



Grid Distributed System and Key Technologies of
Agricultural Product Supply Chain

农产品供应链 网格分布式系统及关键技术

王丽娟 / 著



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

Grid Distributed System and Key Technologies of
Agricultural Product Supply Chain

农产品供应链 网格分布式系统及关键技术

王丽娟 / 著



人民邮电出版社
北京

图书在版编目 (C I P) 数据

农产品供应链网格分布式系统及关键技术 / 王丽娟著. — 北京 : 人民邮电出版社, 2019.5
ISBN 978-7-115-50970-3

I. ①农… II. ①王… III. ①网格—应用—农产品—供应链管理—研究 IV. ①F724.72

中国版本图书馆CIP数据核字(2019)第047755号

内 容 提 要

本书主要以网格技术为底层技术支撑, 对农产品供应链分布式 (Agricultural Product Supply Chain Distributed, APSCD) 系统平台进行设计, 并对系统中涉及的网格环境下的分布式仿真平台构建、异构数据集成和信息共享技术、嵌入式移动数据库关键技术以及系统稳定技术等进行研究。

全书分为 4 部分, 共 13 章, 主要包括: 农产品供应链系统网格平台框架、网格环境下农产品供应链分布式仿真平台构建、农产品供应链系统网格资源管理、农产品供应链系统中异构数据无模式数据存储、农产品供应链系统中异构数据简单模式存储、农产品供应链系统中异构数据复杂模式存储、移动数据库复制和缓存技术概述、移动事务处理的异构性研究、基于语义合并的同步复制技术的研究、农产品供应链系统信任模型以及农产品供应链系统中增量监测点机制等内容。

本书适合作为高等院校相关专业的大学生和研究生的教材, 也可供政府部门、科研机构的科技工作者以及从事相关专业的管理人员参考。

| | |
|-----------------|---|
| ◆ 著 | 王丽娟 |
| 责任编辑 | 李彩珊 |
| 责任印制 | 彭志环 |
| ◆ 人民邮电出版社出版发行 | 北京市丰台区成寿寺路 11 号 |
| 邮编 | 100164 |
| 网址 | http://www.ptpress.com.cn |
| 三河市祥达印刷包装有限公司印刷 | |
| ◆ 开本: | 720×960 1/16 |
| 印张: | 14 |
| 字数: | 274 千字 |
| | 2019 年 5 月第 1 版 |
| | 2019 年 5 月河北第 1 次印刷 |

定价: 89.00 元

读者服务热线: (010) 81055493 印装质量热线: (010) 81055316

反盗版热线: (010) 81055315

前言

我国是一个农业大国，建立我国农产品整个产业链新的管理模式，即农产品供应链，对我国农业的发展起着举足轻重的作用。建立合理的农产品供应链体系，优化农产品结构，增强农产品市场竞争力，搭建农产品供应链信息平台，增加农产品在流通过程的价值，也是利用经济解决好“三农”问题的稳健策略。

农产品供应链是指在农产品生产、流通过程中涉及的生产者、中间商和消费者组成的网络体系。由于严重的信息不对称，目前我国农产品供应链成员之间较难形成“利益共享、风险共担”的合作关系，缺乏一个合理、有效的信息综合管理平台。研究农产品供应链网格平台以及关键技术，探讨分布式系统中异构环境下农产品供应链信息集成以及共享，提升农产品的价值，提高整个供应链绩效，是目前亟待解决的关键问题。

作者——华北水利水电大学王丽娟教授采用多学科交叉的方法，运用系统工程理论、技术和方法，借助于计算机、网格和人工智能等技术，依托中国博士后科学基金面上项目、河南省高等学校青年骨干教师项目、河南省科技攻关项目以及郑州市重点科技攻关等项目，对农产品供应链（Agricultural Product Supply Chain, APSC）系统网格平台构建、APSC 系统中异构数据集成和信息共享技术、APSC 系统中嵌入式移动数据库关键技术、网格环境下 APSC 系统稳定关键技术等进行了研究。本书的主要内容是作者近几年来在上述科研项目中的研究成果。

本书由 4 部分组成，共 13 章，第 1 章为绪论，主要介绍相关技术的国内外研究现状和发展趋势；第 2 章为相关技术知识，主要介绍农产品供应链、网格技术、嵌入式数据库、异构数据存储等基本概念和技术；第 3 章为农产品供应链系统网格平台框架，主要介绍农产品供应链网格体系结构、农产品供应链系统框架设计、农产品供应链框架的具体功能以及农产品供应链框架的实现流程；第 4 章为网格环境下农产品供应链分布式仿真平台构建，主要介绍将网格技术和仿真技术融合的研究应

用于农产品供应链系统，将网格技术和仿真技术相融合，建立 GDSF 仿真网格的整体结构模型等内容；第 5 章为农产品供应链系统网格资源管理，主要介绍网格资源的分类组织、负载感知的自主演化、计算资源的动态性维护以及容错机制等内容；第 6 章为农产品供应链系统中异构数据无模式数据存储，主要介绍无模式 XML 数据存储方法设计、XML 文档树的建立、无模式 XML 数据存储算法；第 7 章为农产品供应链系统中异构数据简单模式存储，主要介绍农产品供应链系统中简单模式数据的约束分析、简单模式数据存储策略设计、ASCM 系统中简单模式到关系模式的转换、ASCM 系统中 XML 数据到关系数据的转换等；第 8 章为农产品供应链系统中异构数据复杂模式存储，主要介绍复杂模式数据的存储策略设计、E-Schema 的生成、生成 DOM 树、关系模式映射过程、改写等价模式、加载 XML 文档到关系数据库等技术；第 9 章为移动数据库复制和缓存技术研究，主要介绍数据同步技术、保持移动设备上的数据与服务器上数据的一致性技术等；第 10 章为移动事务处理的异构性，主要介绍移动事务要访问的分布异构数据库系统的解决方案，提出了对异构数据首先进行融合然后进行分类的方法，进而消除冗余，并能达到信息之间互补的效果，分别给出了基于语义融合的异构数据分类模型和算法、多特征 Vague 融合模型；第 11 章为基于语义合并的同步复制技术，主要介绍基于语义的事务关联结果集合及优化同步复制算法设计、基于语义的事务关联结果集合及优化冲突检测及策略等；第 12 章为农产品供应链系统稳定措施，主要介绍农产品供应链系统中使用的信任模型，包括信任的定义、直接信任值的计算、间接信任值的计算等；第 13 章为农产品供应链系统中增量检查点机制，主要介绍检查点机制、Condor 的检查点机制以及系统对 Condor 检查点机制的改进等内容。

在本书的撰写过程中，感谢研究生陈姗、张光辉、马志伟、周伟强、马刚、吕途、毛媛媛、张东等在资料收集和文献整理等方面做的工作。

在全书的撰写和相关技术研究过程中，虽然作者投入了大量精力，然而受知识水平所限，书中不妥之处在所难免，恳请读者批评指正。

作 者

2018 年 11 月 16 日

|| 目 录 ||

| | |
|---------------------------------|-----|
| 第 1 章 绪论 | 001 |
| 1.1 研究的背景及意义 | 002 |
| 1.2 国内外研究现状 | 002 |
| 1.2.1 网格技术 | 002 |
| 1.2.2 网格应用平台 | 003 |
| 1.2.3 网格资源管理 | 005 |
| 1.2.4 网格环境分布式仿真 | 007 |
| 1.2.5 农产品供应链系统中的异构数据存储技术 | 009 |
| 1.2.6 农产品供应链系统中的移动数据库关键技术 | 013 |
| 第 2 章 相关技术知识 | 031 |
| 2.1 供应链 | 032 |
| 2.1.1 集成供应链 | 032 |
| 2.1.2 虚拟供应链 | 032 |
| 2.1.3 农产品供应链 | 033 |
| 2.2 网格平台 | 034 |
| 2.2.1 网格技术 | 034 |
| 2.2.2 网格体系结构 | 036 |
| 2.3 网格仿真 | 042 |
| 2.3.1 网格分布式仿真体系结构 | 042 |
| 2.3.2 基于 GDSF 系统的仿真应用流程 | 048 |
| 2.4 XML 数据存储 | 051 |
| 2.4.1 XML 模式 | 051 |

| | |
|---|------------|
| 2.4.2 XML 解析 | 054 |
| 2.4.3 XML 到关系数据的映射方法 | 055 |
| 2.5 嵌入式移动数据库 | 061 |
| 2.5.1 数据复制与缓存技术 | 061 |
| 2.5.2 移动事务处理技术 | 062 |
| 2.5.3 数据广播技术 | 063 |
| 2.5.4 位置相关查询优化 | 063 |
| 2.5.5 数据的安全性 | 064 |
| 2.5.6 容错和恢复技术 | 064 |
| 2.6 ASCM 系统异构数据 | 064 |
| 第 3 章 农产品供应链系统网格平台框架 | 071 |
| 3.1 供应链网格体系结构 | 072 |
| 3.2 供应链系统框架设计 | 073 |
| 3.3 供应链框架的具体功能 | 074 |
| 3.4 供应链框架的实现流程 | 075 |
| 第 4 章 网格环境下农产品供应链分布式仿真平台构建 | 077 |
| 4.1 农产品供应链分布式仿真平台集成框架 | 078 |
| 4.2 农产品供应链分布式仿真平台中的域结构 | 079 |
| 4.3 基于网格服务实现 HLA/RTI 向网格环境扩展 | 082 |
| 4.4 基于语义的仿真服务的描述 | 083 |
| 4.5 基于语义的仿真服务的匹配算法 | 087 |
| 4.6 系统结构的论证 | 093 |
| 4.7 ASCM 系统网格仿真平台服务调度算法的性能分析 | 093 |
| 第 5 章 农产品供应链系统 网格资源管理 | 097 |
| 5.1 农产品供应链管理系统面临的问题 | 098 |
| 5.2 农产品供应链系统设计 | 099 |
| 5.3 网格资源的分类组织 | 100 |
| 5.3.1 网格资源的分类 | 100 |
| 5.3.2 VO 内网格资源的组织 | 101 |
| 5.3.3 基于 RCT 的网格资源组织结构 | 104 |
| 5.4 RCT 的自组织机制 | 104 |
| 5.4.1 初始化 | 105 |

| | |
|---|------------|
| 5.4.2 负载感知的自主演化 | 106 |
| 5.4.3 计算资源的动态性维护 | 109 |
| 5.4.4 容错机制 | 110 |
| 5.5 资源发现算法与性能分析 | 110 |
| 5.5.1 资源发现 | 110 |
| 5.5.2 性能分析 | 112 |
| 第 6 章 农产品供应链系统中异构数据无模式数据存储 | 113 |
| 6.1 无模式 XML 数据存储方法设计 | 114 |
| 6.2 XML 文档树的建立 | 116 |
| 6.3 无模式 XML 数据存储算法 | 119 |
| 第 7 章 农产品供应链系统中异构数据简单模式存储 | 121 |
| 7.1 ASCM 系统中简单模式数据的约束分析 | 122 |
| 7.2 简单模式数据存储策略设计 | 123 |
| 7.3 ASCM 系统中简单模式到关系模式的转换 | 124 |
| 7.3.1 简化 DTD | 124 |
| 7.3.2 DTD 内联图到关系模式的转换 | 128 |
| 7.4 ASCM 系统中 XML 数据到关系数据的转换 | 129 |
| 7.4.1 XML 树的建立 | 130 |
| 7.4.2 节点编码 | 131 |
| 7.4.3 XML 文档信息到关系表的转换过程 | 132 |
| 第 8 章 农产品供应链系统中异构数据复杂模式存储 | 135 |
| 8.1 复杂模式数据的存储策略设计 | 136 |
| 8.2 E-Schema 的生成 | 137 |
| 8.3 生成 DOM 树 | 140 |
| 8.4 关系模式映射过程 | 142 |
| 8.5 改写等价模式 | 143 |
| 8.5.1 内联改写 | 143 |
| 8.5.2 choice 合并改写 | 144 |
| 8.6 加载 XML 文档到关系数据库 | 145 |
| 第 9 章 移动数据库复制和缓存技术 | 147 |
| 9.1 移动计算与嵌入式移动数据库 | 148 |

| | |
|---------------------------------------|------------|
| 9.1.1 移动计算 | 148 |
| 9.1.2 嵌入式移动数据库 | 149 |
| 9.2 移动数据库中数据缓存技术的研究与应用 | 156 |
| 9.2.1 数据缓存的工作过程 | 156 |
| 9.2.2 数据缓存中的 3 个关键问题和解决方案 | 157 |
| 9.2.3 移动数据库中数据缓存技术的应用 | 159 |
| 9.3 关键技术 | 163 |
| 9.3.1 J2ME | 163 |
| 9.3.2 RMS | 164 |
| 9.3.3 复制与同步的实现 | 165 |
| 第 10 章 移动事务处理的异构性 | 171 |
| 10.1 异构数据库系统 | 172 |
| 10.2 数据库中的 Vague 融合技术 | 173 |
| 10.2.1 基本流程 | 173 |
| 10.2.2 融合模型 | 173 |
| 10.3 数据库中的语义融合技术 | 174 |
| 10.3.1 融合原理 | 174 |
| 10.3.2 融合步骤 | 175 |
| 第 11 章 基于语义合并的同步复制技术 | 177 |
| 11.1 基于语义的事务关联结果集合并优化同步复制算法 | 178 |
| 11.1.1 传统的增量复制 | 178 |
| 11.1.2 移动数据库同步复制粒度分析 | 180 |
| 11.1.3 基于语义合并的事务关联结果集同步复制算法设计 | 181 |
| 11.1.4 合并优化方法 | 182 |
| 11.1.5 语义合并的事务关联结果集同步的工作过程 | 184 |
| 11.1.6 基于语义的事务关联结果集合并优化同步复制算法描述 | 186 |
| 11.2 基于语义的事务关联结果集合并优化冲突检测及策略 | 188 |
| 11.2.1 冲突原因与类型 | 188 |
| 11.2.2 传统的冲突消解策略 | 189 |
| 11.2.3 基于语义合并的冲突检测与消解算法 | 190 |
| 11.2.4 消解算法设计 | 192 |
| 11.3 原型系统实现及实验 | 193 |
| 11.3.1 测试环境与工具 | 193 |

| | |
|--------------------------------------|------------|
| 11.3.2 系统的总体构架 | 193 |
| 11.3.3 系统角色的实现 | 194 |
| 11.3.4 系统测试及分析 | 196 |
| 第 12 章 农产品供应链系统稳定措施 | 203 |
| 12.1 系统中使用的信任模型 | 204 |
| 12.1.1 信任的定义 | 204 |
| 12.1.2 直接信任值的计算 | 205 |
| 12.1.3 间接信任值的计算 | 207 |
| 12.2 系统中使用的供应链性能指标 | 208 |
| 12.2.1 定性评价指标 | 208 |
| 12.2.2 定量评价指标 | 209 |
| 第 13 章 农产品供应链系统中增量检查点机制 | 211 |
| 13.1 检查点机制 | 212 |
| 13.2 Condor 的检查点机制 | 212 |
| 13.3 系统对 Condor 检查点机制的改进 | 214 |



第1章

绪论

1.1 研究的背景及意义

目前，我国的农业信息化程度还比较低。信息滞后是影响农业经营活动的重要因素，在很大程度上制约着农业生产经营活动的顺利开展。产生信息滞后的一个主要原因是农产品经营活动的各个参与方之间信息交流不够充分、相互配合不够完备以及相互之间的关系比较独立，在这些因素的综合作用下，不仅农产品供应链各个参与者的经营成本增高，同时也严重影响着整个供应链系统的正常运转。

以网格技术为底层技术支撑，探索网格环境下农产品供应链（Agricultural Product Supply Chain, APSC）分布式仿真平台构建技术，并对系统中涉及的异构数据集成和信息共享技术、嵌入式移动数据库关键技术以及系统稳定技术等进行研究，可为农产品供应链系统中各种异构软硬资源的高度整合提供较好的技术支撑。

本书的研究内容可为未来继续开发农产品供应链新的信息综合管理平台与新技术提供重要参考，可进一步促进我国农业信息化建设的发展。

1.2 国内外研究现状

1.2.1 网格技术

网格技术在近年来得到了飞速的发展。作为一个可以整合各种异构硬件、软件资源的计算框架，网格技术的研究主要涉及资源发现、安全控制、任务调度和权限管理。

Foster 作为网格领域的领军人物，在网格技术规范的标准化、网格技术整体架构的制定和安全访问等方面都做出了突出的贡献。陈建刚等^[1]针对网格环境的特点，分析了资源访问过程中所遇到的信任管理问题，通过引入管理域内和管理域间的信任关系，结合信任的主观性特点，提出了基于模糊集合的网格资源访问信任机制。谭一鸣等^[2]就树形网格计算环境下的自适应任务调度算法进行了研究，提出了一种

基于树形计算网格的自适应调度算法，实现了对小粒度独立任务和用户大作业的自适应最优调度。通过对网格环境的实时监测，给出了基于节点负载状况、节点任务执行时间、任务传输时间和任务特性的自适应调度算法，即基于最优任务分配方案的启发式任务调度算法^[3]。Takemiya H 等^[4]使用标准网格中间件构建网格应用进行了研究。石志国等^[5]提出了一种面向网格计算的分布式匿名协作算法，提出了可信计算技术，为提高网格协作安全性提出了一种匿名分组身份验证算法，该算法可以非常可靠地解决网格计算平台之间的身份匿名验证问题，算法中所有涉及的验证过程都是基于匿名机制实现的。除了实现匿名验证机制以外，算法还提供了一套完整标记恶意网络实体的方法，提出了网格计算中虚拟分组的匿名认证平台架构，并在此架构基础上分成 5 步实现匿名验证算法。方存好等^[6]对网格计算环境中的主动服务模型进行了研究。Ranganathan^[7]对数据网格中的计算和数据调度算法进行了深入的研究，并提出了一个通用的可扩展的调度算法。李民等^[8]提出了网格环境下的一种基于无向带权图的矩阵子任务合并方法，可以用来解决网格计算中的任务划分问题。张伟哲等^[9]对支持资源自动发现的网格计算门户原型进行了研究。梅一多等^[10]提出了一种基于对等计算方式的网格资源发现模型，使用多种策略保证了网格资源发现服务的可扩展性和消息的路由效率。赵曦滨等^[11]对网格计算中面向虚拟组织的高可用安全目录服务进行了研究，针对虚拟组织对目录服务的要求，结合检查点归档算法及多机部署方案，设计并实现了基于 Web services 的轻量目录服务，并针对目录服务自身的安全做了基于代理服务器和服务迁移机制的安全扩展。这种目录服务可以在保持服务安全性和可靠性的同时为虚拟组织提供良好的可用性。

1.2.2 网格应用平台

目前，将网格技术应用到农业领域的研究还处在探索阶段。国家数字林业网格（National Digital Forestry Grid, NDFG）是科学技术部“863”计划高性能计算机及其核心软件的专项课题，目标在于配合国家数字林业建设，针对林业与生态环境建设工程应用需求，运用网格技术最新成果，建立面向森林资源与林业生态工程的行业应用网格，从而有效地支持我国森林资源和林业生态工程的管理和决策。李凡等^[12]对国家数字林业网格体系结构进行了深入的研究，论证了林业信息化的发展过程及建立国家数字林业网格的意义，说明了网格技术的特点及采用这项技术进行林业信息化的优势，阐述了数字林业网格设计的目标和作用，提出了数字林业网格体系结构的

建立及先进方法的应用。黄震春^[13]基于 Java 语言设计实现了分布式的网格资源注册元服务，该元服务在数字林业网格中，为数字林业网格提供稳定和高效的资源注册与查找服务。张风霞等^[14]对基于构件网格的农业专家系统做了深入的研究，在农业 Web 服务器中存在大量的智能化构件资源，常规的服务组织方式不利于这些构件之间的协同服务和分发，基于构件网格提出了农业专家系统的模型驱动体系结构，分析了分布式构件间的集成规则，并研究了数据挖掘构件、推理机构件和构件组装的接口问题及专家系统构件组装的实现方案，对构件资源的描述信息采用 Web 服务技术标准，为 Internet 上的这些构件资源提供统一的访问和组装方法。

赵新娟等^[15]针对当今供应链系统的现实情况，把供应链管理技术、网格技术综合地运用在一起，提出了基于网格计算的供应链管理系统模型。陈志圣等^[16]利用网格技术实现节点间的信息共享，使信息在各个网格节点之间高效地传递，从而降低了供应链系统的整体成本。由于供应链设计的区域环境的规模相对较小，因此，更适合使用网格技术来构建供应链系统。王秦等^[17]分析了网格计算对供应链管理的影响，通过模型求解、优化、比较，设计了网格计算环境下供应链管理的最优化模型；并且通过仿真实验，证明了使用网格技术搭建供应链系统平台的优势。石声波等^[18]通过对网格计算和供应链管理特性进行分析，提出了基于网格计算的供应链管理理论，从而创造性地提出了供应网格的概念，同时对供应网格进行了研究和分析，进一步总结出了供应网格的特性及供应网格是供应链技术和网格技术发展的必然趋势这一结论，为供应链管理系统的发展指出了一个新的发展方向。

李绍俊、钟耳顺^[19]提出以多级分区空间索引目录结构为基础的对等导航服务网格体系结构。付建芳^[20]通过将网络通信技术、计算机技术、地理网格技术和 3S (GIS、GPS、RS) 技术进行集成，提出基于网格化来管理基层应急资源。以智慧城市建设为全局视角，提出了一种全新的政府各部门使用网格化管理资源的服务模式，实现地理空间信息资源的共享与交换。

贺雅楠等^[21]基于气象信息综合分析处理系统 MICAPS4 (Meteorological Information Combine Analysis and Process System) 设计并实现了智能网格预报平台，采用模型—视图—视图模型 (Model-Views-Views Model, MVVM) 的设计模式，实现业务逻辑与视图的分离，通过对各子功能模块的划分，降低模块之间的耦合度。林秀玉等^[22]通过一种基于网格管理模型 (Service-Oriented Architecture, SOA) 的空间信息共享平台设计框架，结合空间信息多级网格和 SOA 服务体系框架的优点，在广域网络环境下整合空间信息资源，从而有利于空间信息共享。柴磊等^[23]提出基于分级诊疗的网格化管理平台框架，基于客户机/服务器工作模式，构建由资源库和网

格单元两大模块组成的基于分级诊疗的网格化管理平台。陈祥涛等^[24]结合 Globus 工具包、集群技术和 CSCD 技术，采用 OGSA 的网格体系结构，提出了一种网格支持的产品协同设计框架——GSCD 模型。邹晓亮^[25]提出了基于网格的协同设计平台架构，将网格计算技术融入协同设计领域，利用网格计算的动态资源融合特性，充分发挥了网格应用优势。吕霞等^[26]从网格平台的云架构体系升级入手，阐述了基于云架构非结构化地质数据存储与管理、快速搭建地质调查业务系统中间件和野外地质调查工作管理与安全保障服务等一系列关键技术的研究与实现。阙文枭等^[27]设计了一种新的桌面网格六层架构模型，并且提供了相应的关键技术及实现方法，解决了桌面网格技术中的应用部署难、整合异构计算资源难、隔离性差及扩展性差等问题，并使系统在通用性方面得到了很大的提高。黄廷哲^[28]研究了基于全局弱式无网格法的频域电磁探测数值模拟，详细推导了天然场源大地电磁法和人工源频率域电磁法边值问题的弱式等价形式，介绍了全局弱式无网格法求解的基本原理，在均匀半空间模型下进行了算法验证和参数优化，与有限元法的结果进行了对比，表明了算法的正确性和精确性。赵军等^[29]提出大跨 X 形网格立体共梁张弦结构体系，该结构体系减少了资源、设备的投入，满足了项目的工期要求，确保了施工过程的可控性和可靠性，也达到了设计要求。

1.2.3 网格资源管理

由于网格系统越来越庞大，并且网格的异构性、分布性等特点使得网格资源的发现面临越来越多的技术挑战。近年来，为了解决网格资源的发现问题，人们提出了网格信息服务（Grid Information Service, GIS）^[30]的概念，同时出现了许多有代表性的网格信息服务系统，如 Globus MDS 和 Condor 的群匹配机制等。同时，网格系统中也出现了 P2P 和语义等技术^[31-34]。然而在实验中提到的网格资源发现算法却很少在现实的网格系统中得到应用。在实际的部署中，分布式的拓扑结构是大多数 GIS 系统的首选，GIS 的节点只保存资源的元信息，并且资源的元信息被随机地注册到 GIS 节点上，当用户查询可用的、满足条件的网格资源时，搜索算法就要对所有的 GIS 节点进行遍历，从而极大地降低了网格资源发现的效率。因此，当对网格资源按照网格资源的属性进行分类组织和管理时，在查找满足条件的资源时就可以有效地缩小查询范围，从而降低网格资源发现的代价，提高网格资源的发现效率。第二，当网格系统的规模非常庞大时，网格系统就要有一种自适应和自组织的动态演化机制对资源进行组织和管理。

在现有的网格系统中，一般按照集中式、层次式、P2P 和混合结构等对网格资源进行组织和管理，并且基于各自的网格结构设计了不同的网格资源发现算法。在小规模的网格系统中，通常采用集中式资源组织模式，因为在小规模的网格资源系统中，集中式的网格资源管理方式具有较高的效率，但是随着资源的动态加入和网格资源的动态扩大，网格的效率会越来越差，并且出现单点失效和性能瓶颈的可能性会越来越大。因此集中式的网格资源管理算法在大规模网格资源共享和协同中很少应用。Raman^[35]采用了集中式的资源组织与发现机制，通过声明式的 ClassAd 描述用户查询请求和资源属性，采用 Gang-Matchmaking 的方法将用户的查询匹配到合适的资源。

在分层式的网格资源组织结构中，位于较高层次的节点常常要管理比低层次节点更多的信息，因此，节点的负载会随着管理层次的增多而增加，不利于网格系统的扩展。Iamnitchi 等^[36]定义了 GRIP 和 GRRP 两个协议、GIIS 和 GRIS 两个组件用以构建层次化的网格信息服务。李凡等^[12]对 MDS 和 Condor 进行了全面的性能测量与分析。

P2P 是另外一种网格资源的组织方式，目前较常见的 P2P 结构可分为两种，分别为结构化的 P2P 和无结构化的 P2P，结构化的 P2P 一般通过分布式散列表（Distributed Hash Table，DHT）技术，构建描述资源的属性键值和资源位置的映射关系。较常见的结构化的 P2P 项目有 Chord 和 CAN 等^[37-39]，它们都能对网格资源进行快速的定位，但是难以支持多属性的网格资源查询和网格资源的范围查询，而且结构化的 P2P 网格资源组织方式容易出现查询热点引起负载不均衡的问题。而无结构化的 P2P 系统由于没有对资源和资源的存储位置进行强制性的限制，从而降低了网格资源的维护成本，但是由于在资源搜索时采用泛洪式搜索，从而增加了网络的开销^[40]。

在网格的资源组织中，还有一种混合型的网格资源组织结构，它针对具体的需求对网格资源采取特定的网格资源组织方式。这 4 种网格资源的组织方式主要是为了解决将用户的查询请求路由到目标节点的问题，并没有考虑具体应用的特点，另外采用泛洪式的搜索方式进行盲目搜索，增加了网络的负载，降低了资源发现的效率。

在网格的资源发现方面有很多研究成果，Cheema 等^[41]提出了一种基于 DHT 的 P2P 网格资源管理方法，该方法把资源的属性分为静态部分和动态部分，并分别和资源 ID 进行绑定，并用 Pastry 系统中的一个键值来表示，并用 Pastry 环中的层叠的弧来表示网格资源，网格资源静态的属性表示弧的起点，网格资源动态属性的取值

范围表示弧长。同样，这种网格资源的查询方法也没有提供负载均衡机制用以解决查询热点导致的瓶颈问题。Padmanabhan 等^[42]提出了一种自组织组（Self-Organizing Group）的层次与 P2P 的混合网格资源组织结构，通过用网格资源的具体特征描述网格，并根据网格资源特征的相似度组织网格，其中特征相似的网格资源组成一个组，并通过 Gossip 协议选举出一个组的代表，负责组内资源的对外交互任务。这种自组织机制中的“组”的概念类似于本文提出的基于主属性聚类 RCT 的网格资源组织方式，它们的区别主要是基于主属性聚类 RCT 网格资源组织方式是以主属性的值为依据组成的二叉排序树，且 RCT 在定义时充分考虑了具体应用对网格资源的动态需求。

在网格资源的范围查询和多属性查询方面，也有许多重要的研究成果。Gao 等^[43]提出了一种支持 DHT 结构上的范围查询的逻辑范围查询树。由于逻辑范围查询树所有属性的值域在运行之前都被分成了 $2n$ 个子区间，并且每个节点不需要其他节点的协助就可以知道整个树的结构，因此逻辑范围查询树不需要动态地维护整个树的结构。另外，逻辑范围查询树能够根据网格资源节点的查询和注册负载动态地调整逻辑范围查询树的结构。在逻辑范围查询树中，网格资源同时向多个网格节点发送注册信息进行注册，另外由于网格的资源随着时间的变化动态地变化，因此当网格资源发生变化时，就需要向所有注册自身信息的网格节点发送更新信息，这样就给网络带来了巨大的额外开销。Bharambe 等^[44]将所有的网格资源按照属性注册到不同的节点上，每个节点负责管理网格资源的一个属性，并把这种节点称为路由 hub，通过把这种路由 hub 组织成一个层叠的环形网来支持多属性和范围查询。由于网格资源需要将自身的资源注册到所有包含自身属性的网格节点上，因此当网格资源和网格资源的属性分量很多时，就增加了网格资源注册时的开销。基于网格资源主属性聚类的 RCT 由于只采用少数有代表性的属性作为网格资源的主属性进行注册，极大地降低了网格资源注册时的开销。Li 等^[45]提出了通过先建立逻辑 R-Tree，然后将网格资源的数据注册到 R-Tree 上，建立基于跳图的层叠网资源组织方式。在跳图层叠网的基础上，他们又提出了基于跳图层叠网的路由算法，该算法支持在多维属性上的区间查询。和本节提出的基于网格资源主属性聚类的 RCT 相比，跳图层叠网不支持资源动态自动调整的自适应机制。

1.2.4 网格环境分布式仿真

在网格环境分布式仿真研究方面，洪流等^[46]针对网格环境中典型的交互密集型