

· 专题二 ·

高速信息网络与并行处理

与会专家提出如下优先发展领域与建议:

1. 高速信息系统理论与体系结构:大系统理论,网络的可靠性、安全保密性,多种通信系统的综合,网络模型、控制与管理,协议与标准,神经网络理论应用,人与网的行为影响研究。
2. 信息传输:同步数字系列及高速光纤通信系统,高速宽带用户环路及信息压缩技术,低轨道个人卫星通信系统,移动通信新技术研究。
3. 交换与网络:ATM与B-ISDN技术,光交换高速计算机网络,异构网的互连技术,信息网业务提供与智能网技术。
4. 信息网应用系统:大容量应用系统基础理论,教育科研示范网及试验网,智能信息处理与决策系统,CATV网的综合利用技术。
5. 高速网络上的并行计算技术。

高速网络并行计算软件环境与算法研究

孙家昶

(科学与工程计算国家重点实验室,中国科学院计算中心,北京中关村,100080)

1 美国“高性能计算与通讯”

美国为保持其在高性能计算与计算机通讯领域中的世界领先地位,增强国民经济、国家安全、教育及整体环境的竞争能力,于1991年度开始执行“高性能计算与通讯”(High Performance Computing and Communication, HPCC)——美国总统科学战略项目。该项目由四部分组成:(1)高性能计算机系统(HPCS),内容包括今后几代计算系统的研究、系统设计工具、先进的典型系统及原有系统的评价等;(2)先进软件技术与算法(ASTA),内容有巨大挑战问题的软件支撑、新算法设计、软件分支与工具、计算技术及高性能计算研究中心等;(3)国家科研与教育网(NREN),内容有中接站及十亿位级传输的研究与开发;(4)基本研究与人类资源(BRHR),内容有基础研究,研究参与及培训,教育、培训及课程教材。

HPCC具体目标是:(1)对一大批重要的应用课题,计算性能要达到每秒万亿次运算(10^{12} ,或称teraops);(2)发展有关的系统软件工具,对一大批问题改进算法;(3)国家研究网能力要达到每秒十亿位(10^9 位,或称gigabit);(4)充分保证计算机科学领域与工程领域中博士生与其他训练有素的科研人员的需求。

已具备的条件有:已宣布可扩展到100 gigaosp能力的计算系统;90年代中期可达到teraops;已开发 10^{15} 位内存系统;已开发操作系统技术,可有效地支持异构分布式应用,系统配

国家自然科学基金及中国科学院“八五”重点项目

置中可包括高性能系统与工作站。

1993年 HPCC 的“巨大挑战”问题有：(1) 磁记录技术。要在一平方米的磁盘表面上压缩记录十亿位数据；(2) 新药研制。特别是癌症与爱滋病药的研制；(3) 高速城市交通。新型低噪音飞机的研制，空气动力学的计算；(4) 催化剂设计。改变迄今为止多数催化剂靠经验设计的习惯，转向计算机辅助设计；主要发展分析这些复杂系统的大规模量子化学模型；(5) 燃料燃烧原理。研制新型发动机；(6) 海洋模型模拟。整体海洋模拟；(7) 臭氧层空洞；(8) 数字解剖。如三维 CT 扫描图象处理，人脑主体模型，三维生物结构与四维时间结构；(9) 空气污染。计算模拟能提供有效控制污染的传播，揭示其物理与化学机理；(10) 蛋白质结构设计。促进计算生物学研究；(11) 金星图像分析。麦哲伦空间发回的有关金星资料已达 3×10^{12} 位需要处理；(12) 密码破译技术。

美国国家自然科学基金会(NSF)在 HPCC 中的分工：(1) HPCS 基本体系结构研究：典型实验系统；(2) ASTA 软件工具与数据库研究：“巨大挑战”问题，计算机 Access；(3) NREN：Facilities coordination and deployment, gigabit 研究；(4) BRHR：基础研究，教育、培训及课程。

应该指出，在 1993 年度 HPCC 项目预算中“先进软件技术与算法”投资已大大超过硬件投资，这是因为“高性能计算”的主要关键技术已转向先进软件与算法。

2 高速网络上的并行计算环境

硬件环境有各种类型的先进工作站(高分辨率彩色显示器、高性能网络、分散的局部高性能计算系统。软件环境与软件工具有系统软件(支持异构配置)和应用软件。软件库有编程环境与工具、高级语言、先进编译技术、优化工具、并行化工具、数据管理、可视化工具、调试与运行测试工具、多功能软件工具的集成。

2.1 高性能计算机技术与市场发展趋势 传统大型机与向量巨型机技术已经接近极限，成本高，3—5年内将逐渐被淘汰。大规模并行处理系统 MPP 和可扩展(Scalable)并行处理系统 SPP；技术将日趋成熟，性能价格比好，但缺乏应用软件。MPP 的峰值与应用程序测试距离极大，可扩展工作站机群。所有石油公司都表示，未来的计算机密集型工作将主要由 MPP/SPP 和工作站机群在承担。大量基于 MPP/SPP 工作站机群的应用软件正在开发。

2.2 工作站并行机群系统是并行处理系统发展的一个重要趋势 它具有高度的灵活性，能充分利用现有计算机资源以及性能价格比高等优势。由于工作站性能迅速增长，加上高速光纤网(FDDI)以及即将推出的高速以太网(Fast Ethernet)和异步传输模式(ATM)，极大地增加了网络带宽，缓解了网络通信瓶颈。同时，由于一批异构并行虚拟处理平台(如 PVM, MPI, Express, LINDA 等)及其基于这些应用平台的一批并行软件问世，可以预料，工作站机群在近期内可能成为大型科学与工程计算的一个重要手段。

可扩展的并行机群系统适合我国国情，在国际上即将成为并行处理与应用的重要发展方向，有着广泛的应用价值与市场前景。目前急需研制开发一批基于国际公认标准平台上的可扩展的应用并行软件。这项工作国际上也刚起步，这对我国是一个很好的机遇，也是一个挑战。

国际上从 90 年代初陆续推出异构计算机网络计算环境(HeNCE)及相应的并行处理开发平台，并在继续完善和更新。现在 PVM(Parallel Virtual Machine)愈来愈被诸多计算机公司承认为公用平台。Oak Ridge National Laboratory 目前已发布 PVM 3.2.5 版。一些国际杂志如《Parallel Computing》等从 1993 年下半年起已开始刊登有关工作站机群并行的应用分析文章。

我国一些部门也已对工作站机群上并行应用软件表示出强烈的兴趣。它对于我国的大型科学与工程计算,部分转向以先进工作站机群为主的并行系统将起重要作用,使能充分利用现有工作站资源,便于进一步升级扩充,为国家节约大量资金。同时,对于并行软件的模块化、标准化、系列化、与国际标准接轨,对于今后开发能适应快速多变的计算机市场的并行软件商品,有重要价值。

3 消息传递环境(Message Passing Environment)

高速网络上的通讯与多数的 MPP(Massive Parallel Processing)系统都是消息传递型(M. P.),所有的 MP 系统在逻辑上都是相似的。由于同步化机制不同,数据交换不同,各公司产品不兼容。

3.1 可移植的消息传递编程环境 其主要做法是选择特定环境,在所有硬件平台上实现,使之在一个可移植编程环境下写的程序可以在所有的消息传递环境下运行。异构环境对于网络计算是至关重要的,高速网络配置多种硬件环境。

3.2 消息传递标准环境(MPI) 不同的可移植平台有不同的用户界面,其中多数可移植平台只支持部分用户系统功能,所以在 1992 年成立了 MPI(Message Passing Interface)委员会,以确定消息传递界面的新标准,目的在于支持最佳的可移植平台。目前已发布 MPI 定义, MPI 的可移植平台正在实现中,很有可能将形成一个类似于 FORTRAN 77 那样的国际消息传递标准环境平台。《Parallel Computing》今年四月份专门刊登有关消息传递环境的专辑,这是一个很值得我们重视的技术动向。

4 高速网上的应用软件与算法研究

目前网络并行计算技术的发展主要受两方面因素的制约:一是网络通信速率的限制;二是编制应用程序比较麻烦。第一个困难随着网络通信技术的发展会在一定程度上得到缓解。解决第二个困难的方法之一则是开发出面向普通用户的、基于标准网络计算平台之上的、具有标准接口的基本并行程序包(如线性代数、快速 Fourier 变换等等)作为应用开发平台,从而提高应用并行程序的开发和运行效率,并且简化已有串行软件的移植。

4.1 应用软件 缺少有效的可移植的并行应用软件是目前网络并行计算的一个瓶颈。各个并行机公司都已宣布正在大力研制应用软件,已有的并行软件涉及到的主要应用领域有:航天航空、汽车设计、化学、生物学、计算机图形与图像处理、电子学/电子工程及石油工业等。

4.2 高质量的并行数学库 在开发应用软件以前,必须先建立高质量的并行数学库,包括:(1)基本数学库,如 FFT、BLAS、基本函数;(2)基本通讯功能库(处理各种网络结构下的通讯,负载平衡与动态处理器分配等);(3)并行通用数学库,如 LINPACK、非线性方程组求解、微分代数方程组、随机过程、统计计算及偏微分方程数值解;(4)算法研究所涉及到的数学分支,如计算数论、计算代数、计算几何及计算方法等;(5)应用研究。

4.3 可扩展、可移植的高效算法 解决重大“挑战问题”最重要的两个关键问题是设计出可扩展的体系结构和设计出可扩展的算法。

4.4 基础研究 包括并行计算模型研究和高观点的并行计算方法研究。有两种可供选择的做法,即在现有串行算法上的并行化设计技术,或直接形成可扩展、可移植的高效算法。后者从计算模型出发,直接面向并行系统、计算方法与设计算法。

5 几点看法与建议

并行处理与高速信息网络密不可分。为适合我国国情,“九五”期间可优先发展可扩展的高速网络工作站机群。我国的高速网络要为高性能计算服务,要改变以往的“重硬轻软”的倾向,软件建设(包括软件环境与算法研究)应在我国的高速网络项目中处于重要位置。以下向国家自然科学基金会提出几点建议。

关于有限目标的建议。通过建设我国的高速信息网络,逐步把“计算”作为我国的自然科学研究的三个重要手段(理论,实验,计算)之一,力争在“九五”期间,对于科学与工程中重大的“挑战性”问题做到解决一批(计算数学,计算物理,计算力学,石油模拟,CAD,计算机图形学等),推动一片(计算化学,计算生物,石油勘探,环境科学等)。

“九五”期间“有所赶”的优先资助领域,建议包括:网络并行处理开发工具;网络并行测试系统;可移植的网络标准数学库建设;分布式并行计算方法基础研究(并行计算模型,并行迭代,并行差分格式理论,分布式数据库等);若干重要应用并行算法与软件的研制;分布式并行计算的教学环境与培训。

在措施上,参与国际网,利用国际公共网资源(如 NETLIB, XNETLIB 等),追踪国际先进技术。逐步建立我国的公共网资源,减少重复研究。强调软件研制的可扩展性与可移植性,便于国际交流与商品化。尽快建设我国的科研教育高速网,提供必要的并行环境,培养人才,利用国家实验室已有的和即将引进的高性能工作站及中关村地区网上的中高档工作站或并行机,进行几十个到几百个节点的中、大规模并行算法的研究、测试,以测定机群并行系统的可扩展性与实用效果。加强网络并行计算的基础研究。

参考文献

- [1]McBryan O. An overview of message passing environments. *Parallel Computing*, 20(1994), April, 417-444
- [2]Grand Challenges 1993: High Performance Computing and Communications
- [3]Kuck D J. What do users of parallel computer systems really need? *International Journal of Parallel Programming*, 22 (1994), Feb., 99-127

PARALLEL COMPUTING SOFTWARE ENVIRONMENT AND ALGORITHM RESEARCH ON HIGH SPEED NETWORKS

Sun Jiachang

(The Key Laboratory of Science and Engineering Computing, Computing Center, CAS, Beijing 100080, China)