

HEAVY MACHINERY

重型机械行业网格 体系结构及其关键技术

曹丽君 张忠平 著



科学出版社

重型机械行业网络 体系结构及其关键技术

曹丽君 张忠平 著

科 学 出 版 社

北 京

内 容 简 介

本书针对重型机械网格体系结构及其资源管理、调度等关键技术进行了深入研究。书中将网格技术和重型机械行业相结合,提出了基于多层覆盖网络结构的网格资源发现机制及资源发现算法、基于模糊聚类的网格资源发现机制、基于最小代价函数的网格资源发现模型及其算法、基于复制和插入的启发式任务调度算法、决策树型层次网格资源组织模型及调度算法,为建立行业网格的研究提供了一种新的理论方法和途径,并通过对资源管理、资源发现、任务调度等关键技术的研究,为其他领域研究基于网格技术的应用提供借鉴意义。

本书具有定位准确、内容先进、取舍合理、体系得当、风格鲜明等特点,适合于各级计算机专业学生及相关专业学者对网格技术的学习和研究。

图书在版编目(CIP)数据

重型机械行业网格体系结构及其关键技术/曹丽君,张忠平著. —北京:科学出版社, 2014

ISBN 978-7-03-041344-4

I. ①重… II. ①曹… ②张… III. ①计算机网络-应用-重工业-机械工业-研究 IV. ①F407.42

中国版本图书馆CIP数据核字(2014)第147889号

责任编辑:郭丽娜 责任校对:刘玉靖

责任印制:吕春珉 / 封面设计:耕者设计工作室

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

http://www.sciencep.com

厚 诚 则 铭 印刷

科学出版社发行 各地新华书店经销

*

2014年7月第 一 版 开本:B5(720×1000)

2014年7月第一次印刷 印张:13

字数:260 000

定价:65.00元

(如有印装质量问题,我社负责调换<厚诚则铭>)

销售部电话 010-62134988 编辑部电话 010-62138978-2011

版权所有,侵权必究

举报电话:010-64030229; 010-64034315; 13501151303

前 言

网络是把整个因特网整合成一台巨大的超级虚拟计算机,实现互联网上所有资源的互联互通,完成计算资源、存储资源、通信资源、软件资源、信息资源、知识资源等智能共享的一种新兴的技术。

虽然网格技术已经在科学研究的诸多领域中得到应用,但由于网格所提供的强大功能和本身固有的复杂性,必然为网格系统的理论研究和应用实现带来前所未有的挑战。因此,本书针对重型机械网格体系结构及其关键技术进行了深入研究,其目的是整合现有重型机械行业的资源,实现重型机械领域的资源共享和信息服务。

本书共 8 章,分别为绪论,重型机械行业网格体系结构设计、基于多层覆盖网络结构的网格资源发现机制、基于多层覆盖网络结构的资源发现算法、基于模糊聚类的网格资源发现机制、基于最小代价函数的网格资源发现模型及其算法、基于复制和插入的启发式任务调度算法、决策树型层次网格资源组织模型及调度算法。

第 1 章为绪论。阐述了网格技术的起源、发展和基本特点,分析了国内外的网格项目、网格体系结构、网格资源管理和网格应用领域的研究现状;论述了重型机械行业信息化面临的挑战和行业网格建设的意义。

第 2 章为重型机械行业网格体系结构设计。重型机械行业中数据密集、计算密集、资源分布、异构等问题制约了行业应用和信息化发展,网格是新一代高性能计算环境和信息服务基础设施,能够实现动态跨地域的资源共享和协同工作。结合网格技术和重型机械行业特点,给出重型机械行业网格设计的理论原则和方法,并在此基础上设计重型机械行业网格体系结构。面向重型机械行业网格(Heavy Machine Industry Grid, HIMG)平台是基于网格技术,以“服务”为中心,集成、处理重型机械行业数据并向社会提供信息服务的平台。HMG 构建了从信息获取、集成、处理到应用的完整的 7 层体系结构,从数据库层、算法工具层、分析逻辑层到应用系统层 4 层资源管理逻辑模型,同时结合行业应用需求,着重研究资源管理、资源交易、服务链等关键技术,实现重型机械领域的资源共享和信息服务。

第 3 章为基于多层覆盖网络结构的资源发现机制。网格资源发现机制是目前网格领域中的一个热点研究问题。对现有的网格资源发现方法进行了分析,指出了这些方法存在的不足,提出了一种基于多层覆盖网络结构的资源发现机制,将具有同一类型资源的网格信息结点组织在一起形成域,把结点数量较多的域通过类型粒度组织成层次结构。利用类型匹配路由的资源搜索技术将资源发现请求转发到资源所在的域,采用 Topk 技术在相应的域内找出与请求匹配的最佳的 k 个资源,从而改善了资源发现的性能,提高了用户的满意度。

第4章为基于多层覆盖网格结构的资源发现算法。本章提出了网格资源发现系统的资源组织模型,提出了一种基于MLON结构的资源发现算法,针对目前资源发现方法主要是面向资源属性精确查询和资源单属性查询的缺点,根据MLON资源组织模型,提出了一种新的网格资源发现算法——MLON-SSA资源发现算法;通过GridSim实现相关算法的仿真。

第5章为基于模糊聚类的网格资源发现机制。首先分析了模糊聚类分析基础知识和网格资源发现机制所涉及的网格资源管理的基础理论知识,然后对资源发现进行了论述和深入研究,提出了网格资源发现系统的资源组织模型和一种基于网格重组模型的资源发现算法。

第6章为基于最小代价函数的网格资源发现模型及其算法。从网格资源的动态性和可重复使用性的角度共同出发提出了一种新的网格资源分类方法;提出了一种基于最小代价函数的网格资源发现模型,构造了基于最小代价函数的网格资源发现算法、资源路由结点选择算法、资源注册算法和结点退出算法;通过GridSim实现相关算法的仿真。

第7章为基于复制和插入的启发式任务调度算法。针对网格环境中复杂的计算密集型的并行分布式应用,提出了一种新的基于复制和插入的任务调度算法(Duplication and Insertion Based on Scheduling, DIBS)。该算法可以同时执行多个应用程序,利用决定路径对任务进行排序,缩短了应用程序总的执行时间。实验结果表明,该算法更加符合网格的复杂环境,能够更好地满足不同用户的实际需要。

第8章为决策树型层次网格资源组织模型及调度算法。针对网格资源具有动态性、异构性、分布性等特征,提出了决策树型层次网格资源组织模型,并在该模型的基础之上执行网格资源发现过程发现资源。通过资源发现的结果集研究网格资源调度算法,针对Min-Min调度算法存在的不足,提出了OPT-Min-Min调度算法。

本书由河北科技师范学院学术著作出版基金资助,由河北科技师范学院曹丽君副教授统稿并编写第3~8章(共21万字),燕山大学张忠平教授审稿并编写第1章和第2章(共5万字)。

对于本书的编写,作者尽量体现自己的科研思维和成果,虽竭尽全力,但因能力和水平有限,书中难免存在疏漏和错误之处,希望广大读者批评指正。

曹丽君

于河北科技师范学院

张忠平

于燕山大学

2014年6月

目 录

第 1 章 绪论	1
1.1 引言	1
1.1.1 网格的起源与发展	1
1.1.2 网格的基本特点	4
1.2 网格的研究现状	5
1.2.1 国内外的网格研究项目	5
1.2.2 网格体系结构的研究现状	9
1.2.3 网格资源管理的研究现状	12
1.2.4 网格的应用	15
1.3 重型机械行业网格的研究内容和研究意义	17
1.3.1 研究意义	17
1.3.2 研究内容	18
第 2 章 重型机械行业网格体系结构设计	20
2.1 重型机械行业网格设计理论原则与方法	20
2.1.1 设计理论原则	20
2.1.2 设计方法	21
2.2 重型机械行业网格概念与技术	25
2.3 重型机械行业网格框架与体系结构	25
2.4 重型机械行业网格关键技术	26
2.4.1 管理技术	27
2.4.2 资源信息管理	27
2.4.3 资源交易服务	30
2.4.4 服务链模型	31
2.5 本章小结	32
第 3 章 基于多层覆盖网络结构的网格资源发现机制	34
3.1 资源发现机制的重要地位	34
3.1.1 网格资源发现研究的意义	34
3.1.2 网格资源发现研究现状	35
3.2 网格中的资源管理	38
3.2.1 网格环境中的资源	38
3.2.2 网格资源管理	40

3.2.3	网格资源发现与定位	43
3.3	多层覆盖网络资源组织机制	44
3.3.1	资源发现及资源组织模型	45
3.3.2	多层覆盖网络资源组织机制	53
3.3.3	多层覆盖网络的构造与维护	56
3.3.4	模型分析	61
3.4	本章小结	61
第 4 章	基于多层覆盖网络结构的资源发现算法	62
4.1	资源发现相关工作	62
4.1.1	资源发现机制的原则	62
4.1.2	欧式空间	63
4.1.3	Top- k 查询	64
4.2	基于多层覆盖网络结构的资源发现算法	65
4.2.1	问题描述	65
4.2.2	匹配策略	65
4.2.3	多维属性查询转化	66
4.2.4	MLON-SSA 算法	68
4.2.5	算法分析	70
4.3	基于 GridSim 的仿真实验	70
4.3.1	GridSim 仿真包及结构	71
4.3.2	仿真结果及评价	74
4.4	本章小结	80
第 5 章	基于模糊聚类的网格资源发现机制	81
5.1	网格资源重组模型设计	81
5.1.1	设计目标	81
5.1.2	基于等价关系的网格资源模糊聚类分析	81
5.1.3	GRRM 模型设计	89
5.1.4	模型分析	93
5.2	基于模糊聚类的网格资源发现机制	93
5.2.1	资源发现与资源查找	94
5.2.2	基于 GRRM 的网格资源发现方法	95
5.2.3	算法分析	98
5.3	GRRM-SSA 资源发现算法仿真实验	99
5.3.1	仿真实现过程	99
5.3.2	仿真结果分析	103
5.4	本章小结	106

第 6 章 基于最小代价函数的网格资源 发现模型及其算法	108
6.1 基于最小代价函数的网格资源发现模型	108
6.1.1 模型中的基本定义	108
6.1.2 模型中的主要函数及其功能	110
6.1.3 模型的工作原理	114
6.2 基于最小代价函数的网格资源发现算法	117
6.2.1 资源路由节点选择算法	117
6.2.2 资源注册算法	120
6.2.3 节点退出算法	122
6.2.4 资源发现算法	123
6.3 本章小结	128
第 7 章 基于复制和插入的启发式任务调度算法	130
7.1 调度算法的重要地位	130
7.2 网格中的任务调度	134
7.2.1 网格任务调度原理	134
7.2.2 网格任务调度的体系结构	136
7.2.3 网格任务的调度过程	137
7.3 协同分配算法	139
7.3.1 关联任务的调度算法	140
7.3.2 问题描述	140
7.3.3 动态协同资源算法	142
7.3.4 性能分析	144
7.4 基于复制的网格任务调度算法	145
7.4.1 基于 DAG 模型的任务调度算法	145
7.4.2 静态调度中的任务复制技术	151
7.4.3 DIBS 调度算法	154
7.5 LHCNF 算法及 DIBS 调度算法的仿真实验	159
7.5.1 仿真包的使用方法	159
7.5.2 仿真实现过程及结果分析	159
7.6 本章小结	163
第 8 章 决策树型层次网格资源组织模型及调度算法	165
8.1 决策树型层次网格资源组织模型	165
8.1.1 相关技术	165
8.1.2 决策树型层次网格资源组织模型	169
8.1.3 模型中的主要函数	171
8.1.4 模型维护算法	172

8.1.5	网格资源发现	175
8.2	Min-Min 算法改进	176
8.2.1	网格资源调度	176
8.2.2	Min-Min 网格资源调度算法	179
8.2.3	OPT-Min-Min 调度算法	182
8.3	OPT-Min-Min 算法实现和性能分析	187
8.3.1	实验基准和实验环境	187
8.3.2	ETC 矩阵的生成	189
8.3.3	算法性能评估	190
8.3.4	综合性能分析	193
8.4	本章小结	194
参考文献		195

第 1 章 绪 论

1.1 引 言

网格（Grid）一词在 20 世纪 90 年代中期首次被用来描述科学和工程分布计算的基础设施。这种基础设施把计算资源、数据存储设施、广域网络、仪器设备等连成有机的整体，方便用户使用这个基础设施中的任何资源^[1]。网格是继万维网之后出现的一种新型网络计算平台，目的是为用户提供一种包括网页在内的各种资源全面共享的基础设施。网格将分布的资源集成为一台能力巨大的超级计算机，提供计算资源、存储资源、数据资源、信息资源、知识资源、专家资源、设备资源的全面共享。资源共享是网格的根本特征，消除信息孤岛是网格的奋斗目标，整合网络中的信息孤岛，在世界范围内形成一个开放的、一体化的新型网络计算平台，即用全球大网格（Great Global Grid, GGG，也称为 3G）来全面共享各类资源和支持协同工作，以满足人类日益增长的服务需求。

在当今社会，网格作为一种计算和数据管理的基础设施，能为商业活动、政府行为、科学研究以及大众娱乐等诸多领域提供信息化支持。《网格：一种未来计算基础设施蓝图》一书将网格描述为“网格是构筑在互联网上的一组新兴技术，它将高速互联网、计算机、大型数据库、传感器、远程设备等融为一体，为科技人员和普通老百姓提供更多的资源、功能和服务。互联网主要为人们提供电子邮件、网页浏览等通信功能，而网格则能提供更多、更强的功能，它让人们透明地使用计算资源、存储资源和其他资源，旨在支持虚拟组织的资源共享与协同工作^[2]。”网格基础设施能够将各种资源动态地组织起来，帮助人们解决大规模、资源密集、分布式的应用问题。在英国，人们提出了一个新术语 e-Science（大意就是基于新一代信息基础设施的科学研究），它代表了通过全球协作方式进行科学研究的新观点。从对网格的描述中可以看出，网格技术就代表了一种资源共享和协同工作的理念。

1.1.1 网格的起源与发展

生活在这个时代的每个人都比较熟悉与我们生活息息相关的供电系统，通过输电线路，电力资源被送到我们的家里、办公室、教室以及其他的工作和生活场所。

当我们需要使用电力的时候，只要把电器的电源插头接入电网的插座就可以了。不管我们使用的电器是冰箱、空调、电视机，还是充电器、电炉、电热毯，

不管电器的内部设施简单还是复杂，它们都可以自动使用电力资源，为人类提供它们自己所具有的功能。至于电网中的电力资源来自于什么地方，是采用什么原料发电的，经过什么电路输送到用户的插座等问题，用户根本不需要知道，用户要做的事情就是把设备的插头插入插座，付费用电而已。因此，网络就是把电网作为比拟的对象开始建立和发展的，希望成熟的网络如同电网一样，用户只要把设备的插头插入网络的“插座”，就可以使用网络中的资源。网络是借鉴电力网（Electric Power Grid）的概念提出来的，网络的最终目的是希望用户在使用网络计算能力时，就如同现在使用电力一样方便。电力网和网络组成的简单对比如图 1-1 所示。

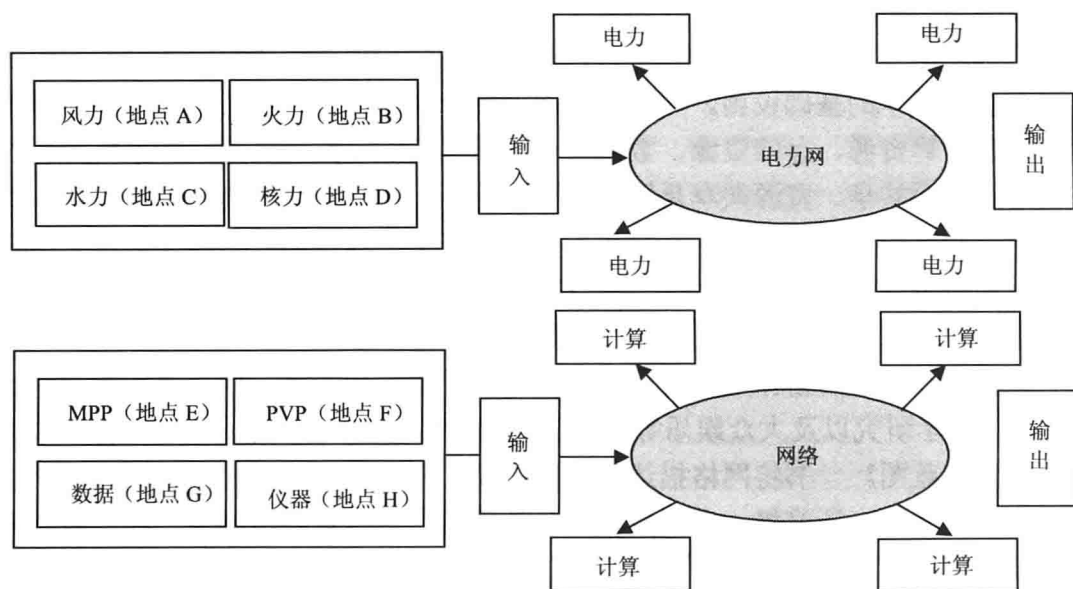


图 1-1 电力网与网络组成的对比

网络把分布在不同地理位置上、用通信媒介连接起来的各种资源形成一体，为用户提供恰到好处的服务。1990 年左右出现的网络操作系统和分布式操作系统的目标就是为用户提供可以使用各种资源的操作平台。1993 年出现了异构计算，人们期望能够跨平台使用计算资源。1995 年出现了并行分布式计算，产生了在计算机界具有广泛影响的并行虚拟机（Parallel Virtual Machine, PVM）等，人们可以用网络连接起来的多台计算机或工作站共同计算一个问题。1996~1998 年，元计算研究非常热点，元计算的研究最终导致了现在的网络。网络概念是在 I-WAY（Information Wide Aear Year）^[3]项目中提出来的。I-WAY 是一个为期一年的计划，实验用异步传输模式（Asynchronous Transfer Mode, ATM）宽带网络连接全国范围的超级计算机中心、虚拟现实研究机构和应用开发机构，在平台上面实验了约 60 个应用项目。

网络发展经历了三个阶段：第一阶段是网络的萌芽阶段，开始于 20 世纪 90

年代早期,研究内容是关于千兆位网实验床以及一些元计算方面的工作;第二阶段是一些早期的实验,时间大概从20世纪90年代中期到晚期,出现了一些比较重要的开创性和奠基性的研究项目,如I-WAY、Globus和Legion等;目前是第三阶段,也是迅速发展阶段。关于网格的研究、开发和应用项目大量涌现,出现了影响很大的全球网格论坛(Global Grid Forum, GGF)。同时,网格计算也不再仅仅局限于科学研究,工业界与学术界开展了广泛联盟,正致力于使网格计算在更广泛的领域得到推广和应用。

David等人把网格按照技术发展的顺序分为三代,如图1-2所示。

第一代网格(1990~1995年)的典型代表是FAFNER(Factoring via Network-Enabled Recursion)和I-WAY两个项目。这个时期的网格都是把超级计算机节点连接起来,为高性能应用提供计算资源,元计算的概念就是在这个时期被提出来的。两个代表项目尽管在多个方面存在着差异,但是要克服的困难却是大致相同的,都要解决通信、资源管理和远程数据管理的问题。FAFNER发展之后出现了SETI@home和Distributed.NET等项目,I-WAY发展之后出现了Globus和Legion等项目。

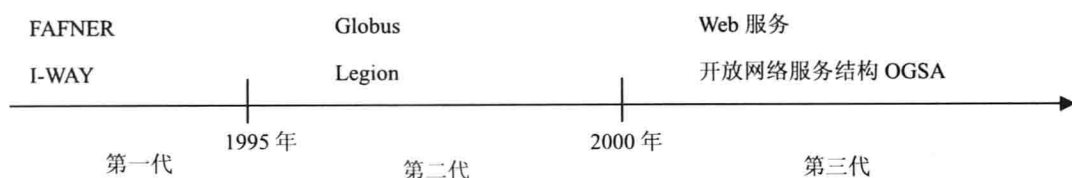


图 1-2 网格技术的发展阶段

第二代网格(1995~2000年)的典型代表是Globus和Legion等项目,尤其是Globus工具集的开发和应用,对网格的研究和发展起了显着的推动作用。这个时期的网格要解决的问题涉及网格的异构性、可扩展性、适应性等,网络安全也被提到了很高的高度。这个时期的主要工作是支持大规模数据和计算的中间件的开发,人们逐渐认识了计算网络,建立了面向协议和模块的网格体系结构。

第三代网格(2000年至今)的典型代表是万维网服务和开放网络服务结构。万维网服务是面向商业领域的计算框架,作为万维网服务框架中的基本技术,简单对象协议(Simple Object Access Protocol, SOAP)、万维网服务描述语言(Web Service Description Language, WSDL)和统一描述、发现和集成(Universal Description, Discovery and Integration, UDDI)规范获得了业界的普遍支持。这个阶段,网格的主要特点是面向服务的网格结构,2002年2月发布的开放网络服务结构(Open Grid Service Architecture, OGSA)^[4]把万维网服务和计算网络结合起来,建立了以服务为核心的网格结构。

从网格的总体发展来看,其趋势有以下几点。

(1) 与企业计算融合的趋势

要实现网格的最终目标,在标准与技术方面与企业融合是至关重要的。随着开放网格服务基础架构(Open Grid Service Infrastructure, OGSI)、Web 服务资源框架(Web Service Resource Framework, WSRF)和 Web 服务查询(WS-Inspection)文档规范的出现,标志着网格和企业计算在目标、理念和实现等方面走向融合。

(2) 标准化的趋势

WSRF 实现了网格应用与 Web 服务的融合,网格应用不再需要另外的 Web 服务体系。目前,实现 WSRF 和 OGSA 基本服务的 GT4 将成为未来网格的新标准。

(3) 合作的趋势

网格的研究已从美国和欧洲推广到了世界各地,各个国家和地区都投入了大量的资金进行网格技术研究和国家网格建设。为了将有限的资金投入庞大的网格系统建设中,在国际与地区间研究机构和商业机构开展合作是非常必要的。GGF、国家网格论坛以及网格间协作组织、协作成员之间的合作,都会推动网格的开发与研究进程。

1.1.2 网格的基本特点

网格是一个在地理位置上广泛分布的基础设施,用户数量巨大,资源千差万别,需求种类繁多,应用功能各异。网格需要为用户提供安全、高效、高质量的资源和服务,必须具备如下特点。

(1) 可扩展性

网格系统初期的规模较小,可扩展性能够保障随着资源节点的不断加入,系统的规模随之扩大,提供的功能和服务不断完善。

(2) 分布性与共享性

网格中的资源规模较大,通常分布在地理位置互不相同的地方,但它们却可以被充分共享,即网格上的任何资源都可以提供给网格中的任何用户,多个用户共享使用一个资源,一个用户可使用多个资源。网格的分布与共享特性分别描述了网格硬件在物理上的特征和网格软件支持下实现的逻辑特征。

(3) 虚拟性

网格中的资源和用户都要经过抽象,将实际的用户和资源虚拟化为网格用户和网格资源。由于资源信息对外提供的只是一个虚拟化的接口,实际的用户和资源是相互不可见的。因此,网格使用开放式标准协议和框架可以访问网格中的资源,为相互协作和集成提供依据。

(4) 动态性与多样性

组成网格系统的各虚拟组织的自治性导致了动态性。由于资源提供者拥有相应资源的最高权限,它可以随时决定将资源加入或退出,以及替换或升级资源,因而资源的状态、可获得性和资源的负载等信息是动态变化的;在网格环境中可

以有不同的体系结构的计算机系统和类别不同的资源，导致了网络的多样性。

（5）自治性与管理的多重性

自治性是指网络允许资源拥有者对其资源有自主的管理能力。同时，网络资源也必须接受网络系统的统一管理，以便在不同资源之间建立联系，实现共享和互操作，并作为一个有机的整体协调不同位置的资源用户，为更多用户提供服务。因此，网络管理的多重性体现在：允许资源拥有者对资源的自治管理，又要求资源必须接受网络系统的统一管理。

（6）容错性

随着网格中参与协同任务的资源数目的增加，系统出错概率也会随之增加。与单任务的分布式环境不同，网格系统由于其地域分布和资源共享所造成的系统行为和系统性能经常发生变化，因此需要网格系统具有容错机制，以保证协同任务的顺利完成。

（7）协商性

网格支持资源的协商使用，可以确保资源请求者和资源提供者通过协商得到不同质量的服务，满足不同的实际需求，如系统响应时间、数据带宽、资源可用性、安全性等要求。通过协商，请求者和提供者之间还可以建立专用的服务接口，改善服务质量（Quality of Service, QoS）。

1.2 网络的研究现状

目前，网络的研究已从美国和欧洲推广到世界各个大洲。各个国家和地区都投入了大量的资金进行网络技术研究和网络基础设施的建设。美国政府用于网络技术基础研究的经费已经超过 5 亿美元；英国政府已经投资超过 1 亿英镑，建设英国国家网格；欧洲联盟也投入巨资建设欧洲数据网格和欧洲网格；亚洲的中国、日本、韩国、马来西亚等国家也开始了网格研发和建设工作。美国军方目前正在实施“全球信息网格”计划，预计在 2020 年完成。

网络技术的出现使得许多行业具有挑战性的问题可以在新的平台上完成和实现。国内、国外均开展了大量的针对网格项目以及网格实现过程中所涉及的体系结构、资源管理、通信与安全、性能分析等关键问题的研究。

1.2.1 国内外的网格研究项目

1. 国外的网格研究项目

网络技术研究发源于美国，美国是目前网格研究走在最前列的国家。网格的研究最初是从科研机构开始的。美国和欧洲、亚洲的多家研究机构开展了与网格相关的研究工作，开发了 Globus、Legion、Condor 等比较有影响的软件与工具。

(1) Globus

Globus 是世界上第一个获得广泛使用的网格软件, 由美国阿岗 (Argonne) 国家实验室研发, 有十几所大学和研究机构参与了该项目。它旨在为计算网格创建基本的基础设施以及高级服务。该项目研究了网格计算的关键理论, 开发了支持网格计算的工具软件和网格应用程序, 规划组建了大型的网格实验平台。关键理论包括资源管理、网络安全、信息服务、数据管理等。

(2) Legion

Legion 是弗吉尼亚大学的一个中间件项目计划, 它是为了网格应用而设计的基于对象的元系统软件。该项目的目标是通过为处理器、数据系统、文件系统等提供标准的对象表示, 进而推动分布式系统软件的原则性设计。在 Legion 内, 所有的实体都是对象, 包括硬件实体和软件实体。每个对象由它的类对象定义和管理。该软件支持透明调度、数据管理、容错、站点自治和多种安全机制。

(3) Condor

Condor 是威斯康星-麦迪逊大学的研究项目, 它是一个专用的计算密集型的作业负载管理系统, 提供作业队列机制、作业调度策略、优先级方案、资源监控及资源管理等功能。Condor 使用一些特有的属性, 如检查点、任务迁移、远程系统调用, 实现有效的利用空闲和专用资源的能力; 支持用户提交串、并行作业, 并监控作业的运行过程。

(4) 欧洲数据网格

欧洲数据网格 (European Data Grid) 是欧洲联盟支持的一个项目, 目标是建设一个强度大、共享超过 PB 级数据的大规模分布式数据库下一代计算基础设施。它把来自欧洲联盟成员国、美国、日本、俄罗斯等国家的不同学科的研究和开发人员结合在一起, 基本思想是把海量数据分散到全球的计算机上进行处理, 并由全球的物理学家共同分析。其主要功能是负载调度和管理、数据管理、网格监控、构造层管理和海量存储管理。它有 12 个工作小组, 研究内容涉及中间件、基础设施、应用和管理四个方面。

(5) UNICORE

UNICORE (Uniform Interface to Computer Resource) 是德国联邦教育和研究部资助的一个项目, 主要合作者是德国的五家科研机构。其目的是提供一套软件, 允许用户给远程高性能计算机提交作业, 而不需要知道远程机器的操作系统、数据存储格式、管理策略和过程。它最大可能地使用已有的万维网技术, 为终端用户和计算中心提供强大的运行网格的功能。

(6) Nordu

Nordu 网格是北方四国——挪威、丹麦、瑞典、芬兰用网格中间件建设一个网格基础设施。Nordu 网格的主要构成部分有计算元素、存储元素、复制目录、信息系统、网格管理器用户接口等。

(7) 亚太网格

亚太网格是亚洲和太平洋地区网格研究的基础设施。它把重点放在资源共享、网格技术开发、网格技术推广使用等方面,它不局限于少数的发达国家,也不局限于特定的网格,更不对研究组织进行限制。它是亚洲和太平洋地区各国研究人员进行技术交流、交换信息的场所。

(8) Ninf

Ninf 是日本正在实施的全球计算基础设施项目,它允许用户访问包括硬件、软件和数据在内,分布在广域网上的资源;Bricks 是日本的一个关于性能评价系统的项目,允许在典型的高性能全球计算环境上分析和评价不同的调度策略。

(9) N*Grid

N*Grid 是韩国情报和通信部支持的一个项目,目标是建立韩国国家网格,该项目包括计算网格、数据网格、访问网格和应用网格四部分。它将韩国的超级计算机和高性能集群连接在一起,建立应用实验床、应用门户,开发具体的应用程序。

(10) e-Science

e-Science 是英国的一个对全球网格计算具有重要影响的庞大研究计划,目标是利用网格等新一代分布式技术创建新的科学研究环境。该研究计划已进入第二阶段。从第一阶段的评估结果可以看出,有两个重要的变化值得关注;一是提出了“Web Service Grids”的概念;二是提出了重视对成熟网格中间件的规范与推广工作,以进一步扩大网格技术研究成果的应用领域,为此,还成立了开放中间件基础设施研究院(Open Middleware Infrastructure Institute, OMII)。

2. 国内的网格研究项目

相比之下,网格在我国的研究起步较晚,1996年中科院计算技术研究所开始网格技术的研究。1999~2001年,在教育部的支持下,李三立院士带领清华大学网格研究组进行了先进计算基础设施(Advanced Computational Infrastructure, ACI)的研究。1999~2001年,中科院计算所联合十几家科研单位,承担了“863”重点项目——国家高性能计算环境(National High Performance Computing Environment, NHPCE)。2001年提出了织女星网格(Vega Grid)计划(包括GridOS、信息网格和知识网格三个层次)——“服务网格”的理念。2002年12月24日,上海市宣布将投入2亿多元人民币,建设E研究院(E-Institute),其中“上海教育科研网格”是其中的一大块,由多所上海的大学参加。2003年10月,上海市科委设立重大攻关项目“信息网格及其典型应用研究(ShanghaiGrid)”,并投入900万元人民币,带动资金5000万元人民币。同时,上海市政府启动了科教兴市重大项目——“上海城市网格”的建设。2003年10月,教育部投资2000万元人民币正式启动“中国教育科研网格项目(ChinaGrid)”一期工程,全国共有12所知名高校参加。目前,国内网格的研究集中在以下五个项目的研究,被称之为中

国“五大网格”。

(1) 中国国家网格 (China National Grid, CNGrid)

该项目由国家“863”高技术研究发展计划资助,旨在建立面向企业、高等院校、科研机构、政府部门的国家高性能计算环境,主节点采用自行研制的、面向网格的高性能计算机。若干工业、服务业、科学研究以及资源环境领域的网格应用投入运行,实现资源共享、协同工作。该项目的主要任务包括:建设中国国家网格实验床(CNGrid),建立具有4万亿次以上计算能力的网格主节点,支持网格应用以及维护网格系统运行的网格软件,开发具有代表性的网格生产性应用。

(2) 国家空间信息网格 (China Spatial Information Grid)

该项目由国家“863”高技术研究发展计划信息获取与处理主题资助。该主题将面向国家重大需求,结合宽带信息网、高性能计算等信息基础设施的建设和发展,系统性地开展对观测数据的定量化处理、网络化共享与集成应用等关键技术研究,构建我国国家空间信息网格,从整体上提高我国空间信息获取、处理与应用能力,实现多维多源精细空间信息一体化获取、空间信息的定量化处理、空间信息的网络化共享与应用三个方面的重大突破,推动我国遥感信息资源的社会化与应用产业化。

(3) 国家自然科学基金委网格 (NSFC Grid)

国家自然科学基金委的《以网络为基础的科学活动环境研究重大研究计划》定位在网络计算环境的基础科学理论、综合试验平台、典型应用示范三个层次中的基本科学问题和关键技术研究,同时重点建立高能物理、大气和生物信息等网络计算环境实验应用系统。

(4) 中国教育科研网格 (ChinaGrid)

ChinaGrid 是中国教育科研网格的英文名称,是教育部“十五”“211 工程”公共服务体系建设的重大专项。ChinaGrid 由国内12所大学首批参与,将充分利用中国教育科研网(CERNET)和高校的大量计算资源和信息源,开发相应的网格软件,配合网络计算机的使用,将分布在CERNET上自治的分布异构的海量信息资源集成起来,实现CERNET环境下资源的有效共享,消除信息孤岛,提供有效的服务,形成高水平低成本的计算机服务平台,将高性能计算送到教育与科研网用户的桌面上,成为国家科研教学服务的大平台。ChinaGrid 主要由公共网格服务支撑平台、校园网格平台及生物信息学网格、图像处理网格、远程教育网格、流体力学网格和海量信息处理网格这五大专业应用网格构成。

ChinaGrid 具有三个之最:

① 最大的网格。最大的网格是指中国教育科研网格未来的覆盖面将从12所学校开始逐步扩展到全国大中小学以及其他教育科研机构。

② 最先进的网格。一个角度是指充分跟踪国际前沿的技术进展,参加国际网格发展GGF活动,参加讨论,争取与它一起制定相关的标准。另一个角度,也要