(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 1890627 B
(45)授权公告日 2010.06.16

(21)申请号 200480036121.6

(22)申请日 2004.12.23

(30)优先权数据
10/750,013 2003.12.31 US

(85)PCT申请进入国家阶段日
2006.06.05

(86)PCT申请的申请数据
PCT/US2004/043687 2004.12.23

(87)PCT申请的公布数据
WO2005/066827 EN 2005.07.21

(73)专利权人 英特尔公司
地址 美国加利福尼亚州

(72)发明人 D·J·施密特 D·M·普夫 S·科塔姆雷迪 L·雷诺

(74)专利代理机构 上海专利商标事务所有限公司 31100
代理人 陈斌

(51)Int.Cl.
G06F 5/06(2006.01)

(56)对比文件
EP 1152573 A2,2001.11.07, 全文.
US 5272728 A,1993.12.21, 全文.
US 4740962 A,1988.04.26, 全文.
US 2002/0009049 A1,2002.01.24, 全文.

审查员 王燕

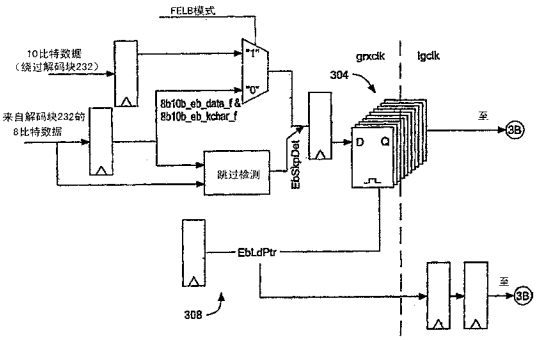
权利要求书 4 页 说明书 9 页 附图 11 页

(54)发明名称

集成电路设备、系统和方法

(57)摘要

在第一集成电路(IC)设备中接收多个符号，其中这些符号已由第二IC设备传送并经串行点到点链路接收。这些符号包括已由第二设备插入数据序列的非数据序列。这些符号被载入缓冲器。数据序列和部分非数据序列根据改变着的卸载指针从缓冲器中卸载。为了防止缓冲器的上溢并响应于检测非数据序列，卸载指针改变一个以上条目，从而载入缓冲器的非数据序列的非数据符号在从缓冲器卸载时被跳过。在另一实施例中，为了防止缓冲器的下溢，卸载指针在卸载时在包含非数据符号的缓冲器条目上停止。还描述和声明了其它实施例。



CN 1890627 B

1. 一种用于集成电路的方法，包括：

a) 在第一集成电路设备中接收多个符号，所述符号是由第二集成电路设备传送的并且是经串行点到点链路接收的，其中所述多个符号包括由所述第二集成电路设备根据预定方法插入到一数据序列中的一非数据序列；

b) 根据载入指针将所述多个符号载入缓冲器；

c) 根据一指向缓冲器中不同条目的改变的卸载指针，从缓冲器中卸载数据序列和部分非数据序列，其中所述卸载指针在每卸载一个符号时改变一个条目；以及

d) 为了防止缓冲器的上溢并响应于 (i) 在缓冲器的入口处检测到所述非数据序列以及 (ii) 将指出这种检测的指示符传送通过缓冲器，将卸载指针改变一个以上条目，从而载入缓冲器的非数据序列的一非数据符号在 c) 中卸载时被跳过。

2. 如权利要求 1 所述的方法，其特征在于，通过检测到第一非数据符号后跟所述非数据序列中一不同的第二非数据符号的组合，来检测到所述非数据序列。

3. 如权利要求 2 所述的方法，其特征在于，传送所述指示符包括：

响应于检测到所述非数据序列的第一和第二非数据符号来产生一标志，并在 b) 中将所述标志与所述非数据序列一起载入所述缓冲器时，使所述标志与所述第一非数据符号对齐。

4. 如权利要求 3 所述的方法，其特征在于，响应于在缓冲器的出口处检测到所述标志，在 d) 中改变所述卸载指针，从而跳过载入缓冲器的第二非数据符号。

5. 如权利要求 1 所述的方法，其特征在于，所述非数据序列是包括非数据符号 COM 后跟非数据符号 SKP 的 PCI Express 序列。

6. 一种用于集成电路的方法，包括：

a) 在第一集成电路集成电路设备中接收多个符号，所述符号是由第二集成电路设备传送的并是经耦合所述第一和第二集成电路设备的串行点到点链路接收的，其中所述多个符号包括由所述第二集成电路设备插入到一数据序列中的一非数据序列；

b) 根据载入指针将所述多个符号载入缓冲器；

c) 根据一改变的卸载指针从缓冲器中卸载所述数据序列和部分非数据序列，其中所述卸载指针在每卸载一个符号时改变缓冲器的一个条目；以及

d) 为了防止缓冲器的下溢，并响应于 (i) 在缓冲器的入口处检测到所述非数据序列以及 (ii) 将指出这种检测的指示符传送通过所述缓冲器，在 c) 中卸载时将所述卸载指针停止在包含一非数据符号的缓冲器条目上。

7. 如权利要求 6 所述的方法，其特征在于，通过检测到第一非数据符号后跟所述非数据序列中一不同的第二非数据符号的组合，来检测到所述非数据序列。

8. 如权利要求 7 所述的方法，其特征在于，传送所述指示符包括：

响应于检测到所述非数据序列的第一和第二非数据符号来产生一标志，并在 b) 中将所述非数据序列载入所述缓冲器时，使所述标志与所述第一非数据符号对齐。

9. 如权利要求 8 所述的方法，其特征在于，响应于在缓冲器的出口处检测到所述标志，将所述卸载指针停止在包含所述非数据序列的所述第二非数据符号的缓冲器条目上。

10. 如权利要求 6 所述的方法，其特征在于，所述非数据序列是包括非数据符号 COM 后跟非数据符号 SKP 的 PCI Express 序列。

11. 一种集成电路设备，包括：

缓冲器，具有接收由另一集成电路设备经串行点到点链路传送的多个符号的输入，所述缓冲器具有多个条目；

检测逻辑，具有接收所述多个符号的输入，以及向缓冲器的所述输入馈入非数据符号序列标识符的输出；

第一指针逻辑，提供分别将所述多个符号依次载入所述缓冲器的多个条目的第一指针；

第二指针逻辑，提供分别将所述多个符号从所述缓冲器的多个条目中依次卸载的第二指针；

比较逻辑，比较第一和第二指针；以及

指针控制逻辑，具有与所述第二指针逻辑耦合的输出，

其中所述指针控制逻辑响应于 a) 所述标识符在缓冲器的输出处显现，以及 b) 所述比较逻辑指示出缓冲器不如预定阈值满，将所述第二指针停止在包含一非数据符号的条目上。

12. 如权利要求 11 所述的集成电路设备，其特征在于，所述多个符号是要根据第一时钟信号接收的，所述第一时钟信号是要从所述另一集成电路设备的传送时钟导出的。

13. 如权利要求 12 所述的集成电路设备，其特征在于，所述第一时钟信号是要从内嵌于包含所述多个符号并要由所述另一个集成电路设备传送的信息流中的传送时钟导出的。

14. 如权利要求 12 所述的集成电路设备，其特征在于，所述第二指针逻辑根据从所述集成电路设备的本地时钟导出的第二时钟信号来推进第二指针，

且其中所述第一指针逻辑根据所述第一时钟信号来推进所述第一指针。

15. 一种用于集成电路的系统，包括：

处理器；

主存储器；以及

集成电路设备，与所述处理器和主存储器通信耦合，并向所述处理器提供 I/O 访问，所述集成电路设备具有支持串行点到点链路的链路接口电路，所述电路包括：

缓冲器，具有接收在所述链路上传送的多个符号的输入，所述缓冲器具有多个条目；

检测逻辑，具有接收所述多个符号的输入，以及向缓冲器的所述输入馈入非数据符号序列标识符的输出；

第一指针逻辑，提供分别将所述多个符号依次载入所述缓冲器的多个条目的第一指针；

第二指针逻辑，提供分别将所述多个符号从所述缓冲器的多个条目中依次卸载的第二指针；

比较逻辑，比较第一和第二指针；以及

指针控制逻辑，具有与所述第二指针逻辑耦合的输出，其中所述指针控制逻辑响应于 a) 所述标识符在缓冲器的输出处显现，以及 b) 所述比较逻辑指示出缓冲器不如预定阈值满，将所述第二指针停止在包含一非数据符号的条目上。

16. 如权利要求 15 所述的系统，其特征在于，所述多个符号是要根据第一时钟信号接收的，所述第一时钟信号是要由所述集成电路设备从另一设备的传送时钟导出的。

17. 如权利要求 16 所述的系统，其特征在于，所述第一时钟信号是要从内嵌于包含所述多个符号并要由所述另一个设备传送的信息流中的传送时钟导出的。

18. 如权利要求 15 所述的系统，还包括图形元件；以及

其中所述集成电路设备是将所述处理器通信耦合到主存储器和图形元件的存储器控制器集线器 MCH。

19. 如权利要求 15 所述的系统，其特征在于，所述集成电路设备是将所述处理器通信耦合到外围设备的 I/O 控制器集线器 ICH。

20. 一种用于缓冲器管理的方法，包括：

在弹性缓冲器的入口处检测预定义的非数据符号序列；

将表示对所述序列的检测的标识符传送通过所述弹性缓冲器；以及

在所述弹性缓冲器的出口处处理所述标识符，以在弹性缓冲器中避免上溢和下溢情形之一。

21. 如权利要求 20 所述的方法，其特征在于，所述序列是 PCI Express SKP 有序组。

22. 如权利要求 20 所述的方法，其特征在于，所述处理被设计成将所述弹性缓冲器保持为半满状态。

23. 一种集成电路设备，包括：

缓冲器，具有接收由另一集成电路设备经串行点到点链路传送的多个符号的输入，所述缓冲器具有多个条目；

检测逻辑，具有接收所述多个符号的输入，以及向缓冲器的输入馈入非数据符号序列标识符的输出；

第一指针逻辑，提供分别将所述多个符号依次载入所述缓冲器的多个条目的第一指针；

第二指针逻辑，提供分别将所述多个符号从所述缓冲器的多个条目中依次卸载的第二指针；

比较逻辑，比较第一和第二指针；以及

指针控制逻辑，具有与所述第二指针逻辑耦合的输出，其中所述指针控制逻辑响应于 a) 所述标识符在缓冲器的输出处显现，以及 b) 所述比较逻辑指示出缓冲器比预定阈值满，将所述第二指针停止在包含一非数据符号的条目上。

24. 如权利要求 23 所述的集成电路设备，其特征在于，所述多个符号是要根据第一时钟信号接收的，所述第一时钟信号是要从所述另一集成电路设备的传送时钟导出的。

25. 如权利要求 24 所述的集成电路设备，其特征在于，所述第一时钟信号是要从内嵌于包含所述多个符号且是要由另一个集成电路设备传送的信息流中的传送时钟导出的。

26. 如权利要求 24 所述的集成电路设备，其特征在于，所述第二指针逻辑根据从所述集成电路设备的本地时钟导出的第二时钟信号来推进第二指针，且其中所述第一指针逻辑根据所述第一时钟信号来推进所述第一指针。

27. 一种用于集成电路的系统，包括：

处理器；

主存储器；以及

集成电路设备，与所述处理器和主存储器通信耦合，并向所述处理器提供 I/O 访问，所

述集成电路设备具有支持串行点到点链路的链路接口电路，所述电路包括：

缓冲器，具有接收经所述链路传送的多个符号的输入，所述缓冲器具有多个条目；

检测逻辑，具有接收所述多个符号的输入，以及向缓冲器的输入馈入非数据符号序列标识符的输出；

第一指针逻辑，提供分别将所述多个符号依次载入所述缓冲器的多个条目的第一指针；

第二指针逻辑，提供分别将所述多个符号从所述缓冲器的多个条目中依次卸载的第二指针；

比较逻辑，比较第一和第二指针；以及

指针控制逻辑，具有与所述第二指针逻辑耦合的输出，其中所述指针控制逻辑响应于 a) 所述标识符在缓冲器的输出处显现，以及 b) 所述比较逻辑指示出缓冲器比预定阈值满，使所述第二指针推进一个以上条目从而跳过包含非数据符号的条目。

28. 如权利要求 27 所述的系统，其特征在于，所述多个符号是要根据第一时钟信号接收的，所述第一时钟信号是要由所述集成电路设备从另一设备的传送时钟中导出的。

29. 如权利要求 28 的系统，其特征在于，所述第一时钟信号是要从内嵌于包含所述多个符号且要由所述另一个设备传送的信息流中的传送时钟导出的。

30. 如权利要求 27 所述的系统，还包括图形元件；以及

其中所述集成电路设备是将所述处理器通信耦合到主存储器和图形元件的存储器控制器集线器 MCH。

31. 如权利要求 27 所述的系统，其特征在于，所述集成电路设备是将所述处理器通信耦合到外围设备的 I/O 控制器集线器 ICH。

集成电路设备、系统和方法

技术领域

[0001] 本发明的一个实施例一般涉及适合电子系统的通信耦合元件的串联、点到点互连技术，尤其涉及具有依照PCI Express Base Specification快速基本规范1.0a(2003年10月7日的勘误表)(“PCI Express”)。也可描述其它实施例。

背景技术

[0002] 电子系统由被设计成在系统的输入/输出(I/O)互连上彼此通信的若干元件组成。例如，现代计算机系统可包括以下元件：处理器、主存储器、和系统接口(也称为系统芯片组)。元件可包括一个或多个集成电路(IC)设备。例如，系统芯片组可具有使处理器能与系统存储器和图形元件通信的存储器控制器集线器(MCH)设备。此外，可提供通过MCH将处理器和存储器连接到计算机系统的其它元件，诸如大容量存储设备和外围设备的I/O控制器集线器(ICH)设备。在该情形中，诸如由PCI Express限定的单独、点到点链路可用来允许一对设备之间的双向通信，例如处理器和MCH、MCH和图形元件、ICH和大容量存储设备。

[0003] PCI Express点到点链路可具有能同时运行的一条到多条通道。每条通道都具有可同时运行的双重单向路径。每条路径可具有单组发送器和接收器对(例如设备A端口的发送器、设备B端口的接收器)。在该情形中，发送器和接收器可驱动并感测传输介质，诸如印刷电路板中可穿过板与板接头的一对金属迹线。或者，可提供诸如光纤的其它传输介质。

[0004] 点到点链路用来在设备之间传送各类信息。然而在所谓“较高层”上，两个设备(也称为请求者和完成者)中对等体之间的通信可使用事务(transaction)来引导。有将数据往返传送到存储器映射位置的存储事务。根据PCI Express，有传送各种消息并可用于像中断信令、出错信令、和功率管理的功能的消息事务。

[0005] “建立”一事务可有三个抽象层。第一层可以是事务层，它开始将来自设备内核的请求或完成数据转换成用于事务的数据分组的过程。第二个构架建立层称为数据链路层；它确保在链路上往返的分组被适当地接收(通过诸如差错控制编码的技术)。第三层称为物理层。该层负责分组在链路上的实际传送和接收。给定设备中的物理层与一侧的其数据链路层(同一设备)交互，并与另一侧的作为链路一部分的金属迹线、光纤或其它传输介质交互。物理层可包含用于发送器和接收器、并行到串行和串行到并行转换器、频率和相位控制电路、以及阻抗匹配电路的电路。还包含用于其初始化和维护所需的逻辑功能的电路。通过例如允许重新使用基本相同的事务和数据链路层、同时升级物理层(例如增大收发时钟频率)，分层架构可允许较简便的升级。

[0006] 现在给出物理层的一动作示例。一旦通电，设备A和设备B上的物理层就负责初始化链路，并使其对各种事务作好准备。该初始化过程可包括确定应对该链路使用多少通道，并且链路应在什么数据速率上运行。在适当初始化链路之后的某时，在设备A上发出存储器读取请求。最终，包括该读取请求的分组到达设备A的物理层，包括由更高层添加的报头、误差控制信号和序列号。然后物理层取得该数据分组，并将其转换成串行数据流(可能

在向其添加成帧数据之后），并使用具有预定义定时规则的例如电差分信号来传送该流。

[0007] 一旦设备 B 的物理层看到信号在其接收器输入端出现，它就采样该信号以恢复该数据流，并将该流重建成数据分组（例如在去除帧信号之后）。然后分组被往上传送给设备 B 的数据链路层，该层剥去报头并检错；如果没有差错，则该分组被往上传送给传输层，在该层上存储器读取请求被提取并发送给适当的逻辑功能，以访问在请求中指定的位置。

附图说明

[0008] 本发明的各个实施例通过示例、而并非通过附图中的限制来进行说明，在附图中类似标号表示相似元件。应注意，对本发明“一”实施例的引用并非必然是同一实施例，并且它们表示至少一个。

[0009] 图 1 示出通过串联点到点链路彼此耦合的一对集成电路设备。

[0010] 图 2 示出用来在集成电路设备中实现串联的点到点链路的部分链路接口电路的框图。

[0011] 图 3A 和 3B 示出可用来实现点到点链路的物理层中的缓冲器管理。

[0012] 图 4 示出在图 3 的缓冲器管理电路中非数据符号检测标志可如何对齐的时序图。

[0013] 图 5 是示出指针比较运算示例的示例时序图。

[0014] 图 6 示出用于管理缓冲器以防止上溢的示例时序图。

[0015] 图 7A 示出用于管理缓冲器以防止上溢的示例时序图。

[0016] 图 7B-7C 示出缓冲器的示例启动情形的时序图。

[0017] 图 8 标识多媒体台式个人计算机的各个元件，其中一部分通过 PCI Express 虚拟信道 (VC) 彼此通信耦合。

[0018] 图 9 示出企业网络的框图。

具体实施方式

[0019] 本发明的一个实施例涉及通过非数据符号处理用于点到点链路的缓冲器管理。图 1 示出通过串联点到点链路彼此耦合的一对集成电路设备。IC 设备 104（设备 A）和 108（设备 B）可以是包含处理器 112 和主存储器 114 的计算机系统的一部分。在该示例中，串行点到点链路 120 被用来使设备 B 的内核与设备 A 的内核通信耦合。该链路 120 具有双重单向路径 122，其中链路接口 124 用来与每一相应设备 A 和 B 的设备内核交接。

[0020] 在该实施例中，设备 B 被称为计算机系统的根联合体 (root complex)，并向例如设备 A 中的图形元件提供具有 I/O 访问的处理器 112。该根联合体可被分成图形和存储器控制器集线器 (GMCH)、以及 I/O 控制器集线器 (ICH)。ICH 将用作 GMCH 和系统的其它 I/O 设备，包括非易失性大容量存储设备、诸如跟踪板或鼠标的定位设备、以及网络接口控制器（未示出）之间的另一个接口。点到点链路 120 可被复制，以使设备 B 与处理器 112 和主存储器 114 通信耦合。其它表征点到点链路 120 的其它平台架构也是可能的。

[0021] 图 1 的接口 124 可被视为实现用于串行点到点链路的多层架构（如以上背景技术中所述）。接口 124 的部分细节在图 2 中示出。接口 124 支持传输介质 122 和其相应设备 104、108 的数据链路层之间的独立收发路径。在传送路径中，数据分组形式的信息从数据链路层到达，并被分成由编码块 208 编码的符号。块 208 编码的目的是要嵌入时钟信号，从而

独立的时钟信号无需被传送到传输介质 122 中。该编码可以是众所周知的 8B-10B，其中 8 比特数量被转换成 10 比特数量；其它编码方案也是可能的。在某些情形中，诸如独立选通或时钟信号在介质 122 中传送的情形，并不需要这种编码。

[0022] 在框 208 中编码之后，数据单元（在此称为符号）由模拟前端 (AFE) 传送块 214 的并行 - 串行块 212 处理，以产生比特流。注意，在此使用的“比特”可表示两种以上不同状态，例如二进制比特、三进制比特等。术语“比特”在此仅为方便起见而使用，并非旨在受限于二进制比特。然后比特流被导入传输介质 122。如上在背景技术中所述，该传输介质可以是在印刷线路板上形成的一对金属迹线。也可使用其它形式的传输介质 122，诸如光纤。

[0023] 框 208-214 序列可用作点到点链路 120（图 1）的单通道。一般而言，在点到点链路 120 中可有一条以上通道，使从数据链路层接收的分组可在多个传输通道上剥除。

[0024] 现在参看图 2 所示接口 124 的接收侧，每条通道具有其相关联 AFE 接收块 224，用来例如通过采样传输介质 122 中的信号来从传输介质 122 中接收信息流。AFE 接收块 224 在传输介质 122 的信令和 IC 设备 104 的信令（例如芯片上、互补金属氧化物半导体、CMOS、逻辑信令）之间转换。如以下将要描述的，信息流表示已由设备 B 在串联点到点链路 120（参见图 1）上传送的 M 比特符号（其中 M 是大于 1 的整数）序列。

[0025] 由 AFE 接收块 224 提供的比特流被馈入用来对齐或锁定已接收符号的符号对齐逻辑 228。换言之，并如以下所述，符号对齐逻辑 228 将划分所接收比特流内的准确符号边界，由设备 104 中物理层的后续部分使用。

[0026] 然后符号对齐的比特流可被馈入解码块 232，该解码块 232 解除由编码块 208 执行的编码（例如，10B-8B 解码、以产生由 8 个二进制比特组成的信息符号）。

[0027] 然后解码后的符号被馈入弹性缓冲器 EB 234。EB 234 用来补偿符号用设备 B 和设备 A 的本地时钟信号 (local_clk) 传送的速率容差中的任何差异。local_clk 用来从 EB 234 中卸载符号，并且在一些情形中如下所述地运行通道到通道偏斜消除电路 238 的各个部分（在链路 120 由一个以上通道组成的情形中）。应注意，解码块 232（如果提供）可置于更下游，例如在 EB 234 的输出处或偏斜消除电路 238 的输出处。

[0028] 一部分 EB 234 的示例框图在图 3A 和 3B 中示出。在该示例中，EB 234 具有通过解码块 232（参见图 2）从对齐逻辑 228 接收 8 位符号的输入（图 3A 的左边）。以下将描述的可选方案是远端环回模式 (FELB)，其中符号因为已绕过解码块 232 而为 10 比特宽。或者其它符号宽度是可能的。

[0029] 符号可以是表示源自数据链路层、事务层、或诸如设备内核的一些其它更高层的一些有效载荷的“数据”符号。或者，符号可以是“非数据”符号，例如由物理、数据链路、或传输层之一产生的特定符号，以实现对在串行点到点链路上传输的信息上的一些类型控制。这些非数据符号的若干示例将如下给出为 PCI Express 特定符号。

[0030] PCI Express 定义被添加到所传送分组的众多特定符号。例如，可添加特定符号以标记分组的开始和结束。这通过让接收设备知道一个分组在哪里开始并在哪里结束来实现。不同的特定符号被添加到源自事务层而非数据链路层的分组中。此外，有要由物理层用于补偿两个通信端口的运行数据速率中的较小差异的称为“SKP”（跳跃符）的特定符号。还有要用于物理层的通道和链路初始化的称为“COM”（逗号）的特定符号。

[0031] 到达 EB 234 的输入处的符号要根据由载入指针逻辑 308 提供的载入指针 EbLdPtr

顺序载入缓冲器 304(可具有先入先出架构,也称为队列)的一系列条目。由卸载指针逻辑 312 提供的卸载指针 EbUldPtr 用来从缓冲器 304 顺序卸载符号。如图 3A 所示,有通过缓冲器 304 的竖直虚线。这表示由 EB 234 在接收时钟 grxclk 和本地时钟 lgclk 之间执行的时钟跨越。符号根据 grxclk 载入,并且它们根据 lgclk 卸载。尽管这两个时钟域可被设计成在频率方面彼此尽可能地接近,但每个时钟域被允许有频率上的一些容差或极小变化,常常指定为百万分之几。grxclk 可从(已传送符号的)另一 IC 设备的传送时钟导出,其中该传送时钟可被嵌入由其它设备传送的信息流中,或者在诸如源同步情形中它可在独立的时钟或选通信号中提供。根据 PCI Express,grxclk 可具有 +/-300ppm 的容差。相同容差可被分配给设备 A 的本地时钟 lgclk。

[0032] 为了说明 EB 234、特别是缓冲器 304 的上溢和下溢问题,假设在启动时缓冲器 304 的载入和卸载指针被分开约缓冲器的一半深度。取决于 grxclk 和 lgclk 的频率之间的实际差异,这些指针可开始彼此漂离或者可开始彼此漂近,从而随着时间的流逝指针会碰撞,即上溢或下溢。EB 234 的理想情形是载入和卸载指针总是被分开缓冲器 304 的一半深度。如下所述,通过将卸载指针调节或控制为 a) 检测符号的特定或非数据符号序列以及 b) 缓冲器的濒临上溢或下溢情形的功能,而不调节更新载入指针的缺省方式,可实现该理想情形。

[0033] EB 234 的卸载指针可被受控(使用例如图 3B 中的卸载指针逻辑 312 和指针控制逻辑 314)为使用已由设备 B(参见图 1)插入数据序列的符号的预定义、特定或非数据序列,避免上溢和下溢情形。简言之,为了防止缓冲器的下溢,卸载指针可响应于检测到非数据序列在包含非数据符号的缓冲器的入口处停止。这在根据改变卸载指针卸载数据序列时实现。这使载入指针离开卸载指针,从而防止下溢。

[0034] 另一方面,为了防止缓冲器的上溢,卸载指针可被改变一个以上条目,从而跳过非数据序列的非数据符号(因为它目前载入到缓冲器中),同时从缓冲器中卸载符号。再一次,这响应于检测非数据序列来实现。这使卸载指针离开载入指针,再一次避免碰撞。用于实现上溢和下溢避免能力的示例技术的细节给出如下。

[0035] 现在参看图 3A 和 3B,EB 234 的缓冲器 304 可被设计成在每个条目中不仅存储一个符号(诸如 8 比特或 10 比特字符),还存储该符号的一个控制位,该控制位表示符号是数据符号还是非数据符号 (8b10b_eb_kchar_f),以及预定义的非数据序列指示符 (EbSkpDet)。kchar_f 控制比特可由解码块 232 产生,同时 EbSkpDet 可由如图所示的 EB 234 产生。后一指示符用于 PCI Express 实施例的特定示例,其中所使用的特定非数据序列为 SKP 有序组。或者,可使用另一预定义的非数据序列。EbSkpDet 非数据序列指示符可如下所述由 EB 234 用于卸载指针的管理。

[0036] 为了适当地调整 EB 234 的卸载和载入指针,在有序组的 PCI Express COM 情形中 SKP 有序组检测标志用所接收的非数据符号在缓冲器的输入处产生并对齐。在有序组中 COM 符号在一个或多个 SKP 符号之前。指示符通过 EB 234 传送,从而可在 lgclk 域中相对于有序组采取校正动作(图 3A 中所示竖直线的右边)。如在图 4 的时序图中所示,有序组指示符可以是当非数据符号 COM 后面是非数据符号 SKP 时对 grxclk 的一个周期赋值的信号。在图 4 中,波形 8b10b_eb_data[7:0] 表示所接收的符号(在该情形中包括插入表示为 Dx.x 序列的数据序列的 SKP 有序组)。所接收符号和有序组指示符 EbSkpDet 在存储到缓冲器 304 的条目中之前下跳。注意,COM 符号和 EbSkpDetin 的赋值在同一 grxclk 的周期

中进行。换言之，检测标志 EbSkpDet 以及在该情形中的 8 比特符号 EbDataIn[7:0] 被赋值并载入到缓冲器中（作为 EbSkpDetin）。

[0037] 参看图 3B，比较逻辑 316 能采样卸载和载入指针相对于彼此的位置，从而指针的适当调节可在检测到非数据序列时实现。这表示，在本实施例中指针之一可能需要跨过时钟域，以确定两个指针在队列内的位置。在本实施例中，在 grxclk 域内的载入指针将跨越到 lgclk 域上。注意，使用灰度来表示指针可提供比简单二进制更准确和有效的实现。

[0038] 在 lgclk 域内，有生成以表示缓冲器 304 情形，即比 1/2 满多或比 1/2 满少的两个指示符。或者，可定义仍然可使 EB 234 避免上溢和下溢情形的其它情形（诸如比预定阈值多或比预定阈值少）。在该示例中，比 1/2 满多的指示符是 EbMrHlfFull，并表示 grxclk 域比 lgclk 域“更快”。当赋值该指示符时，并且当已接收到非数据序列时，应从有序组中去除非数据符号（在该情形中为 SKP），以尝试将缓冲器带回其理想的 1/2 满的情形。

[0039] 另一方面，比 1/2 满少的指示符 EbLsHlfFull 表示相反，即 lgclk 域比 grxclk 域快。在该情形中，当已接收 SKP 有序组时应添加 SKP，以将指针带回其理想的 1/2 满的情形。当然，当两个指示符都未赋值时缓冲器可以为 1/2 满，从而无需对载入和卸载指针采取任何动作。在本发明一实施例中，SKP 实例的添加和去除通过指针控制逻辑 314（图 3B）在卸载指针 EbUldPtr（而非载入指针 EbLdPtr）上动作来实现。其运算可通过图 6 和 7 的示例时序图说明，如下所述。

[0040] 图 5 是根据指针在不同时钟域可如何比较指针的示例时序图。图 5 示出在该示例中 grxclk 更快的 grxclk 和 lgclk 波形。在此，载入指针 EbLdPtr 跨入 lgclk 域，在载入指针的实际位置和同步位置 EbLdPtrSync 之间有 1 ～ 2 个周期的滞后。为了补偿与跨越载入指针相关联的该时间延迟，还可通过将当前值减去 2（在该实例中）产生 EbUldPtrAdj 来调节卸载指针的值。然后可在 EbLdPtrSync 和 EbUldPtrAdj 之间作比较，从而在该情形中如在 lgclk 的周期 4 所示缓冲器比 1/2 满多。注意，在该示例中缓冲器 304 的深度被假定为 10 个条目，尽管其它深度也可工作。

[0041] 仍然参看图 5 的时序图，注意在 lgclk 的前 4 个周期中，同步载入指针 EbLdPtrSync 与经调节的卸载指针 EbUldPtrAdj 相差约缓冲器深度的一半，在本情形中即为 5 个条目。相应地，撤消对 EbMrHlfFull 和 EbLsHlfFull 的赋值。然而，在周期 3 中，同步载入指针往前跳一条目（从条目 8 到条目 0），并且因为两个指针之间的差异大于缓冲器深度的一半，EB 234 被视为多于 1/2 满，因而接近上溢。本领域普通技术人员之一基于本说明书将理解对下溢情形可画出相似的时序图。

[0042] 对于指针比较逻辑 316（图 3），用于确定指针位置的算法可如下。如果经调节的卸载指针大于同步载入指针，则经调节的卸载指针和同步载入指针之间的差异为队列内空闲的条目数量。另一方面，如果同步载入指针大于经调节的卸载指针，则同步载入指针和经调节的卸载指针之间的差异为队列内的条目数量。当然，当同步载入指针等于经调节的卸载指针时指针碰撞，即 EB 234 上溢或下溢。指针碰撞可由于例如缺少所接收的非数据序列，或者 grxclk 和 lgclk 频率之间的差值太高并在设计规范之外。在该情形中，将向后续符号处理块或设备 A 的上层发送指针碰撞的指示，从而开始给定链路的所有通道（参见图 2）中的指针 2 被移回其初始或重置值的复原状态。

[0043] 现在参看图 6 和 7，示出非数据序列可如何处理以避免上溢和下溢情形的示例时

序图。回想如上所述，当已接收到 SKP 有序组时，一标志在缓冲器 304 的入口处产生，并连同有序组的符号通过缓冲器传送。在所述示例中，应用以管理缓冲器的调节在缓冲器的出口处，即在 lgclk 域内发生。特别地，卸载指针取决于缓冲器的状态（例如：1/2 满、比 1/2 满多、或比 1/2 满少）进行调节。图 6 示出在缓冲器比 1/2 满多的情形中调节或控制卸载指针的过程的时序图。注意在该示例中 grxclk 比 lgclk 快，从而可能导致上溢情形。在该情形中，包括 COM 加上单一 SKP 的 SKP 有序组在 EB 234 的入口处接收。然后在周期 1 向缓冲器加载与 COM 对齐的 SKP 检测标志 (EbSkpDetin)，进入缓冲器的条目 9。

[0044] 注意在 lgclk 域内，直到周期 3 指针都分开 5 个条目。在周期 3，同步载入指针从条目 8 移到条目 0，表示载入指针已超过卸载指针一个周期。在周期 7 中，从缓冲器（条目 9）中卸载与 SKP 有序组相关联的 SKP 检测标志 (EbSkpDetOut)。由于缓冲器现在比 1/2 满多，卸载指针 EbUldPtr 将把一额外条目往前移，即不移向条目 0 而是移向条目 1。由于经调节的卸载指针 EbUldPtrAdj 反映卸载指针的移动，EbUldPtrAdj 和 EbUldPtrSync 之间的差异回到 5 个条目，且因为未发出比 1/2 满指示符在周期 9 更新缓冲器状态。因而，将卸载指针改变一个以上条目导致载入缓冲器的非数据符号（在该情形中为 SKP）被跳过，同时符号如 EbDataOut[7:0] 所反映地被卸载。

[0045] 现在参看图 7A，示出用于在缓冲器比 1/2 满少时管理 EB 234 以避免下溢的过程的示例时序图。在该情形中 grxclk 域比 lgclk 域慢，从而使缓冲器的排空比填充快。在示图顶部，有插入如 EbDataIn 所示的到达 EB 234 入口处的数据序列的非数据序列。

[0046] COM 符号以及 SKP 检测标志将在条目 9 中存储，如 grxclk 的周期 1 所示。然后参看 lgclk 域，缓冲器直到周期 3 都为 1/2 满，其中同步载入指针停留在条目 9 达两个周期，而经调节的卸载指针继续递增。这是因为 grxclk 和 lgclk 之间的不匹配或容差差异，如以上在图 5 的时序图中所示。相应地，在周期 4 中，赋值比 1/2 满少指示符。在周期 6 中，SKP 检测标志在赋值 EbLsHlfFull 时从缓冲器中卸载，且卸载指针 EbUldPtr 在赋值 HldUldPtr 时于周期 7 停止。这使得卸载指针在周期 7 时保留在条目 0 上（该条目包含 SKP）。因而，另一个 SKP 被插入该序列，如在 EbDataOut[7:0] 的周期 7 中所见。

[0047] 然后，当从周期 7 转换到周期 8 比较同步载入指针和经调节卸载指针时，指针再次变回分开 5 个条目，使 EB 234 的指针返回到其理想情形。

[0048] 以下提供在以上示例中载入和卸载指针如何运行的另一种描述。对于载入指针，只要 EB 234 活动或启用，该指针就可一直（根据 grxlk）递增 1。然而，对于卸载指针，仅当 EB 234 当前未正在处理缓冲器比 1/2 满多时的非数据序列，且非数据序列在缓冲器比 1/2 满少时未在最后一个循环接收非数据序列时，卸载指针（在初始化后）递增 1（根据 lgclk）。此外，当处理非数据序列且缓冲器比 1/2 满多时卸载指针可递增 2。最后，当非数据序列已在最后一个周期接收并且缓冲器比 1/2 满少时，卸载指针未增加，即停止。

[0049] 用于管理弹性缓冲器的上述方法和装置的一个优点是它是不管收发时钟所允许的容差如何都能保持为串行点到点链路接收的稳定符号流的相对鲁棒的技术。注意，不仅在通电后使链路运行前的初始训练期间，而且在由 IC 设备接收每个分组期间（其中假设每个分组常常都将包括特定、非数据序列的一个或多个实例，以便于使该过程在给定通道的正常运行期间重复），都可执行该过程。在本发明的另一个实施例中，设备 A（参见图 1）可用远端环回模式 (FELB) 运行。使用 FELB，在设备 A 中接收的一系列符号在序列（由 EB

234，参见图 2）已缓冲之后环回到设备 B。相应地，用 FELB，经缓冲序列的符号内容可在设备 A 之外监视，以确定（如由设备 B 传送的）原始序列如何由设备 A 的 EB 234 更改。

[0050] 弹性缓冲器指针的启动

[0051] 本发明的另一实施例为一种启动机制，该启动机制自动调节在 EB 234 中遭遇的异步、时钟跨越延迟，并有助于减少缓冲器 304 所要求的大小。在这种实施例中，EB 234 的载入和卸载指针的启动可基于两种不同准则。Qual_EbActive 项可被定义为在 lgclk 域（卸载指针域）内产生、然后时钟跨越到 grxclk 域（载入指针域）。该条件在赋值时释放载入指针。Qual_EbActive 项可包括以下条件：1）链路接口 124 的链路初始化单元（未示出）表示该 EB 234 的通道可用（例如 gi_gp_laneup 赋值 -lgclk 域）；2）接口 124 的接收时钟被启用（gi_gp_piclken 被赋值 -lgclk 域）；3）EB 234 指针未重置（因为指针碰撞未赋值 gi_gp_ebptrrst-lgclk 域）；4）符号对齐逻辑 228（参见图 2）已获得符号锁（gP_gi_kalignlck-lgclk 域）；以及 5）载入指针已重置。该项可添加于有 PCI Express L0s 进入 / 退出情形（sync_loadrest_done-lgclk 域）的实施例。

[0052] 一旦载入指针已释放，它就被时钟跨越到 lgclk 域。在该时钟域内，载入指针已在连续时钟内改变的事实是现在可释放卸载指针的指示。卸载指针将继续增加，直到任一以上 5 个条件在卸载指针可重置，并且在某时间之后载入指针也将重置（时钟跨越）的情形中变成假。

[0053] 作为示例，卸载指针可被重置成值“000”。相反，载入指针可被初始化为值“001”。其原因是从缓冲器半满情形开始，但在该示例中导致两个时钟的时钟跨越惩罚（使载入指针时钟跨越并离开卸载指针），还导致实际上产生 Ebactive_unload 项的下跳阶段。这表示载入指针可从值“001”开始。注意，卸载指针可仍被减少 2 用于作出检查缓冲器空间的比较。该技术可总是以相同的 EbMrHlfFull 情形开始开始。但是，这并不具有到达 EB 234 的非数据符号 SKP 的第一实例将使缓冲器 304（在此为队列）再次半满的意义。

[0054] 在图 7B 的实例时序图中，内核时钟域的活动指示符 (qual_EbActive) 在内核时钟域 (lgclk) 的周期 1 中被赋值。然后活动指示符发送到 grxclk 域，以产生然后在周期 3 赋值的 sync_EbActive_load 信号。随着 sync_EbActive_load 信号的赋值，载入指针 (ldptr) 从其重置值中释放，并且将开始移动。同时，阻止内核时钟域 lgclk 中的卸载指针 (unldptr) 移动直到 sync_ldptr 开始移动。到周期 6，在卸载指针和经调节卸载指针开始移动之前，同步载入指针已开始移动。这导致赋值比 1/2 满多信号 (EbMrHlfFull)。注意，在该实例中启动机制总是导致 MrHlfFull 在开始时赋值，然后到达的 SKP 第一实例将使队列达 HlfFull 条件。因此，MrHlfFull 的启动条件可被称为瞬变条件。

[0055] 应注意，当 EB 234 的卸载指针在保持其重置状态之后已开始运行时，队列出口处的数据直到卸载指针到达队列的重置载入指针处的条目（即队列的第一条目）时才有效。为了防止非有效数据破坏后续的符号处理阶段（例如图 2 偏斜消除电路 238），来自队列出口处的 SKP 检测标记和 K 字符（非数据符号呈现）比特可用称为 EbOutVld 的有效指示符或标志门控。如图 7C 的示例时序图所示，当阻止卸载指针移动（且认为 EB 234 不活动）时，该指示符可保持未赋值，并且直到卸载指针移到载入指针的重置值（在图 7C 中正好为 ENT0）时才被赋值。以下规则可用来定义该 EbOutVld 的操作：1）EbOutVld 在 EB 234 活动（赋值 qual_EbActive）且卸载指针已移到载入指针的重置状态时被赋值；以及 2）EbOutVld

在EB 234停用（非赋值qual_EbActive）时不被赋值。如上所述，EB 234的出口处的有效标志可防止SKP检测标志EbSkpDetOut和K字符检测标志EbKcharDetOut错误地从存储在队列中的无效符号中被赋值。

[0056] 上述链路接口电路和方法还可用设计成通过提供多媒体等时支持的串行、点到点互连技术通信的IC设备来实现。等时支持是使用确定性和时间相关方法传送数据的特定类型QoS（服务质量）保证。基于平台的等时支持依赖于一种备有证明文档系统的设计方法，使需要对系统资源的恒定或专用级别访问的应用程序能在给定时段内获得所需带宽。

[0057] 一个示例是如图8所示在写报告的同时在桌面上观看源自公司CEO的员工广播。数据从内联网路由到桌面主存储器中，其中应用程序通过内插卡使用数据来创建发送给用户耳机的音频流、并通过图形控制器创建发送给显示器的视频流。如果同步操作在桌面个人计算机(PC)内发生，诸如盘读取、从因特网下载数据、字处理、电子邮件等，则无法确保音频和视频流将真正地无干扰。数据仅根据“最佳尝试”方法传送。用户可经历当多个应用程序为同一资源竞争时的跳过或停止。PCI Express中的等时通过建立确保时间敏感的应用程序能保护充足的系统资源的机制来解决该问题。例如，在图8中，视频时间敏感数据将确保充足的带宽以防止以诸如电子邮件的不重要数据为代价的跳过。

[0058] 上述链路接口电路和方法还可用被设计成通过在从嵌入式应用程序到基于底板的交换系统的通信设备中使用的串行点到点链路技术通信的IC设备来实现。在高级交换中，提供通过交换机构发送对等分组的机制。这些机制也受益于PCIExpress可用的基于服务器类硬件的误差检测。在通信设备内有两类主要使用类型：控制面处理和数据面处理。控制面指系统的控制和配置。串行链路可用作配置和控制大量系统内处理器和卡的接口。基于底板建立的交换机通常具有可插入和使用的各种卡。基于底板的交换机可提供现场可升级能力。大多数交换系统提供开始时仅占据一半底板，当需求或用户数量增加时向卡添加附加端口或更快速连接的能力。串行连接技术可用作控制面互连，以配置和监视在系统内安装的不同类型卡。PCIExpress内的枚举和已建配置协议例如适用于配置卡和服务的低引脚计数、高带宽接口。

[0059] 数据面指数据流动的实际路径。在数据面中，高级交换扩展可定义通过交换机构在对等链路上封装和发送PCI Express数据分组。

[0060] PCI Express内核架构提供用于符合新互连需求的坚实基础。高级交换(AS)架构覆盖在该内核上，并通过使用插在事务层上PCI Express数据分组之前的特定AS报头建立有效的、可升级的和可扩展的交换机构。AS交换机仅检查提供路由信息（向哪里发送分组）、报务类ID（服务信息的质量）、拥挤避免（用于防止报务阻塞）、分组大小和协议封装的报头内容。通过分离路由信息，交换机设计更为简单和成本有效。另外，对分组添加额外的报头能使交换机构封装任何数目的现有协议。

[0061] 上述链路接口电路和方法还可用设计成通过提供用于网络连接（代替例如吉比特以太网）的串行点到点互连技术通信的IC设备来实现。网络连接可由公司移动和桌面计算机用来共享文件、发送电子邮件、以及浏览因特网。服务器以及通信设备可期望来实现这种网络连接。企业网络内的这种网络连接示例如图9所示。

[0062] 尽管以上示例可在组合和时序逻辑电路的环境中描述本发明各实施例，但本发明的其它实施例可通过软件来实现。例如，一些实施例可被提供为计算机程序产品或软件，它

可包括具有存储其上的可用来将计算机（或其它电子设备）编程为根据本发明一实施例执行一过程的指令的机器或计算机可读介质。在其它实施例中，操作可通过包含微码、硬接线逻辑、或通过经编程计算机组件和定制硬件组件的任何组合的特定硬件组件来执行。

[0063] 此外，设计可经过各个阶段，从创建到模拟到制造。表示设计的数据可用多种方式来表示设计。首先，如在模拟中有用的，硬件可使用硬件描述语言或另一功能描述语言来表示。此外，带有逻辑和／或晶体管栅极的电路级模型可在设计过程的某些阶段上产生。此外，大多数设计在某阶段达到了表示各个设备在硬件模型中的物理放置的数据级别。在使用常规半导体制造技术的情形中，表示硬件模型的数据可以是指定不同掩模层上用来产生集成电路的掩模的各种特征的呈现或缺失的数据。在设计的任何表示中，数据可存储在任何形式的机器可读介质中。经调制或以其它方式生成以传送这种信息的光波或电波、存储器、或诸如盘的磁性或光学存储可以是机器可读介质。任一这些介质可“携带”或“指示”设计或软件信息。当表示或携带代码或设计的电子载波被传送到执行电子信号的复制、缓存、重新传送的程度时，制成新的副本。因而，通信供应商或网络供应商可复制表征本发明一实施例的物品（载波）的副本。

[0064] 总而言之，已描述了用于管理串行点到点连接的弹性缓冲器的方法和装置的各个实施例。在前面说明书中，本发明已参照其特定示例性实施例进行了描述。然而，可作各种更改和改变而不背离如所附权利要求所述的更广泛精神和范围是显而易见的。例如，尽管已描述将串行点到点链路用作诸如台式、服务器、或笔记本计算机中的印刷线路板上两个器件之间的芯片到芯片连接的系统实施例，但也可使用具有串行点到点链路的缓冲器管理技术，该串行点到点链路是用于将计算机连接到诸如键盘、监视器、外部大容量存储设备、或照相机的外围的外部总线的一部分。点到点链路可不仅用于计算机系统，而且可用于诸如手机单元、远程通信交换机、和数据网络路由器等的通信产品。因此，说明书和附图在此被视为说明性的，而不是限制性的。

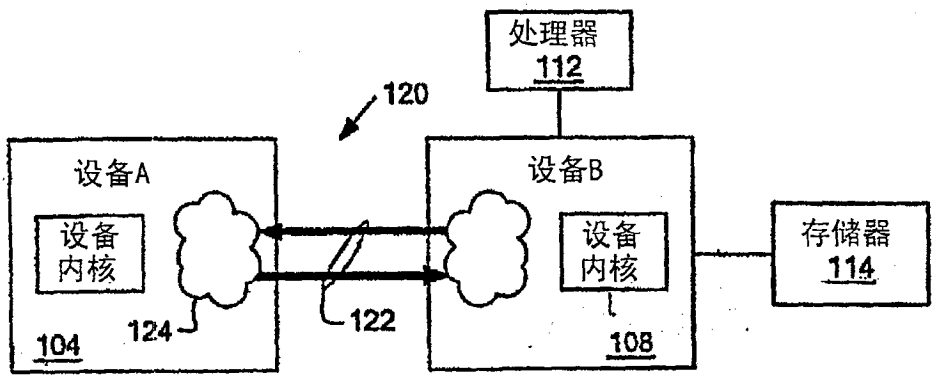


图 1

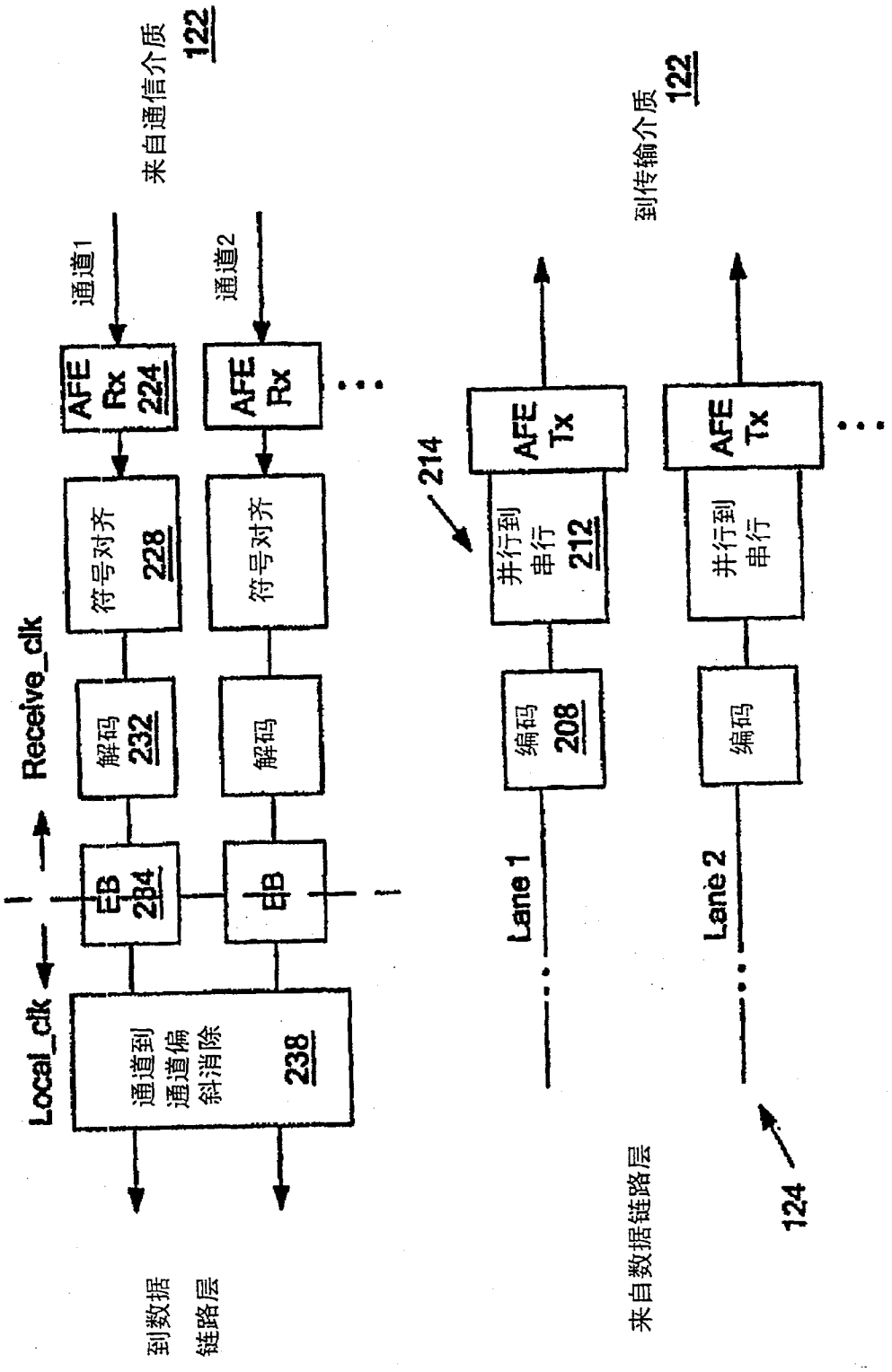

来自通信介质
122
通道1
通道2
AFE Rx 224
符号对齐 228
解码 232
EB 284
通道到通道偏斜消除 238
Receive_clk
Local_clk
到数据链路层
AFE Tx
并行到串行 212
编码 208
214
Lane 1
Lane 2
到传输介质
122
来自数据链路层
124

图 2

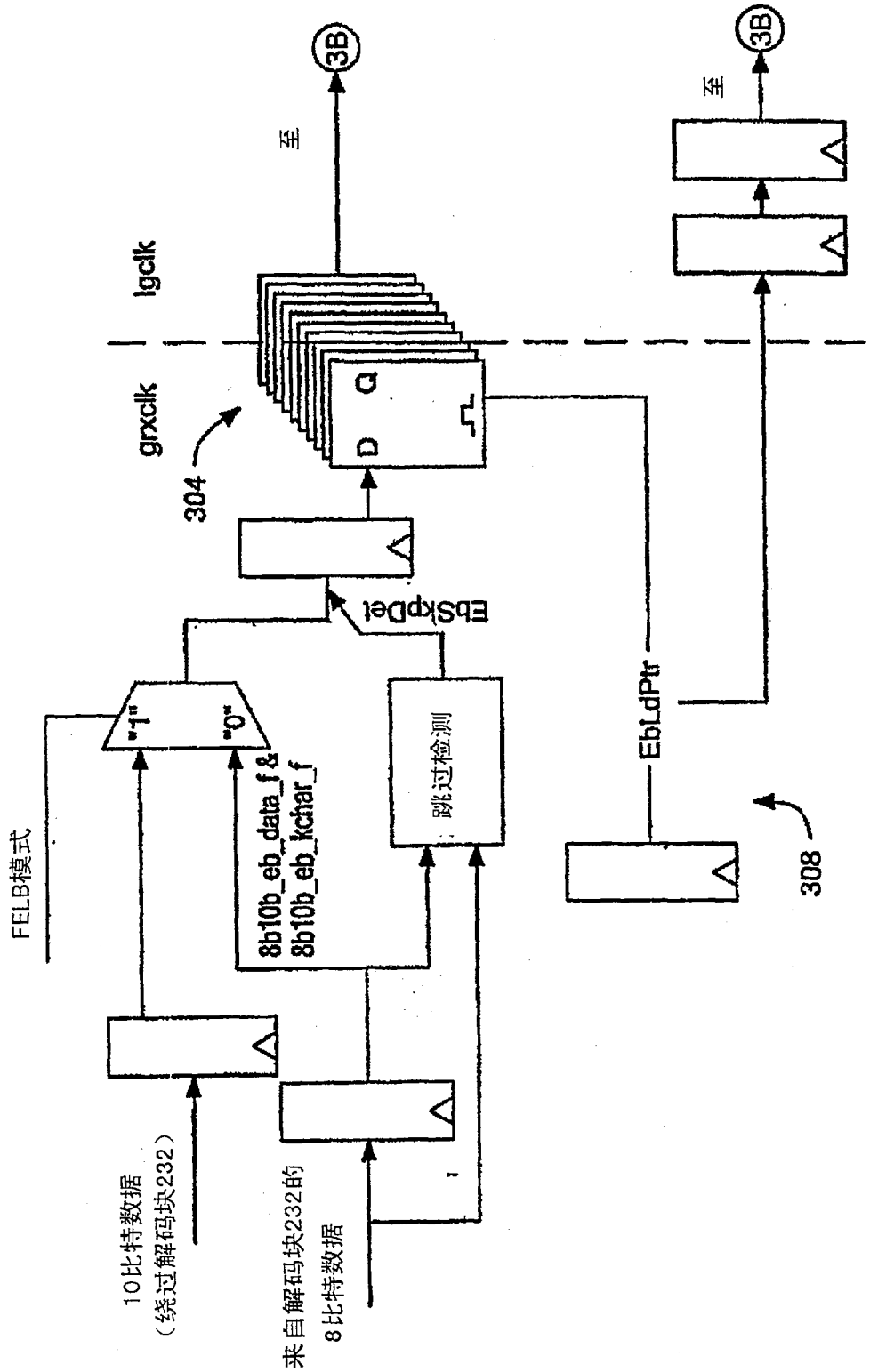

至
3B
至
3B
lgclk
grxclk
304
D
Q
EbSkpDel
"1"
"0"
FELB模式
8b10b_eb_data_f &
8b10b_eb_kchar_f
跳过检测
EbLdPtr
308
10比特数据
（绕过解码块232）
来自解码块232的
8比特数据

图 3A

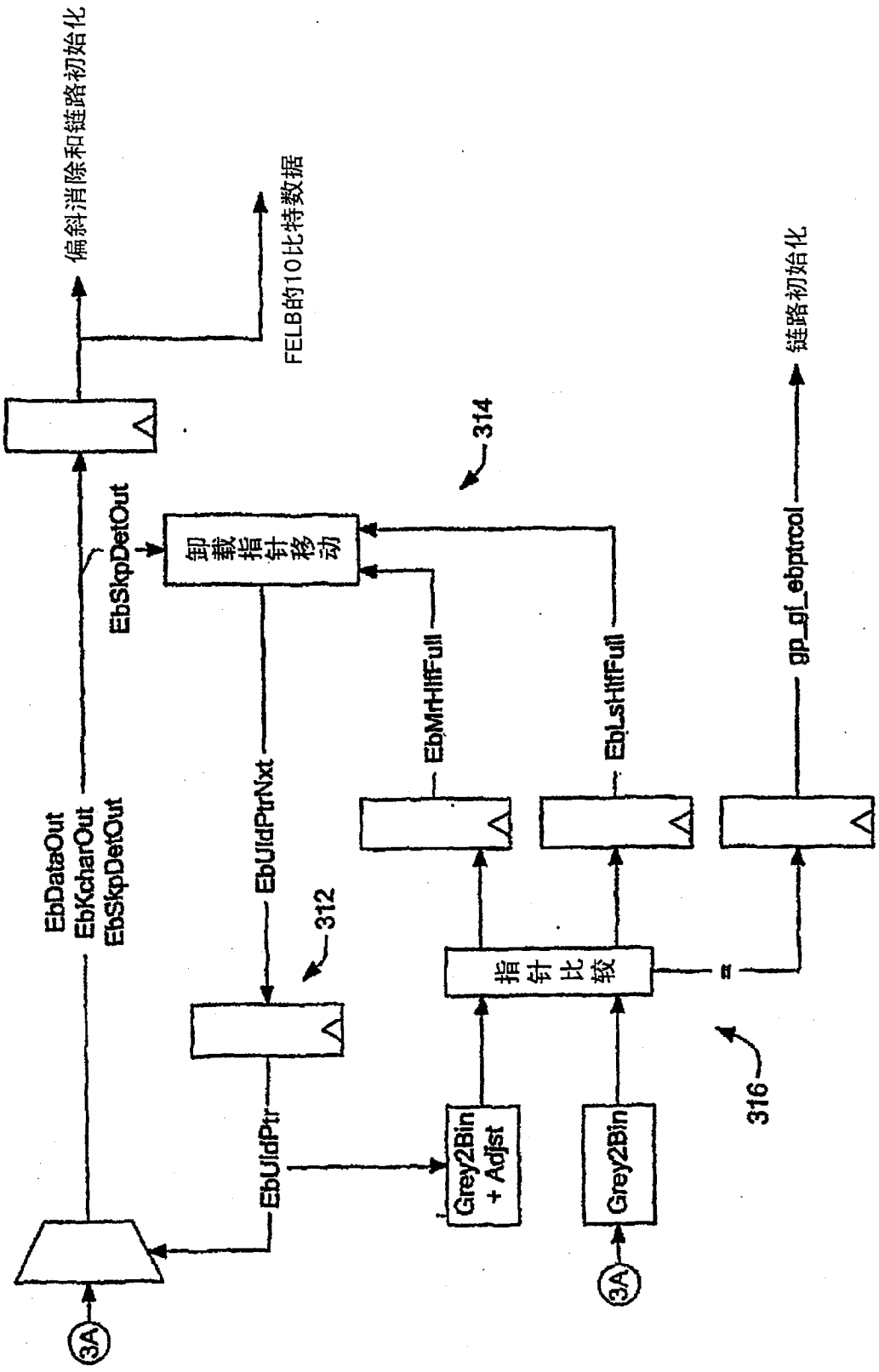

偏斜消除和链路初始化
FELB的10比特数据
314
卸载指针移动
EbSkpDetOut
EbMrHiffFull
EbLsHiffFull
gp_gf_ebptrcol
链路初始化
EbDataOut
EbKcharOut
EbSkpDetOut
EbUldPtrNxt
312
指针比较
=
316
EbUldPtr
Grey2Bin + Adjst
Grey2Bin
3A
3A

图 3B

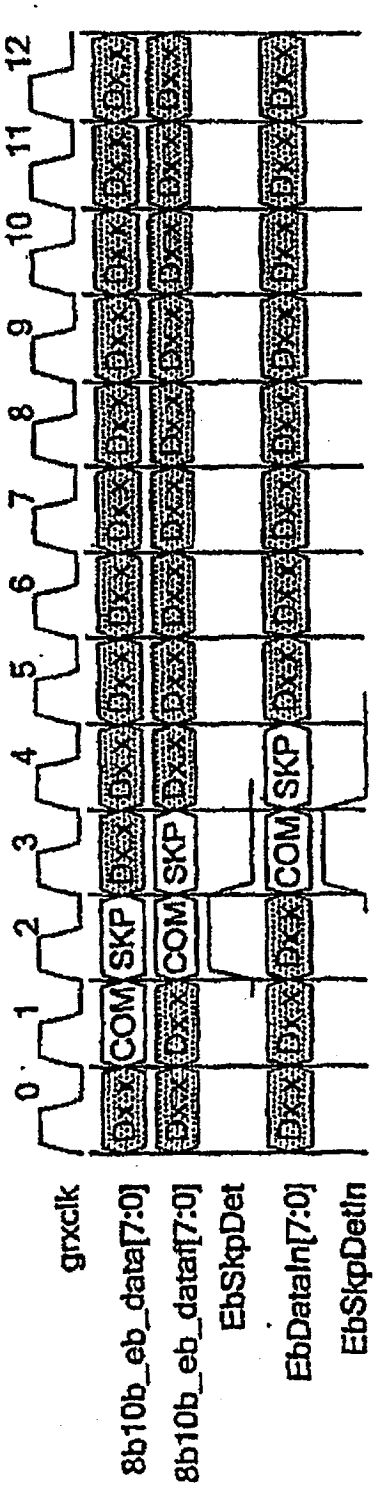

0
1
2
3
4
5
6
7
8
9
10
11
12
COM
SKP
grxclk
8b10b_eb_data[7:0]
8b10b_eb_dataff[7:0]
EbSkpDet
EbDataIn[7:0]
EbSkpDetIn

图 4

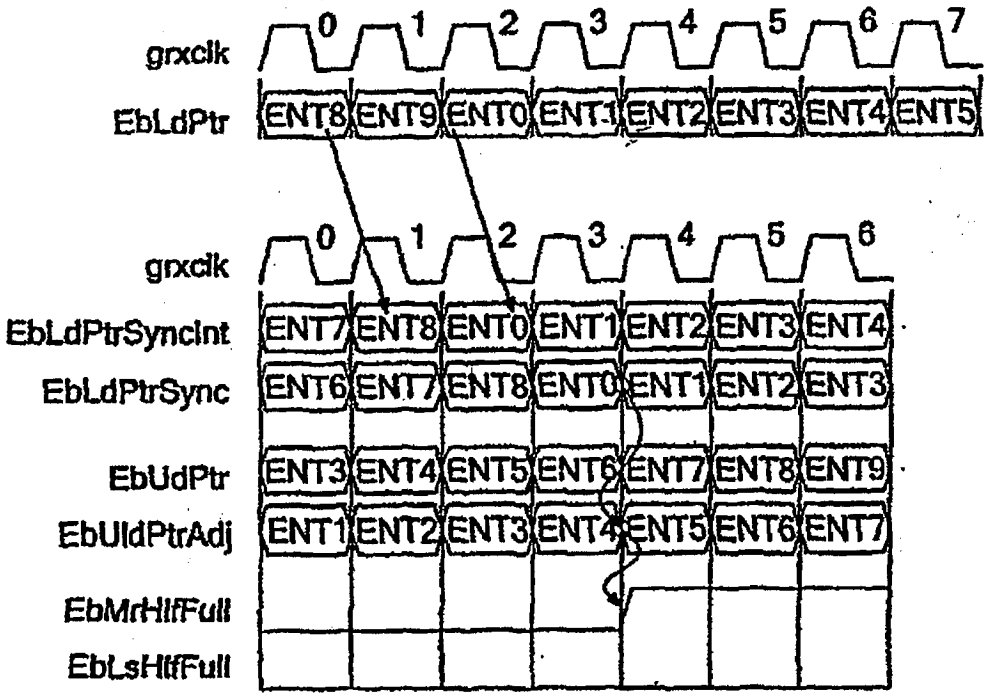

grxclk
0
1
2
3
4
5
6
7
EbLdPtr
ENT8
ENT9
ENT0
ENT1
ENT2
ENT3
ENT4
ENT5
grxclk
0
1
2
3
4
5
6
EbLdPtrSyncInt
ENT7
ENT8
ENT0
ENT1
ENT2
ENT3
ENT4
EbLdPtrSync
ENT6
ENT7
ENT8
ENT0
ENT1
ENT2
ENT3
EbUdPtr
ENT3
ENT4
ENT5
ENT6
ENT7
ENT8
ENT9
EbUldPtrAdj
ENT1
ENT2
ENT3
ENT4
ENT5
ENT6
ENT7
EbMrHlfFull
EbLsHlfFull

图 5

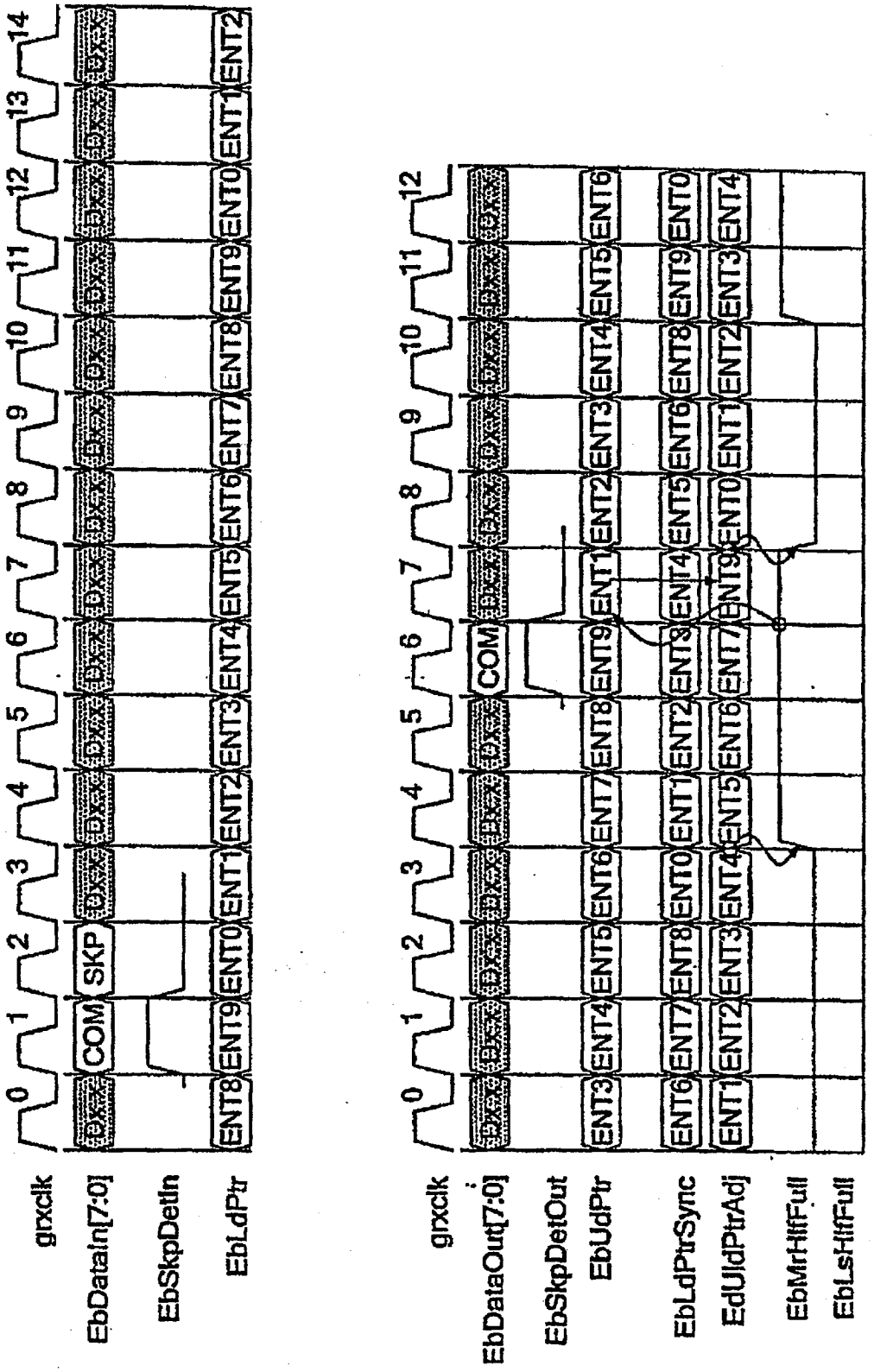

grxclk
EbDataIn[7:0]
EbSkpDetIn
EbLdPtr
EbDataOut[7:0]
EbSkpDetOut
EbUdPtr
EbLdPtrSync
EdUldPtrAdj
EbMrHlfFull
EbLsHlfFull
COM
SKP

图 6

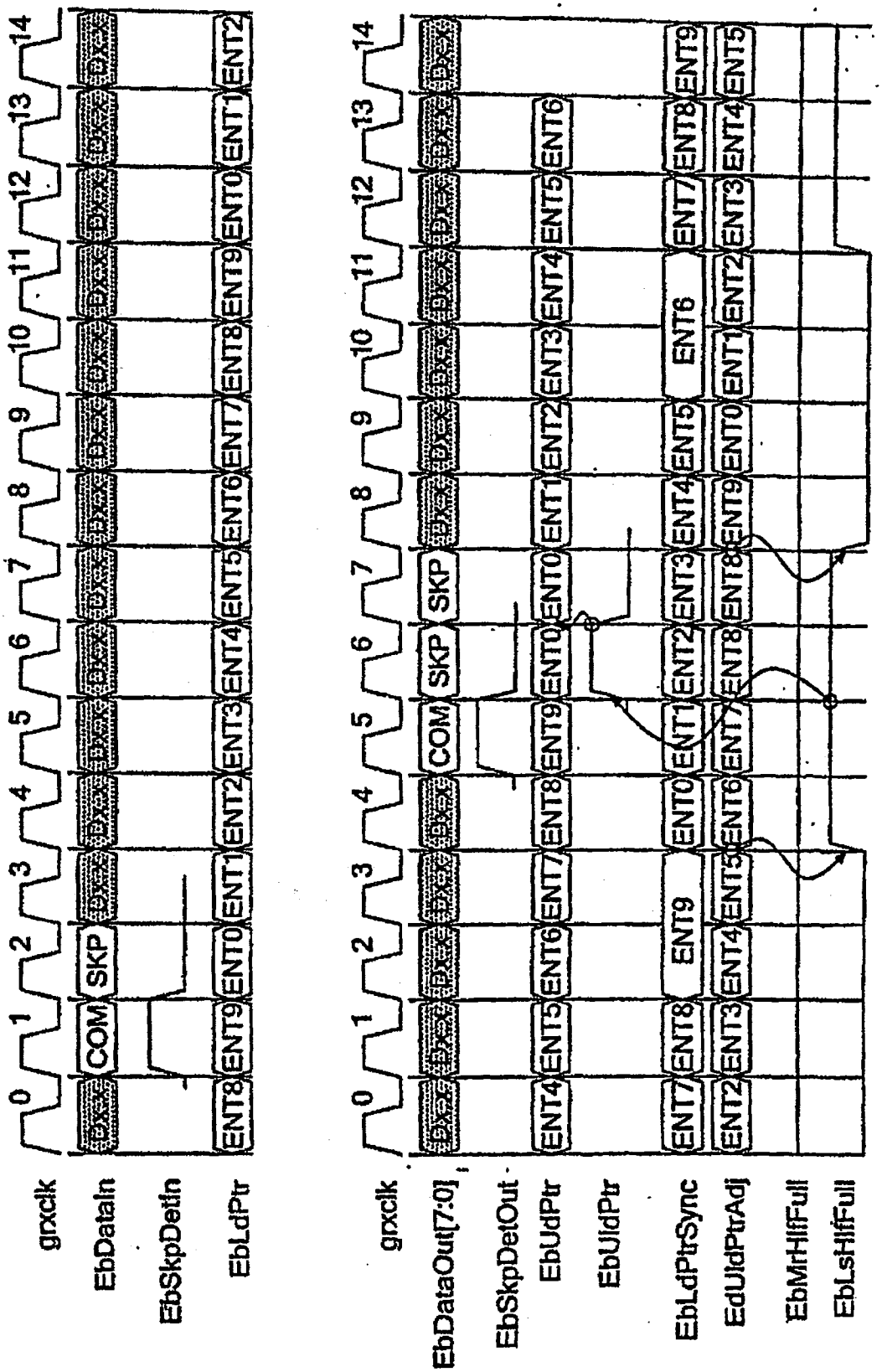

grxclk
EbDataIn
EbSkpDetIn
EbLdPtr
EbDataOut[7:0]
EbSkpDetOut
EbUdPtr
EbUldPtr
EbLdPtrSync
EdUldPtrAdj
EbMrHlfFull
EbLsHlfFull

图 7A

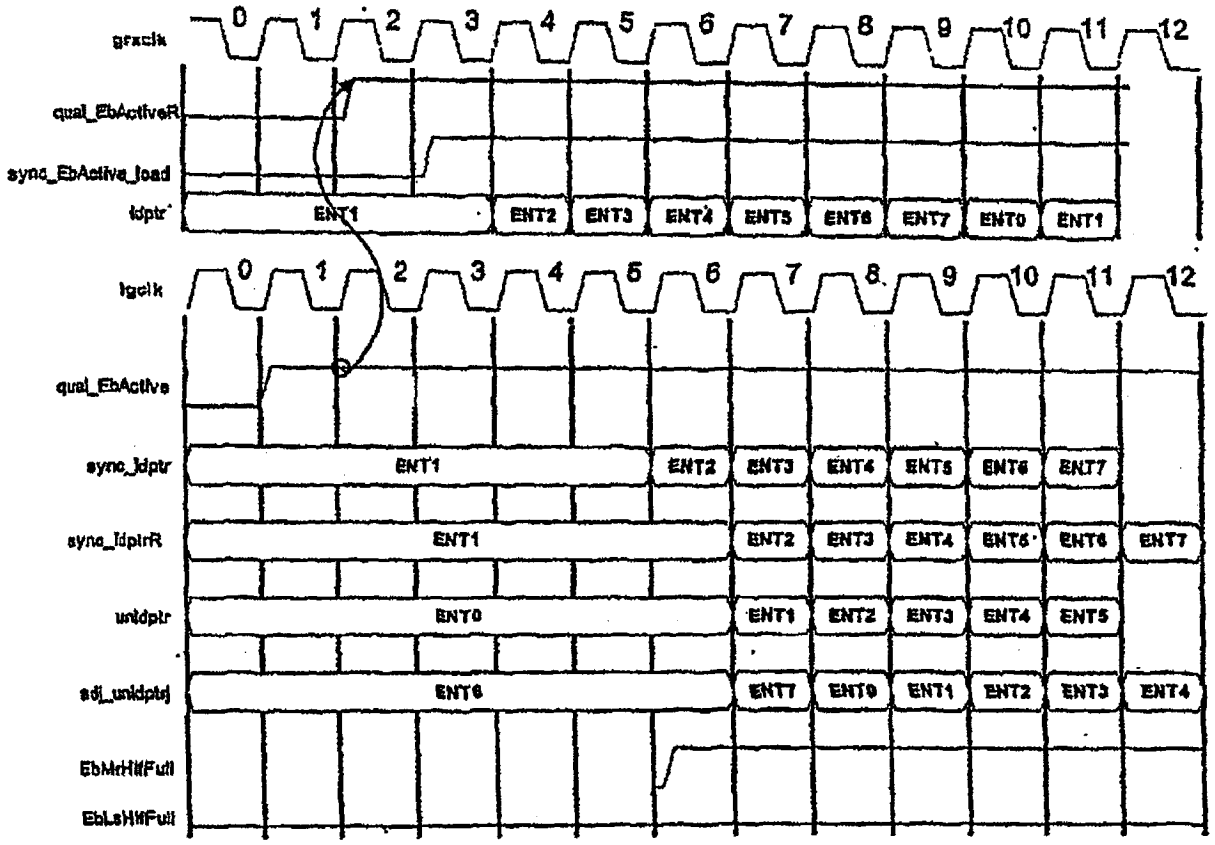

grxclk
qual_EbActiveR
sync_EbActive_load
ldptr
ENT1
ENT2
ENT3
ENT4
ENT5
ENT6
ENT7
ENT0
ENT1
lgclk
qual_EbActive
sync_ldptr
sync_ldptrR
unldptr
adj_unldptr
EbMrHlfFull
EbLsHlfFull

图 7B

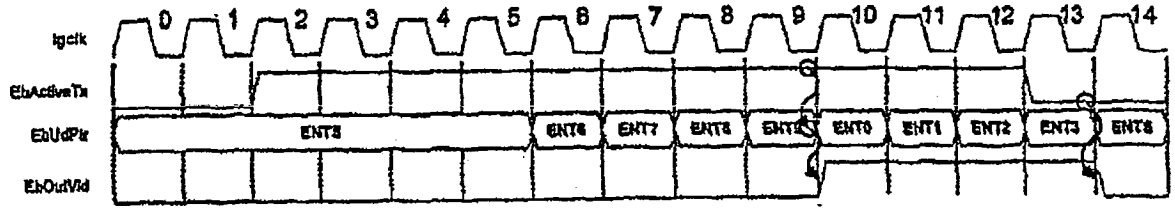

lgclk
EbActiveTx
EbUdPtr
EbOutVld
ENT8
ENT7
ENT0
ENT1
ENT2
ENT3

图 7C

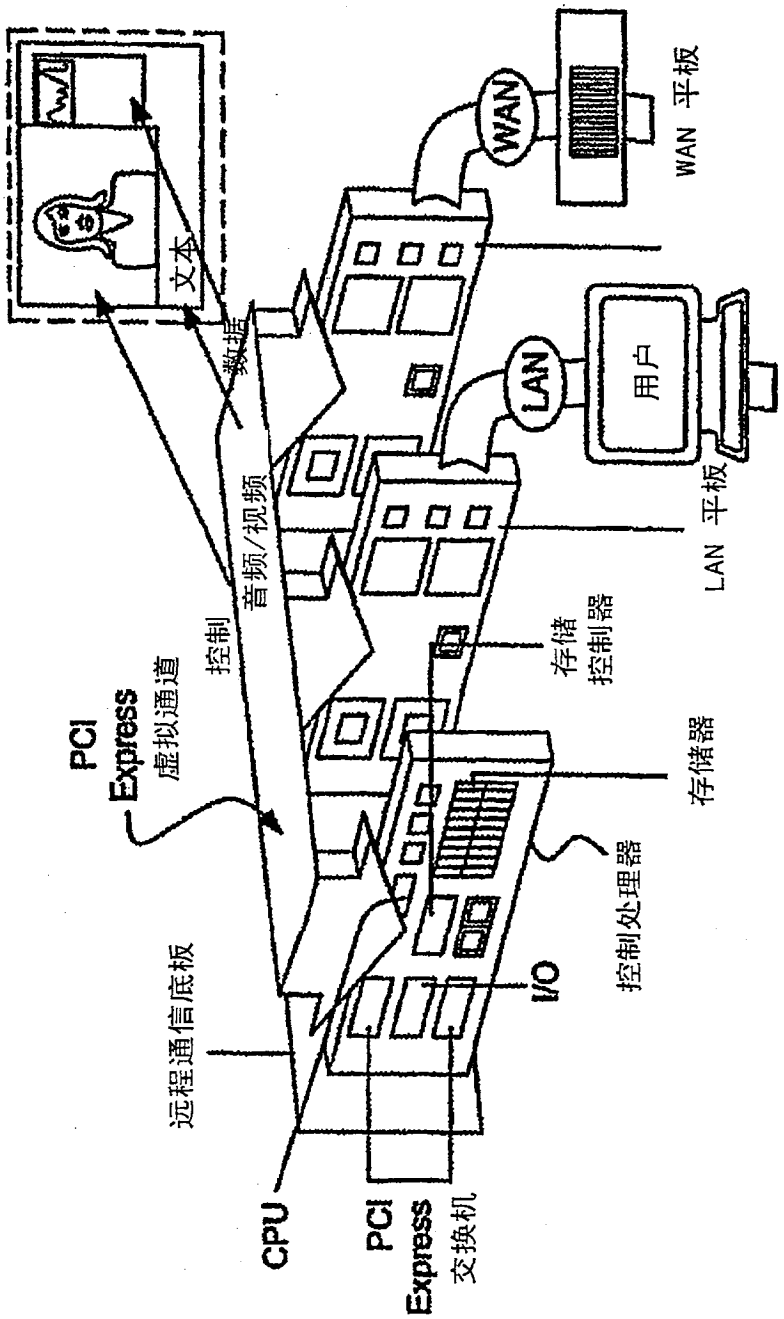

WAN
WAN 平板
文本
数据
LAN
用户
音频/视频
LAN 平板
控制
存储
控制器
PCI
Express
虚拟通道
存储器
控制处理器
I/O
远程通信底板
CPU
PCI
Express
交换机

图 8

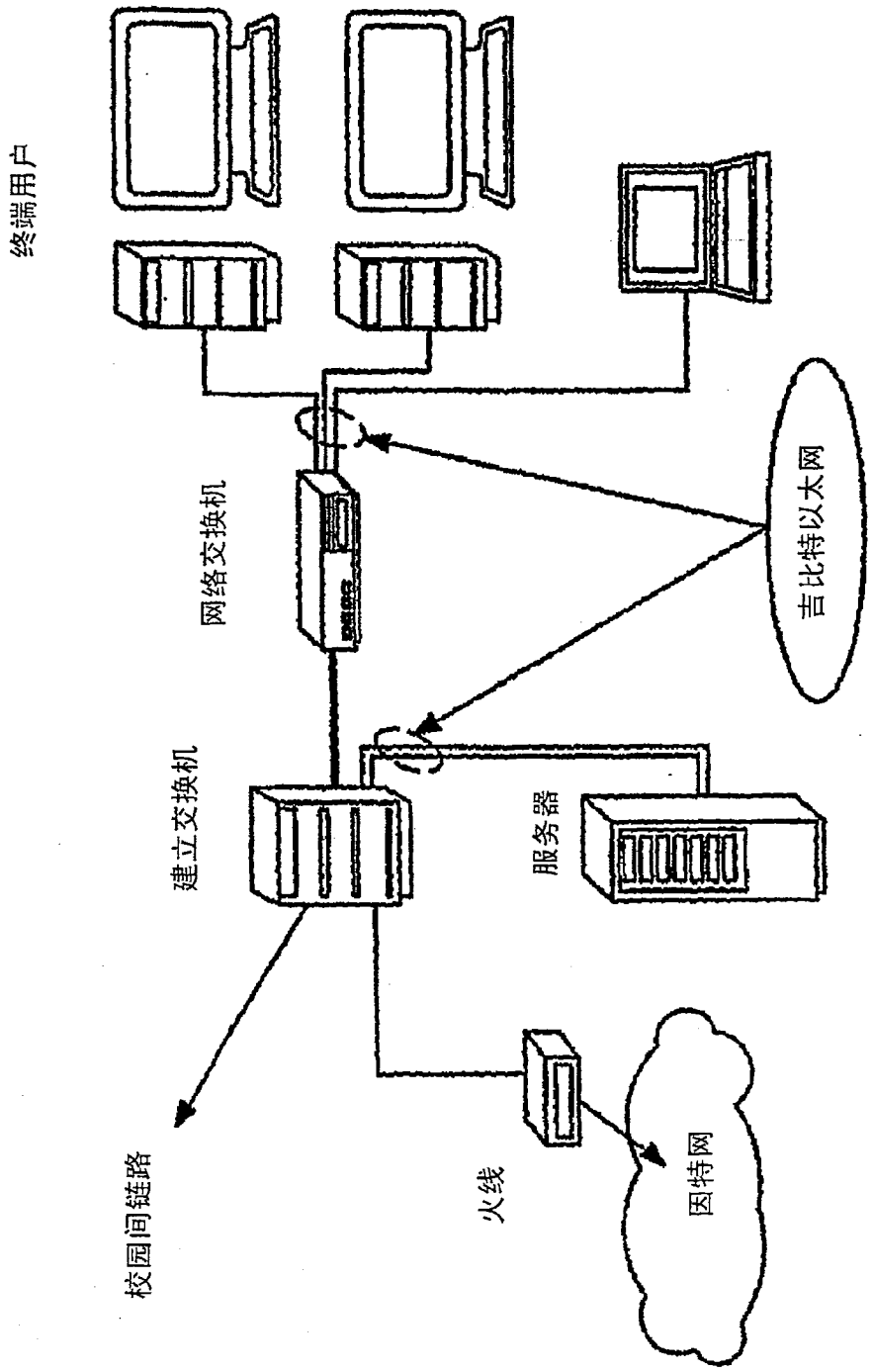


图 9