

网格计算与网格 GIS 体系初探

戴丽君 刘 闯

(中国科学院地理科学与资源研究所,北京 100101)

摘要:本文主要讨论了网格计算的概念,介绍了国内外网格计算与网格体系架构的主要研究进展,并在此基础上研究了基于网格计算的地理信息系统——网格 GIS 的概念。文章在开放网格服务结构 OGSA 的基础上初步提出了包括资源层、中间件层、网格服务层和应用层的网格 GIS 体系结构,并以科学数据中心群网络设计体系为例,对网格 GIS 的体系结构进行深入探究。

关键词: 网格计算; 网格 GIS; 体系架构

中图分类号: P208 **文献标识码:** A **DOI:** 10.3772/j.issn.1674-1544.2008.02.006

Research of Grid Computing and Grid GIS System Structure

Dai Lijun, Liu Chuang

(Institute of Geographic Sciences and Resource Research,
Chinese Academy of Sciences, Beijing 100101)

Abstract: This paper is about Grid computing. It discusses the status and the future of Grid computing. It introduces the definition of Grid GIS, which is based on Grid computing and discusses the structure of Grid GIS based on Open Grid Service Architecture. The Grid GIS structure has resource layer, middleware layer, grid service layer and user layer. It especially talks about the structure of the Science Data Centre Networks.

Keywords: Grid computing, Grid GIS, architecture

地理信息系统 (GIS) 的研究对象是空间信息。空间信息描述的是现实世界中各种空间现象和对象,其空间特征、属性特征和时间特征的特性^[1-2],决定了对地理信息应用的高要求。随着遥感技术的发展,地理数据海量及地理分布广泛造成很大程度上资源与信息的浪费,但是目前 GIS 的数据管理体系结构、方法和技术已经不能满足人们对海量信息的共享、协同分析处理等要求。在现有网格技术发展的基础上,提

出了网格 GIS 的概念^[3-4],希望能解决地理信息系统应用所面临的问题。

1 网格计算与 网格体系结构

在《网格:一种未来的计算基础设施蓝图》中,网格被描述为“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、

第一作者简介:戴丽君(1984-),女,江苏盐城人,博士,主要研究方向是网格计算与地理信息系统。

收稿日期:2008年2月22日。

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,它能让人们共享计算资源、存储资源和其他资源”^[5]。网格实际上是把地理位置上分散的资源集成起来的一种基础设施^[6]。它整合了整个网络上的资源,包括计算机、计算机池、仪器、设备、传感器、数据、软件等资源,形成一个虚拟的大型超级计算机资源,实现计算资源、存储资源、数据资源、知识资源等全面共享和动态发展。

网格技术研究发源于美国,美国的多家研究机构相继开展了与网格相关的研究工作,开发了 Condor、Legion、Globus 等比较有影响的软件 and 工具。美国军方正在实施一个宏大的网格计划,即“全球信息网格(Global Information Grid)”,预计在 2020 年完成。作为这个计划的一部分,美国海军和海军陆战队已启动了一个耗资 160 亿美元、历时 8 年的项目,包括系统的研制、建设、维护和升级。美国能源部(Department of Energy)国家重点实验室 Sandia Livermore 与 Los Alamos 共同承担的 ASCI Grid (Accelerated Strategic Computing Initiative Grid) 先进战略计算创新计划网格计划,是主要用于军事目的的高性能计算发展计划,旨在不进行物理核试验的情况下,通过计算模拟来开展核武器的全方位研究。

欧洲数据网格(European Data Grid)是欧盟支持的项目,目标是为世界范围内分布的科研团体提供计算强度大、共享超过 PB 级数据的大规模分布式数据库的下一代计算基础设施,提高数据分布存储、传输和计算密集型分析处理的能力,以进行科学研究,开展面向高能物理学、地球观测、生物信息学等应用的研究工作。

亚太网格(AP Grid)是亚洲和太平洋地区网格研究的基础设施。参加亚太网格的研究机构有来自澳大利亚、加拿大、中国、印度、日本、马来西亚、新加坡、韩国、泰国、美国等国家的几十个机构,主要研究资源共享、网格技术开发、网格技术推广使用等方面。

我国网格技术的发展从 1999 年开始,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 ACI(Advanced Computational Infrastructure)的研究。1999 年底到 2001 年初,中科院计算所联合十几家科研单位开发了“织女星网格(Vega Grid)”。

网格体系结构是关于如何建造网格的技术,包括对网格基本组成部分和各部分功能的定义和描述,网格各部分相互关系与集成方法的规定,网格有效运行机制的建立等。显然,网格体系结构是网格的骨架和灵魂,是网格最核心的技术^[7]。到目前为止,主流的网格应用的体系结构包括 3 种:一是早期由 Foster 等提出的五层沙漏结构(Five-Level Sandglass Architecture),主要侧重于定性的描述;二是在以 IBM 为代表的工业界的影响下,在考虑 Web 技术的发展及其影响后, Foster 等结合 Web Service 提出的开放网格服务结构 OGSA(Open Grid Services Architecture),其目标是使网格应用的所有服务标准化;三是 Web 服务资源框架 WSRF(Web Services Resource Framework)^[8],是 OGSA 最新核心规范,主要侧重于网格和 Web 服务的融合。

2 网络 GIS 及其体系结构

网格技术给 GIS 的发展带来了机遇和挑战,因此,人们提出了网络 GIS 的概念。例如,“Grid GIS 指基于网格计算的地理信息系统,可以认为网络 GIS 是 GIS 与网格技术的有机结合,是 GIS 在网格环境下的一种应用”^[9]。学者潘显映等给出了网络 GIS 的等式概念,即“Grid GIS = GIS + Grid + Internet + Middleware”。网络 GIS 是建立在网络、网格技术之上的 GIS^[10]。笔者认为,网络 GIS 即 GIS 在网格环境中的应用,是网格技术与 GIS 技术的结合。它采用了网格计算理念,在网格服务环境下构建面向应用的 GIS 平台。GIS 通过网格环境提升了自身功能,实现了继 Web GIS 发展后的再一次质的提高。Grid GIS 可以从网格环境中的任意节点,访问网格上地理信息系统的各种空间信息和海量的数据,并且可以实现空间分析、查询等地理信息系统特有的功能,支

持数据、资源、计算的共享。

网络 GIS 体系结构是基于现有的开放网格服务结构 OGSA (Open Grid Services Architecture) 提出的。开放网格服务结构 OGSA 是 Global Grid Forum 4 的重要标准建议,是继五层沙漏结构之后的一种重要的网格体系结构,是把 Globus 标准与万维网服务 (Web Service) 结合起来,在五层沙漏结构的基础上,结合 Web Service 技术提出的一种新的网格体系结构。OGSA 采用 Web Service 的 WSDL 和 SOAP 规范。OGSA 结构的意义在于它将网格从以科学和工程计算为中心的学术研究领域,扩展到更广泛的以分布式系统服务集成为主要特征的社会经济活动领域^[11]。目前对网络 GIS 的研究有很多,如学者骆剑承、周成虎等提出的基于中间件技术的网络 GIS 体系结构^[3] (利用中间件技术来构建空间信息网格计算环境和服务体系),南京师范大学任建武提出的数据资源层、网格服务层和应用层 3 个层次的网络 GIS 逻辑架构^[12]等。

本文基于对网络计算、网格服务技术和网络中间件的研究,根据地理信息系统应用的需求,提出了基于网络中间件面向地理信息系统应用的空间信息服务的体系结构,包括资源层、网络中间件层、网格服务层和用户应用层。

资源层即当前系统可用的各种资源,包括本地资源和异地可利用资源,如数据资源、计算资源、信息资源等。资源层处于系统的最底层,构成了空间数据与信息共享的物理基础。

网络中间件层包括一系列工具和协议软件,能够用来帮助规划和组建大型的网格试验和应用,网络中间件平台屏蔽网格资源层中计算资源的分布及异构特性,向用户应用层提供透明一致的使用接口,同时提供用户编程接口和相应的环境,以支持网络应用的开发^[13]。

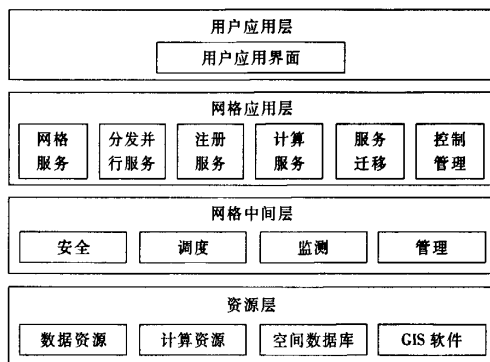
网格服务层是在网络中间件层的基础上结合应用需求进行设计,向上层应用提供统一访问接口,简化应用层的实现,包括故障处理服务、信息服务、安全核心服务、系统运行监控与调度服务等。

用户应用层是由具体的网络 GIS 系统应用界面组成,体现用户需求的软件系统。用户利用

提供的接口和服务完成网格应用的开发,如资源管理、安全监控、故障处理等。

3 基于网络 GIS 体系架构的科学数据中心群网络

科学数据中心群网络是由 UN GAID e-SDDC (促进发展中国家科学数据共享与应用全球联盟) 创建的一个共享数据平台,也是网络 GIS 的一个具体应用。该网络是在 J2EE 环境下,采用网络服务中间件技术,引用 Web Service 及 XML/GML 技术而开发的地理科学数据共享平台。它的体系结构设计基于 Globus Toolkit 中间件架构,充分体现了网络 GIS 的体系结构 (图 1): ①资源层是由网络



节点 N_i 组成,存储数据、计算、信息等资源,包括对地观测数据库 (MODIS、TM、ETM、CBERS、MSS 等卫星影像数据)、自然资源数据库、社会经济统计数据库等。②网络中间件层以 GT4 为基础,实现空间信息的查询索引服务、资源管理、数据管理等功能。其中,查询索引服务是根据用户的需求,自动寻找到合适的资源,自动转换查询和操作参数,自动解释和归纳查询结果,一项索引服务可汇集多个资源信息,并且提供了信息访问接口。资源管理器 GRAM 为用户提供了使用远程系统的简单接口,用来实现作业管理和调度。数据管理模块为用户提供了统一的数据视图及访问接口。③网格服务层实现与数据资源及其应用无关的功能,在安全管理、监控管理的基础上实现数据共享互操作等功能。④用户应用层设计提供身份验证、作业管

理、数据传送、信息服务等服务。用户应用界面是一个利用 JSP、基于 HTML 页面实现的动态图形用户接口。

4 结 语

本文讨论了网格计算的概念,以及网格计算的研究进展,并在此基础上提出了基于网格计算的 GIS,即网格 GIS 的概念,讨论了其体系结构。将 GIS 与网格计算的结合,为解决当前地理信息系统的共享及并行计算等问题提供了很好的帮助。网格 GIS 是一个比较新的研究方向,如何真正把网格计算与 GIS 应用相结合,使网格计算环境成为 GIS 的共享平台,实现理想化的信息共享处理、计算服务等共享及全球意义上的资源共享,还有许多工作要做。

参考文献

- [1] Guting R H. An Introduction to Spatial Database Systems[J]. VLDB Journal: Special Issue on Spatial Database Systems, Chile, 1994, 3 (4): 357 - 399.
- [2] Michael Worboys, Matt Duckham. GIS: A Computing Perspective[M]. CRC Press; Second Edition, 2004.
- [3] 骆剑承,周成虎,蔡少华,等.基于中间件技术的网格 GIS 体系机构[J]. 地球信息科学, 2002(3): 17 - 25.
- [4] 夏曙东,李琦,承继成. 网格技术对地理信息系统应用与发展影响的分析[J]. 地球信息科学, 2003(4): 53 - 56.
- [5] I. Foster and C. Kesselman. The Grid: Blueprint for a New Computing Infrastructure [J]. Morgan Kaufmann Publishers, 1999, 6(3): 417 - 429.
- [6] 徐志伟,冯百鸣,李伟. 网格计算技术[M]. 北京: 电子工业出版社, 2004: 1.
- [7] 董海良,刘强. 基于 Globus Toolkit 4.0 的空间信息服务的实现[D]. 陕西: 电子科技大学, 2007.
- [8] 彭斌,袁亚兵. 网格计算及其应用发展研究[J]. 微计算机信息, 2006, 22(30): 22 - 24.
- [9] 王铮,吴兵. Grid GIS——基于网格计算的地理信息系统[J]. 计算机工程, 2003(3): 38 - 40.
- [10] 潘显映. 网格 GIS 体系研究及其关键技术[D]. 湖北: 武汉理工大学, 2007.
- [11] 黄程林,杨媛媛,赵新国. 面向服务的空间信息栅格体系结构研究[J]. 装备指挥技术学院学报, 2006, 17(5): 47 - 51.
- [12] 任建武. Grid GIS 关键技术研究[D]. 江苏: 南京师范大学, 2003.
- [13] 李伟,万丈高楼平地起——浅谈网格计算基础[EB/OL]. [2008-02-10]. <http://www.iipl.fudan.edu.cn/research/gc.html>.