

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 13/42 (2006.01)

H04L 12/56 (2006.01)

H04L 29/06 (2006.01)



[12] 发明专利说明书

专利号 ZL 200480014227.6

[45] 授权公告日 2008 年 9 月 3 日

[11] 授权公告号 CN 100416534C

[22] 申请日 2004.5.12

[21] 申请号 200480014227.6

[30] 优先权

[32] 2003.5.23 [33] US [31] 10/444,164

[86] 国际申请 PCT/US2004/014851 2004.5.12

[87] 国际公布 WO2004/107713 英 2004.12.9

[85] 进入国家阶段日期 2005.11.23

[73] 专利权人 英特尔公司

地址 美国加利福尼亚州

[72] 发明人 P·蒙圭亚

[56] 参考文献

US5859853A 1999.1.12

US2001/0055317A1 2001.12.27

第三代 I/O 总线结构(3GIO)的研究. 戴浩, 徐美华, 冉峰, 陈章进. 上海大学学报(自然科学版), 第9卷第2期. 2003

总线技术 PCI Express. 个人电脑. 2002

审查员 谢阿涟

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 王岳 王勇

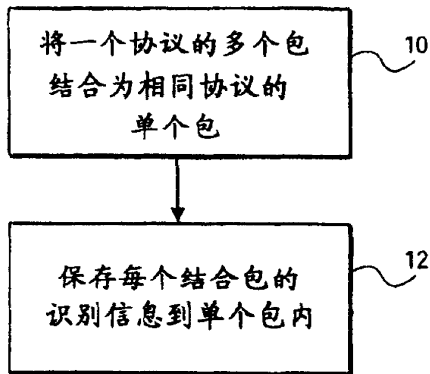
权利要求书3页 说明书6页 附图5页

[54] 发明名称

基于 PCI Express 的包结合方法、设备和系统

[57] 摘要

本发明公开了一种用于结合包的方法和设备。在不同的实施例中，遵守特定协议的多个包可以被结合为遵守相同协议的单个包。所述的单个包还保存每一个结合包的识别信息到所述单个包内。所述识别信息被用来划分和/或发送所述结合包到它们各自的目的地。



1. 一种结合包的方法, 包括:

将遵守一个协议的多个包结合为遵守相同协议的单个包;

保存每一个结合包的识别信息到所述单个包内; 并且

增加一个指示到所述单个包来指示在所述单个包内包含结合包,

其中, (i) 在 PCI Express 链接的再同步之前和(ii)在 PCI Express 链接的再同步期间的至少其中一个时间期间执行所述结合。

2. 如权利要求 1 所述结合包的方法, 其中所述增加步骤包括: 把一个报头放到所述单个包中。

3. 如权利要求 1 所述结合包的方法, 还包括:

发送所述单个包到一个能够进行结合和划分包的设备。

4. 如权利要求 3 所述结合包的方法, 还包括:

所述设备在接收所述单个包之后发送所述单个包到另一个设备。

5. 如权利要求 3 所述结合包的方法, 还包括:

接收所述单个包; 并且

提取所述结合包。

6. 如权利要求 5 所述结合包的方法, 还包括:

发送每一个提取包到合适的目的地。

7. 如权利要求 1 所述结合包的方法, 还包括:

如果未经过预定的固定时间量或者未达到最小结合包尺寸门限值, 则延迟 PCI Express 链接的再同步。

8. 一种结合包的方法, 包括:

接收遵守一个协议的单个包, 所述单个包包含多个结合包和用于指示在所述单个包内包含所述结合包的报头, 每一个结合包遵守相同协议并且具有与该包相关联的识别信息,

其中, 所述单个包(i)在 PCI Express 链接的再同步之前和(ii)在 PCI Express 链接的再同步期间的至少其中一个时间期间被结合。

9. 如权利要求 8 所述结合包的方法, 还包括:

在所述的接收之后发送所述单个包到与总线相耦合的设备。

10. 如权利要求 8 所述结合包的方法, 还包括:

从所述单个包中提取所述结合包。

11. 如权利要求 10 所述结合包的方法, 还包括:

发送所述提取包到一个合适的目的地。

12. 一种结合包的设备，包括：

第一 PCI Express 设备，其能够

将遵守 PCI Express 的多个包结合为遵守 PCI Express 的单个包，所述单个包包含所述单个包内的每个结合包的识别信息和用于指示在所述单个包内包含所述结合包的报头，和

通过 PCI Express 链接发送所述单个包；和

用于接收所述单个包的第二 PCI Express 设备，

其中，所述多个包 (i) 在 PCI Express 链接的再同步之前和 (ii) 在 PCI Express 链接的再同步期间的至少其中一个时间期间被结合。

13. 如权利要求 12 所述结合包的设备，其中第一 PCI Express 设备包括 PCI Express 根设备、PCI Express 转换器设备和 PCI Express 外围设备中的其中一个，以及第二 PCI Express 设备包括 PCI Express 根设备、PCI Express 转换器设备和 PCI Express 外围设备中的其中一个。

14. 如权利要求 12 所述结合包的设备，其中第二 PCI Express 设备能够执行以下至少其中之一：(i) 发送所述单个包到另一个 PCI Express 设备，(ii) 从所述单个包中提取所述结合包，以及 (iii) 发送所述提取包到合适的目的地。

15. 一种结合包的系统，包括：

图形控制器；

与所述图形控制器耦合的第一 PCI Express 设备，所述第一 PCI Express 设备能够

将遵守 PCI Express 的多个包结合为遵守 PCI Express 的单个包，所述单个包包含所述单个包内的每个结合包的识别信息和用于指示在所述单个包内包含所述结合包的报头，和

通过 PCI Express 链接发送所述单个包；和

用于接收所述单个包的第二 PCI Express 设备，

其中，所述多个包 (i) 在 PCI Express 链接的再同步之前和 (ii) 在 PCI Express 链接的再同步期间的至少其中一个时间期间被结合。

16. 如权利要求 15 所述结合包的系统，其中第一 PCI Express 设备包括 PCI Express 根设备、PCI Express 转换器设备和 PCI Express

外围设备中的其中一个，以及第二 PCI Express 设备包括 PCI Express 根设备、PCI Express 转换器设备和 PCI Express 外围设备中的其中一个。

17. 如权利要求 15 所述结合包的系统，其中第二 PCI Express 设备能够执行以下至少其中之一：(i) 发送所述单个包到另一个 PCI Express 设备，(ii) 从所述单个包中提取所述结合包，以及(iii) 发送所述提取包到合适的目的地。

基于 PCI Express 的包结合方法、设备和系统

技术领域

在此公开的实施例通常涉及包传输，并且更具体而言涉及在传输之前的结合包的方法、设备和系统。

背景技术

在一个计算机系统中或多个计算机系统之间（例如，在一个网络上）能以多种不同的形式发送数据。在一些系统中，数据被集合在一起，并且作为一个包发送。PCI Express（外围设备互连专业组的外围设备互连 Express 1.0 规范，公布于 2002 年 7 月 22 日）是一个用于分包和传输数据的协议的例子。不管所用的协议，包的传输需要一定程度的开销和最小的包尺寸。

例如，一些包需要一个报头（例如，开销），所述报头包含关于包的内容和/或目的地的信息。如果在发送包之前需要一个最小的包尺寸，一些系统将在不满足最小尺寸需求的包中“填充”额外的比特。这些填充的比特包含无用的信息并且降低了系统的效率。

下面是 PCI Express 事务处理的一些例子，所述 PCI Express 事务处理典型地需要比特填充来满足最小包尺寸的需求。第一，一个关于 IO 映射（例如，输入/输出映射）设备的简单序列需要一个 8 比特 IO 写事务处理到索引寄存器，随后要求 8 比特 IO 循环到相应的数据寄存器。这转换成两个最小的长度和效率的全 PCI Express 事务处理。

第二，读取 FIFO（例如，先进先出）或门铃寄存器（doorbell register）由每次分离的 PCI Express 事务处理重复地读取相同的地址组成。第三，一些专门的存储器事务处理（例如，编程一个闪存设备）需要字节或字以命令序列次序写入。这些写入中的每一个都需要一个分离的 PCI Express 事务处理。

因此，这三个过程的每一过程都需要多个最小长度的 PCI Express 事务处理，它可能由于比特填充而涉及无用比特的低效率传输。此外，每一个附加的 PCI Express 事务处理都需要系统资源以便形成包、传输包和确认包传输。

在数据通信主要被发送到低带宽外围设备和 IO 设备的情况中，包开

销和最小包尺寸的缺陷是非常显著的，这是因为这些设备的利用率比较低。低的利用率转变成耗散功率，尤其在低功率系统中。对于转换来说，低利用率导致低效率的性能和判优的低效率。

发明内容

为克服以上缺陷，本发明提供一种结合包的方法，包括：将遵守一个协议的多个包结合为遵守相同协议的单个包；保存每一个结合包的识别信息到所述单个包内；并且增加一个指示到所述单个包来指示在所述单个包内包含结合包，其中，(i)在PCI Express 链接的再同步之前和(ii)在PCI Express 链接的再同步期间的至少其中一个时间期间执行所述结合。

根据本发明的另一种结合包的方法，包括：接收遵守一个协议的单个包，所述单个包包含多个结合包和用于指示在所述单个包内包含所述结合包的报头，每一个结合包遵守相同协议并且具有与该包相关联的识别信息，其中，所述单个包(i)在PCI Express 链接的再同步之前和(ii)在PCI Express 链接的再同步期间的至少其中一个时间期间被结合。

根据本发明的一种结合包的设备，包括：第一PCI Express 设备，其能够将遵守PCI Express 的多个包结合为遵守PCI Express 的单个包，所述单个包包含所述单个包内的每个结合包的识别信息和用于指示在所述单个包内包含所述结合包的报头，和通过PCI Express 链接发送所述单个包；和用于接收所述单个包的第二PCI Express 设备，其中，所述多个包(i)在PCI Express 链接的再同步之前和(ii)在PCI Express 链接的再同步期间的至少其中一个时间期间被结合。

根据本发明的一种结合包的系统，包括：图形控制器；与所述图形控制器耦合的第一PCI Express 设备，所述第一PCI Express 设备能够将遵守PCI Express 的多个包结合为遵守PCI Express 的单个包，所述单个包包含所述单个包内的每个结合包的识别信息和用于指示在所述单个包内包含所述结合包的报头，和通过PCI Express 链接发送所述单个包；和用于接收所述单个包的第二PCI Express 设备，其中，所述多个包(i)在PCI Express 链接的再同步之前和(ii)在PCI Express 链接的再同步期间的至少其中一个时间期间被结合。

附图说明

通过实例来说明不同实施例，并且不是通过附图中数字的限制来说

明不同实施例，在附图中相同参考数字表示类似元件。应该指出，在这个公开中涉及的“一，”、“一个，”、“这个，”、“其它，”、“可替换的”，或“不同”的实施例未必是相同的实施例并且这样的参考意味着至少一个。

图 1 是方法的一个实施例的流程图，所述方法用于结合多个包为单个包；

图 2 示出了包含多个结合包的包的一个实施例；

图 3 是方法的一个实施例的流程图，所述方法用于延迟链接的再同步直到完成将多个包结合为单个包为止；

图 4 是与第二 PCI Express 设备耦合的第一 PCI Express 设备的图，根据在此公开的不同实施例，每个设备能够结合和划分包；

图 5 是一个包括能够结合和划分包的多个 PCI Express 设备的系统；

图 6 是基于图 5 系统中设备的结合功能实现的不同数据流的图。

具体实施方式

在下面的描述中，出于解释的目的，阐明了许多特定细节以便对不同实施例提供全面理解。没有某些特定的细节的情况下可以实现这些实施例对本领域的技术人员来说是显而易见的。在其它情况下，忽略或简化某些结构和设备以便避免使得不同实施例的细节模糊不清。

出于说明的目的，下面的描述和附图提供了一些实例。然而，不是在限制的意义上解释这些实例，因为它们没有试图提供所有可能实现的详尽列表。

通过使用机器读取存储在机器可读介质上的指令以及依照这些指令执行功能，可以执行在此公开的不同方法。一个机器可读介质包括任何机制，所述机制通过一个机器（例如，计算机）以一种可读的形式提供（例如，存储和/或传输）信息。例如，一个机器可读介质包括只读存储器 (ROM)；随机存取存储器 (RAM)；磁盘存储介质；光学存储介质；闪存设

备；和传播信号的电，光，声或其它形式（例如，载波、红外信号、和数字信号）。

现在参考图 1，示出了用于结合包的一个方法。在块 10，遵守一个特定协议（例如，PCI Express）的多个包被结合为单个包。所述单个包遵守的协议与多个结合包遵守的协议相同。

在块 12，将每一个结合包的识别信息保存在所述单个包内。在不同的实施例中，识别信息的保存包括将每一个结合包的报头保存在所述单个包内。在一个实施例中，识别信息的保存与包的结合同时发生。

图 2 示出了包含多个结合包的单个包的一个实例。特别地，单个包 14 包括报头 16、结合包 18 和循环冗余校验（“CRC”）20。在不同的实施例中，报头 16 包含一个指示，即单个包 14 包含结合包 18。在其它实施例中，可以将单个包 14 包含结合包 18 的指示设置在单个包 14 内的其它位置（例如，在其它字段）。

每一个结合包 18 包括报头 22、数据字段 24 和 CRC26。结合包 18 彼此相同或彼此有区分。例如，PCI Express 利用不同类型的包，这些包是依赖性的事务处理并且可以以任意的方式结合，这是因为不同实施例对方式没有限制，其中不同的包类型可以被结合在单个包内。

虽然在图 1 中没有示出，但是所述方法的不同实施例可以包括附加的特征。例如，一旦创建了单个包，就能发送所述单个包到能够结合和划分包的一个设备。在一个实施例中，通过 PCI Express 总线能够发送单个包。接收设备可以是例如 PCI Express 根设备、PCI Express 转换器或 PCI Express 外围设备。PCI Express 外围设备的实例其中包括 I/O 设备，例如千兆以太网和加速图形端口。

如果接收设备是一个转换器，能发生几个不同的事件。例如，一旦转换器接收到单个包，转换器可以简单地路由或发送单个包到其它设备。可选择地，转换器能划分单个包（例如，提取结合包）和发送提取包到合适的目的地。

在不同实施例中，转换器利用单个包的报头来确定是否重新发送单个包或者是否开始提取结合包。如果提取这个包，转换器能使用提取包的报头（例如，先前保存的报头）来确定应该发送每一个提取包到何处。

然而，如果接收设备是一个外围设备，所述外围设备能够简单地接收单个包并且提取结合包用于使用/传输。

通过调整结合与通信链接的再同步还能扩展图 1 示出的方法。例如，在 PCI Express 中，基于一些门限值标准（例如，链接的不活动的预定周期）可以将设备之间的链接设置为断电。当用信号通知链接被加电以便接收/传输数据的时候，链接必须再同步。必要地，在应该对链接加电的指示和再同步的完成之间有一个等待时间周期。

不同实施例通过在链接的再同步之前和/或期间结合包来利用这个固有的等待时间周期。在其它实施例中，故意地延迟再同步。由故意延迟造成的附加等待时间有利地允许附加的时间用于包的结合。

在这点上，图 3 是用于故意延迟再同步的方法的一个实施例的流程图。在块 28，将有用的包结合为单个包。在决定块 30，它确定预定时间的固定量是否流逝和/或是否已经满足最小的结合包尺寸门限值。如果不满足所需的条件或状态，在块 28 继续进行包的结合。然而，如果已经满足必需的条件或状态，在块 32 进行链接的再同步。一旦再同步所述链接，就在块 34 发送单个包。

图 4 示出了用来执行上面描述的一些不同方法的两个设备的简图。特别地，第一 PCI Express 设备 36 与第二 PCI Express 设备 38 耦合。正如上面所提及的那样，每一个设备可以是以下的其中一个：PCI Express 根设备、PCI Express 转换器设备和 PCI Express 外围设备。

在一个实施例中，设备 36（和可以是设备 38）能够将遵守 PCI Express 的多个包结合为遵守 PCI Express 的单个包并且包含每一个结合包的识别信息。设备 36 也能够通过 PCI Express 链接发送单个包到设备 38。

设备 38 能够发送从设备 36 接收的单个包到另一个 PCI Express 设备；从单个包中提取结合包；并且发送该提取包到合适的目的地。在不同实施例中，通过单个包的报头和/或在单个包内的每一个结合包的报头来确定合适的目的地。

设备 36 和设备 38 各自包括端口 35，所述端口 35 是一组相关的发射器和接收器。在不同实施例中，端口 35 应该具有匹配能力（例如，如所需要，两个端口 35 应该能够结合和划分上游和下游方向的包）。在所示的实施例中，设备 36 和设备 38 之间有四个路线 37。“路线”是一组不同的发送和接收对。这里使用的术语“链接”（例如，PCI Express 链接）指两个（或多个）端口和它们互联路线。

图 5 示出了包括图形控制器 46、与图形控制器 46 耦合的第一 PCI Express 设备（例如，根 40）和第二 PCI Express 设备（例如，转换器 48）的系统。所述系统还包括与根 40 耦合的中央处理单元（“CPU”）42、与根 40 耦合的存储器 44、与转换器 48 耦合的旧式端点（legacy endpoint）50、与转换器 48 耦合的端点 52。虽然根 40 在图 5 中作为单个元件被示出，但是根 40 可以包括存储器控制器中心（例如，与 CPU 42、图形控制器 46 和存储器 44 耦合的北桥）和一个 IO 控制器中心（例如，与所述北桥和转换器 48 耦合的南桥）。

根 40 和转换器 48 能够执行上面所描述的方法（例如，结合、划分和发送包）。同样地，端点 52 也能够执行先前所描述的方法。但是，旧式端点 50 不能够进行结合和划分包。

基于图 5 系统中设备的结合功能的实现，图 6 示出了不同的数据流。例如，如果在根 40 上不能结合，线 47 上的包流将作为图 6 中的流 0（例如，每个包分别发送）出现。更可能地，如上面描述的，每一个包的分别传输使比特填充成为必要，以便满足传输的最小包尺寸。

然而，如果在根 40、转换器 48 和端点 52 上能够对包结合，则线 47 上的包流将例如作为图 6 中的流 1 出现，所述流 1 具有基于尺寸、目的地和目的地设备的结合能力（和/或实现）适当地结合的多个包。转换器 48 基于它们合适的目的地发送接收的包。在包作为流 1 出现的情况，为旧式端点 50 指定未结合的包 C 和 D，并且为端点 52 指定包含 AB 和 EFG 而结合的包。因此，线 49 上的包流将会作为图 6 中的流 2（例如，没有结合的包）出现以及线 51 上的包流将会作为图 6 中的流 3 出现。

通过利用在此公开的方法和设备，与不进行结合包的常规系统相比，某些 PCI 系统的效率能增加 5-10%。在其它优点之中，当在南桥芯片组中使用在此公开的不同实施例时，端口的利用率显著提高，这允许使用更多带宽用于在相同设备中的外围设备的高通信量。

应该理解的是，尽管在上述描述中已经阐明了不同实施例的许多特征和优点以及不同实施例的结构和功能的细节，但是这个公开仅仅是说明性的。在不脱离不同实施例的范围（所述范围由所附权利要求的条款的广泛全面意义所表达）的情况下，可以详细地做一些改变，尤其对部件的结构和管理进行一些改变。

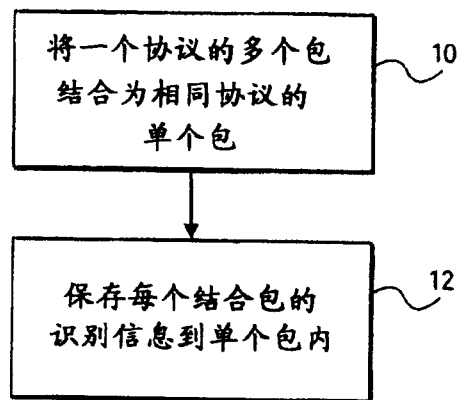


图 1

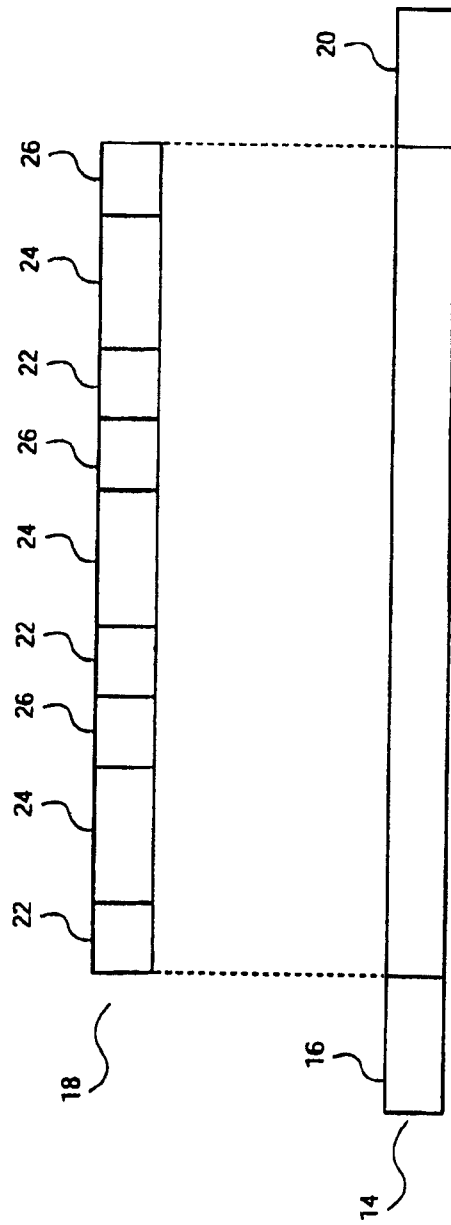


图 2

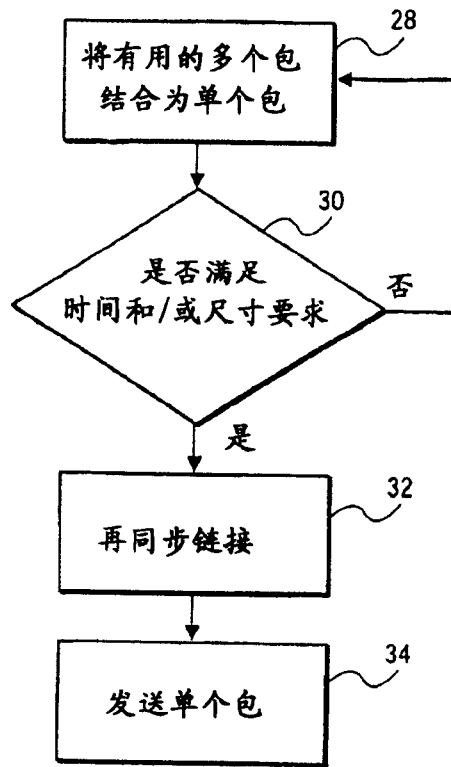


图 3

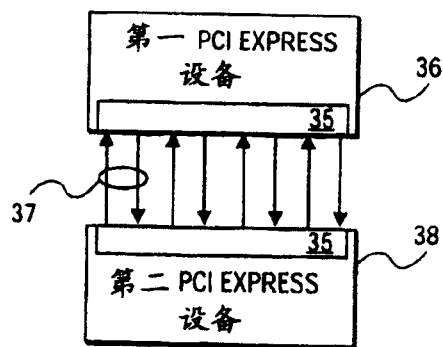


图 4

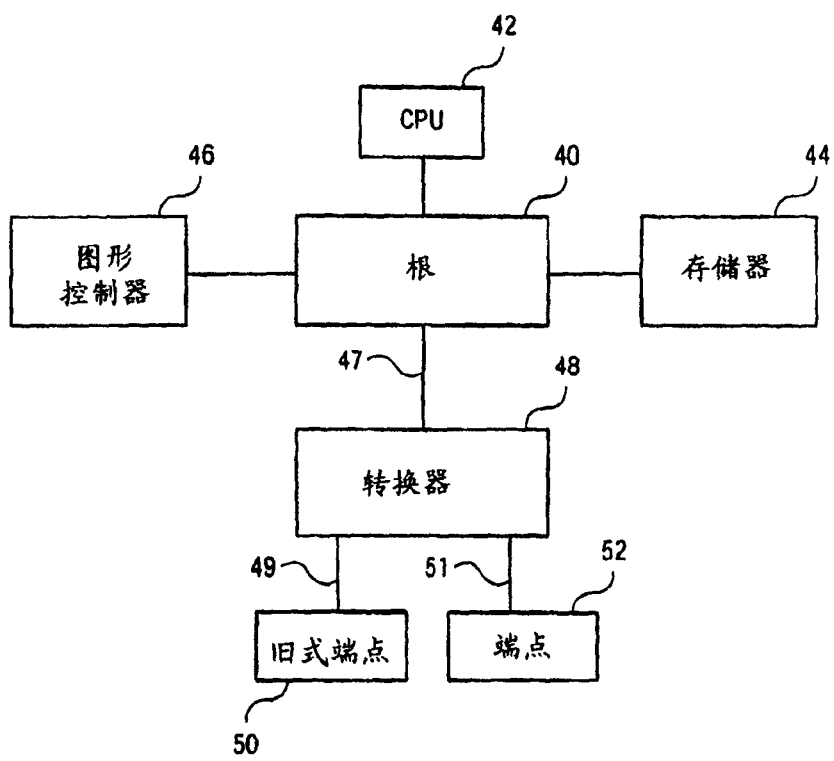


图 5

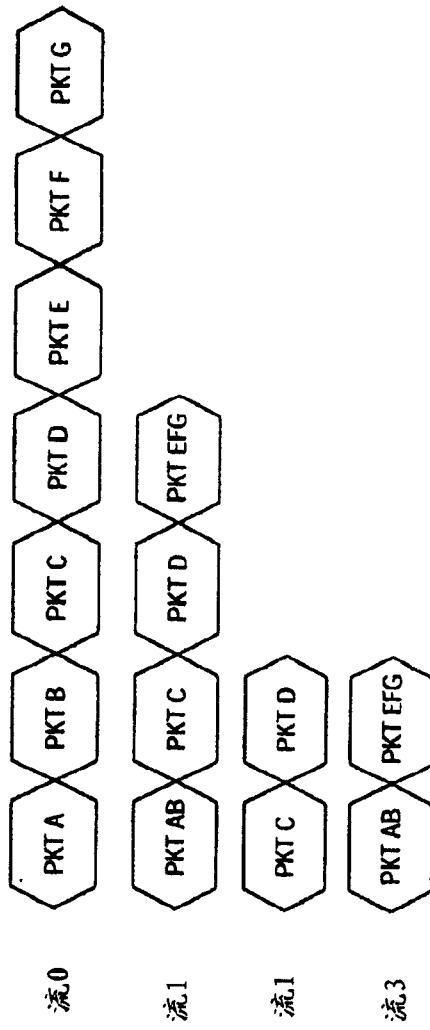


图 6