

都志辉 陈渝 刘鹏 / 编著

李三立 / 审

网格 计算



清华大学出版社

网格计算

都志辉 陈渝 刘鹏 / 编著
李三立 / 审



清华大学出版社

(京)新登字 158 号

内 容 简 介

本书分三篇,第一篇网格原理,介绍了网格的基本概念、特点和相关知识,叙述了网格的结构组成原则以及当前最重要的网格体系结构实例,最后是网格技术所研究的主要方面和内容。第二篇网格实现,以当前最重要的网格计算工具包 Globus Toolkit 为基础,介绍了如何在网格基本原理的指导下,去实现一个具体的网格支撑平台。内容涉及网格资源管理、网格数据管理、网格信息服务以及网络安全等多个方面的网格技术。第三篇网格应用,通过具体的应用实例,展示了网格广泛的应用领域和应用前景。内容包括分布式超级计算、分布式仪器系统、数据密集型计算以及远程沉浸等几个主要的领域。

本书既可以作为高等院校理工科高年级本科生以及研究生教材和教学参考书,又可以作为网格计算领域科研人员的参考书,还可以作为相关领域人员了解网格计算知识的参考材料。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

网格计算/都志辉,陈渝,刘鹏编著. —北京:清华大学出版社,2002
ISBN 7-302-05939-X

I. 网… II. ①都… ②陈… ③刘… III. 网格—计算 IV. O243

中国版本图书馆 CIP 数据核字(2002)第 084035 号

出版者: 清华大学出版社(北京清华大学学研大厦,邮编 100084)

<http://www.tup.tsinghua.edu.cn>

责任编辑: 薛 慧

版式设计: 刘 路

印刷者: 北京市清华园胶印厂

发行者: 新华书店总店北京发行所

开 本: 787×1092 1/16 **印张:** 12 **字数:** 273 千字

版 次: 2002 年 10 月第 1 版 2002 年 10 月第 1 次印刷

书 号: ISBN 7-302-05939-X/TP·3531

印 数: 0001~6000

定 价: 19.00 元

前 言

网 格 计 算

首先感谢李三立院士最早将我们带到网格计算这一研究领域。

如果说 Internet 提供的是四通八达的信息高速公路,而网络提供的则是 Internet 上琳琅满目的资源与服务超市。网格作为一种建立在互联网之上的新一代基础设施,在国内外的学术界和工业界都引起了广泛的关注。国外的网格计算研究最早从 20 世纪 90 年代初开始,而国内大概在 2000 年左右开始这方面的研究。

李三立院士在国内很早就指出网格计算是一个重要的研究方向,并多方呼吁对这一研究领域加强支持。教育部在 2000 年 6 月决定启动教育部重点项目“先进计算基础设施北京上海试点工程”来支持这一研究领域。作者很幸运地成为这一项目的研究成员。在李三立院士的领导下,经过近一年的紧张工作,我们成功地完成了这一项目,并在 2001 年通过了教育部的验收和鉴定,这是国内第一个通过鉴定的网格计算项目。之后我们在该项目研究成果的基础上,继续在网格计算领域开展多方面的深入研究,并且密切关注国内外的最新发展动态。

“信息技术专题”是作者为清华大学计算机系高年级本科生讲授的一门课程,根据我们研究的成果,作者将网格计算这一前沿的研究内容在课堂上向同学们进行介绍,引起了同学们很大的兴趣,取得了较好的效果。同时在研究生的课堂上,作者对这一研究方向也进行了介绍,和研究生共同进行深入探讨,讨论也非常热烈。在我们研究组的内部也经常就这一领域的不同问题进行交流,这一切都是我们在这一研究领域的宝贵积累,有助于更全面、准确、深入地把握这一研究方向。

网格计算权威 Ian Foster 博士的热心帮助是本书问世最直接的原因。由于作者在给同学们上课时需要更丰富的网格计算材料,而关于网格计算的经典著作就是 Ian Foster 和 Carl Kesselman 共同编著的“The Grid: Blueprint for a New Computing Infrastructure”,当时在国内很难找到,Ian Foster 热心地将他的这一著作寄给了作者,通过阅读这一经典著作,我们觉得有义务尽快将这一重要的研究方向以图书的形式,向国内学术界进行介绍,为国内这一领域的研究作出我们的一点贡献。

感谢和我们共同探讨与交流的老师和同学们,感谢 Ian Foster 的热心帮助,感谢清华

大学出版社的大力支持。没有他们,就没有这本书。

本书第一篇由都志辉撰写,第二篇由陈渝撰写,第三篇由刘鹏撰写。

“抛砖引玉”是本书的目的,网格计算是一个崭新的研究领域,欢迎对这一领域有兴趣的朋友和同行为我们多提意见,相互切磋,共同促进国内的网格计算研究。

作 者

2002年8月17日于清华大学

目 录

网格计算

第一篇 网格原理

第 1 章 网格基础	3
1.1 网格内涵	3
1.1.1 网格的概念.....	3
1.1.2 网格的目的.....	4
1.1.3 网格的基本要求.....	4
1.1.4 网格的意义.....	6
1.1.5 网格概念的分歧.....	6
1.2 网格需求	8
1.2.1 计算的重要性.....	8
1.2.2 问题的需求.....	8
1.2.3 相关技术的发展.....	9
1.2.4 网格的应用领域	10
1.2.5 网格的用户群	11
1.3 网格特点.....	12
1.3.1 分布与共享	12
1.3.2 自相似性	13
1.3.3 动态性与多样性	14
1.3.4 自治性与管理的多重性	15
1.4 小结.....	15
思考题	15
第 2 章 网格体系结构	17
2.1 网格体系结构的概念.....	17
2.1.1 定义	17
2.1.2 讨论	18
2.2 五层沙漏结构.....	18
2.2.1 基本思想与概念	18

2.2.2	结构描述	22
2.2.3	与 Globus 的对应关系	25
2.2.4	基于五层结构的应用例子	26
2.3	开放网格服务体系结构	28
2.3.1	基本思想	28
2.3.2	OGSA 的两大支撑技术	31
2.3.3	服务接口与功能机制	33
2.3.4	网络协议绑定	38
2.3.5	高级服务	38
2.3.6	运行环境的作用	38
2.3.7	基于 OGSA 建立虚拟组织	39
2.3.8	基于 OGSA 的应用例子	40
2.4	建造网格的几点建议	42
2.4.1	国家行为	42
2.4.2	从局部到整体	42
2.4.3	利用市场与经济杠杆	43
2.5	小结	43
	思考题	44
第 3 章	网格技术	45
3.1	网格技术分类	45
3.2	网格应用技术	47
3.3	网格编程技术	47
3.3.1	编程支持系统	47
3.3.2	面向对象技术及 Legion	48
3.3.3	基于商品化技术集成的网格编程	49
3.3.4	数值计算编程环境 NetSolve	50
3.4	网格核心服务技术	51
3.4.1	高性能调度技术	51
3.4.2	高吞吐率资源管理技术	52
3.4.3	性能数据收集、分析与可视化技术	53
3.4.4	安全技术	53
3.5	网格底层支撑技术	53
3.6	网格技术发展	54
3.7	小结	56
	思考题	56
	参考文献	57

第二篇 网格实现

第 4 章 Globus 项目介绍	63
4.1 Globus 的起源和发展	63
4.1.1 Globus 项目与 Globus 工具包	63
4.1.2 Globus 对网格计算的理解	64
4.2 Globus 系统结构	64
4.3 小结	66
思考题	66
第 5 章 网络安全基础设施	67
5.1 分布式安全技术	67
5.2 网格计算环境的安全需求和 Globus 的安全目标	69
5.3 网络安全基础设施的安全技术	69
5.3.1 安全认证	70
5.3.2 安全身份相互鉴别	71
5.3.3 通信加密	71
5.3.4 私钥保护	71
5.3.5 安全委托与单点登录	71
5.4 基于 GSI 的任务提交与执行过程描述	72
5.5 GSI 的安全策略设计思想	74
5.5.1 相关术语和安全策略介绍	74
5.5.2 用户与用户代理之间的安全关系	75
5.5.3 Globus 主体与本地资源主体的映射	76
5.5.4 资源代理与资源分配安全	76
5.5.5 进程与用户代理之间的安全鉴别身份转移	77
5.5.6 网格信息服务的安全保证	78
5.6 GSI 的安全策略实现	78
5.6.1 通用安全服务编程接口	79
5.7 小结	80
思考题	80
第 6 章 元计算目录服务	81
6.1 元计算目录服务介绍	81
6.2 元计算目录服务的特点	82
6.3 信息提供者 and 提供的信息类型	83
6.4 元计算目录服务的信息模型和目录信息树	84
6.4.1 LDAP 介绍	84
6.4.2 LDAP 目录信息树	85
6.4.3 MDS 的信息模型	86

6.4.4	MDS 信息层次	87
6.5	GRIS 和 GIIS 介绍	89
6.5.1	网格资源信息服务 GRIS	89
6.5.2	网格目录信息服务 GIIS	90
6.6	使用 MDS	90
6.7	小结	91
	思考题	92
第 7 章	资源分配管理	93
7.1	资源分配管理者 GRAM	93
7.1.1	GRAM 的组成和执行流程	93
7.1.2	GRAM-API	95
7.2	动态协同分配代理 DUROC	95
7.2.1	DUROC 的管理结构	95
7.2.2	DUROC-API	96
7.3	资源描述语言 RSL	98
7.3.1	RSL 语法、标记和属性	99
7.3.2	RSL 语法和标记规则	101
7.3.3	GRAM RSL 参数	102
7.4	小结	106
	思考题	106
第 8 章	数据管理	107
8.1	全局二级存储服务 GASS	107
8.1.1	GASS 简介	107
8.1.2	GASS 缺省的数据移动策略	108
8.1.3	GASS 特殊的数据传输策略	109
8.1.4	与 Globus 工具包其他服务的集成	109
8.1.5	GASS-API	110
8.2	Grid FTP	111
8.2.1	GridFTP 协议功能和实现	111
8.2.2	GridFTP 性能	112
8.3	复制管理服务	114
8.3.1	复制管理服务简介	114
8.3.2	复制管理服务的特点	115
8.3.3	复制管理服务的功能	116
8.3.4	复制目录服务简介	116
8.3.5	复制管理服务 API	118
8.3.6	复制管理服务的处理流程	119
8.4	小结	120

思考题.....	120
第 9 章 网格应用程序开发工具	121
9.1 CoG Kits 简介.....	121
9.2 基于 CoG Kits 的应用.....	122
9.3 Java CoG Kit 设计与实现.....	123
9.4 MPICH-G2 简介.....	127
9.5 小结.....	128
思考题.....	129
参考文献	130

第三篇 网格应用

第 10 章 分布式超级计算	137
10.1 背景.....	137
10.2 应用程序在网格上的分解.....	138
10.3 分布式超级计算的核心技术.....	139
10.3.1 适应性算法.....	139
10.3.2 资源调度策略.....	139
10.3.3 容错.....	141
10.4 SF Express——大规模军事仿真.....	141
10.4.1 背景.....	142
10.4.2 任务的分解.....	142
10.4.3 与 Globus 的融合.....	145
10.5 分布式异构计算环境 Cactus 及其应用.....	147
10.5.1 Cactus 的思想.....	147
10.5.2 Cactus 体系结构.....	148
10.5.3 Cactus 应用举例——模拟黑洞.....	149
10.6 小结.....	152
思考题.....	152
第 11 章 分布式仪器系统	154
11.1 背景.....	154
11.2 分布式仪器系统的核心技术.....	156
11.2.1 基于网络的海量存储系统 HPSS.....	156
11.2.2 分布式监控.....	157
11.2.3 基于策略的访问控制.....	158
11.3 XPort——X 射线设备的科学门户.....	158
11.3.1 背景.....	158
11.3.2 XPort 试验流程.....	159
11.3.3 XPort 的实现.....	160

11.4 小结.....	161
思考题.....	162
第 12 章 数据密集型计算	163
12.1 背景.....	163
12.1.1 生物和医学.....	163
12.1.2 地球观察.....	163
12.1.3 数字天空.....	164
12.1.4 大脑映射.....	164
12.2 CERN 与 DataGrid	164
12.2.1 欧洲原子能研究机构 CERN	164
12.2.2 大型强子对撞机 LHC	165
12.2.3 DataGrid	165
12.3 DataGrid 的设计	166
12.4 DataGrid 的项目管理	168
12.5 小结.....	170
思考题.....	171
第 13 章 远程沉浸	172
13.1 背景.....	172
13.2 应用举例.....	173
13.2.1 虚拟历史博物馆.....	174
13.2.2 协同学习环境 NICE	175
13.2.3 数据可视化协同分析环境 CAVE6D	175
13.3 远程沉浸与网格的结合.....	176
13.4 小结.....	177
思考题.....	177
参考文献.....	178

第一篇

网格原理

都志辉,博士,清华大学计算机系副教授,主要研究方向为 SPMD 集群式并行算法与程序设计、HPF 编译技术、用户层网络通信技术、跨学科高性能计算应用技术与网格计算技术。1998 年北京清华大学计算机系博士毕业,后到清华大学计算机系从事博士后研究。2000 年 5 月博士后出站留在清华大学计算机系从事教学与科研工作。作为项目负责人完成或者作为骨干成员参加的科研项目近 20 项,通过部委级鉴定的科研项目有 6 项,其中一项获得“上海市科技进步一等奖”,一项获得“教育部科技进步二等奖”。

完成了国内第一本关于 MPI 并行程序设计的著作《高性能计算并行编程技术——MPI 并行程序设计》,发表科研论文 30 余篇。

邮件地址:

duzh@tsinghua.edu.cn, duzh@bigfoot.com

个人主页:

<http://hpclab.cs.tsinghua.edu.cn/~duzh/>



第1章 网格基础

“网格”是一个新出现的概念,代表了一种先进的技术和基础设施,是继 Internet 之后又一次重大的科技进步。本章的目的是树立读者关于网格的基本概念,建立宏观的整体认识,抓住网格最主要的特征。为此,本章从网格的内涵、网格的需求以及网格的特点等几个方面对网格进行论述,希望能够回答什么是网格、为什么需要网格以及网格有什么值得注意的特点等重要而基本的问题,为后面深入学习、掌握和运用网格技术建立基本的概念,提供必需的基础性知识。

1.1 网格内涵

什么是网格?在“网格”这两个字的背后,到底代表了一种什么样的技术?网格技术所追求的是一种什么样的前景和应用效果?这些问题,是网格领域研究者需要明确回答的问题,也是目前讨论得非常热烈的问题。本节从基本的概念定义开始,结合网格所希望达到的目的,对网格应该满足的要求、网格的意义以及目前关于网格概念的各种争议等展开论述,使读者可以较为全面地从整体上对网格的基本内涵进行把握。

1.1.1 网格的概念

什么是网格(grid)?网格就是一个集成的计算与资源环境,或者说是一个计算资源池^[22]。网格能够充分吸纳各种计算资源,并将它们转化成一种随处可得的、可靠的、标准的同时还是经济的计算能力。除了各种类型的计算机,这里的计算资源还包括网络通信能力、数据资料、仪器设备甚至是人等各种相关的资源。

什么是网格计算(grid computing)呢?基于网格的问题求解就是网格计算。

这里给出的网格和网格计算的概念是相对抽象的,而且是广义的定义,其实网格计算还有狭义的定义。狭义网格定义中的网格资源主要是指分布的计算机资源,而网格计算就是指将分布的计算机组织起来协同解决复杂的科学与工程计算问题。狭义的网格一般被称为计算网格(computational grid),即主要用于解决科学与工程计算问题的网格。

根据求解问题的特点,人们又提出了多种名称的网格,比如以数据密集型问题的处理为核心的数据网格^[2],以解决科学问题为核心的科学网格^[56],以全球地球系统模型问题求解为主要目的的地球系统网格^[1]等等。此外还有地震网格^[3]、军事网格^[4]、NASA(National Aeronautics and Space Administration)的IPG^[27]等行业网格。

那种认为网格就是仅仅通过网络把计算机、人、仪器、数据等连接起来的观点是过时的,它过分强调了物理的网络和离散的网格资源,而没有将它们作为一个有机的统一整体来看待。另外一种观点就是把网格看做是中间件系统,这种观点也是不全面的。中间件

的确在网格中占有很重要的地位,但是网格决不仅仅只是中间件。这两种观点都存在一定的片面性,第一种观点是过分强调网格物理上的资源组成,第二种观点过分强调网格逻辑上的功能,只有将两者结合起来才是完整的网格系统。物理资源本身和对资源的管理与逻辑上的抽象都是十分重要的,而且两者也是密不可分的,它们是网格环境的两大核心组成要素(如图 1-1 所示)。

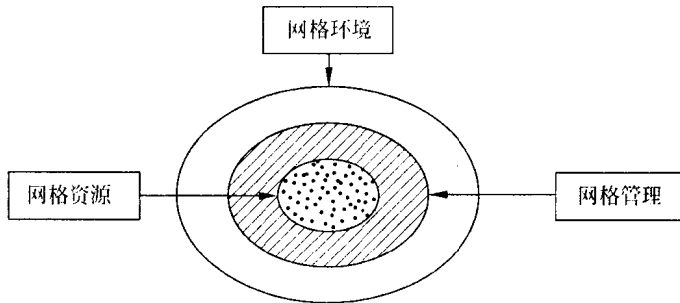


图 1-1 网格资源、网格管理与网格环境三者间的关系

1.1.2 网格的目的

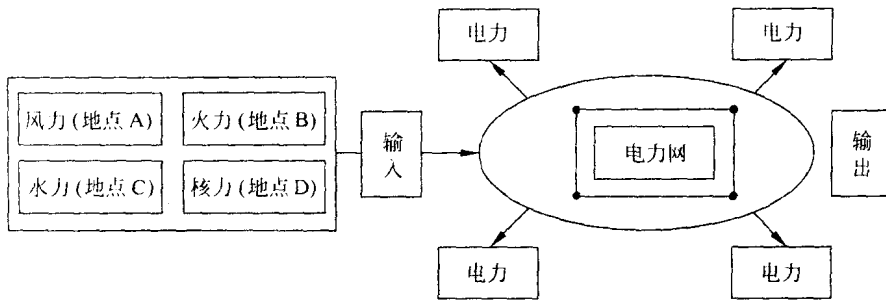
网格是借鉴电力网(electric power grid)的概念提出来的^[22],网格的最终目的是希望用户在使用网格计算能力时,就如同现在使用电力一样方便。我们在使用电力时,不需要知道它是从哪个发电站输送出来的,也不需要知道该电力是通过什么样的发电机产生的,不管是水力发电,还是通过核反应发电,我们使用的是一种统一形式的“电能”。网格也希望给最终的使用者提供的是与地理位置无关、与具体的计算设施无关的通用的计算(广义的计算,可理解为问题求解)能力。图 1-2 是对电力网和网格组成的简单对比示意图。

网格和电力网都有各自资源的消费者和资源提供者,对于电力网来说资源提供者就是发电站,对于网格来说资源提供者是计算机等;对于电力网来说资源消费者就是各种消耗电能的设备,对于网格来说资源消费者就是使用网格计算能力求解问题的用户。不管是电力网还是网格,都有覆盖范围广泛、组成资源多样的特点。正如同电力网中需要大量的变电站等设施对电网进行调控一样,网格中也需要大量的管理结点来维护网格的正常运行。与电力网相比,网格的结构更复杂,需要解决的问题也更多,但是它也会给我们带来更大的便利和帮助。

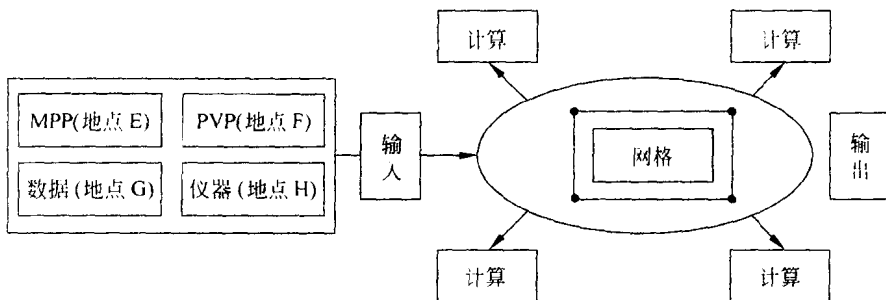
1.1.3 网格的基本要求

对于网格提供的计算能力,有四个基本的要求,分别是:可靠性要求、标准化要求、易访问性要求和价格低廉的要求^[22]。

网格的可靠性是指网格提供的计算能力必须保证是持续、稳定和安全的,不应该因为网格内部个别资源的变化而对网格应用造成影响,即网格内部局部资源的变动对网格应用应该是透明的。就如同我们日常使用电灯时不应该因为个别发电厂临时出现故障而造成整个电网电力供应的中断一样,电力网应该能够保证实时地从别的发电厂或者其他地区的电网引入电力来弥补本地电力的不足,网格也应该能够保证提供持续、稳定的计算能



(a) 电力网构成示意图



(b) 网络组成示意图

图 1-2 电力网与网络组成的对比

力。网格还应该满足各种形式的安全要求,比如数据传输的加密、权限的认证、避免非法入侵和非法使用等,如果没有安全性保障,这种先进的计算服务就不能得到广泛应用。

网格的标准化一方面是指网格资源之间应该有一个统一的可以相互访问的接口或者协议标准,因为只有这样才能实现网格资源的相互操作,从而实现充分的资源共享,标准化是共享的前提;另一方面是指网格对用户提供的计算能力应该满足一定的标准,有一种比较统一的形式,从而便于以一种统一的方式进行访问。对于访问者来说,不能因为时间、地点、具体的访问系统等的不同而要求他不断改变访问形式,访问形式应该有一致性,当然一致性的前提是网格必须提供给用户一个相对稳定的标准化接口。

网格的易访问性是指用户可以在任何时间、任何地点,以自己习惯的统一的形式访问和使用各种网格资源。网格计算能力可以通过网格系统输送到任何角落,随处可得。换句话说,在网格上没有资源处在什么位置的概念,只有“在网格上”或者“不在网格上”的区别,无论你在什么地方,网格资源都在你的旁边。人们以前在解决特定问题时或许不得不到特定的地点来进行,比如到某一个单位去登记和使用特殊的仪器设备等,但是在网格上解决问题时,不应该因为访问者或者资源所在地位置不同而受到限制。

网格费用的低廉性是网格能够被普遍接受和推广的前提。不管网格有多少优点,如果大多数的使用者无法承受其费用,网格就不可能被普及,它的各种优势也就根本无法得到体现。网格技术通过将资源充分共享,最大限度地发挥资源的使用价值,不仅可以将原来闲置和浪费的资源收集起来供网格用户使用,而且可以避免以前由于地理位置限制所带来的各种额外开销,显然网格对使用者存在着很大的降低开销的潜力。

上述四项要求是网格需要解决的问题,也是网格技术发挥作用的地方。网格作为一种新型而重要的基础设施,不是一夜之间就能够奇迹般地突然出现的,需要各个方面联合起来,共同努力才可以实现。

1.1.4 网格的意义

网格概念的提出将从根本上改变人们对“计算”的看法,因为网格提供的是与以往根本不同的计算方式。Randy Bramley 认为网格提供的计算能力是以前所无法得到的,而且也是不能够通过其他的方式得到的。网格概念的核心就是突破了以往强加在计算资源之上的种种限制,使人们可以以一种全新的、更自由、更方便的方式使用计算资源,解决更复杂的问题。

首先是计算能力大小的限制。以前大部分的用户无法得到足够的计算能力,因此许多问题的解决是不能通过计算或者是不能完全靠计算来实现的,对模型以及算法的化简是最常见的近似方法。而网格所提供的计算能力要远远超过以前我们所能够想像的程度,对于大多数用户来说,网格提供给他们计算能力足以满足其计算需求,在这种计算能力的支持下,人们可以做许多以前无法想像和无法完成的工作。

其次是地理位置的限制。计算资源是分布在各处的,有些资源是稀缺或不可复制的,有些资源甚至是无法和特定的地理位置分开的,因此要使用这些资源,在以前许多情况下必须到相应的地方去,这在很大程度上限制了这些资源的使用。而网格把“到资源所在地”对资源进行使用的限制打破了,对资源的使用和使用者所在的位置以及资源所在的位置无关。突破了在使用资源时对位置的限制,是网格具有突出意义的功能。

最后也是非常重要的一点就是网格打破了传统的共享或协作方面的限制。以前对资源的共享往往停留在数据文件传输的层次,而网格资源的共享允许直接控制其他资源,而且共享资源的各方在协作时可以有多种方式更广泛地交流信息,充分利用网格提供的各种功能。比如为了分析臭氧层问题,可以通过网格将各个领域的专家、各种大型专业数据库、大型计算设备、各种模型库和算法库等充分结合起来,协同研究这一问题。网格使得共享与协作的方式和方法更广泛了,而且为这种合作提供了各种控制策略与手段,可以根据需要,动态地与不同的组织或个人建立各种级别的工作关系。

过去人们往往很自然地把计算资源和特定的、有形的计算机等联系起来,而网格就是在剥去了各种具体的计算资源外在的“形”的基础上,将其内在的“神”即计算能力抽取出来,形成一种分布在网上的抽象的计算能力,在实现了“形”和“神”分离的同时,将原来有形的、专用的计算能力转化为一种无形的、更通用的计算能力,正如同电力网将具体的各种类型的发电机发出的电力转化为一种我们认为根本没有什么差别的统一的电力一样(如图 1-3 所示)。

这种观念和使用方式上的改变,是由网格技术支持的,不是凭空产生的。就如同互联网改变了人们传统的通信方式和通信手段一样,网格将改变人们传统的计算方式和计算手段,为人们提供更强大、更方便、更高级的问题求解手段。

1.1.5 网格概念的分歧

到目前为止,关于什么是网格和什么是网格计算,还没有一个普遍接受的定义,关于