的状态的处理的一个实施例；
- [0025] 图 12B 描述根据本发明的一个方面的用于询问控制单元以确定 I/O 操作的状态的处理的一个实施例；
- [0026] 图 13 描述并入了本发明的一个或多个方面的计算机程序产品的一个实施例。
- [0027] 详细的描述参考附图通过例子说明了本发明的优选实施例,以及有益效果和特性。

### 具体实施方式

[0028] 根据本发明的一个方面,方便了输入 / 输出 (I/O) 处理。例如,通过使能快捷地访问与 I/O 处理有关的诸如状态和测量数据的信息,方便了 I/O 处理。此外,在一个例子中,通过减少用于执行 I/O 处理的 I/O 处理系统的组件间的通信,方便了 I/O 处理。例如,减少在诸如信道的 I/O 通信适配器和控制单元间的交换和序列 (sequence) 的数量。这通过从 I/O 通信适配器发送多个命令到控制单元作为由控制单元执行的单一实体以及通过控制单元将从命令得到的数据 (如果有的话) 作为单一实体而发送来完成。

[0029] 多个命令被包含在块中,该块在此被称为传输命令控制块 (TCCB),其地址在传输控制字 (TCW) 中规定。将 TCW 从操作系统或其他应用发送到 I/O 通信适配器,该 I/O 通信适配器又在命令消息中将该 TCCB 转发到控制单元以处理。控制单元通过 I/O 通信适配器处理缺少对与那些各个命令相关的状态的跟踪的每个命令。多个命令也称作信道程序,其在控制单元而不是 I/O 通信适配器中被解析和执行。

[0030] 在示例性实施例中,响应于执行信道程序,控制单元产生包含状态和扩展的状态信息的响应消息。控制单元还可以在有限数量的通信方案下不执行信道程序而产生响应消息,例如以便通知 I/O 通信适配器:将不会执行信道程序。控制单元可以包括一些元素以支持在 I/O 通信适配器和 I/O 设备间的通信,以及支持信道程序的执行。例如,除了便于通信和状态监视的一个或多个队列 (queue)、计时器、和寄存器外,控制单元可以包括控制逻辑以解析和处理消息。I/O 通信适配器解析响应消息,提取状态和扩展状态信息,并且使用提取的信息执行进一步的计算,比如确定扩展的测量字。

[0031] 参考图 1,结合并使用本发明的一个或多个方面,描述 I/O 处理系统的一个例子。I/O 处理系统 100 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、信道子系统 108、一个或多个控制单元 110 和一个或多个 I/O 设备 112,其中每一

个将在以下描述。

[0032] 主存储器 102 存储可以从 I/O 设备 112 输入的数据和程序。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 103。主存储器 102 是直接可寻址的,并且提供用于 CPU 104 和信道子系统 108 对数据的高速处理。

[0033] CPU 104 是 I/O 处理系统 100 的控制中心。其包括用于指令执行、中断反应、计时功能、初始程序装载、和其他机器相关的功能的排序和处理设施。CPU 104 经由诸如双向或单向总线的连接 114 与存储控制元件 106 相耦接。

[0034] 存储控制元件 106 经由诸如总线的连接 116 与主存储器 102 相耦接,经由连接 114 与 CPU 104 相耦接,并经由连接 118 与信道子系统 108 相耦接。存储控制元件 106 控制例如由 CPU 104 和信道子系统 108 做出的请求的排队和执行。

[0035] 如上所述,信道子系统 108 与存储控制元件 106 相耦接,并经由诸如串行链路的连接 120 与每个控制单元 110 相耦接。连接 120 可以实现为采用单模或多模光波导的光链路。信道子系统 108 引导在 I/O 设备 112 和主存储器 102 间的信息流。其减少 CPU 104 直接与 I/O 设备 112 通信的任务,并允许数据处理和 I/O 处理同时进行。信道子系统 108 使用一个或多个信道路径 122 作为在管理到或从 I/O 设备 112 的信息流时的通信链路。作为 I/O 处理的一部分,信道子系统 108 还执行测试信道路径可用性、选择可用信道路径 122 和启动与 I/O 设备 112 的操作的执行的执行的路径管理功能。

[0036] 每个信道路径 122 包括信道 124(在一个例子中,信道 124 位于信道子系统 108 内,如图 1 所示)、一个或多个控制单元 110 和一个或多个连接 120。在另一个例子中,还能够具有一个或多个动态开关(未示出)作为信道路径 122 的一部分。动态开关与信道 124 和控制单元 110 相耦接,并提供物理上互连任何两个附连到开关的链路的能力。在另一个例子中,还能够具有多个系统及因此的附连到控制单元 110 的多个信道子系统。

[0037] 子信道(未示出)也位于信道子系统 108 内。一个子信道被提供用于并专用于通过信道子系统 108 可访问程序的每个 I/O 设备 112。子信道(例如数据结构,比如表格)代表设备对于程序的逻辑状态。每个子信道向信道子系统 108 提供关于相关的 I/O 设备 112 和其附连物的信息。子信道还提供关于 I/O 操作和涉及相关的 I/O 设备 112 的其他功能的信息。子信道是这样的工具,信道子系统 108 通过该工具向在 CPU 104 上运行的操作系统提供关于相关的 I/O 设备 112 的信息,所述操作系统通过执行 I/O 指令获得这些信息。

[0038] 信道子系统 108 与一个或多个控制单元 110 相耦接。每个控制单元 110 提供逻辑以操作并且控制一个或多个 I/O 设备 112,并且通过使用公共设备使每个 I/O 设备的特性适配于信道 124 提供的链路接口。公共设备用于 I/O 操作的执行、关于 I/O 设备 112 和控制单元 110 的状态的指示、在信道路径 122 上的数据传输的定时的控制和某些级别的 I/O 设备 112 控制。

[0039] 每个控制单元 110 经由连接 126(比如总线)附连到一个或多个 I/O 设备 112。各 I/O 设备 112 接收信息或在主存储器 102 和 / 或其他存储器中存储信息。举例来说,I/O 设备 112 的例子包括读卡器和打卡器、磁带单元、直接存取存储设备、显示器、键盘、打印机、定点(pointing)设备、远程处理设备、通信控制器和基于传感器的设施。

[0040] I/O 处理系统 100 的一个或多个上述组件在以下文献中进一步描述:“IBM® z/Architecture Principles of Operation,”公开号 No. SA22-7832-05,第 6 版,2007 年 4 月;



美国专利号 No. 5, 461, 721, 题为“System For Transferring Data Between I/O Devices And Main Or Expanded Storage Under Dynamic Control Of Independent Indirect Address Words (IDAWS),” Cormier 等人, 发表于 1995 年 10 月 24 日; 以及美国专利号 No 5, 526, 484, 题为“Method And System Form Pipelining The Processing Of Channel Command Words,” Casper 等人, 发表于 1996 年 6 月 11 日, 其中每个的全部内容在此通过参考而并入。IBM 是美国纽约阿蒙克的国际商业机器公司的注册商标。此处所用的其他名字可以是国际商业机器公司或其它公司的注册商标、商标或产品名称。

[0041] 在一个实施例中, 为了在 I/O 设备 112 和存储器 102 间传输数据, 使用信道命令字 (CCW)。CCW 指定了待执行的命令, 并包含用于控制处理的其他字段。参考图 2A 描述了 CCW 的一个例子。CCW 200 例如包括: 规定待执行的命令的命令代码 202 (比如读取、反向读取、控制、检测 (sense) 和写入); 用于控制 I/O 操作的多个标志 204; 用于规定数据传送的命令的计数字段 206, 其规定了待传送的、由 CCW 指定的存储区域中的字节的数量; 以及数据地址 208, 其在采用直接寻址时指向包含数据的主存储器中的位置, 或在采用修改的间接数据寻址时指向要处理的修改的间接数据地址字 (MIDAW) 的列表 (比如, 相邻列表)。在 Brice 等人提交于 2006 年 8 月 15 日的题为“Flexibly Controlling The Transfer Of Data Between Input/Output Devices And Memory,” 的美国申请序列号 11/464, 613 中进一步描述了修改的间接寻址, 其全部内容通过参考合并于此。

[0042] 被安排为顺序执行的一个或多个 CCW 形成了信道程序, 在此也称为 CCW 信道程序。例如通过操作系统或其它软件建立 CCW 信道程序。软件建立 CCW 并获得分配给该信道程序的存储器的地址。参考图 2B 描述了 CCW 信道程序的一个例子。CCW 信道程序 210 例如包括定义范围 (define extent) CCW 212, 其具有指向将随定义范围命令使用的定义范围数据 216 的存储器中的位置的指针 214。在这个例子中, 信道中的传送 (TIC) 218 遵循将信道程序交给存储器中的另一区域 (比如, 应用区域) 的定义范围命令, 该另一区域包括一个或多个其它 CCW, 比如具有指向定位 (locate) 记录数据 220 的指针的定位记录 217, 以及一个或多个读取 CCW 221。每个读取 CCW 221 具有指向数据区域 224 的指针 222。数据区域包括用于直接存取数据的地址或用于间接存取数据的数据地址字 (比如 MIDAW 或 IDAW) 的列表。此外, CCW 信道程序 210 包括称为子信道的、由设备地址定义的信道子系统内的预定区域, 用于从执行 CCW 信道程序得到的状态 226。参考图 3 以及参考图 2B 描述了 CCW 信道程序的处理。具体地, 图 3 示出了当执行 CCW 信道程序时, 在信道和控制单元间发生的各种交换和序列的例子。在这个例子中, 用于通信的链路协议是 FICON (光纤连接)。在“Fibre Channel Single Byte Command Code Sets-2 Mapping Protocol (FC-SB-3), T11/Project 1357-D/Rev. 1.6, INCITS (2003 年 3 月) 中描述了关于 FICON 的信息, 其全部内容通过参考合并于此。

[0043] 参考图 3, 信道 300 打开与控制单元 302 的交换, 并且向控制单元 302 发送定义范围命令和与其相关的数据 304。从定义范围 CCW 212 (图 2B) 取回命令, 并且从定义范围数据区域 216 获得数据。信道 300 使用 TIC 218 查找定位记录 CCW 和读取 CCW。信道 300 从定位记录 CCW 217 (图 2B) 取回定位记录命令 305 (图 3) 并且从定位记录数据 220 获得数据。从读取 CCW 221 (图 2B) 取回读取命令 306 (图 3)。每个被发送到控制单元 302。

[0044] 响应于信道 300 的交换的打开, 控制单元 302 打开与信道 300 的交换 308。这可以

发生在定位命令 305 和 / 或读取命令 306 之前或之后。随着打开交换, 响应 (CMR) 被转发到信道 300。CMR 向信道 300 提供控制单元 302 是活动的并正在工作的指示。

[0045] 控制单元 302 向信道 300 发送所请求的数据 310。此外, 控制单元 302 将状态提供给信道 300, 并关闭交换 312。响应于此, 信道 300 存储该数据, 检查各状态并关闭交换 314, 这向控制单元 302 指示: 状态已被接收。

[0046] 读取 4K 数据的上述 CCW 信道程序的处理需要打开和关闭两个交换并需要七个序列。通过将信道程序的多个命令压缩 (collapse) 成一个 TCCB, 减少了在信道和控制单元间的交换和序列的总数。信道, 例如图 1 的信道 124, 使用 TCW 以识别 TCCB 的位置以及用于存取和存储与执行信道程序有关的状态和数据的位置。TCW 由信道解释 (interpret), 而不被控制单元发送或可见。

[0047] 参考图 4 描述用于读取 4k 数据的如图 2B 中包括 TCCB 而不是分离的各个 CCW 的信道程序的一个例子。如所示, 信道程序 400 (此处称作 TCW 信道程序) 包括指定 TCCB 404 在存储器中的位置以及数据区域 406 或指向数据区域 406 的 TIDAL 410 (即与 MIDAW 类似的传输模式间接数据地址字 (TIDAW) 的列表,) 和状态区域 408 在存储器中的位置的 TCW 402。TCW、TCCB 和状态将在以下进一步详细描述。

[0048] 参考图 5 描述 TCW 信道程序的处理。用于这些通信的链路协议例如是光纤信道协议 (FCP)。具体地, 使用 FCP 链路协议的三个阶段, 允许使用支持 FCP 执行由 CCW 控制的数据传送的主机总线适配器。FCP 及其阶段在 T10 Project 1560-D, Revision 4, 2005 年 9 月 13 日的 “Information Technology-Fibre Channel Protocol for SCSI, Third Version (FCP-3)” 中进一步描述, 其全部内容通过参考合并于此。

[0049] 参考图 5, 信道 500 打开与控制单元 502 的交换并且向控制单元 502 发送 TCCB 504。在一个例子中, TCCB 504 和序列发起 (sequence initiative) 在称为 FCP\_CMND 信息单元 (IU) 或传输命令 IU 的 FCP 命令中被传送到控制单元 502。控制单元 502 执行 TCCB 504 的多个命令 (例如, 作为设备控制字 (DCW) 的定义范围命令、定位记录命令、读取命令), 并且例如经由 FCP\_Data IU 将数据 506 转发到信道 500。控制单元 502 还提供状态并关闭交换 508。作为一个例子, 在 FCP 状态帧中发送最终状态, 该 FCP 状态帧具有例如在也称为传输响应 IU 的 FCP\_RSP IU 的有效载荷的字节 10 或 11 中有效的位。FCP\_RES\_IU 有效载荷可用被用于将 FICON 结束状态与附加状态信息一起传输, 该附加状态信息包括支持扩展测量字的计算和通知信道 500 由控制单元 502 支持的打开交换的最大数的参数。

[0050] 在另一个例子中, 为了写入 4k 的客户数据, 信道 500 使用如下的 FCP 链路协议阶段:

[0051] 1. 在 FCP\_CMND IU 中传送 TCCB。

[0052] 2. 向控制单元 502 传送数据的 IU 和序列发起。

[0053] (FCP 传送准备好被禁用)

[0054] 3. 在 FCP 状态帧中发送最终状态, 该 FCP 状态帧具有例如在 FCP\_RSP IU 的有效载荷的 10 或 11 字节中有效的位。FCP\_RES\_INFO 字段或检测字段被用于将 FICON 结束状态与附加状态信息一起传输, 该附加状态信息包括支持扩展测量字的计算并通知信道 500 由控制单元 502 支持的打开交换的最大数的参数。

[0055] 通过执行图 4 的 TCW 信道程序, 只有一个交换被打开和关闭 (还见图 5), 而不是对

于图 2B 的 CCW 信道程序的两个交换（还见图 3）。此外，与用于 CCW 信道程序的七个序列（见图 2B-图 3）相比，对于 TCW 信道程序，存在三个通信序列（见图 4-图 5），。

[0056] 对于 TCW 信道程序，即使将另外的命令添加到该程序，交换和序列的数量仍然相同。例如，将图 6 的 CCW 信道程序的通信与图 7 的 TCW 信道程序的通信相比较。在图 6 的 CCW 信道程序中，以单独的序列将每个命令（例如，定义范围命令 600、定位记录命令 601、读取命令 602、读取命令 604、读取命令 606、定位记录命令 607 和读取命令 608）从信道 610 发送到控制单元 612。此外，以单独的序列将每个 4k 的数据块（例如数据 614-620）从控制单元 612 发送到信道 610。该 CCW 信道需要打开和关闭两个交换（例如，打开交换 622、624 和关闭交换 626、628）并需要 14 个通信序列。这可与用于图 7 的 TCW 信道程序的三个序列和一个交换相比，后者完成了与图 6 的 CCW 信道程序相同的任务。

[0057] 如图 7 中所述，信道 700 打开与控制单元 702 的交换，并向控制单元 702 发送 TCCB704。如上所述，TCCB 704 包括 DCW 中的定义范围命令、两个定位记录命令和四个读取命令。响应于接收到 TCCB 704，控制单元 702 执行命令，并以单个序列向信道 700 发送 16k 的数据 706。此外，控制单元 702 向信道 700 提供状态，并且关闭交换 708。因此，TCW 信道程序需要少得多的系统开销（overhead）来传送与图 6 的 CCW 信道程序相同的数据量。

[0058] 现在转到图 8，更详细地描绘支持 TCW 信道程序执行的图 1 的控制单元 110 和信道 124 的一个实施例。控制单元 110 包括 CU 控制逻辑 802，用于解析和处理包含经由连接 120 从信道 124 接收的诸如图 7 的 TCCB 704 的 TCCB 命令信息的命令消息。CU 控制逻辑 802 可以在控制单元 110 处接收的 TCCB 中提取 DCW 和控制数据以控制设备，例如控制经由连接 126 的 I/O 设备 112。CU 控制逻辑 802 向 I/O 设备 112 发送设备命令和数据，并从 I/O 设备 112 接收状态信息和其它反馈。例如，由于目标为 I/O 设备 112 的前一预定（reservation）请求，I/O 设备 112 可能是忙碌的。为了管理在控制单元 110 接收到存取同一 I/O 设备 112 的多个请求时会发生的潜在的设备预定竞争问题，CU 控制逻辑 802 保持追踪设备忙碌消息和相关数据，并将其存储在设备忙碌队列 804 中。

[0059] CU 控制逻辑 802 可以存取和控制控制单元 110 中的其它元件，比如 CU 计时器 806 和 CU 寄存器 808。CU 计时器 806 可以包括多个计时器功能以追踪 I/O 操作的序列花费多长时间完成。CU 计时器 806 还可以包括一个或多个倒数计时器以监视和中断没有在预定时段内完成的 I/O 操作和命令。CU 寄存器 808 可以包括提供配置和状态信息的固定值，以及在命令被 CU 控制逻辑 802 执行时所更新的动态状态信息。控制单元 110 还可以包括其它缓冲器或存储器元件（未示出）以存储与在信道 124 和 I/O 设备 112 之间的通信有关的多个消息或状态信息。CU 寄存器 808 可以包括最大控制单元交换参数，其定义了控制单元 110 支持的打开控制单元交换的最大数量。

[0060] 信道子系统 108 中的信道 124 包括用于支持与控制单元 110 的通信的多个元件。例如，信道 124 可以包括 CHN 控制逻辑 810，其与 CHN 子统计器 812 和 CHN 子系统寄存器 814 接口。在示例实施例中，CHN 控制逻辑 810 控制在信道子系统 108 和控制单元 110 之间的通信。CHN 控制逻辑 810 可以经由连接 120 直接与 CU 控制逻辑 802 接口，以发送命令和接收响应，比如诸如传输命令和响应 IU。或者，发消息（messaging）接口和 / 或缓冲器（未示出）可以被放置在 CHN 控制逻辑 810 和 CU 控制逻辑 802 之间。CHN 子统计器 812 可以包括多个计时器功能以追踪除了由控制单元 110 追踪的时间之外、I/O 操作序列花

费多长时间完成。CHN 子系统计时器 812 还可以包括一个或多个倒数计时器以监视和中断在预定时段内未完成的命令序列。CHN 子系统寄存器 814 可以包括提供配置和状态信息以及动态信息的固定值,以及在传输命令并接收到响应时更新的动态状态信息。

[0061] 来自控制单元的 FICON 命令响应 (CMR) 帧不是光纤信道扩展 (Fibre Channel Extension, FCX) 传输模式协议的一部分。从传输模式协议中移除 CMR 有助于改善 FCX 的性能。FICON 中的 CMR 通知信道:控制单元已经收到命令,并且信道发送执行该命令。当 FICON 信道接收到 CMR,信道将子信道标记为“子信道和有效设备”。

[0062] 在所有的计算环境中,可能发生在不同 I/O 设备处的中断。如果请求 I/O 设备处的操作的 OS 未能检测中断,则这可能导致数据处理系统中的操作减慢并最终停止。错过的中断处理机 (MIH) 是例如被包括在 OS 103 中的机制,其在通过对正在进行的 I/O 操作计时并确定由 I/O 设备执行操作所花的时间是否已经超过为该操作的执行所分配或设置的预定的“正常”时间量来检测丢失的中断时是有用的。如果达到 MIH 时间,并且 I/O 设备还没有完成操作的执行,则这是如下的指示:可能已经错过中断、链路故障出现、适配器故障出现、控制单元错误出现或由共享系统持有的预留超过预期。

[0063] 当对于 FICON 操作系统 MIH 时间已到时,即达到 MIH 时间时,要看子信道是否被标记了“子信道和设备活动”以确定接下来采取什么动作。对 FCX,子信道在整个操作期间停留在“开始挂起 (start pending)”。这样,通过 FCX,当 MIH 时间到时,I/O 操作系统不能告知 I/O 操作的状态,因为对于整个操作,子信道状态停留在“开始挂起”。

[0064] 根据本发明的一个方面,就在错过的中断超时之前,例如,在达到 MIH 时间前一秒,操作系统使用询问命令来确定控制单元处的 I/O 操作的状态。可以在分配给完成执行 I/O 操作的时间用尽之前并且 I/O 操作还没有完成时利用取消子信道指令来发起询问命令。

[0065] 存在询问命令的几个益处。例如,当 MIH 超时将要发生时,执行询问命令,由此除去了对关于每个 I/O 操作的 CMR 的需要。除去对关于每个 I/O 操作的 CMR 的需要通过减少构造流量 (fabric traffic) 以及信道和适配器系统开销而提高了 FCX 性能。同样,询问命令向控制单元传送关于 OS 的信息,用于如果超时出现、则由控制单元记录 (log)。另一益处是,控制单元向 OS 提供回关于 I/O 操作的详细状态信息,而用于 FICON 的 CMR 仅指示控制单元当前正在执行 I/O。同样,如果 I/O 操作丢失,则通过询问命令而交换的信息对问题确定非常有用。

[0066] 在此所述的询问的实现涉及取消子信道指令和 TCW 中的询问 -TCW 地址字段,并且从信道子系统角度、来自信道子系统的询问命令和响应角度、以及来自控制单元的询问命令和响应角度来描述。这些的每个描述如下。

[0067] 在图 9 中描绘了传输控制字 (TCW) 900 的示例实施例。TCW 900 由信道 124 利用来建立 I/O 操作,并且不被发送到控制单元 110。图 9 中所绘的 TCW 用于从信道子系统角度的询问的实现。

[0068] 在图 9 所绘的示例 TCW 900 中,例如等于“00b”的格式字段 902 指示接下来的是 TCW 900。TCW 900 也包括标志字段 906。标志字段 906 的前五位预留给将来使用并且被设置为零。标志字段 906 的第六位在数据地址字段 914 包含 TIDAL 的地址时被设置为 1。如果 TIDAL 数据地址标志被设置为零,那么数据地址字段 914 包括数据地址。标志字段 906

的第七位是 TCCBTIDAL 标志。在示例实施例中, TCCB TIDAL 标志在 TCCB 地址字段 922 包括 TCCB TIDAL 的地址时被设置为 1。如果 TCCB TIDAL 标志被设置为零,那么 TCCB 地址字段 922 直接寻址 TCCB。标志字段 906 的第八到第二十四位被预留将来使用。字段 907 可以被预留将来使用。

[0069] TCW 900 还包括 TCCB 长度字段 910,其直接表示 TCCB 的长度,并且可以被用来确定 TCCB 的实际长度。R/W 910 字段包括读 / 写位,其被用来指示作为执行 TCW 900 的结果、数据是否正被读取和 / 或写入。在示例实施例中,读 / 写位的读位被设置为 1,以指示作为执行 TCW 900 的结果、输入数据正从 I/O 设备 112 传送到主系统 101 中的系统存储器(例如主存储器 102)。读 / 写位中的写位被设置为 1,以指示作为执行 TCW 900 的结果、输出数据正从主系统 101 中的系统存储器(例如主存储器 102)传送到 I/O 设备。字段 912 可以预留将来使用。

[0070] 地址字段 914 可以包括每个标志字段位 6 的直接地址或间接地址。地址字段 914 的内容可以是输出数据的 TIDAL 的地址(传输模式间接数据地址字的列表)或者输出数据的实际地址。地址字段 914 的内容可以是输入数据的 TIDAL 的地址或者输入数据的实际地址。在示例实施例中,输出数据地址和输入数据地址被包括在单个字段 914 中,并且字段 916 被预留将来使用。或者,可以在字段 914 和 916 之间分开输出数据地址和输入数据地址。

[0071] TCW 900 还包括传输状态块地址字段 920。在该地址处存储对 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)。TCW 900 中的 TCCB 地址字段 922 包括 TCCB 位于系统存储器中的地址。这是对于 TCW900 的要执行 DCW 所存在的控制块。同样如标志字段第 7 位中所述, TCCB 地址字段 922 的内容可以是对于 TCCB 的 TIDAL 地址或 TCCB 的实际地址。TCW 900 中的数据字节计数字段 924 指示对于输出操作要由 TCW 传送的输出数据量或对于输入操作要由 TCW 传送的输入数据量。字段 926 可以预留将来使用。或者,可以在字段 924 和 926 之间分开输出数据计数和输入数据计数信息。在 TCW 900 中预留了几个另外的字段:预留字段 928、预留字段 930 和预留字段 932。

[0072] 根据本发明的一个方面,将 TCW 900 例如从 32 字节扩展到 64 字节,以给将来的功能留出更多的空间。一个这样的功能是通过询问 -TCW 地址字段 934 使得实现的询问功能,该询问 -TCW 地址字段 934 包括指示如果 I/O 未能在分配的时间段内完成、是否应该执行询问的询问值。询问 -TCW 地址字段 934 可以包括另一 TCW 的地址,并且可以由信道 124 使用以在取消子信道 I/O 指令的发起下询问操作的状态,详细说明如下。

[0073] TCW 900 可以由软件建立以由信道使用来驱动 I/O 操作。图 9 绘出的 TCW 是如何配置命令字的一个例子。在包括另外的字段和 / 或不包括在图 9 中所绘的字段的情况下,其他配置也是可能的。

[0074] 根据本发明的一个方面,如果 FCX 开始子信道已经被发送到信道,那么执行取消子信道指令以确定控制单元的状态。如果子信道是“开始挂起”,并且开始子信道已经被发送到该信道,并且 TCW 中的询问 TCW 地址字段的值不为零,那么取消指令将询问命令排队在可以由信道子系统发送到控制单元的子信道中。在数据传送阶段或在传输响应 IU 的扩展状态部分中,该询问命令可以致使关于正被询问的操作的状态的信息从控制单元返回。可以如下实现询问操作的协议。

[0075] 如果 FCX I/O 操作是活动的,则对于该 I/O 操作的 TCW 中的询问 TCW 地址用于指向询问 TCW。如果信道遇到零值的询问 TCW 地址,则信道将不发起询问。在丢失中断 (MIH) 时间到之前,如果 OS 想要询问 I/O 设备,则 OS 利用询问 TCW 的地址来更新 TCW 中的询问 TCW 地址字。如果 OS 仅想要发送取消指令而不引导询问,那么 OS 使 TCW 中的寻址 TCW 地址设置为零。然后利用取消指令将询问发起传递到信道子系统。取消指令执行当前构造的取消 Cancel,但是如果子信道是具有 FCX 开始子信道的“开始挂起”,并且开始 Start 已经被传送到该信道,那么发起被给予信道子系统以询问控制单元。

[0076] 根据本发明的一个方面,如果子信道是空闲的、具有原来的或警告的状态的中断挂起、或者仅是设备活动的,那么放弃发出询问命令的发起。如果信道在“开始挂起”的子信道处接收到询问发起,并且开始仍然在信道中排队,那么信道放弃询问发起。如果信道接收询问发起,并且该信道已经具有正在进行的询问操作,那么该信道放弃新的询问发起。

[0077] 如果信道接收对处于开始挂起并具有对控制单元打开的交换的 UA 的询问发起,那么信道子系统执行询问。根据本发明的一个方面,在执行询问中,信道子系统进行以下。信道子系统取回当前 TCW 中的询问 TCW 地址以得到用于取回询问 TCW 的指针值。如果指针是全零,那么信道放弃询问。在这种情况下,OS 想要进行取消指令而不是询问。如果指针有效,那么信道子系统继续。信道子系统打开新的交换,并且在传输命令 IU 中将由询问 TCW 寻址的 TCCB 内的询问 DCW 发送到控制单元。该操作由信道对完成计时。如果询问操作在信道设置的时间量内未完成,那么信道中止询问操作和正被询问的操作两者。然后子信道被返回到具有接口控制检查状态的 OS。

[0078] 控制单元接收并执行询问命令,基于传输命令 IU 中的命令在传输响应 IU 中将关于 UA 的询问信息传送到信道子系统或者作为数据 IU。在正被询问的 UA 上活动的原始操作不受影响。I/O 子系统可以产生向 OS 报告询问完成的、询问完成位被设置为 1 的中间状态中断。

[0079] 根据本发明的一个方面,询问命令是唯一可以对支持 FCX 的所有控制单元类型都相同的命令。用于询问的传输命令 IU 仅包括一个 DCW。该询问 DCW 命令可以具有与其相关的达到 232 字节的控制数据,这可以是被传递给控制单元的指示为什么正在执行询问的信息。

[0080] 询问命令传输在询问 DCW 1010 中包括的信息以及作为询问命令 1000 的一部分的询问控制数据 1020,如图 10 所示。控制单元用将关于 I/O 操作的信息返回给 OS 的传送响应 IU 来响应。根据本发明的一个方面在传送响应 IU 中返回的信息的格式如图 11 中的项 1100 所示。

[0081] 图 10 描述了根据本发明的一个方面的 DCW 1000 的一个实施例。在示例实施例中,DCW 1000 是八字节长度加上控制数据计数字段 1014 的长度。DCW 包括命令字段 1011、标志字段 1012、预留字段 1013、控制数据 (CD) 计数字段 1014 以及数据字节计数字段 1015、1016、1017 和 1018。DCW 命令字段 1011 是一字节长度,与在 CCW 中利用的 CCW 命令字节相同(但是可能包括未被 CCW 利用的另外的命令代码)。标志字段 1012 包括八位。在示例实施例中,第二位是对 TCCB 中的下一 DCW 1000 的链(chain)命令。当该标志位被设置为零时,其指示这是 TCCB 中的 DCW 程序的最后的 DCW1000。标志字段 1016 的其他位被预留并且被设置为零。询问 DCW 1000 具有包括 40 的十六进制数的 DCW 命令代码字段 1011。控制数

据计数字段 1014 指示随 DCW 包括的询问控制数据的量。字段 1015-1018 包括可以由询问 DCW 传送的读取数据的 4 字节的数据量。

[0082] 如果询问 DCW 的控制数据计数 1014 大于零,那么在 DCW 中指定询问控制数据。发送到控制单元的询问控制数据 1020 用于依赖设备的记录目的,并且被用来帮助排除 I/O 超时。

[0083] 根据本发明的一个方面,询问控制数据 1020 具有如下所述的格式。参考图 10,询问控制数据 1020 的字 0 的字节 0 {Fmt 字段 1021} 包括定义了询问数据的布局或格式 (FMT) 的无符号整数值。字 0 的字节 1 (RC 字段 1022) 包括指示由 OS 发起询问操作的原因的无符号整数值或原因代码 (RC)。RC 值的含义可以如下:

[0084] 0 未指出询问原因。

[0085] 1 超时:正被询问的操作被程序检测超时。

[0086] 2-225 预留。

[0087] 询问控制数据的字 0 的字节 2 (RCQ 字段 1023) 包括指示关于发起询问操作的原因的附加信息的无符号整数值,称作原因代码限定符 (RCQ)。当 RC 字段 1022 包括值 1 时,RCQ 值的含义可以如下:

[0088] 0 未指出询问原因限定符。

[0089] 1 主要:由主要程序检测到超时。

[0090] 2 次要:由次要程序检测到超时。

[0091] 3-225 预留。

[0092] 当 RC 字段 1022 不包括值 1 时,RCQ 没有意义。

[0093] 字 0 的字节 3 {LPM 字段 1024} 包括在正被询问的操作由开始子信道命令发起时所使用的逻辑路径掩码 (LMP)。

[0094] 参考询问控制数据的字 1,字 1 的字节 0 {PAM 字段 1025} 包括在发起询问操作时的路径可用掩码 (LMP) 的值。字 1 的字节 1 (PIM 字段 1026) 包括在发起询问操作时的路径安装掩码 (Path-Installed Mask, PIM) 的值。字 1 的字节 2-3 (超时字段 1027) 对指示分配用于完成 I/O 操作的时间的超时值的指示。当 RC 字段 1022 包括值 1 并且 RCQ 字段 1023 包括值 1 或 2 时,字 1 的字节 2-3 包括无符号整数秒中由程序使用的超时间隔。

[0095] 参考询问控制数据 1020 的字 2,字节 0 (标志字段 1028) 包括具有关于询问的信息的标志。每个标志位的含义可以给出如下:

[0096] 位 0 多路径模式。

[0097] 位 1 程序路径恢复。在路径恢复期间通过程序发出询问。

[0098] 位 2 关键的。设备是用于程序的关键设备。

[0099] 位 3-7 预留。

[0100] 如图 10 所示,字 2 的字节 1-3 { 字段 1029} 和字 3 的字节 0-3 { 字段 1030} 也可以预留给将来使用。字 4-5 (时间字段 1040) 可以包括关于发起询问操作的时间的信息。字 6-7 { 字段 1050} 可以包括标识发起询问操作的程序的程序标识符。该字段的内容可以是依赖于程序的。字 8-N ( 字段 1060) 可以包括依赖于程序的信息。

[0101] 用于询问命令的结束状态信息可以由控制单元在传输响应 IU 有效载荷中设置,如图 11 所示。在传输响应 IU 1100 中,字 0-7 可以包括状态有效载荷 1110,其包括结束状

态和状态标志。字 8-23 可以包括扩展状态有效载荷 1120,其可以按照对于询问的 TCW 中的每个传输状态块地址而被存储在 z 存储器中的状态块地址处。

[0102] 如图 11 所示,字 8 字节 0(ES 字段 1121) 包括包含扩展状态 (ES) 长度,其指示 ES 有效载荷的大小。字 8 字节 1(ES 标志字段 1122) 包括 ES 标志。标志位 5-7 指示类型代码。类型代码定义了 ES 有效载荷的状态区域的格式。对于询问的类型代码是 3。定义状态区域的三位编码可以给出如下:

[0103] 0 类型代码 0。状态区域中没有信息。

[0104] 1 类型代码 1。有效的结束 I/O 状态。

[0105] 2 类型代码 2。错误终止的状态

[0106] 3 类型代码 3。扩展状态是询问响应,并且格式如图 11 所示。

[0107] 4-7 预留

[0108] 字 8 的字节 2 和 3{ 状态字段 1123} 和字 ( 状态字段 1124) 包括指示 I/O 操作的状态的信息。对于 I/O 操作,在这些字段中的值对于询问可以为零。

[0109] 传输扩展状态有效载荷 1120 的字 11-23 可以是询问状态区域。字 11 的字节 0( 格式字段 1126) 包括定义了询问状态区域的布局的无符号整数值。如果该字段的值为零,那么询问状态的内容是无意义的。当格式字节设置为 (01h) 时,询问状态区域的下列定义适用。

[0110] 字 11 的字节 1( 标志字段 1127) 包括关于询问状态区域的信息。每个标志位的意义可以给出如下:

[0111] 位 0 控制宽度状态有效:当位 0 是 1 时,控制单元状态字段包括有意义的信息。当位 0 是零时,控制单元状态字段没有意义。

[0112] 位 1 设备状态有效:当位 1 是 1 时,设备状态字段包括有意义的信息。当位 1 是零时,设备状态字段没有意义。

[0113] 位 2 操作状态有效:当位 2 是 1 时,操作状态字段包括有意义的信息。当位 2 是零时,操作状态字段没有意义。

[0114] 位 3-7 预留

[0115] 字 11 的位 2( 控制状态字段 1128) 包括指示 I/O 设备的控制单元的当前状态的 8 位无符号整数。每个值的意义可以给出如下:

[0116] 0 忙碌:控制单元忙碌,并且依赖于设备的数据字段可以包括关于忙碌状态的附加信息。

[0117] 1 恢复:控制单元正在执行恢复处理,并且依赖于设备的数据字段可以包括关于恢复状态的附加信息。

[0118] 2 询问最大值:控制单元正执行其支持的最大数目的询问操作。

[0119] 3-127 预留。

[0120] 128-255 依赖于设备的含义。

[0121] 字 11 的字节 3( 设备状态字段 1129) 包括指示 I/O 设备的当前状态的 8 位无符号整数。该字节的意义可以给出如下:

[0122] 0 路径组标识:依赖于状态的信息字段包括标识设备已预留的路径组的信息。

[0123] 1 长时间忙碌(long busy):控制单元处于长时间忙碌状态。长时间忙碌的含义



是依赖于设备的,并且依赖于设备的字段可以包括关于长时间忙碌状态的附加信息。

[0124] 2 恢复:设备正执行恢复处理。

[0125] 3-127 预留。

[0126] 128-255 依赖于设备的含义。

[0127] 字 12 的字节 0(操作状态字段 1030)包括指示 I/O 操作是否出现在设备处以及当出现时操作的状态的 8 位无符号整数。该字节值的含义可以给出如下:

[0128] 0 没有 I/O 操作出现。

[0129] 1 I/O 操作出现并且正在执行

[0130] 2 I/O 操作出现,并且正在等待由另一配置发起的 I/O 操作的完成。

[0131] 3 I/O 操作出现,并且正在等待从同一设备范围发起的 I/O 操作的完成。

[0132] 4 I/O 操作出现并且正在等待执行依赖于设备的操作。

[0133] 5-127 预留。

[0134] 128-255 依赖于设备的含义。

[0135] 字段 1031 可以预留给将来使用。

[0136] 字 13-15(字段 1140)可以包括依赖于状态的信息。该字段的内容是依赖于设备的。该字段是否具有意义分别由 CS、DS 和 OS 字段 1128、1129 和 1130 指定。

[0137] 字 16(字段 1150)可以包括标识设备的完成级别的设备级标识符或标记。

[0138] 字 17-23(字段 1160)可以包括依赖于设备的信息。该字段是否有意义可以分别由 CS、DS 和 OS 字段 1128、1129 和 1130 指定。

[0139] 根据示例实施例,根据对于询问的在传输响应 IU 中返回的控制单元状态、设备状态和操作状态信息,OS 103 可以做出对关于花费比分配的时间更长的时间来完成的 I/O 操作采取什么动作的有根据的(informed)决定。

[0140] 图 12A 图示了操作系统用于根据本发明的方面来确定决定什么时候从控制单元请求 I/O 操作的状态的方法。在步骤 1265,当计时器弹出出现时,在步骤 1270,操作系统错过中断处理机(MIH)接收控制。例如,MIH 在发生计时器弹出后,每秒接收一次控制。在步骤 1271,MIH 从头到尾扫描已由在操作系统中运行的所有应用、中间件和子系统所发出的每个活动的 I/O 操作。在步骤 1272 确定是否 I/O 操作即将过期,例如是否在 I/O 操作的时间限的一秒内。如果是,在步骤 1273 确定是否 I/O 操作是用于传输模式(FCX)。如果 I/O 操作是用于传输模式,那么在步骤 1250,建构询问 DCW,并且更新活动的 TCW 以指向询问命令。否则,处理从步骤 1274 进行到步骤 1278,在步骤 1278,执行常规的“传统(heritage)”MIH 处理。从步骤 1250,MIH 处理在步骤 1271 继续,或直到发生计时器弹出时、例如在步骤 1265 后一秒,在步骤 1290 终止。如果在步骤 1272,确定 I/O 操作还没有要过期,那么在步骤 1276 进行检查以查看 I/O 操作时间是否已经超过其分配的时间。如果 I/O 操作还没有超过其分配的时间,那么在步骤 1278 发生惯例的传统(常规的)MIH 处理。如果 I/O 操作时间确实已经过期,那么在步骤 1276 进行检查以查看是否已经发出询问命令。如果还没有发出询问命令,则这是命令不是传输模式命令(FCX)的指示,并在步骤 1278 执行传统 MIH 恢复处理。然而,如果,在步骤 1276 确定询问命令已经发出并且已经成功完成,那么在步骤 1279 检查询问结果以确定是否该设备被预定用于另一系统。如果询问信息指示设备被预定用于另一系统 1140,那么不存在错误,并且未发生错过的中断。因此,在步骤 1290 处理终止。如果设

备没有被预定用于某个其他系统,则在步骤 1277 将由询问返回的信息放置在待写入到用于诊断 (diagnostic) 目的的系统 LOGREC 数据表的记录中。然后在步骤 1278 处理以传统 MIH 处理而继续。

[0141] 图 12B 图示了根据本发明的方面的用于确定 I/O 操作的状态的方法。在步骤 1205 将用于发起 I/O 操作的请求从 OS 103 发送到信道子系统 108,并在步骤 1207 发送到控制单元 110。在步骤 1210,在控制单元 110 处接收和处理该请求。如果 I/O 操作接近其分配的执行时间的终点,如前面段落所述,在步骤 1250,由操作系统发起询问操作,如图 12A 和图 12B 两者中所示。

[0142] 询问开始于步骤 1250,在其处 OS 103 在系统存储器 102 中建立询问控制块,并且向信道子系统 108 发送取消子信道。信道子系统 108 确定询问是否将被发送到控制单元。如果没有满足进行询问的上述条件,则传统 MIH 处理结束 I/O 操作,并且将错误模拟回到 I/O 的发起者。如果满足询问条件,则在步骤 1225,从信道子系统 108 向控制单元 110 发送询问请求。在步骤 1230,控制单元 110 接收询问请求。在步骤 1235,控制单元 110 向信道子系统 108 发送询问响应,指示 I/O 操作、控制单元 110 以及执行询问 I/O 操作的 I/O 设备 112 的状态。在步骤 1240,在信道子系统 108 处接收询问响应,该信道子系统 108 产生中断给 OS。在步骤 1245,OS 接收该询问响应,建立 LOGREC 项以将状态信息记录在控制单元处,然后以如上所述的传统 MIH 处理而继续。

[0143] 应该理解,不是必须执行图 12 所示的所有步骤以确定 I/O 操作的状态。此外,图 12A 和图 12B 中的步骤的顺序是可以如何执行处理的例子。同样地,也可以执行图 12A 和图 12B 中未示出的另外的步骤。

[0144] 如上所述,可以按计算机实现的处理和用于实践那些处理的装置的形式来具体化各实施例。在示例实施例中,本发明以由一个或多个网络元件执行的计算机程序代码来具体化。各实施例包括在计算机可用介质 1302 上的如图 13 所示的计算机程序产品 1300,该计算机可用介质 1302 带有包含体现在实体介质上的指令的计算机程序代码逻辑 1304,作为一件产品。计算机可用介质 1302 的示例产品可以包括软盘、CD-ROM、硬盘驱动器、通用串行总线 (USB) 快闪驱动器或任何其他计算机可读存储介质,其中,当计算机程序代码逻辑 1304 被载入计算机并由其执行时,计算机变为用于实践本发明的装置。实施例包括计算机程序代码逻辑 1304,例如无论是被存储在存储介质中、被载入计算机中和 / 或由计算机执行、或在一些传输介质上、比如在电线或电缆上传输、通过光纤传输或经由电磁辐射传输,其中,当计算机程序代码逻辑 1304 被载入计算机中并由其执行时,该计算机变成实践本发明的装置。当在通用微处理器上实现时,计算机程序代码逻辑 1304 段 (segment) 配置微处理器以建立具体逻辑电路。

[0145] 虽然已经参考示例实施例描述了本发明,但本领域的技术人员将理解,在不背离本发明的范围内,可以做出各种改变,并且可以由等效物替换本发明的元件。此外,在不背离本发明的实质范围内,可以做出许多修改以使具体情况或材料适应于本发明的教导。因此,意要本发明不限于作为预期实现本发明的最佳模式的公开的具体实施例,而是本发明将包括落入所附权利要求的范围内的所有实施例。此外,术语第一、第二等的使用不是表示任何顺序或重要性,而是使用术语第一、第二等将一个元素与另一元素区分开。此外,术语“一个”等的使用不是表示对量的限制,而是表示存在所指项的至少一个。

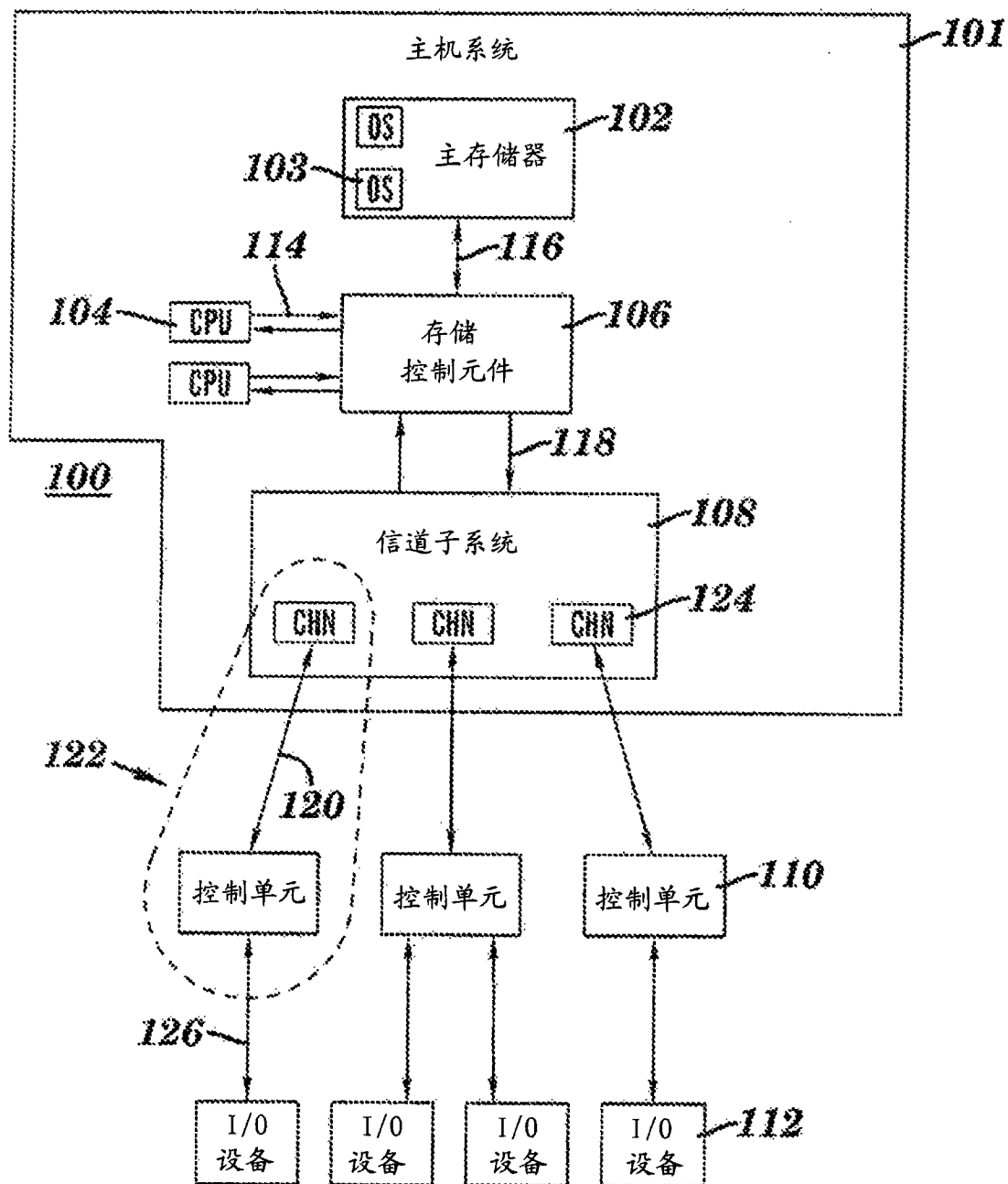


图 1

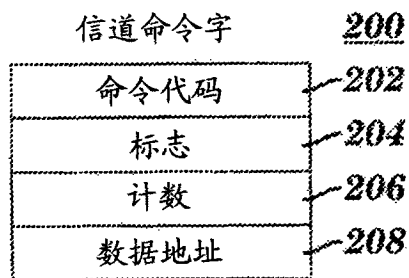


图 2A 现有技术

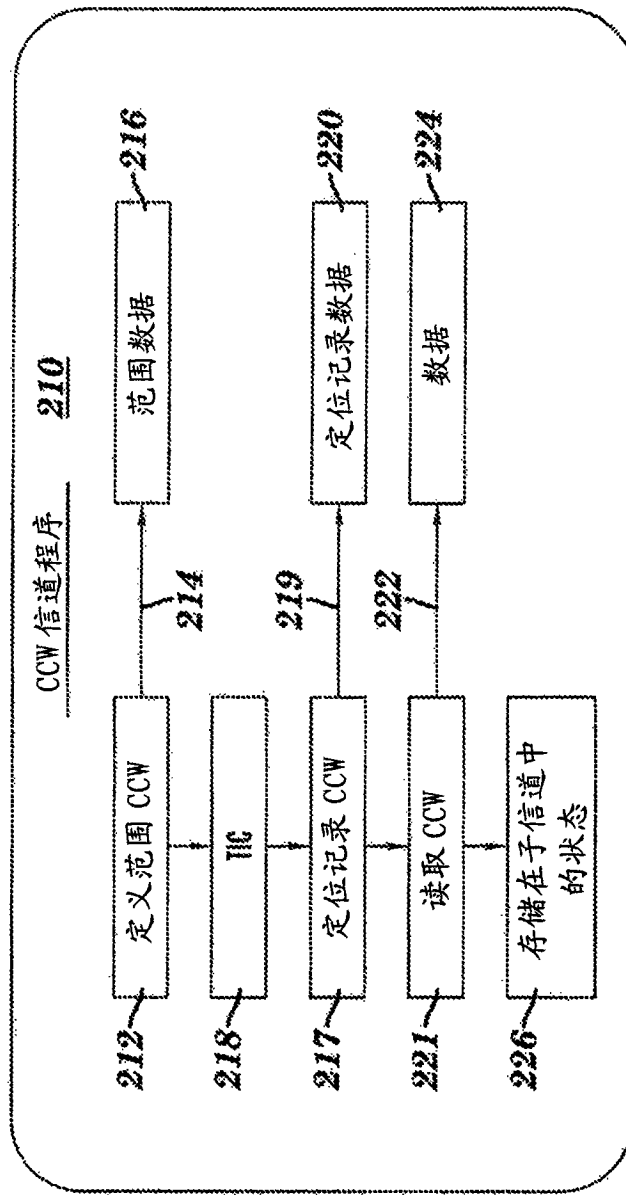


图 2B 现有技术

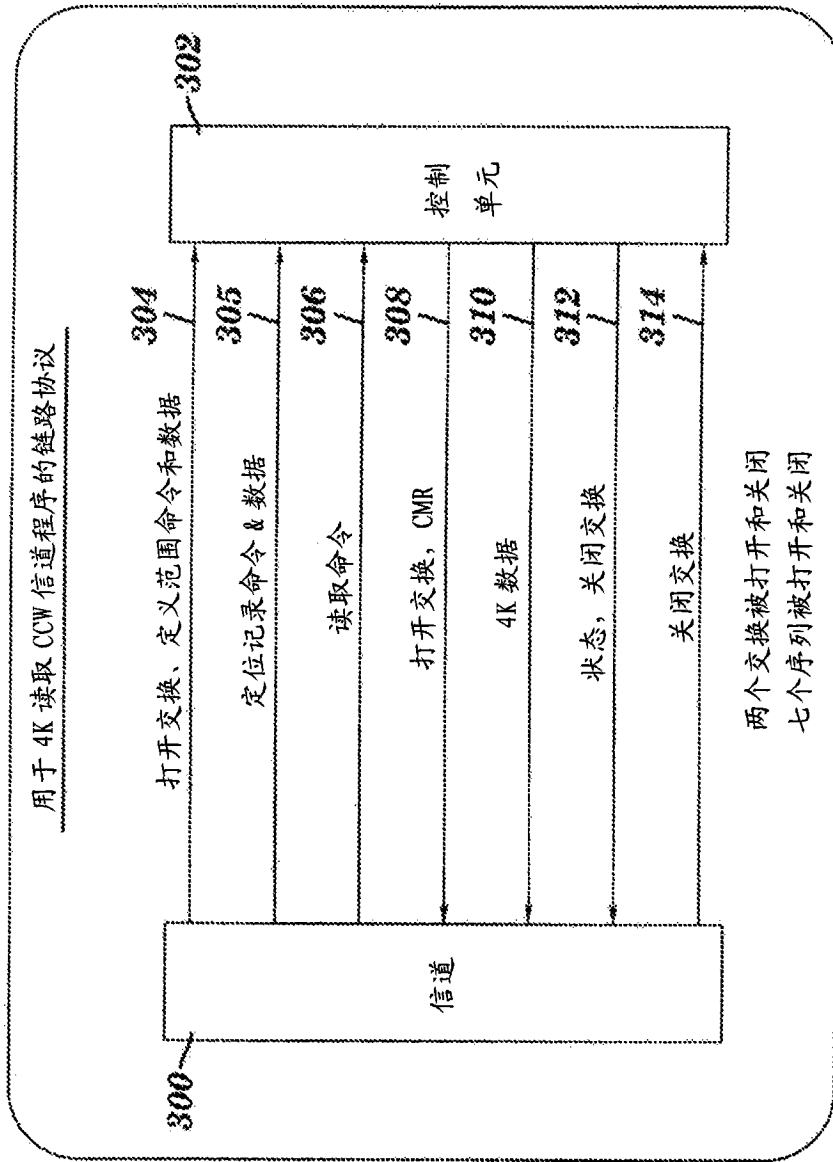


图 3

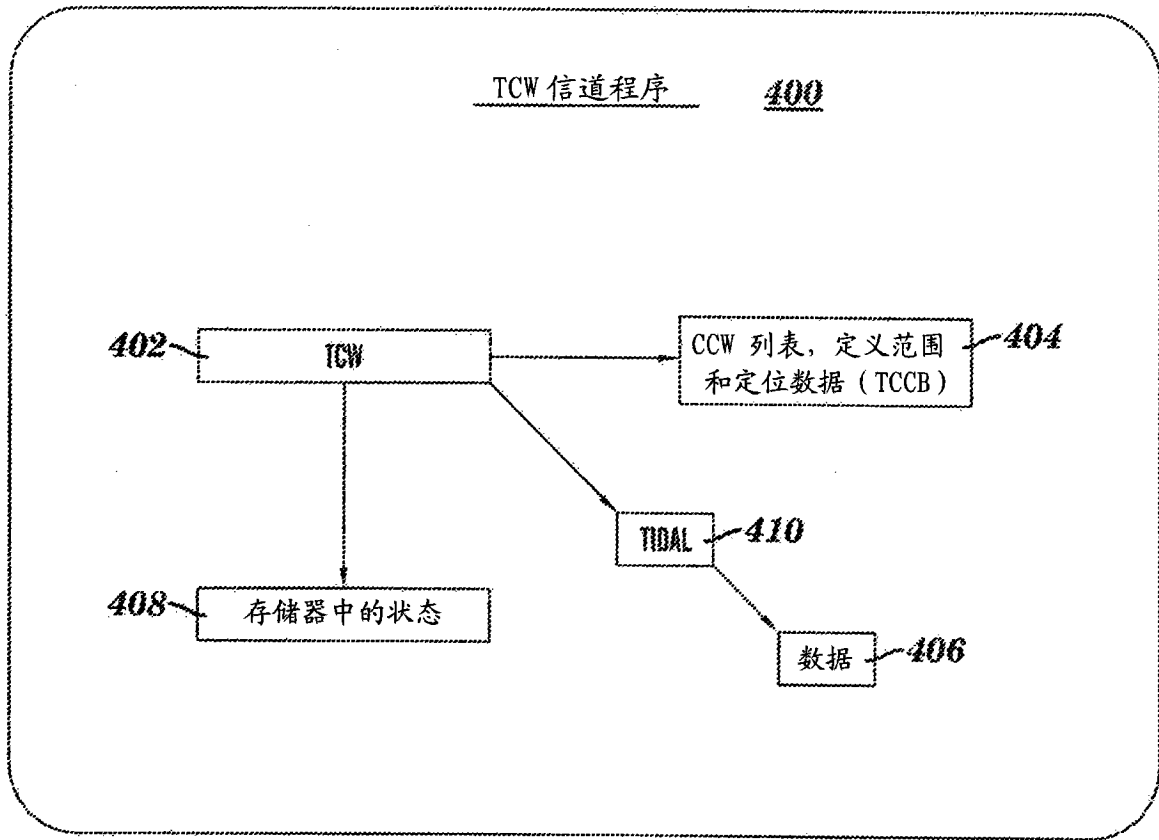


图 4

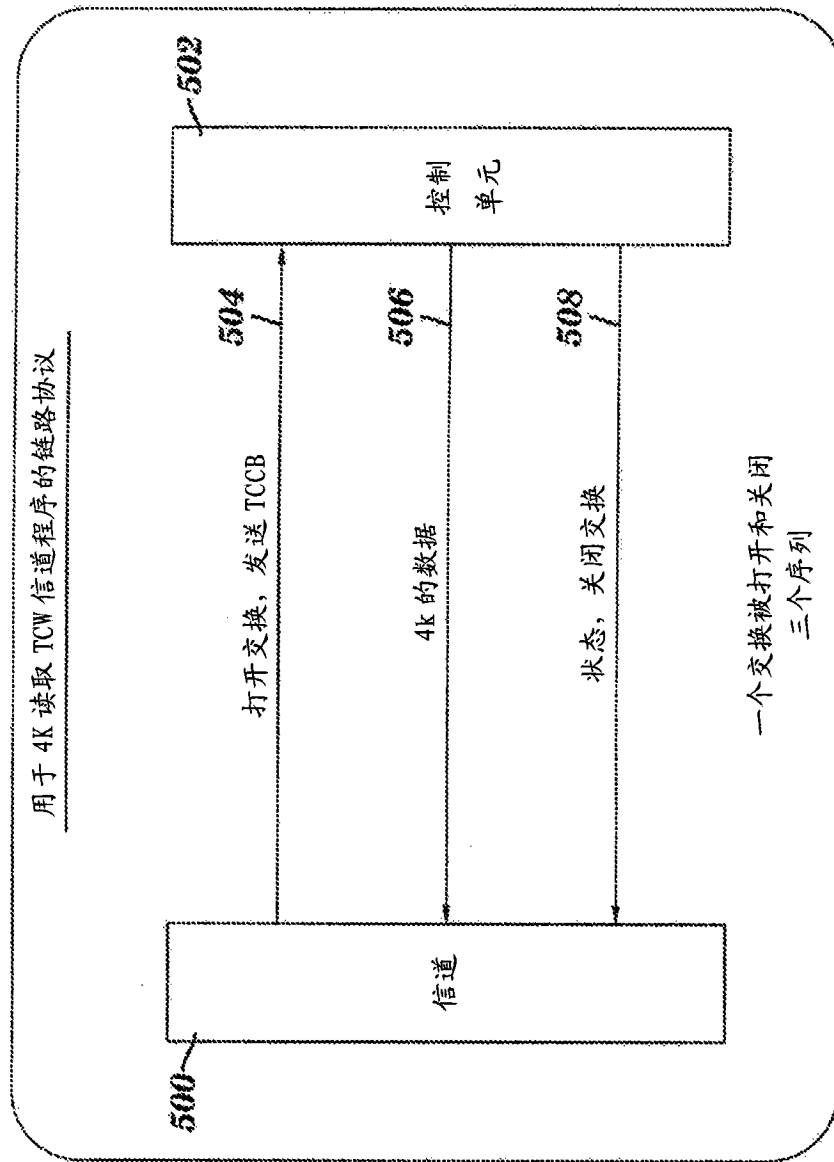
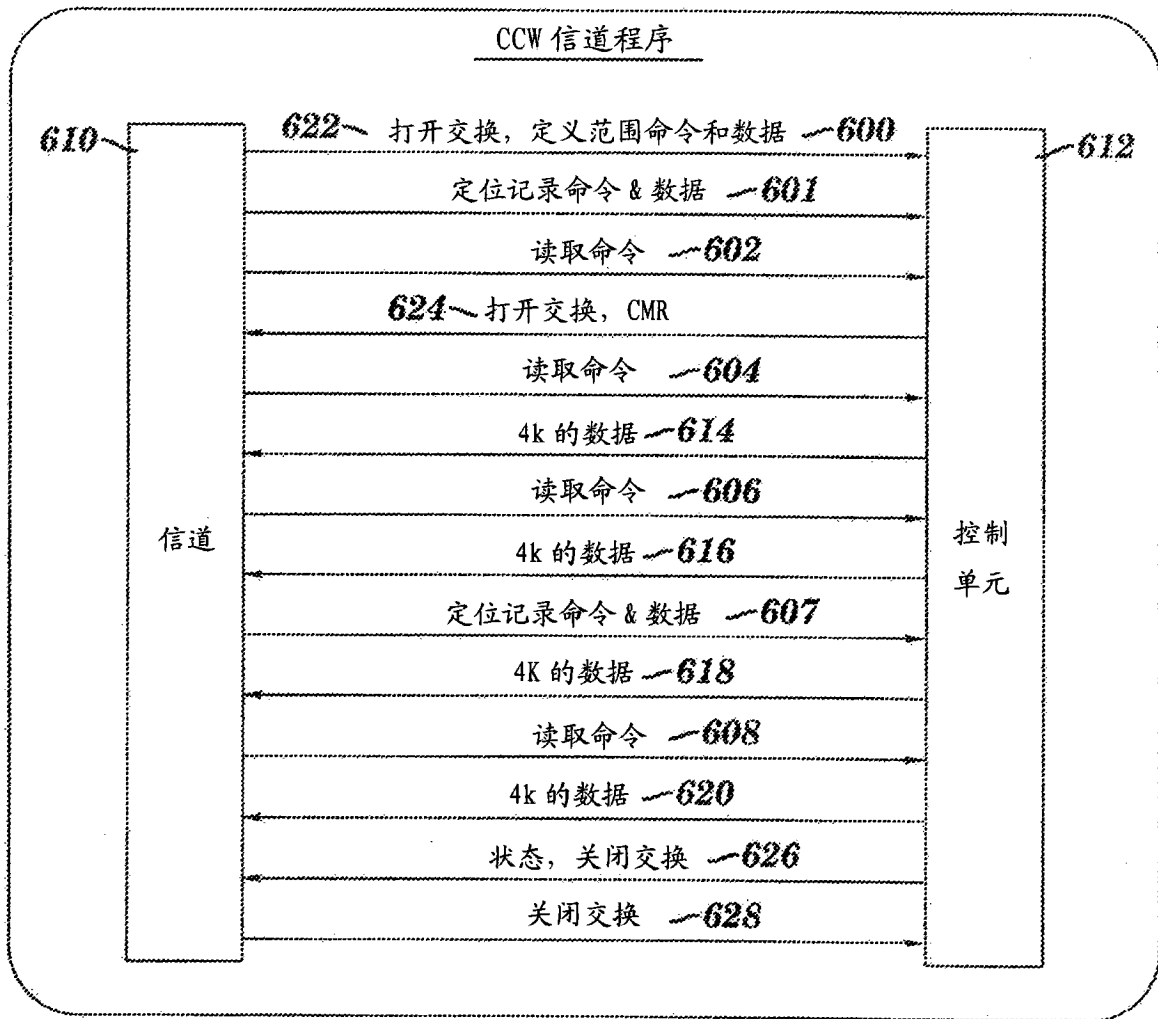


图 5



现有技术

图 6



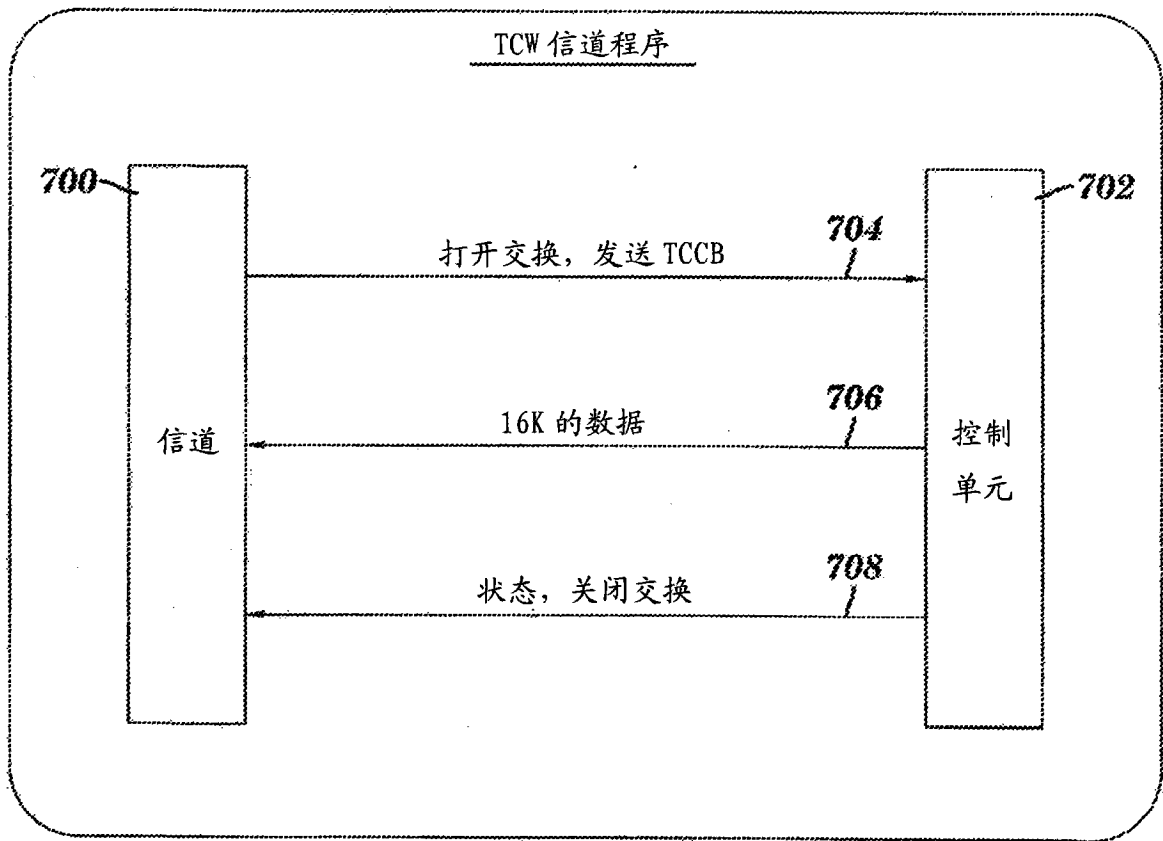


图 7

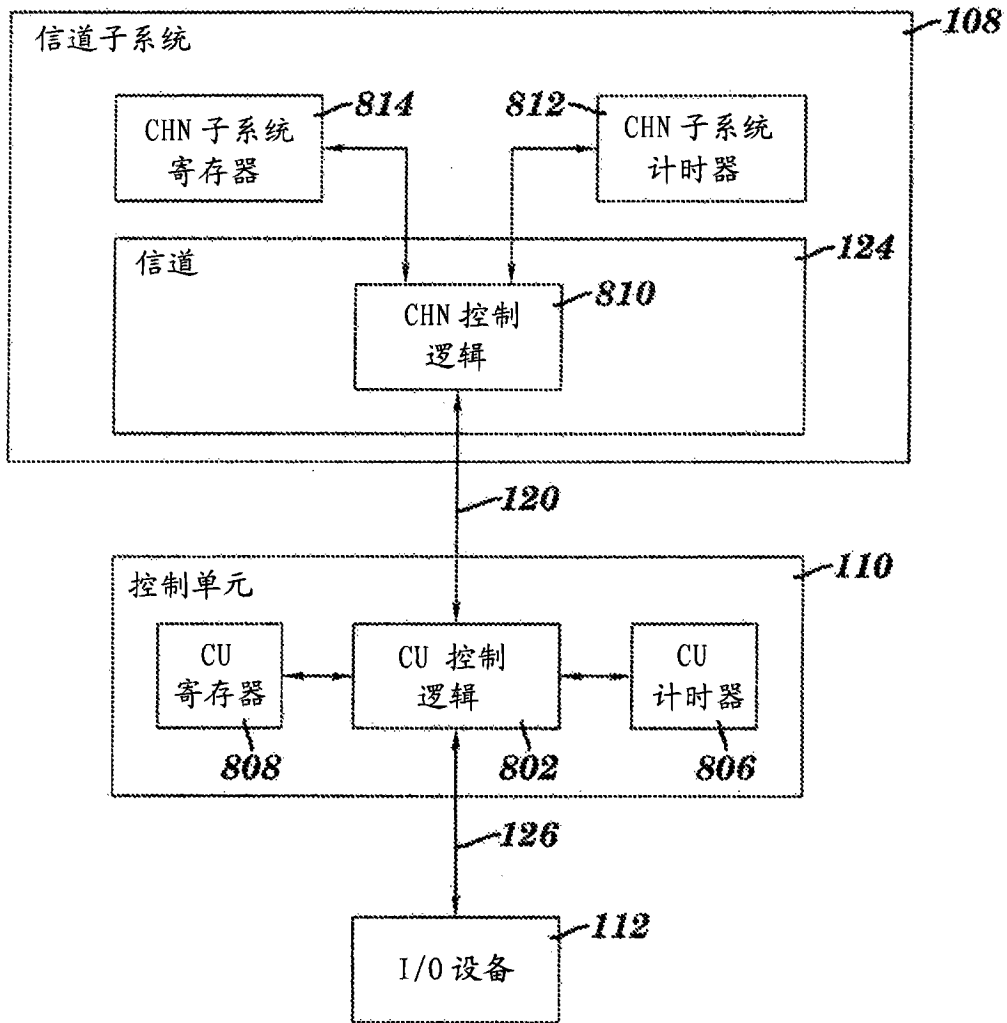


图 8

900 ↙

904	0	1	2	3	4	5	6	7
DW	格式		标志			长度	PW	
1	直接或间接数据地址							
2	914							
3	传输状态块的地址 (其中存储了扩展状态)							
4	916							
5	传输命令 IU 的地址							
6	920							
7	922							
8	924 数据字节计数							
9	926							
10	928							
11	930							
12	932 询问-TCW 地址							
13	934							

图 9

1000 ↙

名称	字	0	1	2	3
DCW	0	<u>1011</u> 40	<u>1012</u> 00	<u>1013</u> 00	<u>1014</u> 控制数据计数 ≤232
	1	<u>1015</u> 00	<u>1016</u> 00	<u>1017</u> 00	<u>1018</u> 00
	0	<u>1021</u> FW	<u>1022</u> RC	<u>1023</u> RCR	<u>1024</u> LPM
	1	<u>1025</u> FWM	<u>1026</u> PMM	超时	<u>1027</u>
	2	<u>1028</u> 标志	预留	预留	<u>1029</u>
询问 控制数据	3	预留	预留	预留	<u>1030</u>
	4	预留	时间	预留	<u>1040</u>
	5	预留	预留	预留	预留
	6	预留	程序标识符	预留	<u>1050</u>
	7	预留	预留	预留	预留
	8	预留	依赖于程序的数据	预留	预留
	10	预留	依赖于程序的数据	预留	预留
	11	预留	依赖于程序的数据	预留	<u>1060</u>

图 10

1100

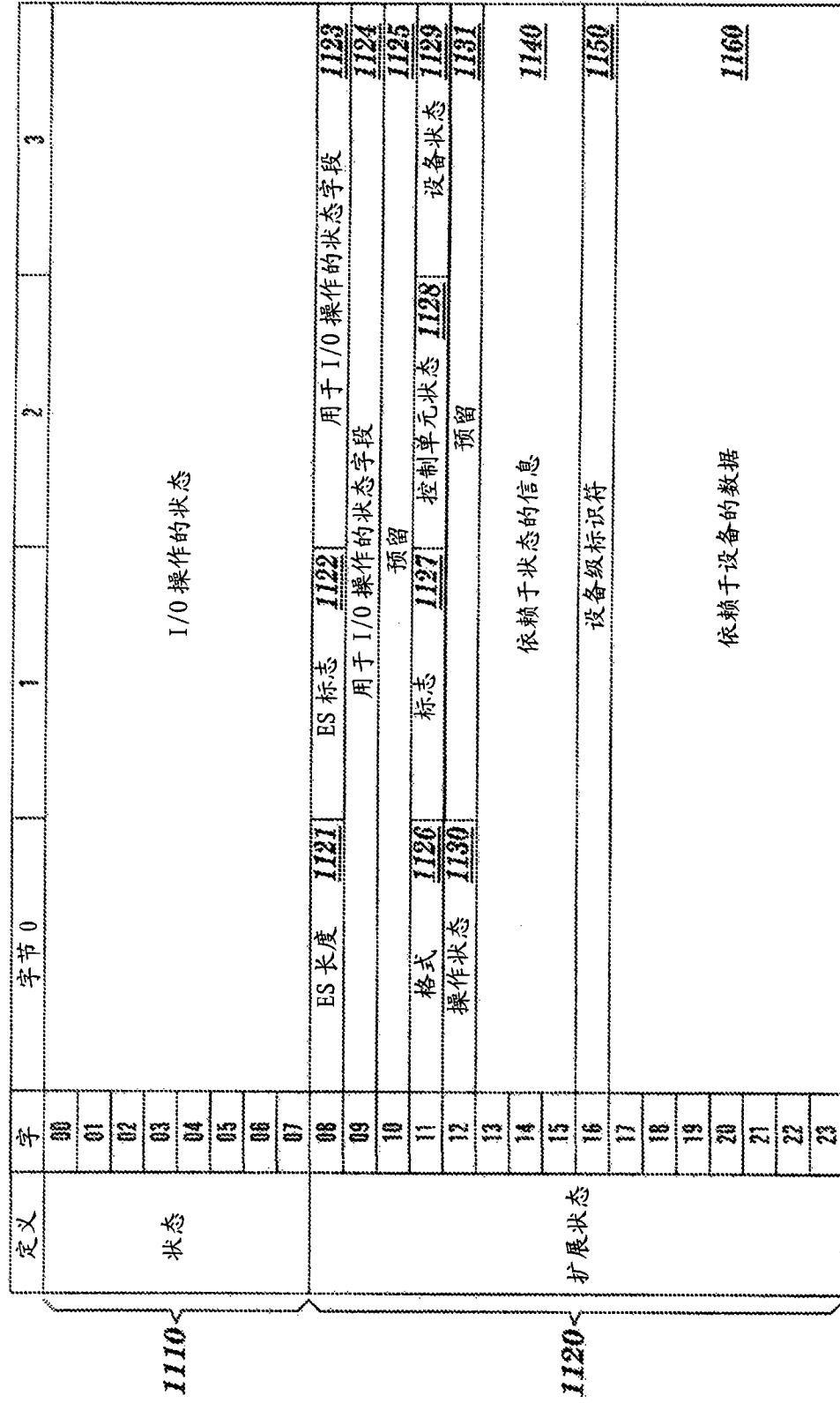


图 11

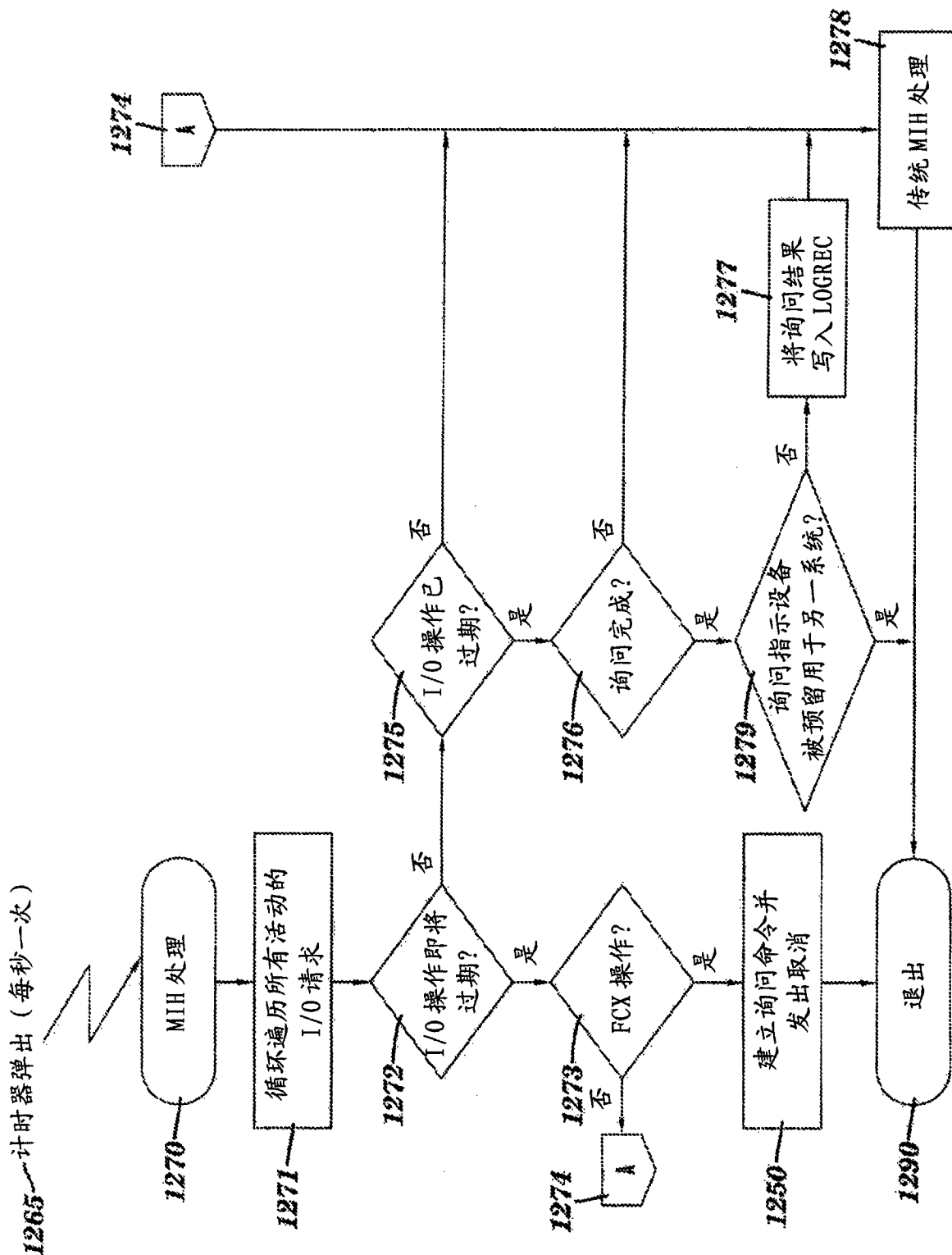


图 12A

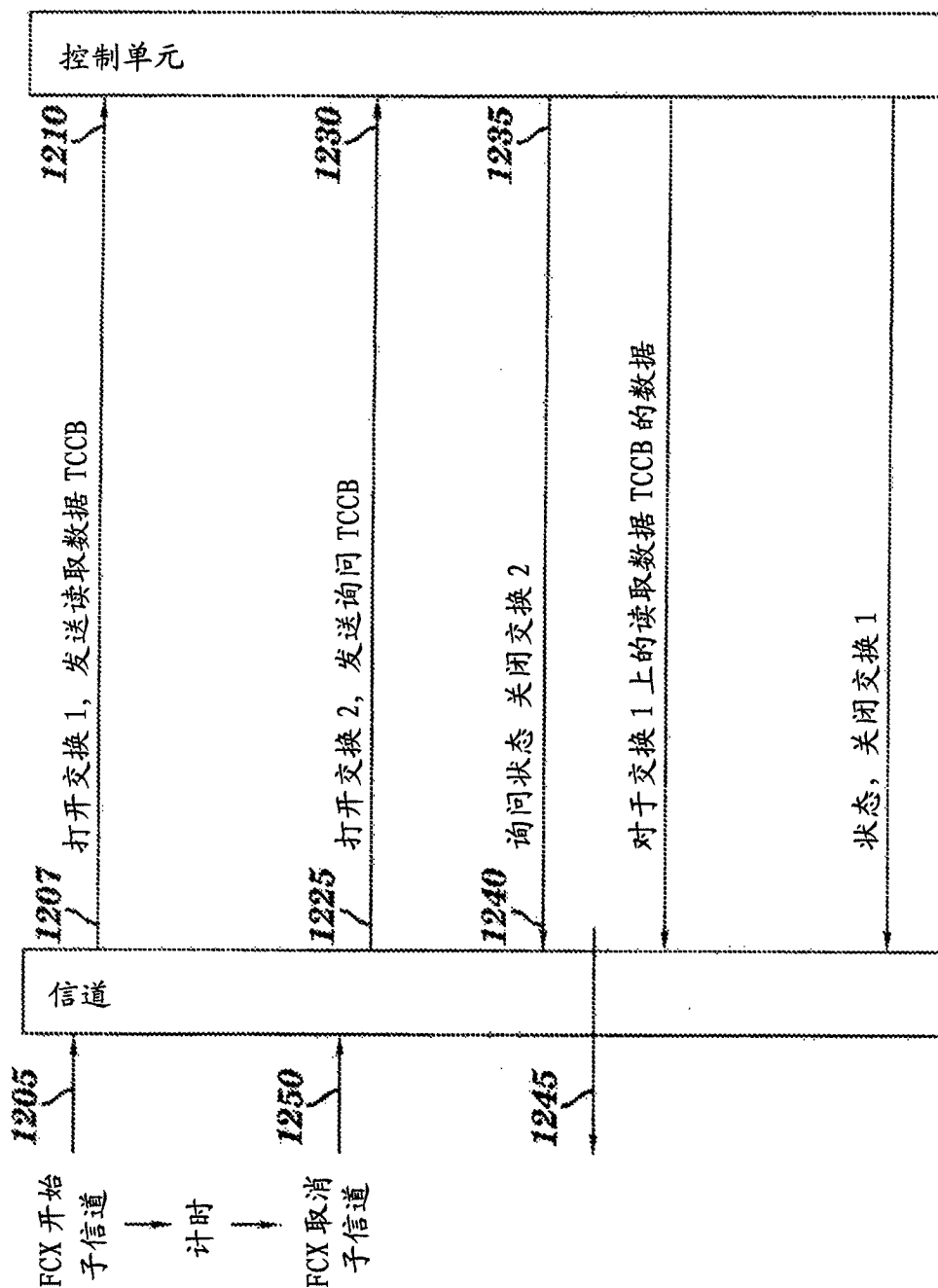


图 12B

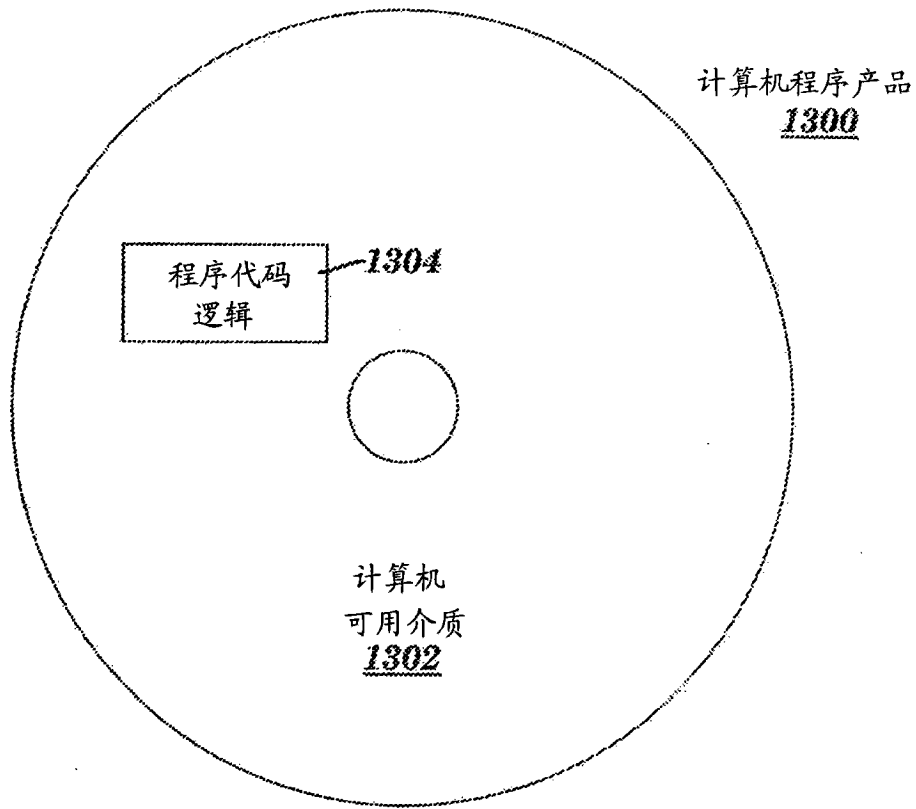


图 13