

PCIE2.0 的超远距离传输实现

袁 焱,李晋文,曹跃胜,胡 军
(国防科技大学 计算机学院,湖南 长沙 410073)

摘 要:PCIE2.0 作为用于芯片间和板间互连的、高性能、点对点、基于报文互换的新型 I/O 互连技术,已被公认为行业的标准,在计算机系统中得到了广泛应用。PCIE2.0 在物理层采用基于 SERDES 的串行通信技术,数据传输速率可达 5Gbps,最多支持 32 通道。随着信号频率的增加,信号完整性问题变得日益突出,衰减、串扰和抖动的共同作用导致信号严重失真,传输距离受到限制。采用一种高效能的中继芯片,对 PCIE2.0 总线高速串行信号进行中继,实现了远距离传输,并在实际系统中得到了验证。

关键词:PCIE2.0;信号完整性;中继

中图分类号:TN915

文献标识码:A

文章编号:1673-629X(2011)10-0150-04

Achieve Ultra-Long-Distance Transmission PCIE2.0

YUAN Yan, LI Jin-wen, CAO Yue-sheng, HU Jun

(School of Computer Science, National University of Defense Technology, Changsha 410073, China)

Abstract:PCIE2.0 as between chips and boards used for interconnects, high-performance, point-based message exchange in the new I/O interconnect technology, has been recognized as the industry standard in the computer system has been widely used. PCIE2.0 at the physical layer of serial communication based SERDES technology, data transmission rates up to 5Gbps, the total number of channels up to 32 channels. With the increase in signal frequency, signal integrity issues become increasingly prominent, attenuation, crosstalk and jitter jointly lead to serious distortion of the signal transmission distance is limited. A high-performance relay chip, high-speed serial bus for PCIE2.0 to relay the signal to achieve the long-distance transmission, and in the actual system has been verified.

Key words:PCIE2.0; signal integrity; repeater

0 引言

PCI Express 作为面向芯片间和板间互连的、高性能、点对点、基于报文交换的新型 I/O 互连技术,已被公认为行业的标准^[1~3]。在数据传输模式上,PCIE 总线采用独特的双通道传输模式,即独立的发送和接收,与全双工模式相似,数据传输速度得到明显提高。在传输速度上,PCI Express2.0 采用基于 SERDES 的串行双向通信技术,单 lane 数据传输速率可达 5Gbps,通道数最多可达 32 lane,支持芯片与芯片和背板与背板之间的通信。

由于信号频率的增加,信号完整性问题变得日益突出^[4,5],衰减、串扰和抖动的共同作用导致信号严重失真,传输距离受到限制。PCIE2.0 物理层的高速串行 SERDES 接口采取了多项措施来解决传输距离的

问题^[6]。首先通过用差分信号传输来代替单端信号传输,达到增强抗噪声、抗干扰能力的效果;其次运用时钟和数据恢复技术取代同时传输数据和时钟,解决限制数据传输速率的信号时钟偏移问题;最后编码技术被用来降低信号失真,提高信号抗干扰能力;预加重(pre-emphasis)和均衡(equalization)技术^[7]补偿传输介质的码间干扰(ISI)和高频损耗,高速串行信号的恢复与传输能力得到增强。预加重是克服传输线损耗的有效手段,不过较高的驱动强度将产生电磁干扰(EMI),并且会使系统容易遭受近场的串扰。张开接收器眼图的一种替代方案或互补方案是利用接收器均衡技术。然而即使采用了以上技术,信号的传输距离也超不过 20inch。文中使用中继芯片 GN1406,首先介绍了它的基本功能,对在高速串行链路中使用 GN1406 对信号中继做了仿真实验,信号的传输距离可达 40inch,并在真实系统中得到了成功应用。

1 SERDES 接口

随着对信息流量需求的不断增长,传统的并行接口技术成为进一步提高数据传输速率的瓶颈,而基于

收稿日期:2011-03-08;修回日期:2011-06-18

基金项目:国家自然科学基金(60873212)

作者简介:袁 焱(1982-),男,硕士研究生,研究方向为高速信号仿真;李晋文,副研究员,博士,研究方向为计算机系统结构、高速信号传输。

SERDES 的串行通信技术逐步取代传统并行总线。基于 SERDES 技术的高速串行接口正在成为一种通用的 IO 接口标准,从而取代传统并行总线而成为高速接口技术的主流,近年来世界上有多个标准组织已经或正在制订从 1G 到 10G 的高速串行接口标准。SERDES 产品为当前高速串行接口行业标准,其中 2.5G/3.125G 为第一代产品,5G/6.25G 为第二代产品^[8]。

SERDES(SERializer(串行器)/DESerializer(解串行器)的简称)。是一种时分多路复用(TDM)、点对点的通信技术,即在发送端多路低速并行信号被转换成高速串行信号,经过传输媒体(光缆或铜线),最后在接收端高速串行信号重新转换成低速并行信号,从而完成总线数据的传输。这种点对点的串行通信技术充分利用传输媒体的信道容量,减少所需的传输信道和器件引脚数目,从而大大降低通信成本。

PCIE 物理层使用了 SERDES 技术,具体结构由逻辑物理层和电气物理层组成,两个子部分都为相互间独立的发送逻辑和接收逻辑,从而能够进行双单工通信^[9]。物理层的逻辑和电气子部分结构如图 1 所示。

逻辑子层与数据链路层进行数据交换,由发送逻辑 TX 和接收逻辑 RX 组成。

●发送过程为:

(1)物理层从数据链路层获得 TLP 或者 DLLP,定时后送入发送缓冲区。基于多路复用器,物理层使用开始(Start)和结束(End)字符将 TLP 和 DLLP 封装成帧。这些字符是成帧字符,接收器设备用它来检测数据包的开始和结束。

(2)已成帧的数据包通过多路复用器被发送到字节拆分逻辑,PCIE 链路可以由多个通道组成,该逻辑将数据包的各个字节多路复用到各条通道上。在 PCIE 链路的不同 Lane 中传递的数据可能存在偏斜,字节拆分逻辑还有一个重要功能即消除这个偏斜。

(3)数据进入到各自通道的加扰部件,“加扰”后进行 8b/10b 编码,最后通过并转串逻辑将数据发送到 PCIe 链路中。

●接收过程为:

(1)物理层从 PCIe 链路的各个通道获得串行数据,并通过 8b/10b 解码和解扰部件,发送到字节反拆分逻辑。

(2)字节反拆分逻辑将来自不同通道的数据合并,进行 De-skew 操作,然后取出物理层的前后缀并进行边界检查后,将数据放入接收缓冲区中。

(3)物理层将在接收缓冲区中的数据传递到数据链路层。

2 电气物理层

每条通道的电气子部分提供到链路的物理接口。该子部分包含差动驱动器和差动接收器。发送器串行化每条通道上的电气信号并生成串行比特流,将其解串行化为符号,然后将该符号流以及从接收串行比特流中所恢复的时钟提供给逻辑物理层。

3 高速信号多板传输

多板系统^[10]一般由多个子板通过互连背板组合而成,本文研究基于背板构建的高性能多板系统,具体

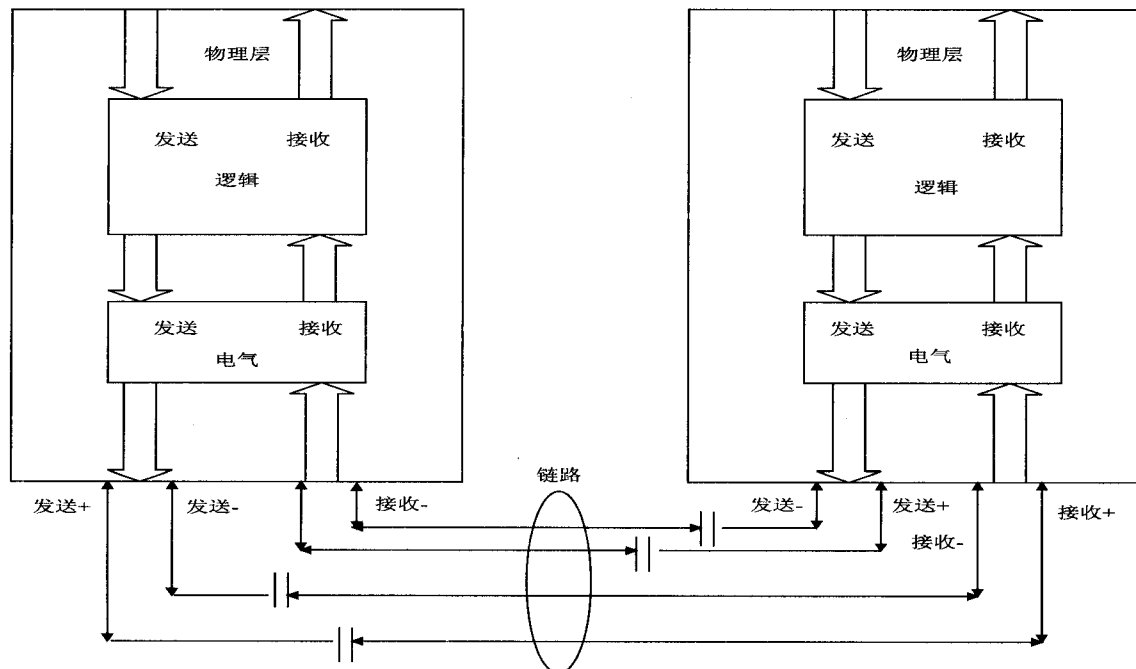


图 1 物理层的逻辑和电气子部分

结构如图 2 所示,系统由页交换板、电源模块、背板和子板组成。当高速差分串行信号跨越两个(或以上)PCB 板时,传输距离远,信号速率快。5Gbps 速率的高速串行差分信号对多板传输链路上的各种要素提出了更高的性能要求,包括器件封装、连接器、传输线、过孔等。精确的要素分析与建模,串扰、衰耗、反射和延时等信号完整性问题的解决,是多板 SERDES 传输能否高速率和高可靠的实现前提。

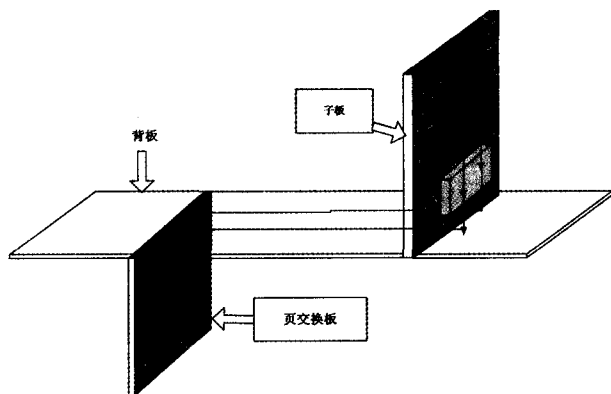


图 2 基于背板的多板系统

在以上所示的系统中,根据仔细的计算,PCIE2.0 总线信号最远需要传输的距离达到 1m 左右。首先使用 Ansoft 公司的 Siwave^[11]对子板、背板和页交换板分别提取了互连线的 S 参数,然后使用 Synopsys 公司的 hspice 工具对全链路进行了仿真,链路由“发送端 serdes-子板-连接器-页交换板-连接器-接收端 serdes”构成,仿真频率设为 2.5GHz。

仿真结果如图 3 所示,发现眼图已经全部闭合,显然无法实现远距离传输。

(BER)来衡量链路传输的可靠性。误码率是最终的测量结果,眼图可分析形成误码的原因。要有效地提高链路的传输性能,在工程实现上,将面临许多严重的技术挑战。其中,影响最大的是频率相关的传输损耗。传输损耗会直接减少接收端眼图的高度,降低噪声容限;而由传输损耗产生的符号间干扰(ISI),也会减少眼图的宽度,降低时序容限;从而导致数据传输误码率的上升,降低链路传输的可靠性。在文中的多板系统中,各板传输线和多个连接器的损耗以及信号的抖动是导致眼图闭上的主要原因,为实现远距离的有效传输,必须使用中继芯片。



图 3 仿真眼图

GN1406 是 Gennum 公司研制的一款中继芯片^[12],它完成物理链路的连接,对衰减的信号进行放大,实现发送数据的恢复与高保真。中继芯片主要完成 PCIE2.0 物理层的功能,负责将收发节点的物理层按位传递信息,完成信号的复制、调整和放大功能,以此来延长高速信号传输距离。该芯片可支持 2.5Gbps、3.125Gbps、5.0Gbps 和 6.25Gbps 四种速率的通信,扩展了在传输介质损耗和重置抖动影响下的高速串行链路的到达和稳定性。

GN1406 的内部结构如图 4 所示,每个芯片包含 4 个互相独立的通道,每个高速输入接口包括一个集成

4 高速串行信号的中继

在高速串行链路设计中,一般采用眼图和误码率

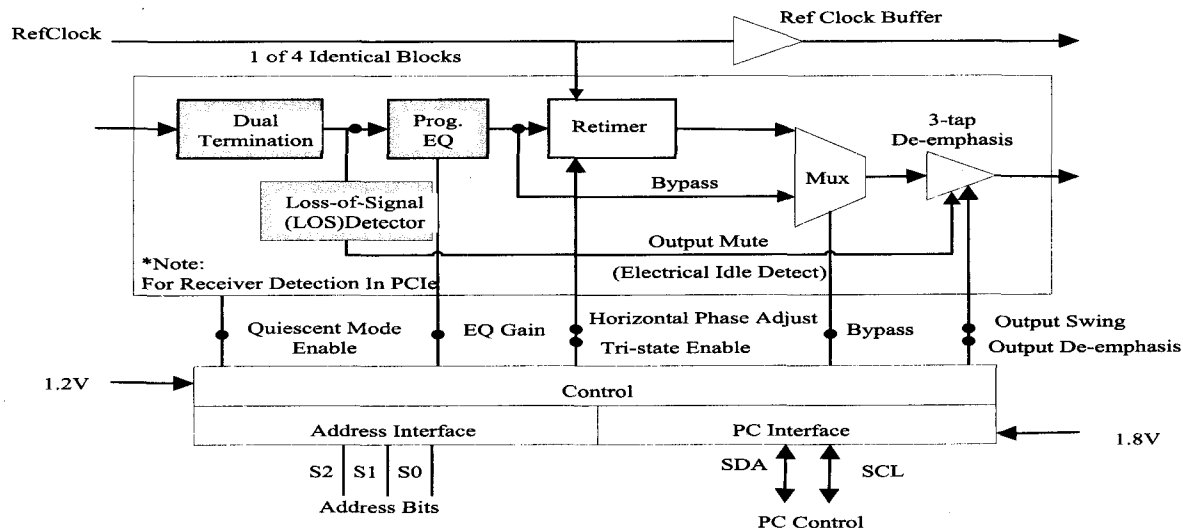


图 4 GN1406 芯片内部逻辑结构图

的高输入可调均衡抖动容忍重定时器电路。该 GN1406 芯片高速串行输出接口具有非常低的抖动和幅度的调整,通过一个 I2C 接口来静态或运行时配置,静态配置方式上电时从片外 EEPROM 读取。去加重水平也可以通过设置来补偿传输过程中的损耗。GN1406 芯片中的每个通道都可以独立配置,甚至支持独立关断某个通道以实现物理链路层的低功耗。

GN1406 芯片具有稳定高效的性能,也具备了支持从 1.25Gbps 至 8.0Gbps 任何数据传输速率的中继操作模式。PCI Express 体系要求链路必须完全透明,并非所有的均衡器或中继器支持这些功能,GN1406 芯片放置在发送和接收装置的中间位置,满足对 PCIE2.0 协议透明的要求,从而确保 PCIE 链路正常运行。电气空闲保护同样也是 PCI Express 协议中的一项关键要求,发射器在休眠和空闲状态下都有相同的差分输出电平,GN1406 通过快速检测其接收输入的电气状态并正确地判断其输出处于空闲状态,并将结果写入信号丢失(LOS)寄存器。

5 模拟与实际系统验证

在系统的 PCIE 链路的发端和收端都使用中继芯片,具体结构如图 5 所示。

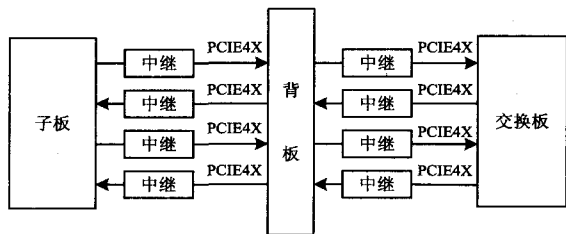


图 5 基于中继的高速链路

传输距离可达 1.1m,通过全链路仿真来评估可行性,仿真结果如图 6 所示。

仿真实验中,发送端预加重参数设置:Ob_preemp_main[4:0] 为 11101,Ob_preemp_pre[4:0] 为 00011,Ob_preemp_post[4:0] 为 00011,在发送端的眼高 585mV,抖动 24ps,而到达接收端眼高 226mV,抖动 95ps。仿真结果表明,通过在链路中使用两级中继,眼图已经睁开了,达到了预期的目标。

6 结束语

由于日益严重的信号完整性问题,给高速信号的远距离特别是跨背板传输带来了相当大的挑战,通过采用基于 GN1406 的两级中继,实现了 PCIE2.0 5Gbps 信号的超远距离传输,距离可达 1m 以上。该技术已在实际系统中得到了应用,实测误码率几乎为 0。

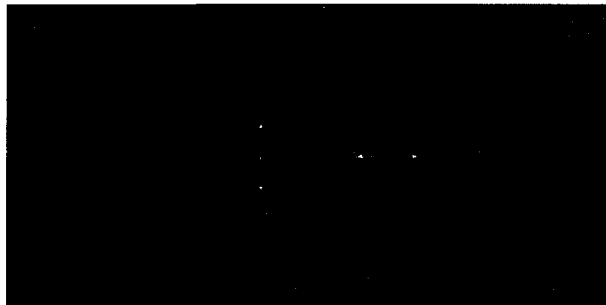


图 6 中继后的眼图

参考文献:

- [1] Budruk R, Anderson D, Shanley T. PCI Express 系统体系结构标准教材[M]. 北京:电子工业出版社,2005.
- [2] 马鸣锦,朱剑冰,何红旗,等. PCI、PCI-X 和 PCI Express 的原理及体系结构[M]. 北京:清华大学出版社,2007.
- [3] 许 军,李玉山,贺占庄,等. PCI-Express 总线技术研究[J]. 计算机工程与科学,2006,28(5):141-143.
- [4] Bogatin E. Signal Integrity: Simplified[M]. [s. l.]: Prentice Hall PTR,2003.
- [5] Johnson H, Graham M. 高速数字设计[M]. 北京:电子工业出版社,2004.
- [6] 卓 沛,严国萍. 高速数字系统中的时序分析与设计[J]. 计算机技术与发展,2007,17(7):171-174.
- [7] 赖万玖. 预加重和均衡技术在高速背板设计中的应用[J]. 电子产品世界,2008(11):100-101.
- [8] 曹跃胜,胡 军,刘焯铭. 高速 SERDES 的多板传输技术与 SI 仿真[J]. 计算机工程与科学,2008,30(8):139-143.
- [9] 王 齐. PCI Express 体系结构导读[M]. 北京:机械工业出版社,2010.
- [10] 胡 军,李晋文,曹跃胜. Gbps 串行链路信号完整性分析与设计[J]. 计算机工程与科学,2009,31(7):102-105.
- [11] user's guide-SIwave[M]. [s. l.]: Ansoft Corporation,2009.
- [12] Preliminary Data Sheet-GN1406 Quad Multi-rate Repeater/Retimer[M]. [s. l.]: Gennum Corporation,2009.

(上接第 149 页)

- [7] 刘 静,耿国华. 基于 USB2.0 的高速大容量数据采集存储系统[J]. 计算机技术与发展,2011(2):38-40.
- [8] Counsell S, Loizou G, Najjar R. Quality of manual data collection in Java software[J]. An Empirical Investigation,2007,12(3):33-35.
- [9] Zhao Jianjun. Static analysis of Java bytecode[J]. Wuhan University Journal of Natural Sciences,2001,6(1):45-49.

- [10] Chen Zhirui, Tan Hongzhou. Logic structure of programmable instructions for Java[J]. Journal of Electronics (China), 2009,26(5):44-46.
- [11] 李 刚. 疯狂 Java 讲义[M]. 北京:电子工业出版社,2010:33-56.
- [12] Little G, Miller R C. Keyword programming in Java[J]. Automated Software Engineering,2009(3):32-35.