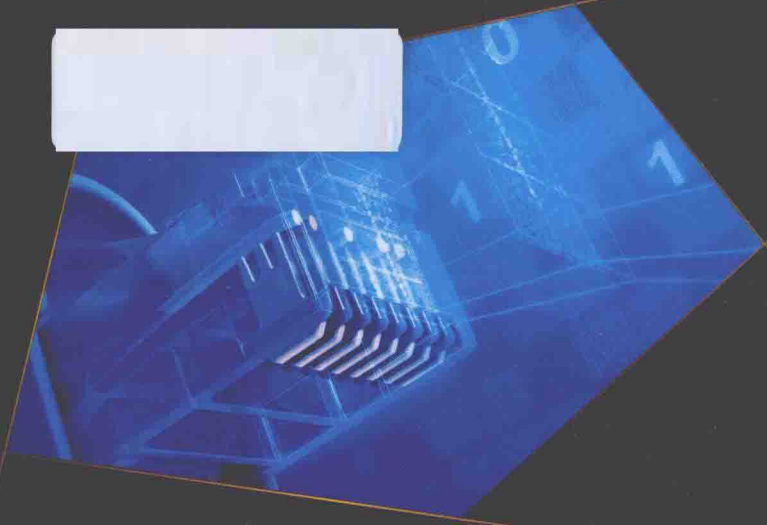
以网络为基础的
科学活动环境研究系列

网络计算环境：数据管理

程耀东　单志广　姜进磊　著



科学出版社

以网络为基础的科学活动环境研究系列

网络计算环境：数据管理

程耀东　单志广　姜进磊　著

科学出版社

北京

内 容 简 介

本书系统讲述以网络为基础的科学活动环境中的数据管理技术。全书由概论、非结构化数据管理、结构化数据管理、应用实例四大部分组成，包括数据管理背景、数据管理需求与挑战、数据管理体系结构、数据存储、元数据管理、数据传输、存储资源管理、数据管理标准、OGSA-DAI、异构数据库整合、高能物理网格数据管理、虚拟天文台数据管理 12 章。

本书取材广泛，内容系统，集成了多种网络数据管理技术，反映了国内外前沿技术发展，可供广大网络计算及相关领域的科研和技术人员阅读参考。

图书在版编目 (CIP) 数据

网络计算环境：数据管理/程耀东，单志广，姜进磊著．—北京：科学出版社，2014.10

（以网络为基础的科学活动环境研究系列）

ISBN 978-7-03-042157-9

Ⅰ．①数… Ⅱ．①程…②单…③姜… Ⅲ．①数据管理 Ⅳ．①TP274

中国版本图书馆 CIP 数据核字(2014)第 237655 号

责任编辑：任 静 / 责任校对：胡小洁

责任印制：徐晓晨 / 封面设计：迷底书装

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

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2014 年 10 月第 一 版 开本：720×1 000 1/16

2014 年 10 月第一次印刷 印张：14 3/4

字数：275 000

定价：72.00 元

（如有印装质量问题，我社负责调换）

序

近年来，以网络为基础的科学活动环境已经引起了各国政府、学术界和工业界的高度重视，各国政府纷纷立项对网络计算环境进行研究和开发。我国在这一领域同样具有重大的应用需求，同时也具备了一定的研究基础。以网络为基础的科学活动环境研究将为高能物理、大气、天文、生物信息等许多重大应用领域提供科学活动的虚拟计算环境，必然将对我国社会和经济的发展、国防、科学研究，以及人们的生活和工作方式产生巨大的影响。

以网络为基础的科学活动环境是利用网络技术将地理上位置不同的计算设施、存储设备、仪器仪表等集成在一起，建立大规模计算和数据处理的通用基础支撑结构，实现互联网上计算资源、数据资源和服务资源的广泛共享、有效聚合和充分释放，从而建立一个能够实现区域或全球合作或协作的虚拟科研和实验环境，支持以大规模计算和数据处理为特征的科学活动，改变和提高目前科学研究工作的方式与效率。

目前，网络计算的发展基本上还处于初始阶段，发展动力主要来源于“需求牵引”，在基础理论和关键技术等方面的研究仍面临着一系列根本性挑战。以网络为基础的科学活动环境的主要特性包括：

(1) 无序成长性。Internet 上的资源急剧膨胀，其相互关联关系不断发生变化，缺乏有效的组织与管理，呈现出无序成长的状态，使得人们已经很难有效地控制整个网络系统。

(2) 局部自治性。Internet 上的局部自治系统各自为政，相互之间缺乏有效的交互、协作和协同能力，难以联合起来共同完成大型的应用任务，严重影响了全系统综合效用的发挥，也影响了局部系统的利用率。

(3) 资源异构性。Internet 上的各种软件/硬件资源存在着多方面的差异，这种千差万别的状态影响了网络计算系统的可扩展性，加大了网络计算系统的使用难度，在一定程度上限制了网络计算的发展空间。

(4) 海量信息共享复杂性。在很多科学研究活动中往往会得到 PB 数量级的海量数据。由于 Internet 上信息的存储缺少结构性，信息又有形态、时态的形式多样化的特点，这种分布的、半结构化的、多样化的信息造成了海量信息系统中信息广泛共享的复杂性。

鉴于人们对于网络计算的模型、方法和技术等问题的认识还比较肤浅，基于 Internet 的网络计算环境的基础研究还十分缺乏，以网络为基础的科学活动环境还存在着许多重大的基础科学问题需要解决，主要包括：

(1)无序成长性与动态有序性的统一。Internet 是一个无集中控制的不断无序成长的系统。这种成长性表现为 Internet 覆盖的地域不断扩大，大量分布的异构的资源不断更新与扩展，各局部自治系统之间的关联关系不断动态变化，使用 Internet 的人群越来越广泛，进入 Internet 的方式不断丰富。如何在一个不断无序成长的网络计算环境中，为完成用户任务确定所需的资源集合，进行动态有序的组织和管理，保证所需资源及其关联关系的相对稳定，建立相对稳定的计算系统视图，这是实现网络计算环境的重要前提。

(2)自治条件下的协同性与安全保证。Internet 是由众多局部自治系统构成的大系统。这些局部自治系统能够在自身的局部视图下控制自己的行为，为各自的用户提供服务，但它们缺乏与其他系统协同工作的能力及安全保障机制，尤其是与跨领域系统的协同工作能力与安全保障。针对系统的局部自治性，如何建立多个系统资源之间的关联关系，保持系统资源之间共享关系定义的灵活性和资源共享的高度可控性，如何在多个层次上实现局部自治系统之间的协同工作与群组安全，这些都是实现网络计算环境的核心问题。

(3)异构环境下的系统可用性和易用性。Internet 中的各种资源存在着形态、性能、功能，以及使用和服务方式等多个方面的差异，这种多层次的异构性和系统状态的不确定性造成了用户有效使用系统各种资源的巨大困难。在网络计算环境中，如何准确简便地使用程序设计语言等方式描述应用问题和资源需求，如何使软件系统能够适应异构动态变化的环境，保证网络计算系统的可用性、易用性和可靠性，使用户能够便捷有效地开发和使用系统聚合的效能，是实现网络计算环境的关键问题。

(4)海量信息的结构化组织与管理。Internet 上的信息与数据资源是海量的，各个资源之间基本上都是孤立的，没有实现有效的融合。在网络计算环境下如何实现高效的数据传输，如何有效地分配和存储数据以满足上层应用对于数据存取的需求，以及有效的数据管理模式与机制，这些都是网络计算环境中数据处理所面临的核心问题。为此需要研究数据存储的结构和方法，研究由多个存储系统组成的网络存储系统的统一视图和统一访问，数据的缓冲存储技术等海量信息的组织与管理方法。

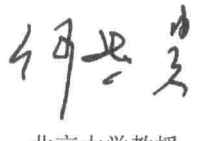
为此，国家自然科学基金委员会于 2003 年启动了“以网络为基础的科学活动环境研究”重大研究计划，着力开展网络计算环境的基础科学理论、体系结构与核心技术、综合试验平台三个层次中的基本科学问题和关键技术研究，同时重点建立高能物理、大气信息等网络计算环境实验应用系统，以网络计算环境中所涉及的新理论、新结构、新方法和新技术为突破口，力图在科学理论和实验技术方面实现源头创新，提高我国在网络计算环境领域的整体创新能力和国际竞争力。

在“以网络为基础的科学活动环境研究”重大研究计划执行过程中，学术指导专家组注重以网格标准规范研究作为重要抓手，整合重大研究计划的优势研究队伍，

推动集成、深化和提升该重大研究计划已有成果，促进学术团队的互动融合、技术方法的标准固化、研究成果的集成升华。在学术指导专家组的研究和提议下，该重大研究计划于 2009 年专门设立和启动了“网格标准基础研究”专项集成性项目(No.90812001)，基于重大研究计划的前期研究积累，整合了国内相关国家级网格项目平台的核心研制单位和优势研究团队，在学术指导专家组的指导下，重点开展了网格术语、网格标准的制定机制、网格标准的统一表示和形式化描述方法、网格系统结构、网格功能模块分解、模块内部运行机制和内外部接口定义等方面的基础研究，形成了《网格标准的基础研究与框架》专题研究报告，研究并编制完成了网格体系结构标准、网格资源描述标准、网格服务元信息管理规范、网格数据管理接口规范、网格互操作框架、网格计算系统管理框架、网格工作流规范、网格监控系统参考模型、网格安全技术标准、结构化数据整合、应用部署接口框架(ADIF)、网格服务调试结构及接口等十二项网格标准研究草案，其中两项已列入国家标准计划，四项提为国家标准建议，十项经重大研究计划指导专家组评审成为专家组推荐标准，形成了描述类、操作类、应用类、安全保密类和管理类五大类统一规范的网格标准体系草案，相关标准研究成果已在我国三大网格平台 CGSP、GOS、CROWN 中得到初步应用，成为我国首个整体性网格标准草案的基础研究和制定工作。

本套丛书源自“网格标准基础研究”专项集成性项目的相关研究成果，主要从网络计算环境的体系结构、数据管理、资源管理与互操作、应用开发与部署四个方面，系统展示了相关研究成果和工作进展。相信本套丛书的出版，将对于提升网络计算环境的基础研究水平、规范网格系统的实现和应用、增强我国在网络计算环境基础研究和标准规范制订方面的国际影响力具有重要的意义。

是以为序。



北京大学教授
国家自然科学基金委员会“以网络为基础的科学活动环境研究”
重大研究计划学术指导专家组组长

2014 年 10 月

前　言

随着科学研究规模的不断扩大，模拟实验与科学仪器产生了越来越多的海量数据。针对海量数据问题，包括数据采集、数据存储、数据传输、数据共享、数据分析和数据可视化等，构成了完整的科学研究周期。海量数据也催生了新的科研探索，由软件处理各种仪器或模拟实验产生了大量数据，并将得到的信息或知识存储在计算机中，科研人员只需要从这些计算机中查找数据。例如，在天文学研究中，科研人员并不直接通过天文望远镜进行研究，而是从数据中心查找所需数据进行分析研究，数据中心存储有海量的、由各种天文设备收集到的数据。海量数据的管理是以网络为基础的科学活动环境中重要的组成部分。

众多的科学和工程应用计算都需要处理大量的数据，需要处理的数据量级达到TB 或 PB。位于欧洲核子研究中心的大型强子对撞机(large hadron collider，LHC)每年产生 25PB 的数据，美国宇航局的卫星每天将处理或生成超过 2TB 的数据，全球气候变暖模拟实验也产生 TB 数量级的数据。例如，天气预报的计算、飞机模型的计算、流场计算等领域都是把连续变量离散化，用差商来代替微商进行计算的。计算问题的精度要求越高，变量离散的区间越小，计算的数据量也就越大。这类问题的求解一般都需要访问和存储大量的数据。应用领域中不仅一个程序需要访问大量的数据，不同的程序之间也需要传输大量的数据。数据密集型的科学计算和工程应用需要在系统之间传输的数据量达到了 TB 甚至更高数量级。一些数据分析应用程序和可视化显示的应用程序需要访问在地理位置上广泛分布的大量数据。

数据规模不断扩大，给数据管理带来新的挑战，大规模数据管理需要高效存储、放置、调度 PB 级甚至 EB 级的数据，同时在数据计算和处理过程中能够保证中间数据的容错，以避免计算任务的失败，缩短计算任务的完成时间。

随着现代高科技的发展，以网络为基础的科学活动环境成为科学研究中必不可少的一部分，相关技术一直是国际计算机科学领域的研究热点，从并行计算、网格计算和效用计算到当前的云计算，都给科学研究乃至人们的生活带来巨大的变革。

本书介绍了以网络为基础的科学活动环境中的数据管理技术，集中介绍了网络数据管理的主要技术，对目前数据管理的情况进行了综述，相信本书对开始从事网络数据管理的研究人员、工程技术人员和希望了解网络数据管理的普通读者都会有所帮助。

本书作者们的研究工作得到了国家自然科学基金项目“网格标准基础研究”

(No.90812001)的资助，并得到了国家自然科学基金委员会“以网络为基础的科学活动环境研究”重大研究计划学术指导专家组的悉心指导，在此表示深深的谢意！

中国科学院高能物理研究所的汪璐、黄秋兰、伍文静也参加了本书的编写，汪璐主要参与第 3 章和第 4 章，黄秋兰主要参与第 5 章和第 6 章，伍文静主要参与第 7 章、第 11 章和第 12 章。本书编写过程中还得到了中国科学院高能物理研究所陈刚研究员的大力支持，在此表示感谢。

由于作者水平所限，加之网络计算环境下数据管理和大数据技术的研究仍处于不断的发展和变化之中，书中的疏漏和不足之处恳请读者批评指正。

作　者

2014 年 8 月

目　录

第二篇　非结构化数据管理

第三篇　结构化数据管理

第四篇 应 用 实 例

第一篇　概　论

第 1 章

数据管理背景

1.1 数 据 增 长

人类探索世界的脚步永无止境，而科学研究的方式也在不断发展。远古时期，人们依靠观察和思辨来认识和探索世界。17 世纪以来，随着牛顿经典力学基本运动定律的发表，科学家逐渐把实验与理论作为科学研究的基本手段。然而，随着人类探索世界的不断深入，许多科学问题的实验研究和理论研究变得越来越复杂，甚至难以给出明确的结论。近半个世纪以来，随着电子计算机的诞生与快速发展，计算机仿真模拟变成第三种不可或缺的科学研究手段，以帮助科学家去探索实验与理论难以解决的问题，如宇宙的起源、汽车碰撞、天气预报等。而在当前社会，各个学科领域的研究不断向纵深发展，无论实验装置还是计算机仿真模拟的规模都变得越来越大，产生了越来越多的数据，从而催生了围绕海量数据获取、存储、共享和分析的科学研究手段。来自科学仪器或者计算机仿真模拟的实验数据被收集和存储起来，并通过先进高速的网络分享给处于不同的国家或机构的合作者。依靠分布式计算技术和协同工作环境，科学家不仅共享数据，还共享软件、模型、计算、专家知识甚至人力等资源，从而加快科学成果的产出。现代科学研究，特别是粒子物理、生命科学、能源环境、先进材料与纳米科学等新兴或交叉领域的发展要进行跨国家、跨地域的协作与交流，而以网络为基础的科学活动环境的发展与完善正在对其产生深远的影响。

在“纸笔研究”时代，科学家的数据记录在笔记本上，帮助分析数据的工具可能是一把尺子。在今天，科学研究成果的获得不仅取决于科学家的智慧和勤奋，还取决于海量科学数据的处理能力。基于海量数据处理的科学探索已经成为一种新的科学研究方法，也是科研信息化的重要内容之一。

科学仪器和电脑仿真产生的新数据以每年一倍的速度急速扩张，超过了 CPU 处理能力的增长速度(摩尔定律：CPU 处理能力每 18 个月翻一番)。1946 年，美国军方的 ENIAC(electronic numerical integrator and computer)被称为世界上第一台“电