



中国计算机学会文集

China Computer Federation

Proceedings

CCFP 0008

# 中国计算机科学技术 发展报告 2007

庆祝“973”计划实施十周年

中国计算机学会 主编

清华大学出版社



中国计算机学会文集

China Computer Federation

Proceedings

CCFP 0008

# 中国计算机科学技术 发展报告 2007

庆祝“973”计划实施十周年

中国计算机学会 主编

清华大学出版社  
北京

## 内 容 简 介

本书是由中国计算机学会学术工作委员会组织编写的、权威性的计算机科学技术年度发展报告。本报告汇总了2003年以来,关于可信网络与普适服务、可视媒体、大型应用软件、微处理芯片、新一代互联网、无线传感网络、网络存储、虚拟计算环境、计算系统虚拟化、语义网格等主题的11个“973”项目,分别由活跃在各个方向上的科研人员撰写项目进展报告,详细介绍了对应项目在理论研究和实现等方面取得的进展。从一定角度反映了我国计算机科学与技术工作者当前的研究进展,对学术研究有重要参考价值。

本书可作为广大计算机科学技术人员了解当前计算机科学技术发展动态的一个渠道,适合本领域决策人员和科研人员来参考,并可作为高等院校硕士生、博士生的参考书。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

## 图书在版编目(CIP)数据

中国计算机科学技术发展报告 2007: 庆祝“973”计划实施十周年/中国计算机学会主编.  
—北京: 清华大学出版社, 2008. 11

ISBN 978-7-302-18814-8

I. 中… II. 中… III. 计算机科学—发展—研究报告—中国—2007 IV. TP3

中国版本图书馆 CIP 数据核字(2008)第 166451 号

责任编辑: 薛 慧

责任校对: 赵丽敏

责任印制: 李红英

出版发行: 清华大学出版社

地 址: 北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编: 100084

社 总 机: 010-62770175

邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者: 北京市世界知识印刷厂

装 订 者: 三河市新茂装订有限公司

经 销: 全国新华书店

开 本: 185×260 印 张: 31.75 字 数: 762 千字

版 次: 2008 年 11 月第 1 版 印 次: 2008 年 11 月第 1 次印刷

印 数: 1~8000

定 价: 65.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话: (010)62770177 转 3103 产品编号: 031327-01

# 序 言

信息技术是过去数十年发展最快、应用最为广泛的科学技术,是高技术的代表,并将继续作为未来人类社会经济发展的重要驱动力。

国家重点基础研究发展计划(“973”计划)是以国家重大需求为导向,对我国未来发展和科学技术进步具有战略性、前瞻性、全局性和带动性的基础研究发展计划,是国家科技计划体系的重要组成部分。“973”计划实施十年来,针对我国信息产业自主创新能力薄弱、核心技术缺乏等突出问题,着力从微电子、光电子、下一代信息网络与通信、高可信高效率软件、量子通信等多个方面进行前瞻性的部署,通过科学家们的辛勤工作,取得了一批在国际学术界具有重要影响的创新成果,为我国信息产业的跨越式发展发挥了重要作用。

值此“973”计划实施十周年之际,欣闻中国计算机学会将“973”计划信息领域众多与计算机科学紧密相关的项目汇集于此,塑成《中国计算机科学技术发展报告 2007》一书,这是对我国计算机及信息科学技术发展历程的一次记录,也从一定角度反映了我国信息科学领域发展的现状,具有重要参考价值。



科技部基础研究司司长

2008年10月16日

# 前 言

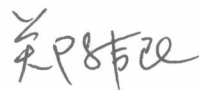
近年来,信息技术发展速度惊人,已成为改造传统产业、振兴国民经济的重要支柱。在我国,信息产业正以高于全国 GDP 三倍的速度快速增长,已成为经济发展中增长最快、最具活力的产业之一。然而,我国信息技术的科技自主创新能力薄弱、核心技术缺乏,导致我国信息高科技产品整体上仍以仿制为主,缺乏自己的创新特色。为了根本改变这一现状,必须大力加强基础研究,全面提高我国信息领域的创新能力。

“十一五”期间,信息领域基础研究的发展目标是借鉴国际信息技术的发展趋势,结合我国的国情和发展状况,着重从信息获取、处理、传输、存储、再现、安全、利用等方面开展系统深入的研究,为我国信息产业的跨越式发展奠定坚实的理论技术基础。

本报告汇总了 2003 年以来、关于可信网络与普适服务、可视媒体、大型应用软件、微处理芯片、新一代互联网、无线传感网络、网络存储、虚拟计算环境、计算系统虚拟化、语义网格等主题的 11 个“973”项目,分别由活跃在各个方向上的科研人员撰写项目进展报告,详细介绍了对应项目在理论研究和技術实现等方面取得的进展。

这一报告是我国计算机科学技术发展历程的一个记录,从一定角度反映了我国计算机科学与技术工作者当前的研究进展,对学术研究有重要参考价值。它对进一步推动我国计算机科学技术的振兴和发展,推动我国信息化进程将起到重要作用。同时,也可以作为广大计算机科学技术人员了解当前计算机科学技术发展动态的一个渠道,适合本领域决策人员和科研人员参考。

值此报告行将付梓之际,谨向为本报告贡献稿件的所有专家学者表示感谢。报告中的观点虽仅代表撰稿人的个人意见,但仍具有非常重要的参考价值。在内容安排上,各个项目进展报告以立项时间为序。由于策划组稿时间仓促,本报告在主题选择、形式和内容安排上一定存在可以进一步改进的地方,期望得到广大专家和读者的批评指正。



中国计算机学会副理事长

2008 年 7 月 18 日

# 目 录

|                                         |            |
|-----------------------------------------|------------|
| 语义网格的基础理论、模型与方法研究 .....                 | 诸葛海 等 (1)  |
| 1 语义网格 .....                            | (1)        |
| 2 核心科学问题 .....                          | (2)        |
| 3 课题研究进展 .....                          | (2)        |
| 3.1 语义资源空间模型的理论、模型与方法研究 .....           | (2)        |
| 3.2 语义网格的过程语义及其验证理论与服务集成技术 .....        | (4)        |
| 3.3 基于语义网格的语义关联存储模型及管理通信平台 .....        | (7)        |
| 3.4 语义网格资源描述模型、形式化理论和支撑技术 .....         | (15)       |
| 3.5 语义网格环境下支持产品创新设计的知识供应理论与技术 .....     | (18)       |
| 3.6 语义网格在中医药知识共享与服务中的应用研究 .....         | (20)       |
| 3.7 基于语义的网络信息检索、集成与应用 .....             | (22)       |
| 3.8 基于语义的资源服务协同和管理研究 .....              | (25)       |
| 4 总结 .....                              | (28)       |
| 代表论著 .....                              | (29)       |
| 新一代互联网体系结构理论研究 .....                    | 吴建平 等 (32) |
| 1 新一代互联网发展面临的主要矛盾和基本科学问题 .....          | (33)       |
| 1.1 下一代互联网及其基本特征 .....                  | (33)       |
| 1.2 新一代互联网络发展面临的主要矛盾 .....              | (34)       |
| 1.3 四个基本的科学问题 .....                     | (35)       |
| 参考文献 .....                              | (39)       |
| 2 新一代互联网体系结构基础理论 .....                  | (39)       |
| 2.1 互联网和下一代互联网的基础理论研究逐步受到人们的重视 .....    | (40)       |
| 2.2 新一代互联网体系结构的基本科学问题及内在联系 .....        | (42)       |
| 2.3 多维可扩展的新一代互联网体系结构 .....              | (42)       |
| 2.4 多维可扩展的新一代互联网体系结构中的五个基本要素 .....      | (43)       |
| 2.5 基于 IPv6 真实地址的可信任新一代互联网 .....        | (44)       |
| 参考文献 .....                              | (45)       |
| 3 新一代互联网路由交换理论 .....                    | (45)       |
| 3.1 引言 .....                            | (45)       |
| 3.2 主要研究成果 .....                        | (46)       |
| 3.3 代表性成果 1——基于 ForCES 的开放路由器体系结构 ..... | (46)       |
| 3.4 代表性成果 2——可扩展的交换设备模型 .....           | (50)       |

|     |                                         |           |
|-----|-----------------------------------------|-----------|
| 3.5 | 代表性成果3——域间路由优化理论与技术 .....               | (52)      |
| 3.6 | 结论和下一步的工作 .....                         | (54)      |
|     | 参考文献 .....                              | (54)      |
| 4   | 网络动态行为和传输控制理论 .....                     | (55)      |
| 4.1 | 引言 .....                                | (55)      |
| 4.2 | 主要研究成果 .....                            | (55)      |
| 4.3 | 代表性成果1——高性能传输系统的性能分析与优化设计 .....         | (58)      |
| 4.4 | 代表性成果2——超点检测研究 .....                    | (62)      |
| 4.5 | 结论和下一步的工作 .....                         | (66)      |
|     | 参考文献 .....                              | (67)      |
| 5   | 可信任的互联网安全体系结构和安全监控理论 .....              | (67)      |
| 5.1 | 引言 .....                                | (67)      |
| 5.2 | 主要研究成果 .....                            | (69)      |
| 5.3 | 代表性成果1——真实 IPv6 源地址寻址体系结构及验证方法 .....    | (71)      |
| 5.4 | 代表性成果2——基于源地址认证的组播安全认证 .....            | (73)      |
| 5.5 | 结论和下一步的工作 .....                         | (76)      |
|     | 参考文献 .....                              | (76)      |
| 6   | 新一代互联网服务模型和服务管理理论 .....                 | (77)      |
| 6.1 | 引言 .....                                | (77)      |
| 6.2 | 主要研究成果 .....                            | (77)      |
| 6.3 | 代表性成果——分布式环境中服务组合的迭代选择算法 .....          | (78)      |
| 6.4 | 结论和下一步的工作 .....                         | (89)      |
|     | 参考文献 .....                              | (89)      |
| 7   | 新一代网络远程实验室 DRAGON-Lab .....             | (90)      |
| 7.1 | 引言 .....                                | (90)      |
| 7.2 | 主要研究成果 .....                            | (91)      |
| 7.3 | 代表性成果1——DRAGON-Lab 网络数据共享平台的设计与实现 ..... | (91)      |
| 7.4 | 代表性成果2——地址随机化算法研究 .....                 | (93)      |
| 7.5 | 结论和下一步的工作 .....                         | (97)      |
|     | 参考文献 .....                              | (98)      |
| 8   | 总结与展望 .....                             | (98)      |
|     | 现代设计大型应用软件的共性基础 .....                   | 孙家广 (101) |
| 1   | 概述 .....                                | (102)     |
| 2   | 研究工作主要进展 .....                          | (104)     |
| 2.1 | 智能设计理论与方法研究 .....                       | (105)     |
| 2.2 | 产品定制设计与优化设计研究 .....                     | (111)     |
| 2.3 | 产品设计中的几何计算理论与方法 .....                   | (119)     |
| 2.4 | 从二维信息重构相应三维形体理论与方法研究 .....              | (130)     |

|                                  |       |
|----------------------------------|-------|
| 2.5 产品协同设计理论与方法研究 .....          | (137) |
| 2.6 协同设计大型应用软件体系结构及形式化机理研究 ..... | (147) |
| 3 今后的研究工作 .....                  | (150) |
| 参考文献 .....                       | (151) |

下一代互联网信息存储的组织模式和核心技术研究 ..... 冯 丹 (160)

|                                   |                   |
|-----------------------------------|-------------------|
| 1 基于对象的海量存储系统关键问题研究 .....         | 冯 丹 等 (164)       |
| 1.1 引言 .....                      | (164)             |
| 1.2 主动存储对象模型 .....                | (165)             |
| 1.3 高可用性研究 .....                  | (173)             |
| 1.4 高效元数据管理 .....                 | (176)             |
| 1.5 结论 .....                      | (179)             |
| 参考文献 .....                        | (179)             |
| 2 下一代网络存储的高速通道研究 .....            | 韩冀中 贺 劲 韩承德 (182) |
| 2.1 引言 .....                      | (182)             |
| 2.2 下一代网络存储的高速通道 .....            | (184)             |
| 2.3 分布式缓存的研究 .....                | (186)             |
| 2.4 轻量级通信协议及相关通信机制 .....          | (187)             |
| 2.5 应用加速引擎机制 .....                | (188)             |
| 2.6 总结与未来展望 .....                 | (189)             |
| 参考文献 .....                        | (190)             |
| 3 海量数据网络存储系统的研究 .....             | 曹 强 谢长生 (191)     |
| 3.1 引言 .....                      | (191)             |
| 3.2 海量数据网络存储系统总体结构 .....          | (192)             |
| 3.3 存储系统的软件结构 .....               | (195)             |
| 3.4 数据的组织方式 .....                 | (198)             |
| 3.5 存储安全技术研究 .....                | (200)             |
| 3.6 原理性研究 .....                   | (202)             |
| 参考文献 .....                        | (205)             |
| 4 P2P 存储系统关键问题研究 .....            | 杨 智 田 敬 代亚非 (208) |
| 4.1 引言 .....                      | (208)             |
| 4.2 P2P 存储系统需要研究的关键问题 .....       | (208)             |
| 4.3 深入理解 P2P 环境的动态性 .....         | (210)             |
| 4.4 数据的持久性研究 .....                | (210)             |
| 4.5 数据的安全性研究 .....                | (212)             |
| 4.6 UPStore: 一个开放的 P2P 存储平台 ..... | (213)             |
| 参考文献 .....                        | (214)             |
| 5 网络存储对象化及其动态部署的核心技术研究 .....      | (215)             |
| 5.1 引言 .....                      | (215)             |

|     |                                                  |       |
|-----|--------------------------------------------------|-------|
| 5.2 | 总体思路及主要研究内容 .....                                | (216) |
| 5.3 | 服务动态部署技术 .....                                   | (218) |
| 5.4 | 虚拟存储技术 .....                                     | (221) |
| 5.5 | 数据备份和恢复技术 .....                                  | (224) |
| 5.6 | 未来工作 .....                                       | (226) |
|     | 参考文献 .....                                       | (226) |
| 6   | 大规模网络存储系统服务质量的研究..... 舒继武                        | (227) |
| 6.1 | 引言 .....                                         | (227) |
| 6.2 | 主要研究内容与进展 .....                                  | (227) |
| 6.3 | 进一步的工作 .....                                     | (236) |
|     | 参考文献 .....                                       | (236) |
| 7   | 分布式存储系统的应用——多媒体网络 GIS 及其<br>关键技术的研究..... 喻占武 李忠民 | (237) |
| 7.1 | 引言 .....                                         | (237) |
| 7.2 | 基于双重集群的网络 GIS 体系结构 .....                         | (239) |
| 7.3 | 按 PB 级处理能力设计大型 GIS 应用系统 .....                    | (240) |
| 7.4 | 基于对象存储的空间数据存储模型 .....                            | (241) |
| 7.5 | 地形数据的分布策略 .....                                  | (242) |
| 7.6 | 所取得的进展 .....                                     | (242) |
| 7.7 | 总结 .....                                         | (243) |
|     | 参考文献 .....                                       | (243) |
|     | 延长摩尔定律的微处理芯片新原理、新结构与新方法研究..... 李国杰 胡伟武           | (246) |
| 1   | 研究背景 .....                                       | (247) |
| 2   | 主要研究内容 .....                                     | (249) |
| 2.1 | 可扩展、可重构片上并行体系结构及原型芯片设计 .....                     | (249) |
| 2.2 | 片上并行系统的编程模型与支撑环境 .....                           | (250) |
| 2.3 | 低功耗体系结构、电路设计与功耗管理 .....                          | (250) |
| 2.4 | 高性能处理芯片的可靠性设计 .....                              | (252) |
| 2.5 | 高性能处理芯片的设计验证与测试 .....                            | (253) |
| 3   | 研究进展 .....                                       | (255) |
| 3.1 | 片上并行体系结构研究进展 .....                               | (255) |
| 3.2 | 片上并行编程系统与支撑环境 .....                              | (258) |
| 3.3 | 低功耗体系结构、电路设计与功耗管理 .....                          | (262) |
| 3.4 | 高性能处理芯片的可靠性设计 .....                              | (265) |
| 3.5 | 高性能处理芯片的设计验证与测试 .....                            | (268) |
| 4   | 总结和未来工作 .....                                    | (275) |
|     | 参考文献 .....                                       | (275) |

|                                           |           |
|-------------------------------------------|-----------|
| 虚拟计算环境聚合与协同机理研究 .....                     | 卢锡城 (279) |
| 1 研究背景 .....                              | (280)     |
| 2 科学问题、概念模型与体系结构 .....                    | (282)     |
| 2.1 科学问题 .....                            | (282)     |
| 2.2 概念模型 .....                            | (283)     |
| 2.3 体系结构 .....                            | (286)     |
| 3 虚拟计算环境的关键技术 .....                       | (287)     |
| 3.1 资源虚拟化 .....                           | (287)     |
| 3.2 资源按需聚合 .....                          | (290)     |
| 3.3 自主协同 .....                            | (295)     |
| 3.4 可信保障 .....                            | (297)     |
| 3.5 编程开发和运行环境 .....                       | (300)     |
| 4 实验与验证 .....                             | (303)     |
| 4.1 实验床 .....                             | (303)     |
| 4.2 应用案例分析 .....                          | (305)     |
| 5 相关工作比较 .....                            | (306)     |
| 6 结论和进一步研究 .....                          | (308)     |
| 参考文献 .....                                | (310)     |
| 海量信息的协同性和可生存性的理论与实践研究 .....               | 李 未 (314) |
| 1 绪论 .....                                | (316)     |
| 1.1 引言 .....                              | (316)     |
| 1.2 研究目标 .....                            | (318)     |
| 1.3 工作基础和条件 .....                         | (319)     |
| 1.4 问题划分 .....                            | (324)     |
| 1.5 研究团队 .....                            | (324)     |
| 2 国内外研究现状和发展趋势 .....                      | (325)     |
| 2.1 面向服务的网格计算和对等网计算 .....                 | (326)     |
| 2.2 海量信息协同系统 .....                        | (327)     |
| 2.3 分布式软件系统可生存性研究 .....                   | (327)     |
| 2.4 我国的相关研究工作 .....                       | (327)     |
| 3 海量信息协同性和可生存性的信息科学特征研究 .....             | (328)     |
| 3.1 系统信息更新的基础理论研究 .....                   | (328)     |
| 3.2 海量信息传输的观测试验研究 .....                   | (330)     |
| 3.3 海量信息计算的统计特征研究——RB 模型的研究 .....         | (333)     |
| 3.4 小结 .....                              | (333)     |
| 4 海量信息系统协同性随机模型与可生存性复杂性分析 .....           | (334)     |
| 4.1 Spin Glasses 理论与海量信息协同的算法和复杂性研究 ..... | (334)     |
| 4.2 海量信息系统协同性的随机模型研究 .....                | (335)     |

|                     |                                 |           |
|---------------------|---------------------------------|-----------|
| 4.3                 | 海量信息系统可生存性的代数学研究 .....          | (336)     |
| 4.4                 | 网络环境下协同计算的形式化模型与行为研究 .....      | (337)     |
| 4.5                 | 小结 .....                        | (338)     |
| 5                   | 实时协同可生存的海量信息系统的试验与验证平台 .....    | (339)     |
| 5.1                 | 海量协同服务网络的体系结构 .....             | (339)     |
| 5.2                 | 基于内容的大量协同服务 .....               | (342)     |
| 5.3                 | 可生存的协同服务网络 .....                | (351)     |
| 5.4                 | 面向复杂协同网络的试验平台 .....             | (356)     |
| 5.5                 | 小结 .....                        | (357)     |
| 6                   | 可生存的海量信息系统软件设计理论 .....          | (357)     |
| 6.1                 | 需求分析 .....                      | (357)     |
| 6.2                 | 研究模型与技术 .....                   | (358)     |
| 6.3                 | 小结 .....                        | (359)     |
| 7                   | 海量信息系统可生存性分析方法及软件可生存性增强技术 ..... | (360)     |
| 7.1                 | 软件体系结构与服务分布 .....               | (360)     |
| 7.2                 | 软件运行状态分析 .....                  | (361)     |
| 7.3                 | 可生存性增强技术 .....                  | (361)     |
| 7.4                 | 小结 .....                        | (362)     |
| 8                   | 海量信息系统软件可生存性策略和测评方法研究 .....     | (364)     |
| 8.1                 | 研究范围 .....                      | (364)     |
| 8.2                 | SAP 模型与关键算法 .....               | (364)     |
| 8.3                 | 可生存性策略研究 .....                  | (365)     |
| 8.4                 | 与国内外同类研究工作相比的创新性 .....          | (367)     |
| 8.5                 | 小结 .....                        | (368)     |
| 可视媒体智能处理的研究进展 ..... |                                 | 胡事民 (369) |
| 1                   | 可视媒体智能处理研究的科学意义 .....           | (369)     |
| 2                   | 研究思路和研究内容 .....                 | (370)     |
| 2.1                 | 可视媒体的认知特征 .....                 | (372)     |
| 2.2                 | 可视媒体的计算理论与方法 .....              | (373)     |
| 2.3                 | 可视媒体的表示、分析与利用 .....             | (373)     |
| 2.4                 | 基于内容的可视媒体安全 .....               | (373)     |
| 2.5                 | 可视媒体的交互与融合处理 .....              | (374)     |
| 2.6                 | 可视媒体的高效计算与系统 .....              | (374)     |
| 3                   | 研究成果及展望 .....                   | (375)     |
| 3.1                 | 项目研究工作的突出进展 .....               | (375)     |
| 3.2                 | 项目队伍建设和人才培养 .....               | (380)     |
| 3.3                 | 科研获奖情况 .....                    | (380)     |
| 4                   | 结束语 .....                       | (380)     |

|                          |                   |
|--------------------------|-------------------|
| 参考文献 .....               | (381)             |
| 无线传感网络的基础理论及关键技术研究 ..... | 倪明选 李明禄 薛广涛 (384) |
| 1 项目意义 .....             | (385)             |
| 1.1 国家需求 .....           | (385)             |
| 1.2 科学意义 .....           | (386)             |
| 2 国内外研究现状和发展趋势 .....     | (387)             |
| 2.1 国外研究现状和发展趋势 .....    | (387)             |
| 2.2 国内研究现状 .....         | (391)             |
| 2.3 小结 .....             | (393)             |
| 3 关键科学问题与研究内容 .....      | (393)             |
| 3.1 关键科学问题 .....         | (393)             |
| 3.2 主要研究内容 .....         | (394)             |
| 3.3 课题设置 .....           | (397)             |
| 4 项目进展 .....             | (398)             |
| 4.1 项目进展 .....           | (398)             |
| 4.2 学术影响 .....           | (401)             |
| 4.3 小结 .....             | (407)             |
| 参考文献 .....               | (408)             |
| 一体化可信网络与普适服务体系基础研究 ..... | 张宏科 (413)         |
| 1 前言 .....               | (415)             |
| 2 国内外研究现状和发展趋势 .....     | (417)             |
| 参考文献 .....               | (420)             |
| 3 科学问题、研究内容和研究目标 .....   | (421)             |
| 3.1 关键科学问题 .....         | (421)             |
| 3.2 主要研究内容 .....         | (423)             |
| 3.3 研究目标 .....           | (424)             |
| 4 研究方案和课题设置 .....        | (425)             |
| 4.1 学术思路 .....           | (425)             |
| 4.2 技术途径 .....           | (425)             |
| 4.3 项目的创新点与特色 .....      | (436)             |
| 4.4 课题设置 .....           | (436)             |
| 5 研究基础、工作条件和研究队伍 .....   | (437)             |
| 5.1 总体情况 .....           | (437)             |
| 5.2 研究基础 .....           | (438)             |
| 5.3 工作条件 .....           | (442)             |
| 5.4 研究队伍 .....           | (443)             |
| 6 研究进展 .....             | (447)             |
| 6.1 原型系统 .....           | (449)             |

|                        |                   |
|------------------------|-------------------|
| 6.2 论文 .....           | (452)             |
| 6.3 特邀报告 .....         | (454)             |
| 6.4 专利 .....           | (454)             |
| 6.5 人才培养 .....         | (454)             |
| 计算系统虚拟化：挑战与机遇 .....    | 金 海 廖小飞 邵志远 (456) |
| 1 引言 .....             | (456)             |
| 2 虚拟化技术的发展与分类 .....    | (458)             |
| 2.1 发展历程 .....         | (458)             |
| 2.2 虚拟化技术的分类 .....     | (460)             |
| 3 虚拟化的相关研究内容 .....     | (469)             |
| 3.1 虚拟计算体系结构 .....     | (469)             |
| 3.2 多计算系统的虚拟化 .....    | (470)             |
| 3.3 虚拟用户使用环境 .....     | (471)             |
| 3.4 虚拟化系统的安全可信机制 ..... | (472)             |
| 3.5 虚拟计算系统的性能评测 .....  | (473)             |
| 3.6 虚拟化应用 .....        | (474)             |
| 4 虚拟化研究进展 .....        | (477)             |
| 4.1 虚拟机基础研究 .....      | (478)             |
| 4.2 虚拟机快速部署系统 .....    | (482)             |
| 4.3 虚拟工作环境 .....       | (485)             |
| 4.4 虚拟集群管理与维护 .....    | (486)             |
| 5 总结 .....             | (488)             |
| 参考文献 .....             | (488)             |
| 索引 .....               | (491)             |

# 语义网格的基础理论、模型与方法研究

诸葛海<sup>1</sup>, 蒋昌俊<sup>2</sup>, 章勤<sup>3</sup>, 瞿裕忠<sup>4</sup>, 傅育熙<sup>5</sup>, 姚敏<sup>6</sup>, 杨广文<sup>7</sup>, 肖侬<sup>8</sup>

<sup>1</sup>中国科学院计算技术研究所, 北京 100080

<sup>2</sup>同济大学电子与信息学院, 上海 200092

<sup>3</sup>华中科技大学计算机科学与技术学院, 武汉 430074

<sup>4</sup>东南大学计算机科学与工程系, 南京 210096

<sup>5</sup>上海交通大学电子信息学院, 上海 200030

<sup>6</sup>浙江大学计算机科学与技术学院, 杭州 310027

<sup>7</sup>清华大学计算机科学与技术系, 北京 100084

<sup>8</sup>国防科学技术大学计算机学院, 长沙 410073

## 摘要

国家重大基础研究计划“973”“语义网格的基础理论、模型与方法研究”项目研究互联网信息有效共享的基础科学问题,旨在突破现有数据模型的局限,建立基于语义的网络资源有效共享的模型、理论和方法。历经三年多研究,在建立相关的理论、模型和方法方面取得了重要进展。该项目的实施显著提高了我国在相关领域的学术影响和地位,带动了我国的相关研究。研究成果对解决信息有效共享这一许多领域的共性问题具有重要的科学意义和应用价值。

## Abstract

The national basic research program project on Semantic Grid's fundamental theory, model and method aims at three fundamental scientific issues of future interconnection environment: normal re-organization, semantic interconnection, and intelligent clustering. This article reports the latest research progress of the project. It intends to breakthrough the limitation of traditional data models to form a new semantic data model and relevant theories for efficiently sharing network resources. It is to establish a semantics-rich intelligent interconnection environment. Important progress has been made through three years' effort. This project has significantly raised the impact of our country in relevant field.

## 1 语义网格

作为一种全球范围的信息共享设施,互联网已成为推动经济和社会发展的动力。但目前的互联网应用平台还不能满足人类社会对网络资源(信息、知识和服务)有效共享和管理的需求。一个重要原因是当前的互联环境缺乏机器可理解的语义信息,难以提供有效的智能服务。为此, Tim Berners-Lee 等人提出建立语义互联网,目前已形成一个研究领域,是通向新一代互联网的一条可能的道路。目前该领域的研究取得了许多有意义的进展(包括提出了 RDF 和 OWL 等语义表示语言),但尚缺乏信息有效共享的基础性理论和模型,特别是基础语义模型。

资源的有效管理是计算机和网络有效工作的基础。计算资源管理发展的里程碑包括文件系统、数据库和操作系统。关系数据模型是一个重要里程碑,它实现了数据的规范化和有效管理。数据库领域产生了三个图灵奖,但经典的数据库模型是在集中式背景下提出来的,当时主要是解决对简单和规范数据的管理问题,并不考虑对复杂、异构、海量、分散、动态的网络资源的有效共享和管理。

实现有效信息、知识和服务共享和管理的一条重要途径是建立适用于互联网环境的语义数据模型,能够对内容进行有效的管理。

我们将语义网格定义为基于语义的智能互联应用环境。它需要一套新的理论和方法来实现对地理分布的各种网络资源进行规范和灵活的组织,实现网络资源的有效共享和智能利用。

本项目从互联网的自然特性、人类社会对互联网应用的需求、信息科学和互联网本身的进化需求所产生的科学问题出发,研究新一代互联网应用平台有效信息共享的关键理论、模型和方法。

## 2 核心科学问题

本项目围绕互联网信息有效共享这一主题,重点研究和解决以下三个关键科学问题:

(1) 规范重构。规范组织是有效管理资源的基本方法,通过定义语义空间的各种范式和操作来规范重构各种网络资源,使动态网络环境下的无序资源规范化,从而提供有效的资源操作,实现有效的资源共享。为此,项目提出包括资源空间模型在内的多种解决方案。规范组织的另一面是自组织性,这是互联网应用的要求,因此,如何协调规范性和自组织性成为互联网时代资源组织的关键问题。

(2) 语义互联。研究如何实现分布在全球的各种网络资源在语义层上互联,消除语义孤岛。通过简洁的方式从资源空间映射到基础语义空间,实现异构资源在语义层的互联。本项目提出包括语义链网络模型在内的多种方案来实现自主的语义互联。通过将语义链网络与资源空间模型相结合,形成基本的语义模型,使得各种资源在基本语义模型中得到统一和互联。

(3) 智能聚融。研究如何解决互联网环境这一复杂系统自组织和优化的问题从而使系统变得更有效力。智能耦合使各种资源(信息、知识和服务)动态耦合起来为某种应用提供智能服务。但紧耦合的优化并不符合互联社会所要求的自主性,为此,我们提出包括流模型和语义社区发现在内的多种解决方案来松耦合地实现这种优化。

在解决以上科学问题的同时,本项目还将开展在文献检索、产品创新设计、中医药、语义搜索引擎等领域开展实际的应用,推动相关的应用研究。

## 3 课题研究进展

### 3.1 语义资源空间模型的理论、模型与方法研究

#### 3.1.1 资源空间模型的成果

分类是人类认识世界和管理资源的最基本的方法,它也可以用来帮助我们管理网络资源。为此,我们系统地提出了基于分类语义的资源空间模型,创新点如下:

- 通过规范多维分类空间,提出了资源空间模型的范式理论及其完整性约束,为准确、正确的资源操作提供了理论保障。

- 为处理分类和资源操作的不确定性提出了概率资源空间模型。
- 有机结合对等网的可扩展性和资源空间的规范性,提出了基于对等网的分散式资源空间模型。
- 为了支持有效的资源查询,提出了适合资源空间的物理存储机制。
- 为了深刻理解资源空间模型的内在特性,提出了资源空间搜索的复杂性理论,资源空间模型查询操作的代数和演算理论,资源空间模型查询操作的完备性理论。
- 提出了资源空间模型和语义链网络的集成及其关系理论以及资源空间模型、互联网本体语言(OWL)和关系数据库间的转换和集成方法。

资源空间模型可望成为互联网资源管理的重要基础模型。

### 3.1.2 互联语义计算模型的成果

互联网成功的关键在于其简洁的超链结构设计和使用模式。为在继承的基础上发展互联网,我们提出了基于关系语义的语义链网络模型,提出了语义链的自动发现机制,建立了资源空间模型到语义链网络的映射、语义链的代数模型、语义关系自动发现的方法,以及结合类比、归纳和演绎的语义关系推理方法,从而构建了独特的语义互联网的模型、理论和方法。资源空间模型和语义链网络模型已在敦煌文化展示领域得到应用,并获得多项知识产权。

资源空间模型和语义链网络模型的结合可望成为未来互联网的重要数据模型。

### 3.1.3 语义覆盖网和知识流的成果

通过在大规模动态网络上构建语义覆盖层来支持智能应用是构建未来智能互联环境的一个重要问题,为此本课题提出可扩展的结构化对等语义覆盖网模型和构建方法,提出基于结构化和非结构化两种对等网模型的分布式资源空间模型和方法。为实现自主的知识共享,提出了对等知识流模型,揭示了知识共享的动态本质和社会性。有机结合知识流网络、语义链网络和对等网络,构成了语义网格的高层体系,为解决大规模网络环境中资源的智能聚融和自组织提供了新方法。在该构建对等语义覆盖网方面,本课题取得了重要进展,在 IEEE TKDE 和 IEEE TPDS 等一流国际刊物上发表了多篇论文。

### 3.1.4 研究特色、学术影响与评价

与国内外同类研究工作相比,本研究提出了独特的语义网格的理念、科学问题以解决问题的理论、模型和方法,研究的基础性、前瞻性和系统性更强。

研究成果发表在计算机领域的旗舰刊物 *IEEE Computer* 和 *Communications of the ACM*, 本领域顶级刊物 *IEEE Transactions on Knowledge and Data Engineering*, *IEEE Transactions on Parallel and Distributed Systems*, *ACM Transactions on the Web* 和 *ACM Transactions on Internet Technology* 以及 VLDB 等一流国际会议。出版的两本英文专著 *The Web Resource Space Model* 和 *The Knowledge Grid* 均是相关方向的首部专著。成果还应邀在系列国际会议做特邀报告 10 余次。发起系列国际会议 SKG(目前已成功举办了三届)。组织国际刊物专刊 10 余次,推动了相关领域的研究。

关于智能聚融的研究成果在 *CACM* 发表,之后得到著名刊物 *Science* 的评论。

Web 发明人 Tim Berners-Lee 在其重要论著 *A Framework for Web Science* 中引用了本课题关于语义链网络的论文(唯一完全出自中国的论文)。研究成果还得到 ACM 主席 Wendy Hall, *Concurrency and Computation: Practice and Experience* 主编 Geoffrey Fox, *Web Semantics* 主编 Carole Goble, 英国皇家计算机学会主席、工程院院士 Nigel Shadbolt, 语义网格倡导者之一 David De Roure, 对等计算专家 Karl Aberer 和 IEEE Fellow Toru Ishida 等专家的引用和正面评价。得到 *IEEE Transactions on Knowledge and Data Engineering* 等一流的国际刊物和 WWW 等一流国际会议的引用。得到生物信息领域国际刊物 *IEEE Transactions on Information Technology in Biomedicine* 和国外同行专著的引用。

SCI 检索表明:项目发表的论文数和被引数在知识网格方向均为第一。研究成果还得到国外专家的积极评论和采用。据 Thomson Essential Science Indicators, 论文引用率居领域 Top 1%, 获国家自然科学基金成立 20 周年优秀成果。研究成果获得 2007 年计算机学会王选奖。

本课题的实施还培养了一批青年科研骨干, 从事该课题研究的博士生获全国优秀博士学位论文提名奖、中国科学院院长特别奖和微软学者奖等奖项。

### 3.1.5 下一步工作

下一阶段研究的重点是结合互联网的发展开展资源空间模型系统的研发、完善资源空间模型和语义链网络模型的理论体系, 提出资源空间模型 2.0 版。强化语义网格技术在文化和科学等领域的应用研究。语义网格致力于建立一种基于语义的智能互联应用环境, 有望在原创性的理论、关键技术以及应用方面产生突破, 形成新的研究领域, 产生重要国际影响。

## 3.2 语义网格的过程语义及其验证理论与服务集成技术

### 3.2.1 引言

语义网格按照语义层次可以划分为如下三层: 资源层、网格层和应用层。资源层主要解决资源的语义描述和规范重构, 网络资源间的语义关系及其演算模型和规律, 网络资源的语义层规范化, 其语义是基于本体论的底层语义; 网格层则主要解决网格资源间的协同、合作等, 实现底层语义和高层语义的相互转化和映射; 应用层面向语义网格的服务工作流和网格服务, 主要解决开放、复杂环境下工作流集成和服务组合中涉及的语义验证问题, 包括过程语义、逻辑语义等高层语义, 这些语义的描述和验证将促进语义网格的语义互联和智能聚融, 并为上层应用提供正确性保证。因此, 本课题旨在解决语义网格所涉及的过程语义、并发语义等系统高层语义的描述模型、分析方法和验证理论等。

为此, 本课题提出了网格、Web 服务的 PN 机语义描述模型及其语义规范。针对上述语义描述模型, 构造了工作流到相应 PN 机模型的转换规范、算法, 基于此, 实现了基于 Web 服务语言规范的服务流程自动建模工具。研究了复杂工作流的递阶建模方法, 定义了 PN 机的精练操作, 并讨论了精练操作的行为表达及其动态性质。提出了一套语义一致性验证理论, 实现对系统高层语义的描述、分析和验证, 丰富和发展了已有的工作流、Petri 网等方面的理论工作。研究了语义网格服务工作流和服务组合的验证分析理论及算法。在过程语义