

光互连与并行处理

李之棠 等编著



电子工业出版社
Publishing House of Electronics Industry
www.phei.com.cn

国家自然科学基金和国家高性能计算基金资助项目

光互连与并行处理

李之棠 等编著

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

光互连是一种利用各种光传输介质把计算机系统内各部件或各子系统连接起来并通过光来高速传递信息的技术;是光学、物理学及计算机科学的综合交叉性新领域。而我国从计算机学科的角度来进行研究和应用光互连技术才刚刚开始,介绍光互连技术的书籍更是少之又少。本书收集、整理和总结了二十多年来国际上光互连领域的一些重要研究成果和发展现状,主要介绍了光互连技术的基本原理、构成、实现方法及其在大规模并行处理领域的应用。通过本书可了解到光互连技术领域的最新进展状况及其技术实质和应用前景。

本书内容新颖,实例丰富,图文并茂。可作为计算机科学与工程、光电子科学与工程及物理学等领域科研人员的参考书,也可作为上述专业或相关专业博士生、硕士生及大学生的教材或参考书。

本书版权归电子工业出版社所有,未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,翻版必究。

图书在版编目(CIP)数据

光互连与并行处理 / 李之棠等编著. —北京:电子工业出版社, 2001. 12

ISBN 7-5053-7446-X

I. 光... II. 李... III. ①光学计算机—开放系统互连②光学计算机—并行处理 IV. TP381

中国版本图书馆CIP数据核字(2001) 第095943号

书 名: 光互连与并行处理

编 著 者: 李之棠等

责任编辑: 梁卫红

排版制作: 今日电子公司制作部

印 刷 者: 北京天竺颖华印刷厂

出版发行: 电子工业出版社 www.phei.com.cn

北京市海淀区万寿路173信箱 邮编: 100036

经 销: 各地新华书店

开 本: 787 × 1092 1/16 印张: 12.25 字数: 305千字

版 次: 2001年12月第1版 2001年12月第1次印刷

书 号: ISBN 7-5053-7446-X
TP · 4298

定 价: 25.00元

凡购买电子工业出版社的图书,如有缺页、倒页、脱页、所附磁盘或光盘有问题者,请向购买书店调换。
若书店售缺,请与本社发行部联系调换。联系电话: 88211980 68279077

序 言

VLSI技术的飞速发展和并行处理技术的进步,一方面使单个处理器的时钟速率不断提升,如当前通用的 Pentium 4 芯片就已达到 2GHz;另一方面也使并行处理系统内的处理器数量不断增加,如本书最后一章提到的法国 SYMPHONIE 系统就有 1024 个处理器,有的大规模并行处理系统(MPP)还有更多的处理器。但随着单个处理器速率的不断提高和系统处理器数量的不断增加,构成系统的各单元(如 PCB 板、芯片、处理器、I/O 设备接口及交换或路由开关等)之间经过互连网络进行通信的开销也随之大大增加,以致使并行处理系统互连网络的吞吐率和时延成为阻碍并行系统总体性能提高的主要障碍之一。

传统电互连网络存在着许多严重制约高性能并行计算系统性能进一步提高的障碍。一方面电连接存在的分布参数会随着传输速度的提高而引发带宽变窄、时钟歪斜、电磁场干扰、I/O 能力受限及功耗高等现象,最终限制并行系统的处理能力;另一方面静态电连接网络不仅电结点多、连线数目大,而且大多采用总线固定连接(使用总线的实质是通过竞争或分时来共享传输介质)。当系统负荷及通信量都很大时,系统的处理能力(平均工作处理机数或并行度)并不能随计算机数量的增加而成比例地增加,平均单个计算机的处理能力可能反而还会下降。对总线的这种激烈竞争几乎可使增加处理器个数得到的好处丧失殆尽。

现实世界对计算系统的应用要求比起当前计算系统所提供的能力来说永远是一个无穷大。例如,当前星(或航天器)载计算机、3D 虚拟现实、T 比特交换及天气预报等应用领域不仅需要高速处理和高速通信,而且还需要高密度(或轻小型化)和低能耗。

1983 年美国提出全光数字计算机的构想(当时美苏冷战达到高潮,美制定了“星球大战计划”),1985 至 1990 年间,进行了全光数字处理器件与光互连网络系统的许多探索。1990 年 1 月,AT&T Bell 实验室宣布第一台光学数字处理器研制成功。虽然而后因冷战结束、美国停止执行“星球大战计划”而使全光数字计算机系统的研究暂时失去近期目标的吸引和财力的推动,但其中有关光互连网络的一些研究成果却转化和发展成为今天的“并行处理中的光互连技术”,并成为计算机界和光电子界的热门研究领域。在过去的几十年中,美、日、欧等国家和地区在光互连研究上投入了相当的人力、物力和财力,研究成效显著。现在已经有多种光互连实验原型机问世,如日本 NTT 的 COSINE 系列实验光互连计算机和法国的 SYMPHONIE 系统等。另外,多种光互连器件也已接近实用,如智能点技术及平面微透镜阵列等。同时,集成光互连器件与 VLSI 电子器件的单块集成工艺也有了妥善的解决方案,如倒焊工艺等。部分商用巨型计算机中也开始采用光互连技术,如 Cray T90 等。

为了避开光在处理和存储方面的困难,最近,国际上光互连研究正趋向于采用微光电子和微电子混合集成技术把 CMOS 电路和量子阱自光电效应器件(SEED)进行混合集成,从而充分发挥光、电各自的优势。特别是可以充分利用光在局域内的广播特性和在非局域内的高速传输特性。专家们一般认为“处理由电来做,传输由光来做”将是未来高性能并行计算系统结构的主要途径之一。非并行的高带宽传输带来了当今因特网的爆炸式增长。并行的高带宽传输必将使并行计算机系统的性能有相当显著的提高。未来的光互连技术将在大规模并行处理系统、航空航天星载系统及巨型计算机系统中发挥越来越大的作用。

光互连是一种利用各种光传输介质把计算机系统内各部件或各子系统连接起来并通过光来高速传递信息的技术。从光互连所用的传输介质来看,主要有光纤互连、波导互连及自由空间光互连等技术。从光互连在计算机系统结构中所处的层次来看,它可以在计算机与计算机、背板与背板、平面内芯片与芯片及自由空间内芯片与芯片之间等不同层次实现光互连。光作为信息载波,具有独立传播、无接触连接、极高时空带宽积、极高密度(或体积小、重量轻)、等光程、低延迟、低功耗、I/O能力强及抗电磁场干扰等诸多优点。

学术界注重把光互连技术同计算机系统结构结合起来是从1994年的MPPOI(Massively Parallel Processing using Optical Interconnections)国际会议开始的,在此以前,光计算和光互连领域的主流科研工作者几乎都是由光学科学家、物理学家和电子工程师们组成的。本书所做的主要工作是收集、整理和总结二十多年来国际上光互连领域的一些重要研究成果和发展现状,旨在通过本书把光互连技术介绍给我国并行处理研发领域的科研工作者及工程师们,从而推动我国光互连技术和并行处理技术的进步。

本书可作为计算机科学与工程、光电子科学与工程及物理学等领域科研人员的参考书,也可作为上述专业或相关专业博士生、硕士生及大学生的教材。

本书的编著主要是在国家自然科学基金(69873016、69972017)和国家高性能计算基金(984055)的资助下完成的。感谢上述基金委及有关方面的同志们在几年前就非常敏锐地洞察到该领域的重要性并及时给予支持。

本书由李之棠、李非、黄辉龙、赵振龙及解云峰等共同编写,全书由李之棠修改定稿。

光互连技术不仅是光学、物理学及计算机科学的综合交叉性新领域,而且工程性特别强,这些方面并不是我们的长处。由于从计算机学科的角度来进行研究和应用光互连技术在我国才刚刚开始,加之我们水平有限,因此,书中难免有这样或那样的缺点或错误,恳请广大读者批评赐教。

作者

2001年9月于华中科技大学

目 录

序言	1
第 1 章 光互连与并行处理概论	3
1.1 光互连通信技术	3
1.1.1 光互连技术概论	3
1.1.2 光互连与光交换的必然性	3
1.1.3 光互连技术的发展现状	5
1.2 层次抽象与光互连技术	7
1.2.1 解决问题的抽象层次	7
1.2.2 通用计算机系统的抽象层次	8
1.2.3 通用系统与专用系统	8
1.2.4 通用系统引入光互连	9
1.2.5 讨论与结论	11
1.3 并行处理中的光互连应用	12
1.3.1 光互连并行处理的四种类型	12
1.3.2 并行光互连的典型应用	15
本章参考文献	21
第 2 章 光互连技术	23
2.1 基本光互连技术	23
2.1.1 三种光互连	23
2.1.2 光互连器件	26
2.1.3 波分复用	27
2.2 智能点技术	30
2.2.1 智能点原理	31
2.2.2 双调排序器	35
2.2.3 发展趋势	40
2.2.4 应用前景	42
本章参考文献	42
第 3 章 光互连网络模型	45
3.1 交换技术	46
3.1.1 常见交换技术	46
3.1.2 波分复用	47
3.2 拓扑结构	48

3.2.1 基本定义	48
3.2.2 度和直径	49
3.2.3 有向超图	50
3.2.4 实际拓扑结构	50
3.3 群体通信	56
3.3.1 电网络的通信模型	56
3.3.2 总线网中的群体通信	57
3.3.3 波分复用的群体通信	57
3.4 并行模型中的光通信	61
3.4.1 PRAM 模型	61
3.4.2 光互连并行计算机	61
3.5 无源星形网模型	62
3.5.1 OPS 在 PRAM 中的层次	62
3.5.2 OPS 的自我仿真	63
本章参考文献	63
第 4 章 光互连网络结构	66
4.1 并行系统内的自由空间光互连	66
4.2 低延时异步光总线	68
4.2.1 总线的构成	68
4.2.2 30 ~ 100Gb/s 光总线	71
4.2.3 0.5 ~ 1Tb/s 光总线	72
本章参考文献	77
第 5 章 可重构光互连网络	79
5.1 MILORD 结构	79
5.1.1 基本概念	79
5.1.2 功率预算	80
5.1.3 重构时间	80
5.1.4 实验结果	80
5.1.5 光交叉开关网络	81
5.2 OEDIPE 结构	82
5.2.1 体系结构	82
5.2.2 并行通信	82
5.2.3 互连并行性	83
5.2.4 网络布局和系统设计	84
5.2.5 系统特点	90
5.3 可重构光互连网络的嵌入特性	90
5.3.1 超平面可重构光背板	91
5.3.2 总线光网的图和超图模型	92
5.3.3 可重构和可划分的光互连模型	93

5.3.4 光嵌入的组合优化问题	94
5.3.5 网络的嵌入	98
本章参考文献	99
第 6 章 光阵列逻辑计算	101
6.1 光阵列逻辑	101
6.2 OAL-NC 结构	102
6.3 OAL-NC 的实现	104
6.4 多处理的软实现	105
6.5 可能的 OAL-NC 系统结构	107
本章参考文献	108
第 7 章 光无源星形网	109
7.1 基于 OPS 的网络	110
7.1.1 单跳单 OPS 多波长网络	110
7.1.2 多跳单 OPS 多波长网络	111
7.1.3 单跳多 OPS 单波长网络	111
7.2 栈图	112
7.2.1 嵌入问题	112
7.2.2 栈图的定义	113
7.3 基于 OPS 的网络模型	113
7.3.1 单 OPS	113
7.3.2 多 OPS	114
7.4 模型的应用	115
7.4.1 POPS 的优化嵌入	115
7.4.2 嵌入环通信	115
7.4.3 其他嵌入方法	116
7.4.4 多跳单 OPS 多波长网络设计	117
7.4.5 多跳单 OPS 多波长网络的寻径	119
7.4.6 评注	121
7.5 结论与展望	121
本章参考文献	122
第 8 章 全光互连网的时分复用控制	124
8.1 多处理机全光互连	124
8.2 电路交换技术	126
8.2.1 静态与动态技术	126
8.2.2 时分复用技术	127
8.2.3 TDM 的静态与动态技术	128
8.3 TDM 网络中的通信开销	131

8.4 循环开销与动态复用控制	134
8.5 结论	137
本章参考文献	138
第9章 贪婪热土豆路由算法	139
9.1 算法的发展历史	139
9.2 模型与定义	141
9.2.1 d维网格	142
9.2.2 贪婪热土豆路由算法	143
9.2.3 受限包	144
9.3 受限包的优先级	144
9.3.1 势能函数分析	144
9.3.2 二维网格中的势能函数	147
9.4 隋性包的优先级	149
9.5 二维 Mesh 的最优单目标路由	150
9.6 多维 Mesh 的随机单目标路由	155
9.6.1 循环前进算法及其性质	156
9.6.2 单目标路由	157
9.6.3 超立方体上多目标路由	159
9.6.4 高维网格中的路由	160
9.7 贪婪路由交换	162
9.7.1 近视与 a -有界算法	162
9.7.2 近视 a -有界算法的例子	162
9.7.3 注释和定义	163
9.7.4 构造	164
9.7.5 构造和瓶颈置换的联系	166
9.7.6 行-列算法	167
9.7.7 Kaklamanis 算法	168
9.7.8 非 a -有界的算法	169
9.8 结论	171
本章参考文献	172
第10章 光互连大规模并行计算机	174
10.1 两个光互连并行计算机系统	174
10.1.1 SYMPAT12	174
10.1.2 SYMPHONIE	175
10.2 电互连的局限性	177
10.2.1 延迟缺陷	177
10.2.2 对 I/O 的影响	178
10.3 SYNOPTIQUE 项目	179
10.3.1 使用光互连的动机	179

10.3.2 MCM 间的光互连	180
10.3.3 设计和技术难点	181
10.4 光互连示例	183
10.5 结论与展望	184
本章参考文献	184

序 言

VLSI技术的飞速发展和并行处理技术的进步,一方面使单个处理器的时钟速率不断提升,如当前通用的 Pentium 4 芯片就已达到 2GHz; 另一方面也使并行处理系统内的处理器数量不断增加,如本书最后一章提到的法国 SYMPHONIE 系统就有 1024 个处理器,有的大规模并行处理系统(MPP)还有更多的处理器。但随着单个处理器速率的不断提高和系统处理器数量的不断增加,构成系统的各单元(如PCB板、芯片、处理器、I/O设备接口及交换或路由开关等)之间经过互连网络进行通信的开销也随之大大增加,以致使并行处理系统互连网络的吞吐率和时延成为阻碍并行系统总体性能提高的主要障碍之一。

传统电互连网络存在着许多严重制约高性能并行计算系统性能进一步提高的障碍。一方面电连接存在的分布参数会随着传输速度的提高而引发带宽变窄、时钟歪斜、电磁场干扰、I/O能力受限及功耗高等现象,最终限制并行系统的处理能力; 另一方面静态电连接网络不仅电结点多、连线数目大,而且大多采用总线固定连接(使用总线的实质是通过竞争或分时来共享传输介质)。当系统负荷及通信量都很大时,系统的处理能力(平均工作处理机数或并行度)并不能随计算机数量的增加而成比例地增加,平均单个计算机的处理能力可能反而还会下降。对总线的这种激烈竞争几乎可使增加处理器个数得到的好处丧失殆尽。

现实世界对计算系统的应用要求比起当前计算系统所提供的能力来说永远是一个无穷大。例如,当前星(或航天器)载计算机、3D虚拟现实、T比特交换及天气预报等应用领域不仅需要高速处理和高速通信,而且还需要高密度(或轻小型化)和低能耗。

1983年美国提出全光数字计算机的构想(当时美苏冷战达到高潮,美制定了“星球大战计划”),1985至1990年间,进行了全光数字处理器件与光互连网络系统的许多探索。1990年1月,AT&T Bell实验室宣布第一台光学数字处理器研制成功。虽然而后因冷战结束、美国停止执行“星球大战计划”而使全光数字计算机系统的研究暂时失去近期目标的吸引和财力的推动,但其中有关光互连网络的一些研究成果却转化和发展成为今天的“并行处理中的光互连技术”,并成为计算机界和光电子界的热门研究领域。在过去的几十年中,美、日、欧等国家和地区在光互连研究上投入了相当的人力、物力和财力,研究成效显著。现在已经有多种光互连实验原型机问世,如日本 NTT的 COSINE系列实验光互连计算机和法国的 SYMPHONIE 系统等。另外,多种光互连器件也已接近实用,如智能点技术及平面微透镜阵列等。同时,集成光互连器件与 VLSI电子器件的单块集成工艺也有了妥善的解决方案,如倒焊工艺等。部分商用巨型计算机中也开始采用光互连技术,如 Cray T90 等。

为了避免光在处理和存储方面的困难,最近,国际上光互连研究正趋向于采用微光电子和微电子混合集成技术把 CMOS 电路和量子阱自光电效应器件(SEED)进行混合集成,从而充分发挥光、电各自的优势。特别是可以充分利用光在局域内的广播特性和在非局域内的高速传输特性。专家们一般认为“处理由电来做,传输由光来做”将是未来高性能并行计算系统结构的主要途径之一。非并行的高带宽传输带来了当今因特网的爆炸式增长。并行的高带宽传输必将使并行计算机系统的性能有相当显著的提高。未来的光互连技术将在大规模并行处理系统、航空航天星载系统及巨型计算机系统中发挥越来越大的作用。

光互连是一种利用各种光传输介质把计算机系统内各部件或各子系统连接起来并通过光来高速传递信息的技术。从光互连所用的传输介质来看,主要有光纤互连、波导互连及自由空间光互连等技术。从光互连在计算机系统结构中所处的层次来看,它可以在计算机与计算机、背板与背板、平面内芯片与芯片及自由空间内芯片与芯片之间等不同层次实现光互连。光作为信息载波,具有独立传播、无接触连接、极高时空带宽积、极高密度(或体积小、重量轻)、等光程、低延迟、低功耗、I/O能力强及抗电磁场干扰等诸多优点。

学术界注重把光互连技术同计算机系统结构结合起来是从1994年的MPPOI(Massively Parallel Processing using Optical Interconnections)国际会议开始的,在此以前,光计算和光互连领域的主流科研工作者几乎都是由光学科学家、物理学家和电子工程师们组成的。本书所做的主要工作是收集、整理和总结二十多年来国际上光互连领域的一些重要研究成果和发展现状,旨在通过本书把光互连技术介绍给我国并行处理研发领域的科研工作者及工程师们,从而推动我国光互连技术和并行处理技术的进步。

本书可作为计算机科学与工程、光电子科学与工程及物理学等领域科研人员的参考书,也可作为上述专业或相关专业博士生、硕士生及大学生的教材。

本书的编著主要是在国家自然科学基金(69873016、69972017)和国家高性能计算基金(984055)的资助下完成的。感谢上述基金委及有关方面的同志们在几年前就非常敏锐地洞察到该领域的重要性并及时给予支持。

本书由李之棠、李非、黄辉龙、赵振龙及解云峰等共同编写,全书由李之棠修改定稿。

光互连技术不仅是光学、物理学及计算机科学的综合交叉性新领域,而且工程性特别强,这些方面并不是我们的长处。由于从计算机学科的角度来进行研究和应用光互连技术在我国才刚刚开始,加之我们水平有限,因此,书中难免有这样或那样的缺点或错误,恳请广大读者批评赐教。

作者

2001年9月于华中科技大学

第1章 光互连与并行处理概论

本章包括:

- ◆ 光互连通信技术
- ◆ 层次抽象与光互连技术
- ◆ 光并行处理中的光互连应用

1.1 光互连通信技术

1.1.1 光互连技术概论

随着高性能计算、高速通信及并行处理等技术的飞速发展,各种高性能计算和超高速交换系统对其内部各元素之间或与外部其他系统之间的通信连接都提出了高密度、高带宽和低能耗的要求。例如,星载计算机、3D虚拟现实、T比特交换及天气预报等应用领域的需求。此外,随着硅和镓-砷等现代半导体技术的发展,处理器的处理速度也已经达到GHz的数量级,使得基于金属传导电信号的电互连技术正在成为许多系统的潜在通信瓶颈。光互连新技术正是在这种背景下产生的。

光互连是一种利用光作为传递信息的载体,实现计算机系统结构内各部件之间或各系统之间互连的技术。从光互连所用的传输介质来看,主要有光纤互连、波导互连及自由空间光互连等技术。从光互连在计算机系统结构中所处的层次来看,它可以在计算机与计算机、背板与背板、平面内芯片与芯片及自由空间内芯片与芯片之间等不同层次实现光互连。光互连技术在通信带宽、等程传输、抗电磁干扰及低能耗等方面比电互连有巨大的优势。

1.1.2 光互连与光交换的必然性

电互连是通过电子在导体或半导体中的传导作用来实现信号传递的,电子信息处理系统的硬件基本单位是电互连构成的逻辑门。电子的特征参数表现为电压或电流,一般用电平的高低来代表二进制的“1”和“0”。依赖于当今的大规模集成电路技术,可在一块很小的芯片上集成复杂度极高、包含近万个或更多的逻辑门系统。然而,由于电子在传导过程中永远存在着R.L.C分布参数,从而使电互连在许多应用领域,特别是那些需要高密度、高带宽和低能耗的领域,存在着难以克服的缺陷。这些缺陷主要是:

- ◆ **带宽受限**: 每一段传导电信号的传输电介质都存在着分布的 R.L.C 参数。从等效电路的观点来看, 这段传输导线相当于一个低通滤波器。当含高频的信号通过时将发生信号严重失真。从而限制了实际有效通信带宽。
- ◆ **时钟歪斜**: 在一个电子系统中, 信号都以“0”或“1”的数字形式流动。若没有一个公共的时钟就不能使整个系统建立时间基准。然而系统时钟从时钟源到各个相互协作的功能模块之间很难做到等距传输。即使路径做到相等, 阻抗也不一定相等。这个问题在高速信息处理系统的板级时钟分布传输时尤其严重, 时钟漂移超过某一限度时, 系统就不能正常工作。
- ◆ **串话严重**: 传输介质在传输信号时, 与邻近的电传输介质存在着电磁效应现象。也就是说, 它既是邻近电传输介质的“发射天线”, 也是邻近电传输介质的“接收天线”。它们相互串扰使噪声功率增大, 系统信噪比急剧下降, 从而将导致系统不能正常工作。
- ◆ **功耗高**: 功耗是集成电路芯片的一个重要参数。信号传输过程的能量耗散主要由两部分构成: 一部分是电传输介质上的能量耗散, 这个能量不是一个常量, 它随器件间距离长短而变。另一部分是互连线两端口上阻抗的能量耗散。当电互连线路在 10mm 左右时, 耗散能量可达 500pJ。在高频信号情况下, 电互连路由上所耗能量会急剧增加, 其辐射能量正比于频率的四次方。当用电集成电路芯片来实现互连时, 芯片内数目众多的互连线将占据芯片大部分面积, 芯片复杂度极高, 功耗已成为一大难题。
- ◆ **I/O 能力受限**: 集成电路芯片 I/O 能力受限是指 I/O 端口数目受限。这里有三个方面的原因: 第一是功耗。驱动每个 I/O 端口工作在一定的速率, 必定耗费比较大的功率, 因此 I/O 端口数目不能多, I/O 端口的速率也不能太高 (根本原因是带宽)。第二是引线的互连模式, 大多数集成电路芯片的互连引线是在环绕芯片周长的二维平面上, 因此它的互连引脚数目受限。第三是集成电路芯片互连的物理实现工艺的限制。

光互连以光作为传递信息的载体, 能弥补电互连的以上缺陷, 有望彻底解决高性能计算和超高速交换系统中普遍存在的通信瓶颈问题, 实现通信系统中的大容量、高速率、低能耗的数据交换。光互连的主要优势如下:

- ◆ **超高时空带宽积**: 目前, 实验室一个单波长光通道的数据传输速率已达到 10^{12} b/s (Terabit/s) 数量级, 实用速率在 10Gb/s ($1\text{G}=10^9$)。光除了空分复用、时分复用、码分复用外, 还有独特的波分复用 WDM (Wavelength Division Multiplexing) 方式。在不增加太多硬件投资的情形下, 波分复用可使系统容量以波长数目 N 的倍数增加。
- ◆ **极低延时**: 几百公里长光纤通道的延迟时间一般在 ms 数量级, 波导和自由空间光的延迟几乎可以忽略。
- ◆ **通道等程**: 光传播不存在电传导中的电阻问题, 只要等程, 即可解决时钟歪斜问题。实际上, 许多光互连系统都处在透镜的物平面与像平面的对应列阵之间, 光学成像系统保证了这种等程性。
- ◆ **无干扰**: 光的传播遵循独立传播原理, 上亿束光束在传播过程中无论怎样相交也不会相互干扰。
- ◆ **低能耗**: 用光互连直接互连芯片表面能减少芯片上的耗散功率。光传递信息的机理是量子阻抗变换, 这种变换的光功耗极低, 而且不随信号传输速率的增加而增加。

- ◆ **I/O能力强**: 光 I/O 端口不仅速率高, 而且可以制成很多端口。这是因为, 一方面光的时间带宽高, 可实现大规模的面阵互连。另一方面光的空间带宽高, 传播时相互间无串扰, 又不受外部电磁波的辐射干扰, 故可实现高密度、大规模的互连。

1.1.3 光互连技术的发展现状

在过去的几十年中, 美、日、欧等国家和地区在光互连研究上投入了相当的人力、物力和财力, 研究成效显著。现在已经有多种光互连实验原型机问世, 如日本 NTT 的 COSINE 系列实验光互连计算机和法国的 SYMPHONIE 系统等。另外, 多种光互连器件也已接近实用, 如智能点技术 (Smart-Pixels)、平面微透镜阵列等; 同时, 集成光互连器件与 VLSI 电子器件的单块集成工艺已经有了妥善的解决方案, 如倒焊工艺等。部分商用巨型计算机中也采用了简单的光互连技术, 如 Cray T90 等。可以说, 从目前的发展形势看, 近期内光互连技术在巨型计算机系统和大规模并行处理系统中将得到大量的使用。

光纤是光互连技术中一种很重要的通信物理介质。最早从 1975 年就开始使用, 此后, 它的发展令人惊奇。光纤是世界范围内因特网成功发展的关键原因之一, 也是在世纪之交推动人类进入信息化社会的重要技术动力之一。

通信用光纤产品的全球消费额从 1975 年的 250 万美元增长到 2000 年的 158 亿美元。目前虽然增速变缓, 但这种增长还在继续, 预计在 20 年后将达 7000 亿美元左右。与光纤通信相关的软件也在相应地增长。光纤产品、电子产品以及其他产品加上软件, 将使全球通信设备市场在 2025 年左右达到数万亿美元。表 1-1 列出了全球光器件消费情况及对未来的预测。

有源光电子器件在全球光纤产品总市场份额中所占比重将上升 40% (62.9 亿美元)。这其中包括光发射器、光接收器、光收发模块、光放大器以及其他有源器件。预计到 2025 年将达到 52% (3830 亿美元)。光缆铺设量的增长将变慢, 所占份额将降至 5%。一些目前市场上还微不足道的产品以及很多还未问世的产品, 预计将在 2025 年时占有全球光通信产品消费市场的 24% 的份额。表 1-2 显示了光通信器件全球消费趋势。

表 1-1 全球光器件消费情况一览表

器件类型	1975		2000		2025		平均年增长率 %	
	\$ 百万	%	\$ 十亿	%	\$ 十亿	%	1975-2000	2000-2025
光缆	1.78	71	4.51	28	40.36	5	37	9
有源器件	0.4	16	6.29	40	383.1	52	47	18
无源器件	0.16	6	1.45	9	140.52	19	44	20
其他有源、无源器件	0.15	6	3.59	23	175	24	50	17
总消费额	2.49	100	15.84	100	738.98	100	42	17

表 1-2 光通信器件全球消费趋势

器件 (\$ 百万)	1975	1985	1995	2000	2005	2015	2025
合波器/分波器	0.05	6.2	55	125	289	1 510	5 320
光开关	0	0	390	2 050	4 330	14 100	38 300
滤波器	0	0	40	340	890	5 500	25 300
光放大器	0	3.9	55	505	2 390	15 800	86 500

大规模使用光纤产品是全球宽带通信网络飞速发展的有力基础。网络的扩展又将带来全球性传输业务收入的大增长,这些业务需求包括Internet的蓬勃发展,大量的全球数据传输等。没有带宽的迅速增长,这些大量的收入增长是不可能的。由于传输成本的迅速下降,一系列与带宽有关的产品和业务将很快展开并迅速增长。其中包括:

- ◆ **视频娱乐节目:** 采用速率高达几十兆比特的数字电视提供与实物一样大小的高分辨率、3D、真彩色视频娱乐节目。
- ◆ **可视电话:** 到2025年,全球将有一亿户以上的家庭装有带大型3D彩色屏幕的可视电话。
- ◆ **视频会议:** 通过因特网提供桌面或便携机的一对一型或组对组型电视会议系统。
- ◆ **大量、即时及连续的全球数据传送:** 允许几百万个不同规模的公司以及部门内部的各个相互连接的高端工作站之间进行数据通信。
- ◆ **高速数据采集:** 来自工业、民用、商用和军用的各种检测仪表的大量实时数据采集。
- ◆ **全球网络:** 少数几个全球性的用户到用户通信公司将控制全球通信市场。移动通信网、卫星长途网和有线电视网络将集成为多功能的用户到用户的综合性网络。采用便携式收发机,可以以低廉的成本实现在任何地点、任意时间及对任何人的通信(包括声音、数据、图像和视频)。

在过去10年中,光互连器件的性能取得了重大进展。光纤放大器和DWDM系统已被广泛采用。激光器模块的速率已从622Mb/s倍增至10Gb/s,速率为40Gb/s的激光器模块也已处于初期试用阶段,光纤的性能也有所提高。这种趋势将继续下去,单模光纤的可利用传输带宽将超过400nm波长带宽,能够容纳1000条以上的波长通路,每个波长的调制速率可达160Gb/s。其他的技术开发成果如下:

- ◆ **透明的光交叉连接开关:** 能支持即时的业务流重配置和抗灾害恢复能力,并且能够扩展到 1024×1024 通路的无阻塞互连。
- ◆ **不断地由分立元件演进到集成“超级元件”、混合光电子集成电路(OEIC)及单片OEIC,** 并逐步提高性能和降低成本。
- ◆ **全光网(AON):** 横跨各大洲,高速数据传输、高密度DWDM网状网。
- ◆ **全光分组式交换机,** 含因特网路由器。
- ◆ **通过副载波调制增加承载能力。**
- ◆ **在400nm波长带宽内,** 通过采用色散管理,光纤可支持的信号传输速率可达160Gb/s。
- ◆ **预计在未来的20年左右时间内,** 最具增长活力的光纤器件将是无源光器件。它们的需求将以每年20%的速率递增。这些无源器件包括连接器、合波器/分波器、滤波器和光开关等。

光缆的铺设量在带宽革命的初期年增长率都超过了100%,后来逐渐下降。这种趋势将继续下去,从2000年到2025年的平均增长率预计只有9%。

网络将继续朝着全球用户到用户的方向演进;通过公用光缆,逐渐合并成少数几个全球性的大网;实现任何人、任何地点、任意时间从设备到设备之间的语音、视频和数据业务的

互连;全球通信吞吐量(千兆比特×公里)将比2000年扩大10000倍以上。Petabit链路将变得很普遍。

尽管世界各地的大型研究机构都已演示了OEIC的优异性能(尤其是在20世纪80年代末期),这种技术至今未达到大量生产的程度。因此在光纤器件方面,我们还处在采用分立元件和“固定螺栓”的组装的阶段。然而产品数量和性能的限制将促使人们逐步采用OEIC产品。许多厂家现在正转向大规模生产混合OEIC收发器,这将对今后的产品设计产生重大影响。预计到2025年,2.5Gb/s收发器的价格将降至10美元以下。

光电子集成将由混合型向单片多通路型发展。采用亚纳米技术的高集成度单片光开关,其每一开头点的售价可望低于1美元。高吞吐量的光网络交换机(ATM、数字交叉连接器、因特网路由器)将由光-电-光结构演进为透明的光通路交换机^[6]。

1.2 层次抽象与光互连技术

计算机系统的设计可在几个抽象层上进行,包括从底层的物理设备到系统结构再到较高的软件等几个层次。将计算机系统分为多级抽象层的优势在于,它能使设计问题分解成几个易于解决的分问题。然而,要在多级抽象层中引入光电设备这样的概念和技术十分困难,因为它们与目前系统框架并不很匹配。为此,我们将对计算机的多级抽象层系统进行研究,它为什么以及怎样阻碍光技术的引入?为了让它有效适应光技术又该怎样修正它呢?

在以下三种情况下,计算机系统可望在不久的将来成功引入光电技术:一是目前多级抽象层系统不能满足不断变化的技术和应用的需求并导致严重的瓶颈效应,而引入光技术将消除这些瓶颈,如互连及存取方面的瓶颈等。二是仅含少量抽象层的专业应用将得到认同和支持,如感观及图像处理等。三是能够实现引入光技术抽象层以上各级抽象层,以致光技术能平滑地移植到目前多级抽象层系统中去,如修正通信框架或标准以匹配光纤交换系统的容量及采用并行体系结构以适应光纤分系统所生成的并行信息流。

1.2.1 解决问题的抽象层次

为了设计具有更强大计算能力的计算机,大量的原子计算元素(如交换开关、晶体管、半导体、门及微处理单元等)集合需要以更高的长/宽比来进行互连^[1]。当互连距离增大时,用于将信号传输到另一个端点的时间也随之增长,同时传输速率最高也不超过光速。这使得各种不同互连都会受到限制,而普通电传导互连更是如此。当传输超过一定长/宽比时,信号时延随着长/宽比的平方而增加。线路损耗太大会导致信号不能很好地传输,即使采用转发器,每传输一位数据所需的能量也会随着线路距离增长而增加^[2, 3]。

由于光在非二维空间互连、电绝缘、无串扰和失真及多点连接的阻抗匹配等方面存在的优点,因此在计算机系统中可采用光互连来实现长距离互连,特别是用它替代电子线路后会有很高的长/宽比。

一般说来,用光纤互连来代替电导互连有着毋庸置疑的先进性,许多分析和定量研究以及一些技术权威们也都肯定了采用光互连所带来的先进之处,然而遗憾的是,光技术融合进主流计算系统的进展非常缓慢。