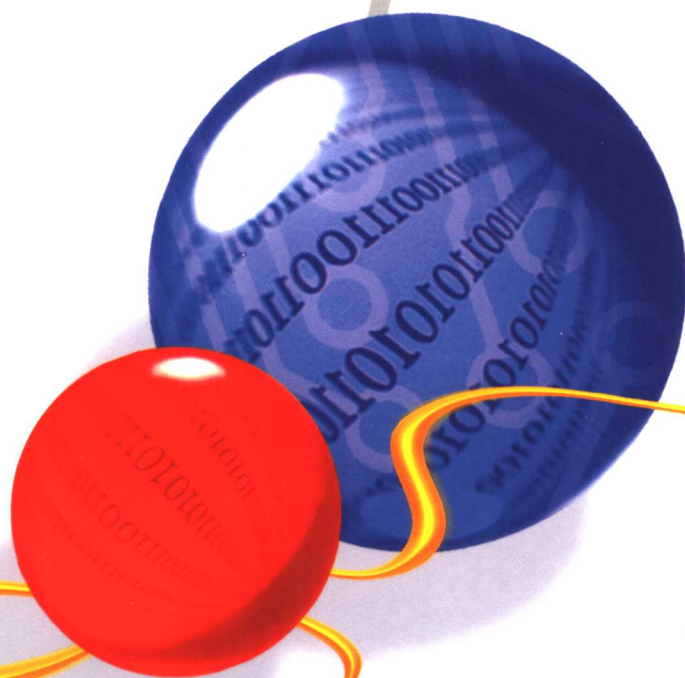


PCI、PCI-X 和

PCI Express

的原理及体系结构

马鸣锦 朱剑冰 何红旗 杜威 编著



清华大学出版社

TP336

23

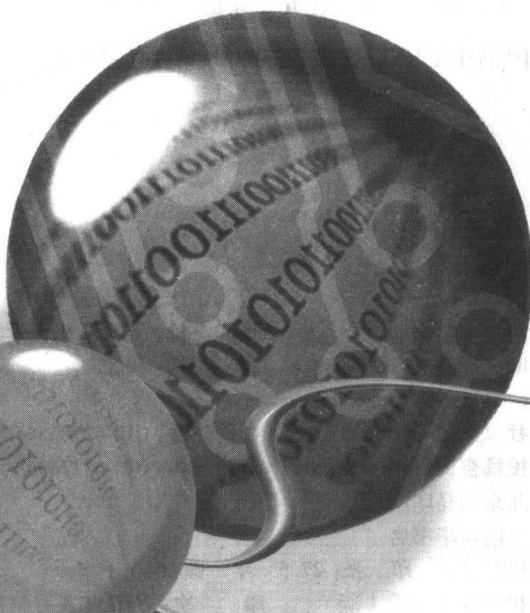
2007

PCI、PCI-X 和

PCI Express

的原理及体系结构

马鸣锦 朱剑冰 何红旗 杜威 编著



清华大学出版社

·北京·

内 容 简 介

PCI Express 是一种新的 IO 互连结构,被称为第三代本地 IO 总线,与其前任总线(PCI/PCI-X)相比,在总线结构上采取了根本性的变革,以满足现在的及将来的处理器和外围设备提出的更高的带宽要求。本书首先对 PCI/PCI-X 总线架构作了全面、简要的介绍;然后在整体描述 PCI Express 架构的基础上,逐一对 PCI Express 所采用的新方法、新思路和新技术进行了深入的描述与讨论;最后完整地介绍了与 PCI 兼容的 PCI Express 配置机制及配置寄存器。使读者在了解新一代 IO 互连结构的过程中获得计算机体系结构方面的理论和知识,提高计算机系统硬件和软件的开发能力。

本书可作为高等院校计算机专业及相关研究领域的研究生教材,可为从事相关专业本科教学的老师提供教学参考,对于从事个人计算机系统设计和应用开发的科技人员来说,也是一本内容翔实,可读性强的参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

PCI、PCI-X 和 PCI Express 的原理及体系结构/马鸣锦等编著. —北京:清华大学出版社,2007.4

ISBN 978-7-302-14438-0

I. P… II. 马… III. 电子计算机—接口设备 IV. TP334.7

中国版本图书馆 CIP 数据核字(2006)第 163970 号

责任编辑:魏江江 李玮琪

责任校对:李建庄

责任印制:何 芊

出版发行:清华大学出版社 地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn> 邮 编:100084

c-service@tup.tsinghua.edu.cn

社 总 机:010-62770175 邮购热线:010-62786544

投稿咨询:010-62772015 客户服务:010-62776969

印 装 者:北京国马印刷厂

经 销:全国新华书店

开 本:185×260 印 张:27.5 字 数:645 千字

版 次:2007 年 4 月第 1 版 印 次:2007 年 4 月第 1 次印刷

印 数:1~3000

定 价:38.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770177 转 3103 产品编号:023847-01

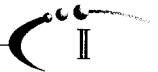
前言

在个人计算机的诸多应用领域中,如商务、计算、通信等,不断出现更新的、更高的、或者说苛刻的应用需求,即市场对更高性能计算机的追求是无止境的,其中更高的速度、更大的容量是首要追求的目标。具体到计算机系统的硬件体系结构上,则是要求有更高的数据传输率和更短的响应时间。

计算机系统硬件结构的基础是系统总线,用系统总线将处理器、主存储器和外围设备连接起来,传输数据、控制和状态信息。系统总线包括处理器总线和IO总线。处理器总线因处理器而异,随着处理器的变化和发展而更新。IO总线是通用总线,用于硬件平台中外围设备的互连和外围设备与主存储器的通信。处理器也经过IO总线对外围设备进行访问。IO总线是PC计算机平台的骨架,它的协议直接决定了计算机的硬件架构与配置,它的性能与指标直接影响着计算机IO子系统的数据传输和处理能力。所以,本书中把总线与计算机硬件架构联系在一起进行论述和讨论,使读者更易理解和掌握。

IO总线的发展历程粗略地可划分为三代。第一代IO总线包括ISA、EISA、MC和VESA。ISA总线的数据宽度为16位,数据传输率约为5MBps,1984年随IBM PC/AT结构微机一起推出,故也称为AT总线,当时系统存储器(主存)和系统外围都挂在此总线上,还没有冠以“IO总线”的称谓。EISA总线名为增强的ISA总线,1989年在ISA总线的基础上推出,仍沿袭PC/AT结构,数据总线宽度和地址总线宽度均扩展为32位,数据传输率约为33MBps。MC总线是随IBM PS/2系列微机推出的MCA——微通道结构总线,总线性能与EISA相当。VESA总线1992年由视频电子标准协会推出,数据和地址总线宽度为32位,总线时钟频率达33MHz,数据传输率可达132MBps。VESA总线是PC微机的第一个局部总线,连到处理器子系统的局部总线上,在基于VESA总线的机器中也可以存在ISA或EISA这样的传统总线,它们的位置在VESA之后,是VESA总线再经IO总线控制器转换后得到的,称为IO扩展总线。第一代IO总线有一个共同的特点,其信号的功能和时序与处理器引脚关系密切,几乎是处理器信号的延伸和扩展,有些信号还与主板上的硬件资源有关系。

第二代IO总线包括PCI、AGP和PCI-X。PCI总线是一个标准的、与处理器无关的局部外围总线,不受制于系统所使用的处理器的种类,通用性更强。PCI将多级总线结构引入个人计算机,不同总线之间通过相应的桥芯片来交接。PCI总线的数据和地址线宽32位,数据线可扩展为64位,1992年和1993年推出的PCI 1.0和PCI 2.0为33.3MHz,数据传输率为133MBps(64位数据线加



倍为 266Mbps), 1995 年推出的 PCI 2.1 总线时钟频率达 66.6MHz, 数据传输率达 266Mbps(533Mbps)。在基于 PCI 的计算机中, 主存储器设备挂在主桥上, PCI 总线专用于外围设备的互连, 开始称之为 IO 总线。AGP 名为加速的图形端口, 是主桥上独用于挂接显示设备的总线, 将 PCI 总线从图形数据传输中解放出来, 使得图形显卡和 PCI 总线上的设备都能获得充足的传输带宽。当市场上出现了要求更高带宽的 IO 设备时, 如 Gbps 以太网、光纤通道、Ultra3 SCSI 和多端口网络接口控制器(NICs), 传统的 PCI 总线无论从总线时钟频率还是总线使用率都显得低了, 于是 1999 年出现了 PCI-X 总线。PCI-X 总线在 PCI 总线的基础上, 进一步提高时钟频率, 改进数据传输协议使之更加合理。PCI-X 1.0 和 PCI-X 2.0(2002 年)的总线时钟频率分别为 66/133MHz、266/533MHz。

以上第一代和第二代 IO 总线都是并行总线, 有多条地址线、数据线和控制线, 挂连多个设备, 称为多下挂式总线(Multi-Drop), 总线带宽由多个设备共享。显然, 在并行总线上挂接多个高速设备, 比如 Gbps 以太网、实时性极强的视频、音频流数据处理设备等, 必须进一步提高时钟频率。但是, 以提高总线频率来提高带宽并不是无限制的, 频率越高总线上能够挂连的电气负载的个数越少。实际上, 频率为 266MHz 以上的 PCI-X 总线只能带一个扩展槽。也就是说, 对于一种给定的总线结构, 存在其频率上限, 超过该上限, 就必然会提出定义新总线结构的要求。

PCI Express 是第三代高性能 IO 总线, 在总线结构上采取了根本性的变革, 主要体现在两个方面: 一是由并行总线变为串行总线; 二是采用点到点的互连。

将原并行总线结构中桥下面挂连设备的一条总线变成了一条链路, 一条链路可包含一条或多条通路, 每条通路由两对差分信号线组成双单工的串行传输通道, 没有专用的数据、地址、控制和时钟线, 总线上各种事务组织成信息包来传送。PCI Express 1.0 支持每条通路在每个方向上的数据传输率达 2.5Gbps, 每字节 10 位编码, 这样两个方向的带宽可达 0.5Gbps, 整个链路的总带宽等于 0.5Gbps 乘以所含的通路数。每条链路的通路数可根据具体设备所需的带宽裁剪, 有效通路数有 7 种可选, 这样最高传输率可达 16Gbps, 大大高于目前任何一种总线, 可满足当前及将来一段时期的高速设备带宽需求。由于总线变为链路, 引脚数大大减少(传统 PCI 总线为 127 个引脚), 每引脚的平均带宽大幅提升, 有助于 PCI Express 成本的降低。

PCI Express 突破传统总线的另一个特点是采用点到点的互连方法, 每个设备都由独立的链路连接, 独享带宽, 这是提高传输率的有效解决方案。

PCI Express 总线结构将一条链路视为一条总线, 从而从地址空间、配置机制及软件上均保持与传统 PCI 总线的兼容。一个 PCI Express 设备占一条总线(链路), 所以在基于 PCI Express 的计算机里, 桥和总线的数目很多。传统 PCI 设备亦可在 PCI Express 平台上运行, 通过 PCI Express 到 PCI(或 PCI-X)的桥, 可引出 PCI(或 PCI-X)总线, 传统 PCI 设备挂在这样的总线上便可融入 PCI Express 结构。

为了反映总线结构的新技术, 凭借对传统 PCI 总线结构的深入了解, 在汲取大量最新资料的基础上编写了本书, 全书共有 17 章和 4 个附录。第 1 章是对 PCI 和 PCI-X 总线结构的综述, 全面、简要地勾画出 PCI 及 PCI-X 总线架构, 也为后文突显 PCI Express 优势及讲述与 PCI 的兼容问题打下基础。第 2 章是 PCI Express 架构概览, 从整体上对



PCI Express 架构进行描述,涉及到后面各章节要详述的主要技术问题,使读者建立起一个较完整的 PCI Express 概念。第 3 章描述事务层包和事务路由,给出所有类型 PCI Express 事务包的格式、内容及路由方式。第 4 章列出各类数据链路层包的格式、内容,并讲述由 ACK/NAK 两种数据链路层包参与的、保证链路传输正确性的链路传输协议。第 5 章讲述事务的传输类别、事务传输所经由的虚通道以及事务的仲裁。第 6 章讲述的流控制协议,是在每一虚通道上实施的、基于信用量的流控制机制,它也是 PCI Express 提高带宽的措施之一。第 7 章讲述事务排序的要求与规则,在一定的条件下允许对事务的传送顺序进行修正,以改善系统性能。第 8 章讲述兼容 PCI 和 PCI-X 的中断机制。第 9 章详述 PCI Express 设备的物理层,包括逻辑物理层和电气物理层。第 10 章讲述系统复位。第 11 章描述由物理层硬件自动执行的链路训练和初始化过程,链路训练使链路能自适应链路上设备端口的要求,能自动从链路错误中恢复。第 12 章讲述电源管理,管理 6 种设备电源状态和 6 种链路电源状态,包括基于硬件的链路节电机制及软件触发的链路电源管理。第 13 章讲述 PCI Express 设备和插槽对热插拔的支持。第 14 章是错误检测和处理,描述 PCI Express 的一套错误检测和报告能力。第 15、16、17 章均是 PCI Express 配置有关的内容,第 15 章是 PCI Express 配置机制及系统枚举,讲述 PCI Express 实施配置的方法;第 16 章是 PCI 兼容的配置寄存器;第 17 章是 PCI Express 专用的配置寄存器。

本书的第 1、2、9~11、14~16 章由马鸣锦编写,第 3、4、5 章由何红旗和马鸣锦编写,第 6、7、8 章由杜威和马鸣锦编写,第 12、13、17 章由朱剑冰和马鸣锦编写。全书由马鸣锦教授定稿并撰写了前言。何红旗、王炜、生拥宏、刘丹、陈烨、刘海雷、张琳等老师和同学参与了本书的校对、文字录入、图表制作等工作,另外,在本书编写过程中得到解放军信息工程大学信息工程学院计算机科学与技术系的大力支持,在此一并表示感谢。

PCI Express 总线及体系结构问世不久,还处在完善、充实和实现的过程之中,书中内容难免有错误和不当之处,作者诚恳希望与同行们对有关问题进行讨论和切磋。我们的通信地址是:郑州市 1001 信箱 715 号,邮编 450002。

作者
2006 年 10 月

目录

第 1 章 PCI 及 PCI-X 架构综述	1
1.1 PCI 架构综述	1
1.1.1 基于 PCI 总线的结构	1
1.1.2 PCI 总线信号的定义和功能	4
1.1.3 PCI 总线的事务类型	6
1.1.4 系统地址空间及 PCI 地址空间映射	7
1.1.5 PCI 总线主要的总线操作	8
1.1.6 改进的 PCI 总线系统	20
1.1.7 PCI 总线结构的局限性	22
1.2 PCI-X 架构简述	23
1.2.1 寄存器-寄存器的信号传送协议	23
1.2.2 PCI-X 提高总线效率的改进措施	25
1.2.3 PCI-X 支持消息通知中断 MSI	27
1.2.4 PCI-X 1.0 总线平台例	27
第 2 章 PCI Express 架构概览	29
2.1 PCI Express 的拓扑结构	29
2.1.1 PCI Express 结构的基本元素	29
2.1.2 PCI Express 系统示例	33
2.1.3 系统拓扑结构示例	35
2.2 PCI Express 事务介绍	36
2.2.1 PCI Express 事务协议简述	37
2.2.2 PCI Express 事务举例	39
2.3 PCI Express 设备的层次	42
2.3.1 设备层及相应的包	42
2.3.2 PCI Express 设备各层的关键功能	45
第 3 章 事务层包和事务路由	54
3.1 事务层包 TLP	54
3.1.1 TLP 的组装与拆解	54
3.1.2 TLP 结构	55



3.1.3 请求事务和完成事务的 TLP	59
3.2 地址空间和事务路由	68
3.2.1 事务层包的路由基础	69
3.2.2 与路由有关的即插即用配置	71
3.2.3 事务层包的路由机制	77
第4章 数据链路层包和 ACK/NAK 链路传输协议	81
4.1 数据链路层包 DLLP	81
4.1.1 DLLP 的组装与拆解	81
4.1.2 DLLP 的类型和结构	82
4.1.3 DLLP 的包格式	83
4.2 ACK/NAK 链路传输协议	84
4.2.1 ACK/NAK 协议概览	84
4.2.2 发送方 ACK/NAK 协议	88
4.2.3 接收方 ACK/NAK 协议	94
4.2.4 包的处理优先序	99
4.2.5 关于 ACK/NAK 协议可靠传递 TLP 的例证	99
4.3 交换开关的直通模式	102
4.3.1 交换开关对直通模式的需求	102
4.3.2 交换开关的直通模式	102
第5章 传输类别、虚通道和仲裁	104
5.1 服务质量的概​​念	104
5.2 传输类别和虚通道	105
5.2.1 VC 配置寄存器的结构	105
5.2.2 VC 分配和 TC 映射	106
5.3 仲裁	108
5.3.1 虚通道仲裁	109
5.3.2 端口仲裁	112
5.3.3 非交换开关的端口仲裁和 VC 仲裁	116
第6章 流控制协议	118
6.1 流控制概念	118
6.2 流控制机构的组成	120
6.2.1 流控制缓冲器	120
6.2.2 流控制包	121
6.2.3 流控制机构的元素及其作用	121
6.3 流控制机构的操作示例	123

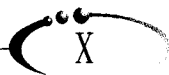
6.3.1	初始化之后的流控制	123
6.3.2	流控制缓冲器填满	124
6.3.3	流控制信用量循环指针的翻转	124
6.3.4	FC 缓冲器溢出错误检查	125
6.4	流控制信用量的通报数量	126
6.4.1	不限定的流控制信用量的通报	126
6.4.2	通报信用量的最少数量要求	126
6.5	流控制初始化和流控制更新	127
6.5.1	流控制初始化的过程与步骤	127
6.5.2	流控制初始化之后的流控制更新	130
第 7 章	事务排序	133
7.1	基本的排序规则	133
7.1.1	生产者/消费者模型	133
7.1.2	PCI Express 的基本排序规则	134
7.2	宽松排序	135
7.2.1	RO 对存储器写和消息事务的影响	135
7.2.2	RO 对存储器读事务的影响	135
7.2.3	基于强排序和 RO 属性的基本排序规则	136
7.3	强排序规则的修正	137
7.3.1	强排序引起的事务阻塞	137
7.3.2	弱排序修正方案	138
7.4	支持 PCI 总线和避免死锁	139
第 8 章	中断	141
8.1	消息信号中断	142
8.1.1	MSI 能力寄存器组及其配置步骤	142
8.1.2	MSI 中断请求的生成	144
8.1.3	使用 MSI 中断的注意事项	145
8.2	虚拟 INT _x 信号中断	146
8.2.1	PCI 的中断提交方法	146
8.2.2	虚拟 INT _x 信号发送	149
8.2.3	INT _x 消息传递相关的规则	151
第 9 章	物理层	153
9.1	逻辑物理层的发送逻辑	153
9.1.1	多路控制逻辑	155
9.1.2	字节剥解	155



9.1.3	乱序器	159
9.1.4	8/10b 编码	161
9.1.5	发送逻辑的其他有关事项	167
9.2	逻辑物理层的接收逻辑	168
9.2.1	Rx 时钟还原与串/并转换	169
9.2.2	符号锁定	170
9.2.3	去除通路间的时差	170
9.2.4	接收器时钟补偿逻辑	172
9.2.5	8/10b 译码器	172
9.2.6	除乱序器	174
9.2.7	解除字节剥解逻辑和过滤器	175
9.2.8	物理层错误处理	175
9.3	电气物理层	175
9.3.1	差分驱动器和差分接收器	176
9.3.2	高速电气信号传送的一些要求	178
9.3.3	探测接收器	179
9.3.4	电气闲	180
9.3.5	减重(预加重)	181
9.3.6	发送器驱动器的特性	182
9.3.7	输入接收器的特性	183
9.3.8	各电源状态下的电气物理层状态	184
第 10 章	系统复位	186
10.1	两类系统复位	186
10.1.1	基本复位	186
10.1.2	带内复位	188
10.2	复位退出	189
10.3	将链路从 L2 低功耗状态唤醒	190
第 11 章	链路训练和初始化	191
11.1	链路训练和初始化的内容	191
11.2	链路训练和初始化期间所使用的有序集	193
11.2.1	TS1 和 TS2 有序集	194
11.2.2	其余有序集	195
11.3	链路训练和状态机	195
11.3.1	Detect 状态	197
11.3.2	Polling 状态	198
11.3.3	Configuration 状态	201



11.3.4	L0 状态	207
11.3.5	L0s 状态	207
11.3.6	L1 状态	210
11.3.7	L2 状态	211
11.3.8	Recovery 状态	212
11.3.9	Hot Reset 状态	215
11.3.10	Disable 状态	215
11.3.11	Loopback 状态	216
11.4	LTSSM 有关的配置寄存器	217
11.4.1	链路能力寄存器	217
11.4.2	链路状态寄存器	218
11.4.3	链路控制寄存器	219
第 12 章 电源管理		220
12.1	电源管理配置软件的基础	220
12.1.1	PCI PM 基础	221
12.1.2	电源管理状态的定义及设备运行环境信息的定义	222
12.1.3	PCI Express 电源管理与 ACPI	224
12.2	功能的电源管理	227
12.2.1	设备的 PM 状态	227
12.2.2	PM 能力寄存器组	232
12.3	链路的活跃状态电源管理	236
12.3.1	活跃状态电源管理概述	238
12.3.2	L0s 状态	239
12.3.3	L1 ASPM 状态	241
12.3.4	ASPM 退出等待时间	246
12.4	软件发起的链路电源管理	248
12.4.1	D1/D2/D3 _{Hot} 和 L1 状态	248
12.4.2	L2/L3 Ready 状态	251
12.5	链路唤醒协议和 PME 的产生	254
12.5.1	PME 消息及相关事宜	255
12.5.2	唤醒非通信链路	256
12.5.3	辅助电源	258
第 13 章 热插拔		259
13.1	PCI Express 的热插拔环境与硬软件组成要素	259
13.1.1	PCI 和 PCI Express 热插拔之间的不同	260
13.1.2	热插拔所涉及的软件	262

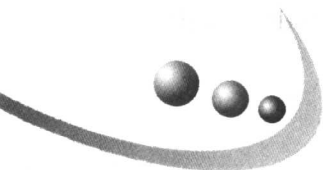


13.1.3	热插拔所涉及的硬件	262
13.2	卡拔出和插入的过程	263
13.2.1	On 和 Off 状态	263
13.2.2	卡拔出步骤	264
13.2.3	卡插入过程	266
13.3	标准化的使用模型	267
13.3.1	标准用户接口	267
13.3.2	标准的热插拔控制信号接口	270
13.4	热插拔控制器编程接口	271
13.4.1	插槽能力寄存器	271
13.4.2	插槽控制寄存器	272
13.4.3	插槽状态寄存器	273
13.4.4	卡插槽与服务器 IO 模块	275
13.5	插槽号	276
13.6	热插拔原语	276
第 14 章	错误检测和处理	278
14.1	PCI Express 错误管理概述	278
14.1.1	PCI Express 错误检测范围	279
14.1.2	错误报告渠道	279
14.1.3	错误分类	280
14.2	PCI Express 的错误源	280
14.2.1	ECRC 的产生和检查	280
14.2.2	数据中毒	281
14.2.3	TC 到 VC 的映射错误	281
14.2.4	链路流控制有关的错误	282
14.2.5	畸形事务层包	282
14.2.6	分割事务错误	283
14.3	错误分类及错误报告方法	285
14.3.1	错误分类	285
14.3.2	错误报告方法	286
14.4	基本的错误检测和处理	286
14.4.1	PCI 兼容的错误报告机制	287
14.4.2	PCI Express 的错误报告机制	288
14.5	高级的错误报告机制	291
14.5.1	ECRC 的产生和检查	292
14.5.2	高级可纠正错误的处理	293
14.5.3	高级不可纠正错误的处理	294

14.5.4	错误日志	296
14.5.5	根复合体的错误跟踪和报告	297
14.6	错误处理流程	299
第 15 章	PCI Express 配置机制及系统枚举	301
15.1	PCI Express 的配置空间	301
15.1.1	配置空间的作用	301
15.1.2	配置空间的层次和大小	302
15.1.3	配置空间的布局结构	302
15.2	PCI Express 的配置访问机制	304
15.2.1	PCI 兼容的配置访问机制	304
15.2.2	PCI Express 增强的配置访问机制	310
15.2.3	关于初始的配置访问的一些时间规定	311
15.3	PCI Express 枚举	312
15.3.1	对具有单个根复合体系统的枚举	313
15.3.2	对具有多个根复合体系统的枚举	317
15.4	与系统枚举和系统拓扑有关的一些问题	319
15.4.1	根复合体和交换开关内的多功能设备	319
15.4.2	嵌入根复合体和交换开关内的端点	321
15.4.3	设备 ID 的获知和记忆	321
第 16 章	PCI 兼容的配置寄存器	323
16.1	类型 0 配置头标区	323
16.1.1	用来识别设备驱动程序的寄存器	324
16.1.2	头标类型寄存器	326
16.1.3	BIST 寄存器	326
16.1.4	能力指针寄存器	327
16.1.5	Card Bus 卡信息结构指针寄存器	328
16.1.6	命令寄存器	329
16.1.7	状态寄存器	330
16.1.8	基地址寄存器	331
16.1.9	扩充 ROM 及扩充 ROM 基地址寄存器	337
16.1.10	中断线寄存器	339
16.1.11	中断引脚寄存器	340
16.2	类型 1 配置头标区	340
16.2.1	总线号寄存器	341
16.2.2	基本的事务过滤机制	342
16.2.3	桥的 IO 过滤器	344



16.2.4	桥的可预取存储器过滤器	348
16.2.5	桥的存储器映射 IO 过滤器	351
16.2.6	桥的命令和控制寄存器	352
16.2.7	桥的状态和次级状态寄存器	355
16.3	PCI 兼容的能力寄存器组	358
16.3.1	底板和插槽	358
16.3.2	底板/插槽编号有关的寄存器	359
16.3.3	两个例子	361
第 17 章 PCI Express 专用配置寄存器		366
17.1	PCI Express 能力寄存器组	366
17.1.1	PCI Express 能力寄存器	367
17.1.2	设备能力寄存器	368
17.1.3	设备控制寄存器	372
17.1.4	设备状态寄存器	374
17.1.5	链路能力寄存器	376
17.1.6	链路控制寄存器	377
17.1.7	链路状态寄存器	379
17.1.8	插槽能力寄存器	380
17.1.9	插槽控制寄存器	381
17.1.10	插槽状态寄存器	383
17.1.11	根控制寄存器	384
17.1.12	根状态寄存器	385
17.2	PCI Express 的扩展能力寄存器组	385
17.2.1	高级的错误报告能力寄存器组	386
17.2.2	虚通道能力寄存器组	391
17.2.3	设备序列号能力寄存器组	397
17.2.4	电源预算能力寄存器组	398
17.3	根复合体寄存器块	404
附录 A PCI Express 插卡连接器的引脚		405
附录 B 分类码		409
附录 C 高级交换		415
附录 D 中英文名词索引		417
参考文献		421



第 1 章 PCI及PCI-X架构综述

PCI Express 不同于以前的 IO 总线,采用串行的、点到点的互连方法来实现两个设备间的通信,但是,使用模型和设计思想与传统 PCI(PCI、PCI-X)相同或相似。

PCI Express 保持了与传统 PCI 相同的存储器、IO 和配置地址空间模型,因此,现有操作系统和驱动程序软件可以不做任何修改地运行于 PCI Express 系统中,即 PCI Express 在软件上向后兼容传统 PCI 系统。PCI Express 所支持的事务,如存储器读写、IO 读写和配置读写也是传统 PCI 中所熟知的。在 PCI Express 系统中,通过 PCI Express 到 PCI(或 PCI-X)的桥,可引出 PCI(或 PCI-X)总线,传统的 PCI 设备挂在这样的总线上亦可在 PCI Express 平台上运行。PCI Express 将每条链路视为仅挂接一个设备的总线,每条链路都由一个桥(物理上的或逻辑上的)引出,保持了传统 PCI 系统中的多级总线拓扑结构的设计思想,因此具有与传统 PCI 类似的配置初始化过程。

鉴于以上原因,为了更好地理解 PCI Express 体系结构,在本章中对 PCI 及 PCI-X 总线结构做一综述。

1.1 PCI 架构综述

1.1.1 基于 PCI 总线的结构

图 1.1 示出了一个最基本的基于 PCI 总线(PCI Bus)的平台。该平台包含三级总线:FSB(Front-Side Bus)、PCI 和 ISA,FSB 是处理器子系统的总线,原称为 Host 总线,总线定义完全取决于系统所使用的处理器;PCI 局部总线是一个完全与处理器无关的总线,不受制于系统所使用的微处理器的种类;ISA 总线亦称为 IO 扩展总线,在实际系统中,也有采用 EISA 或 MC 总线的,以 ISA 总线用得较多。不同的总线之间通过相应的桥芯片来隔离和转接。

平台中两级桥是必须具有的,一是 Host 到 PCI 的桥(常称为主桥——Host 桥),即图中的北桥;另一个是 PCI 到扩展 IO 总线的桥(常称为扩展总线桥),即图中的南桥。

Host 桥应该与系统所使用的处理器相配套,不论何种 Host 桥芯片,其宏观功能是类似的,其中最主要的一个功能就是把处理器的 Host 总线周期转换并组织成 PCI 总线上的周期,同时实现处理器的地址到 PCI 地址空间的映射。系统存储器,也称为主存(Main Memory)与处理器和位于 Host 总线上的处理器片外的高速缓存(Cache)关系密切,直接挂在 Host 桥上与 PCI 总线隔开,图中的主存储器由 SDRAM 芯片组成。所以 Host 桥内

还集成有 Cache 控制器,管理位于 Host 总线上的处理器片外 Cache; 还集成有存储器控制器,管理对系统存储器(主存)的访问。Host 桥上还有 AGP 端口,连出 AGP 总线用于挂接显示设备。

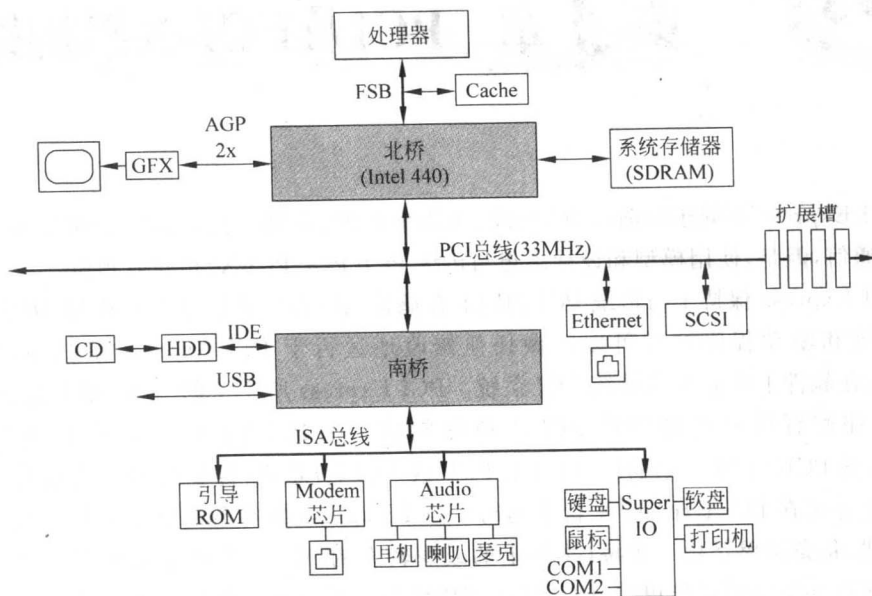


图 1.1 最基本的基于 PCI 总线的平台

为了支持多年积累的、数量庞大的 ISA 产品及其应用程序,也为了挂接较慢的 IO 设备,基本的 PCI 系统中一般设置 ISA 总线。南桥 (PCI-ISA 桥) 在 PCI 总线和 ISA 总线之间提供地址、数据通路,进行两总线周期的相互转换。在 PCI-ISA 桥的内部不仅仅具有两总线间的转换逻辑,还集成有为系统服务的 IO 接口,故这种桥又称为 SIO (System IO)。南桥中的 IO 端口多数保持与原 PC/AT 微机结构兼容,并且为了适应新的微机结构,功能还有所扩展,主要有基于双 8237 DMA 控制器的 DMA 逻辑电路;基于双 8259 可编程控制器的可屏蔽中断控制逻辑电路及 8254 定时/计数器。

基于 PCI 总线的平台中还应包括有其他一些 ISA 兼容的 IO 接口,如键盘接口、鼠标接口、实时时钟电路、异步串行接口、并行接口、软盘控制器等,这些接口常常集成到一个芯片上,称为“Super IO”,挂在 ISA 总线上。

挂在 PCI 上的设备可以是一个 PCI 总线主设备,可以是一个 PCI 总线从设备,也可以既是主设备又是从设备。PCI 总线主设备能够发起 PCI 总线上的事务,访问 PCI 从设备,被访问的从设备就成为该主设备的目标。Host 桥也是一个 PCI 设备,显然是 PCI 总线上最重要的 PCI 设备,当把处理器的一个访问周期推向 PCI 总线、转换为 PCI 访问周期时,从 PCI 总线的角度看,它为 PCI 主设备;当某一 PCI 主设备启动对主存储器的 DMA 访问时,它又为 PCI 目标,接受这个访问并推向主存储器。PCI-ISA 桥既是一个 PCI 主/从设备,又是一个 ISA 主/从设备,当把一个 PCI 主设备的 PCI 事务从 PCI 总线推向 ISA 总线时,在 PCI 总线上它是该 PCI 事务的目标,而在 ISA 总线上它是一个主设

备；反之当把一个 ISA 主设备启动的 ISA 周期推向 PCI 总线时，在 ISA 总线上它为从设备，在 PCI 总线上为主设备。

一条 PCI 总线上的电气负载数是一定的，随总线频率的上升而下降，而且一个扩展槽加上槽内的 PCI 外围设备插卡相当于两个负载。总线上的负载数越多，信号的驱动器将信号驱动至稳定的有效电平所需的时间越长，在规定的总线时钟频率之下，要想保证从驱动器开始驱动一个信号到接收器检测到该有效信号的总时间小于时钟的周期，总线上的电气负载必须在规定数目之下。因此，为了在一个系统中支持多个 PCI 设备，PCI 总线支持多级结构，若实际系统中的 PCI 设备较多，一条 PCI 总线负担不了，可以用 PCI-PCI 桥（亦称为 PCI 桥）扩充下级 PCI 总线，形成一个分级式的 PCI 总线结构，图 1.2 的平台中设置了一个 PCI-PCI 桥，从 PCI 总线 0 引出下级的 PCI 总线 1。继续在下级总线上设置 PCI-PCI 桥，或者在同级总线上设置 PCI-PCI 桥，可以扩展出多条总线，理论上可达 256 条。

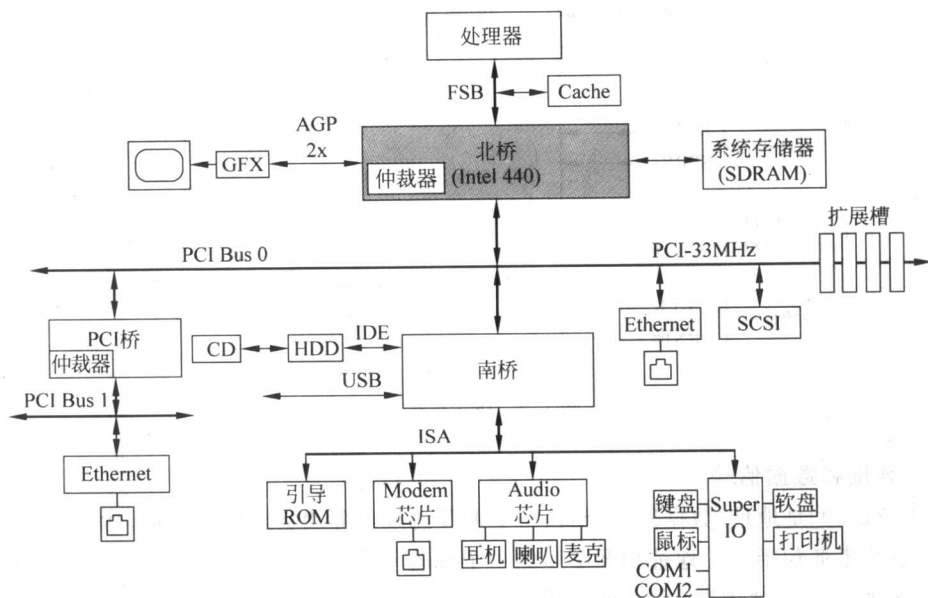


图 1.2 设有 PCI-PCI 桥的 PCI 总线平台

PCI 总线支持多主设备，所以每一条 PCI 总线上都要设置一个总线仲裁器，仲裁器逻辑一般设置在相应的桥中，图 1.2 中，Host 桥里的仲裁器就是 Host 桥次级边 PCI 总线 0 的仲裁器，PCI 桥里的仲裁器就是 PCI 总线 1 的仲裁器。当 PCI 总线主设备欲访问系统中的一个 PCI 目标时，它首先要向欲占用总线的仲裁器提出总线请求，收到仲裁器返回的总线允许信号之后，它便可在该总线上发起一个 PCI 访问事务。若主设备欲访问的目标在同一总线上时，主设备只需要仲裁一条总线（所在的总线）的使用权；若欲访问的目标位于另一条总线上，则对于该主设备与目标之间的每一条总线都需要仲裁其使用权。比如，总线 0 上的主设备欲访问总线 1 上的目标，该总线 0 主设备向 Host 桥里的仲裁器请求总线 0 使用权，然后 PCI 桥以总线 1 主设备的名义向桥中总线 1 仲裁器请求总线 1 的使用权。