

计算机 I/O 总线结构的现状与发展趋势^{*)}

赵 聪 董小社 乔 楠 黄泳翔

(西安交通大学计算机科学与技术系 西安710049)

Current Situation and Development Trend of Computer I/O Bus Architecture

ZHAO Cong DONG Xiao-She QIAO Nan HUANG Yong-Xiang

(Xi'an Jiaotong University, Xi'an 710049, China)

Abstract With the improvement of the performance of all components of computers, Shared Bus, which is the kernel of the traditional computer architecture, has gradually become the bottleneck of the system. This paper presents the limitations of the Shared Bus. From analyzing the new specifications of the bus, which are used to solve the bottleneck brought by shared bus, it sums up the characteristics of the new technologies from their orientations in the level of the architecture. Then by analyzing and comparing the advantages of each new technology, this paper points out the new orientation of the bus technology.

Keywords Bus technology, 3GIO, Infiniband, Switched network, Point to point

1 引言

计算机的总体性能主要依赖于它的五个主要组成部分的性能,即处理器、内存、总线、存储器和显示子系统等。在过去十多年里,计算机的各个组成部分的性能有了不同程度的提高,尤其是作为处理核心的微处理器的性能更是得到了长足的发展;但是,整个计算机的体系结构仍基本沿用了基于共享的并行总线技术。

这种共享总线结构有两个明显的不足:其一,限定了所有的设备只能分享指定的带宽,从而连接的设备越多,争用总线的冲突就越多,每个设备所能得到的带宽就越小;另一个不足就是其“外挂”能力,总线所能挂接外设的最大数目在总线设计时就已经被限定,因而限制了系统的可扩展性。

因此从PC/XT、ISA到PCI,I/O总线的性能越来越成为提高整个计算机系统性能的瓶颈。

为了解决这一日益突出的问题,近期业界在这一技术领域先后提出了一些新的标准,主要用于取代目前已经明显不能适应各个方面要求的PCI总线技术。

2 共享总线技术

2.1 PCI共享总线技术及其局限性

共享总线结构一直是传统计算机体系结构的核心,其中PCI总线无疑是共享总线技术中应用最广泛的一种。PCI总线是由Intel公司推出的一种局部总线,它定义了32位数据总线,可扩展到64位,替代了早期的ISA、EISA和VESA总线,主要用于连接外围附属设备。

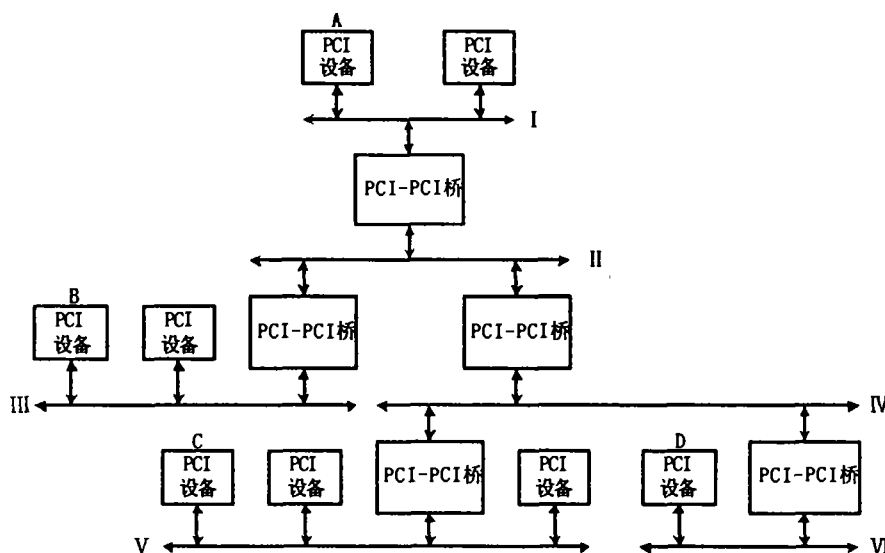


图1 PCI总线系统的分级树结构

PCI总线支持突发读写操作,可同时支持多组外围设备,但是在已有设备占用总线,PCI的共享总线结构要求其他连接的设备等待,也就是说,各个连在总线上的设备要

竞争使用总线,在某一时刻使用总线的只能是一个设备。因此如果有多个设备的话,总线的实际速率必然会下降。而且PCI总线结构不具备很好的容错性能,当总线的某处发生故障时,

^{*)}本文得到教育部留学回国人员基金资助。

将会导致整个系统崩溃,而不能做到将故障隔离^[1]。

PCI 总线的扩展采用的是一种分级树的结构,如图1所示,各级之间通过 PCI-PCI 桥连接。这种分级 PCI 总线使用时会存在下面的问题:①等待时间在分级树中被加长,例如,当 A 设备需要和 D 设备进行通信时,则需要 I、II、IV、V 这四个总线段上完成仲裁;②总线争用加剧,例如,A 和 D 两个设备进行通信时,即使 B 设备和 C 设备所在的总线段 III 和 V 是空闲的,它们也不能参与通信。

2.2 PCI-X 总线技术

PCI-X 总线规范是在与现有 PCI 产品兼容的基础上产生的一种过渡技术,它在一定程度上解决了 PCI 总线的瓶颈问题,如加快加宽了 PCI 芯片组的时钟频率和吞吐量,使时钟频率达到 133MHz,数据采用 64bit 位,总带宽可以达到 1066MB/s;增加了一些智能技术,可以自动去除没有任何数据传输的 PCI-X 设备,以减少设备间的等待周期(在相同的频率下,PCI-X 可以提供比 PCI 高 14-35% 的性能);具有可扩展的频率,一般 PCI 总线的频率是固定的,而 PCI-X 的频率可以随着设备的变化而进行相应调整。但 PCI-X 总线只是 PCI 总线的一种升级,它仍然是基于共享的并行总线结构,性能上无法有质的飞跃^[2]。

传统的基于共享的并行总线提高性能的方法基本上有两种:①依靠增加数据总线的数量:从早期的 8 位、16 位到现今的 32 位、64 位,总线的性能呈倍数增长,只是若再增加总线的数量,就会使总线占据的空间变大,其实现的难度和成本将无法接受;②依靠提高总线的时钟频率:这种并行共享总线是在电路板级实现芯片间的铜布线连接,要提高总线的时钟频率,就要缩短信号的有效距离。当这个距离缩短到以厘米计时,主板的设计将变得十分复杂,随着设备的增加,信号周期的判定和终结也会变得越来越困难,直接影响到了系统的性能和稳定性。也就是说,在 64 位的数据带宽上,133MHz 已经接近了实际应用的极限。因此传统的并行共享总线正在接近它们性能上的极限,现在的 PCI-X 将成为这种传统并行共享总线的最终标准。

3 定位于计算机内部互联的总线标准

3.1 新型 3GIO 总线标准的提出

Intel 公司提出了一种全新的总线技术标准 3GIO (Third Generation Input/Output, 第三代输入输出),不久前经国际 PCI 标准化组织 PCI-SIG 审议通过,被确定为下一代 PCI 总线技术,以后可能命名为 PCI 3.0。3GIO 是一种串行的 PCI 总线技术,它减少了总线的引脚数目,采用点对点的交换结构进行 I/O 互连,其优势在于可以并行地进行 I/O 处理,改变了以往 PCI 总线在一个时钟周期内只能执行一个 I/O 操作的弊端,另外,串行连接可以以很少的信号带来很高的带宽,而且不会像并行连接那样容易受到道间串扰(crosstalk)、电磁干扰或是电容性等问题的影响,而且实现起来成本也相对较低。

3.2 3GIO 总线标准的关键特点

1) 点对点架构:3GIO 采用点对点技术,能够为每一个设备分配一个独享的传输通道,而不需要在设备之间竞争资源,这些高速接入的设备由一个交换器互连起来,交换器作为控制单元,主要是对设备之间的点对点通信进行管理和控制。

2) 多平台互连:3GIO 不仅可以作为系统的内部的连接总线,而且可以作为外部设备与系统连接的总线。通过 3GIO 线缆可以直接将各种外设与系统内的 3GIO 总线连接在一起,这样就可以设计一些与主机系统脱离的高性能的存储控制

器。

3) 向下兼容:符合 PCI 规范的板卡将可以在低带宽的 3GIO 插槽上使用,现有操作系统无须改变就可以启动,与目前 PCI 设备的驱动接口和设置相兼容。

4) 其它:3GIO 还具有 QoS 属性、热插拔和热交换属性,以及电源管理和监控能力^[3]。

3.3 3GIO 总线标准的结构分析

3GIO 采用的是一种点对点的交换式结构,设备之间是通过交叉开关进行互连,可以同时进行数据交换,保证了设备间通信所需的带宽,不存在传统共享总线的设备间分享带宽的问题。这种新的总线技术与共享总线相比,除了在结构上有着质的区别以外,其总线数据的通信形式也不一样,它们采用的是包结构的数据,包中包括数据和路由两种信息。

设计时为了保持兼容性,整个系统就可以看作是用串行连结和交换机实现的 PCI 总线,每一个串行的点对点连接就像是一个 PCI-PCI 桥。为了使这种新的结构对系统软件透明,需要交叉开关保持 PCI 总线结构的分级树概念,以及上游信号(从叶子到根)和下游信号(从根到叶子)的概念,但是它打破了交换结构的通常的对称性,因为交换端口必须被分开定义为上游端口和下游端口,这样使交叉开关执行起来类似于分级树,而在上游端口和下游端口之间就需要有不同的行为和特征,而且在物理实现上两种端口也会有区别,否则互连将不可避免地会出现问题,这里虽然建立的是交叉开关结构,但逻辑结构上仍然是一棵树,适合于单 CPU 控制 I/O 设备。因此 3GIO 是定位于计算机内部互联的总线标准,并不十分适合点到点的外部通信。

3.4 其它的新型总线标准

AMD 公司提出了一种 HyperTransport (HT) 总线技术,它和 3GIO 的基本设计思路类似,也是一种点对点总线技术,允许在两个外设之间直接交换数据而无需计算机处理器的参与。与 3GIO 不同,HyperTransport 是在同一个并行总线中模拟出两个独立数据链进行数据传输,而 3GIO 则是在物理上的串行独立数据链上进行数据传输;另外 HyperTransport 允许以不对称方式与外设之间进行数据传输,比如当系统显示图形信息的时候或者从网上下载大量的数据的时候,都是数据的非对称传输;HyperTransport 定位于多处理器互连总线以及核心逻辑与各种不同 I/O 控制器之间的连接扩展总线,并非像 3GIO 是要作为现有总线的一种升级替换^[4]。

还有一种新的标准,是 Motorola 等公司提出的 RapidIO,它也是基于点对点的交换结构,由于它主要应用于嵌入式网络和通讯市场,因此这里不再讨论。

4 定位于计算机外部连接的 InfiniBand 标准

4.1 InfiniBand 标准的提出

InfiniBand (IB) 是一种用于服务器与 I/O 以及服务器之间的通信的工业标准,它是在 NGIO 以及 FutureIO 两个总线标准的基础上提出来的,是计算机 I/O 连接的一种新的替换架构,在低端功能方面它与 3GIO 有一些重叠之处,但 InfiniBand 主要是用于连结服务器集群等高端任务,应用于大型数据中心。

InfiniBand 结构不仅仅是 PCI 的替代品,它其实是一个可交换的、点对点的通信网络,服务器以及 I/O 设备都可以挂接在上面,当然,在一个节点内部也可以使用 InfiniBand 结构互连多个 CPU (IPC)。它主要由 InfiniBand 连接 (IB Link)、主机通道适配器 (HCA)、目标通道适配器 (TCA)、交换机 (Switch)、路由器 (Router) 构成。InfiniBand 的拓扑结构如图 2

所示^[4,5]。

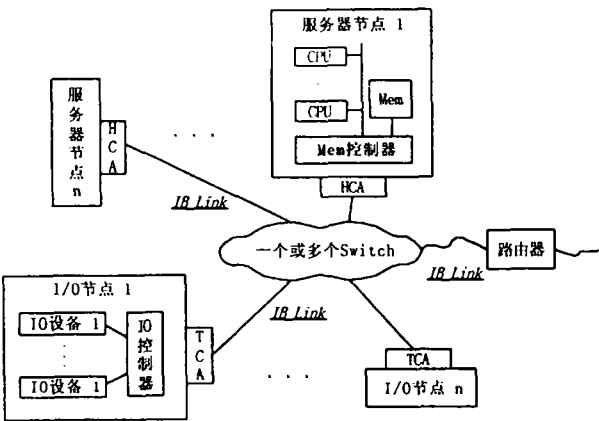


图2 InfiniBand 拓扑结构示意图

4.2 InfiniBand 的关键技术与优势

InfiniBand 这种交换结构的 I/O 技术,其设计思路是通过一组中心交换机构(中心 InfiniBand 交换机/交换机组)在远程存储器、外部网络以及服务器等设备之间建立一个单一的连接链路,并由中心 InfiniBand 交换机/交换机组来指挥流量。InfiniBand 结构的关键技术在于通过采用点对点的交换结构解决共享总线的瓶颈问题,很好地缓解了各硬件设备之间的数据流量拥塞,这种交换结构还专门用于解决容错性和可扩展性问题,这些都是共享总线所没有解决好的问题。

InfiniBand 的点对点交换机采用了与共享总线结构完全不同的连接方式,它通过一个带有源元件的结构(TCA、HCA)与各个设备连接,所有的连接都是点对点的连接,一个设备对应一个终端,不同于传统的 PCI 总线结构,Infiniband 结构中链路的建立和终止得到了很好的控制,所以,采用此结构后,通信系统的性能得以优化,可在高出很多的频率下运行。另外,以往采用共享总线结构的系统一次只能有一个主控

设备,并需要进行总线仲裁,这降低了系统总线结构中数据传输速率;而在 InfiniBand 结构中,可以利用无阻塞交换机将各端点连接起来,每个交换机可以同时有多条链路进行数据发送和接收,因此多个端点相互间可以同时进行通信。此外,仅需要添加一些交换机即可实现系统的扩展,随着交换机数目的增加,系统的总带宽也会相应增加,还可通过 InfiniBand 交换机的互连提供故障转移保护,PCI 总线一直是计算机中的潜在的单点故障点,而 InfiniBand 架构可以方便建立自动的备份与恢复机制,这在传统总线系统上是不可能实现的^[4,5]。

5 计算机 I/O 总线技术发展趋势

所有这些将来的即将实现的互连结构和总线,有一个共同的特征,它们都采用了基于包的点对点的交换互连结构,提供了数倍于 PCI 总线的带宽,下面将几种总线标准的主要性能指标做一比较,如表1所示。

表1 几种总线标准的性能指标(取最高性能参数)

	PCI(2.2)	PCI-X	3GIO	HT	IB
频率(MHz)	66	133		800	
位宽(bit)	64	64		32(双向)	
传输率(Gb/s)	4.2	8.5	2.5~10	12.8	2.5,10,30
引脚数	84,120	90,150	40	55,103,197	4,16,48
传输介质	PCB	PCB	PCB,铜缆	PCB	PCB,铜缆,光纤

总线标准是国际公布或推荐的互连各个模块的标准,它是把各种不同的模块组成计算机系统时必须遵守的规范。随着计算机系统的发展,总线技术也在不断地发展完善,一些总线标准已不能适应当前技术发展的需要,将会被改进或淘汰,随之新的总线标准将会被采纳。根据前面对业界新提出的一些总线标准的探讨,我们相信计算机的并行共享总线将逐渐被基于点对点的 Crossbar 交换技术的串行总线所替代,这将是该技术领域目前的发展趋势,随之传统的计算机结构也将发生巨大的变化,将有效地解决 PCI 等传统 I/O 结构产生的通信传输瓶颈问题,计算机的综合性能将提升到一个全新的水平。

参 考 文 献

1 孙德文. 微型计算机技术. 北京: 高等教育出版社, 2001
2 Cases M, Pham N. Design, Modeling and simulation methodology for PCI-X subsystems. Electrical Performance of Electronic Pack-

aging, IEEE Conf. on, 2000. 33~36
3 Third Generation I/O Architecture White Paper. Available at: <http://developer.intel.com/technology/3gio/index.htm>
4 William T. Futral, InfiniBand Architecture Development and Deployment, USA: Intel Press, 2001
5 Infiniband Architecture Specifications Volume 1, Release 1.0. a, Infiniband Association, 2001. Available at <http://www.infinibandta.org/specs/register/publicspec/>
6 Guerrier P, Greiner A. A generic architecture for on-chip packet-switched interconnections. Design, Automation and Test in Europe Conf. and Exhibition 2000, Proceedings, 2000. 250~256
7 Torres D, Gonzalez J, Guzman M, Nunez L. A new bus assignment in a designed shared bus switch fabric. Circuits and Systems, 1999, ISCAS '99. In: Proc. of the 1999 IEEE Intl. Symposium on, Volume: 1, 1999. 423~426
8 HyperTransport I/O Link Specification, Revision 1.03. Available at: <http://www.hypertransport.org/>