面向服务的分布式空间信息支撑平台

吴信才，吴　亮

中国地质大学信息工程学院，教育部地理信息系统软件及应用工程中心，湖北武汉　430074

摘要：基于 Web 服务技术构建空间信息系统有助于提高其可扩展性和互操作能力，该技术也是这一领域中的热点. 提出了一种面向服务的模型框架，并基于该框架提出了基于 Web 服务的网格层次体系结构，并给出一种 Web 服务与空间信息技术相融合的机制，介绍了基于 Web 服务的空间信息网格支撑环境的总体结构和设计原理，为基于 Web 服务的空间信息系统设计与实现提供了一种新的方法.

关键词：分布式地理信息系统；服务计算；Web 服务.

中图分类号：TP311　　**文章编号**：1000－2383(2006)05－0585－05　　**收稿日期**：2006－05－30

Service-Oriented Distributed Spatial Information Supporting System

WU Xin-cai, WU Liang

Faculty of Information Engineering, China University of Geosciences, GIS Software Research and Application Engineering Center of the Ministry of Education, Wuhan　430074, China

Abstract: Web service-based approach for spatial information system can improve the extensibility and interoperability. In this paper, a service-oriented distributed spatial model is discussed; within the model framework, a web service-based architecture is presented. An approach of integrating web service and spatial information system is proposed. An architecture of a web service-based grid supporting system is introduced, which gives a new approach to web service based spatial information system architecture.

Key words: distributed GIS; service-oriented computing; Web Service.

0　引言

面向服务的计算(service-oriented computing, SOC) 近年来成为软件业界和学术界关注的一个热点话题，概括而言，其核心是通过资源的虚拟化和高层次的抽象，跨越系统底层的异构性，实现资源的交互与共享(Foster and Kesselman, 1997, 1998; Papazoglou and Georgakopoulos, 2003; Erl, 2004). 从广义上讲，地球空间信息学可以回答何时(when)、何地(where)、何种目标(what object)发生了何种变化(what change) 等问题(李德仁等, 2003)，在地球空间信息学中对于如何融合新的分布式计算技术来构建空间模型，采用面向服务，建立空间信息采集、集成、服务为目标的体系结构是值得进一步研究和实践的. 这方面的需求主要表现在：

(1)如何建立具有一致性的空间信息共享与服务体系，形成独立于软件、适用于空间信息互操作的表达模式；目的是提供一个可用于异构和分布式体系下的数据描述机制，使得应用服务机制管理数据成为可能.

(2)如何对异类、异构的网络资源提供无缝的空间信息共享和集成支持；这些资源不仅包括计算、存储、大型仪器等物理资源，也包括网络带宽、软件服务等逻辑资源. 因此，如何建立开放、可扩展的空间服务支撑平台体系结构，以满足对各类网络资源的共享与集成需求，成为空间服务研究的一个核心问题.

基金项目：国家“863”计划：面向网络海量空间信息的大型 GIS(No. 2003AA133010).
作者简介：吴信才(1953－)，男，教授，博士生导师，主要从事地理信息系统研究与应用方向研究. E-mail: wuxincai@mapgis.net

(3)前 2 个需求的侧重在理论研究上,在此基础上,实践中如何结合 Web 服务技术的已有工作基础,将应用建立在自主、可信的服务运行环境上,并在面向服务的体系结构层集中解决空间信息资源的联结、共享,使系统的软件体系结构不光是一个抽象的规则,而且应用者可通过程序级的实施来操作对象,也是研究的一个重点.

以上分析可以得出结论:研究和建立面向空间信息服务模型框架,实现多源、分布式空间数据库的一体化描述和组织,研究空间数据管理与应用服务的体系结构和机制,能有效的构建大型分布式空间信息系统的底层支撑平台,并为其业务应用提供支持.

1 空间信息服务模型

本模型中对于分布式数据管理体系是采取"纵向多级、横向网格"的解决方案,如图 1 所示,在级与级之间,节点与节点之间的连接是采用一种"松耦合"方式;"松耦合"方式是互联网的最佳耦合方式,它受网络环境影响最小. 分布式数据的存取操作采取面向"服务"方式进行,就是把"进行数据存取操作"变为"请求数据存取服务","数据存取服务"是所有"服务"的特例,充分体现"面向服务"的设计思想.

1.1 服务架构

模型中分布式数据管理是跨平台的,按照"面向服务"的思想,每个节点上的数据"管理者"必须提供"服务",在"谁管数据谁提供服务"的基础上,可解决网格节点之间、父节点与子节点之间、不同平台不同系统之间数据不通的问题,因为从用户角度来说,它不考虑数据只要服务. 应用端请求服务而不是直接操作数据,服务端提供服务而不是提供数据内部结构. 各个站点管理数据的软件、提供应用服务功能的软件都可以由不同的服务驱动提供.

具体而言,数据交换模块多级服务器间同步更新的数据传输采用的是基于 SOAP 协议的 Web Service 体系来实现的. 将 Web Service 技术引入网格研究领域,有助于解决网格研究所面临的应用集成、资源共享、系统互操性等问题. 目前,一些研究工作也从不同的角度尝试借助 Web 服务来部分地解决网格系统的互操作性(Mock *et al.*, 2002;

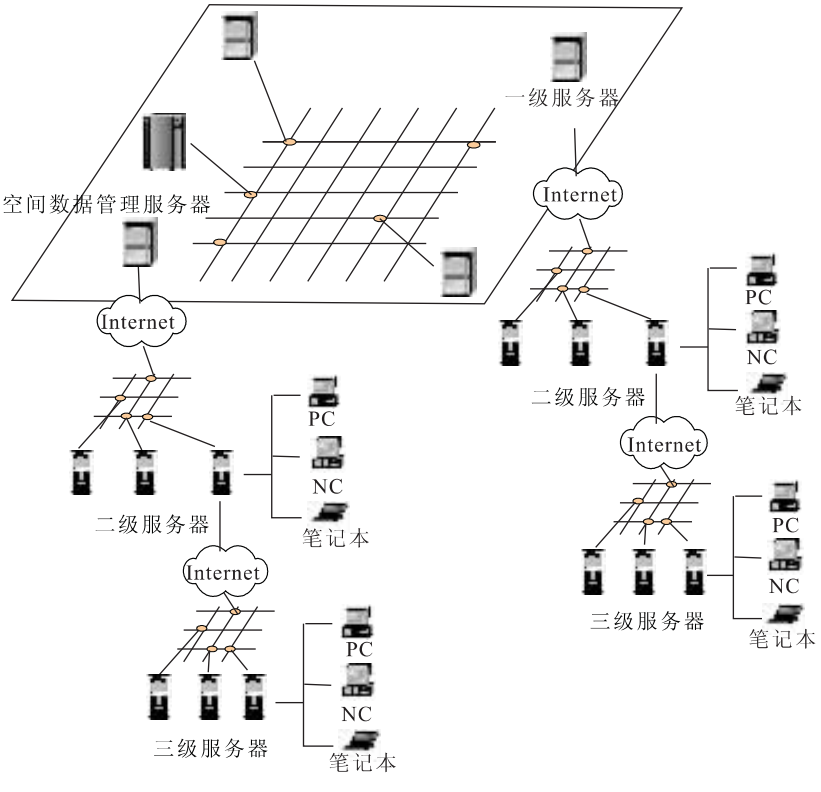

图 1 分布式数据管理原理

Fig. 1 Principle of distributed data management

Sarmenta *et al*.,2002). 在数据交换模块中,分发代理和接收代理分别为一个 Web Service 的站点,2 者间的数据传输采用 HTTP 协议,这样可以保证在正常情况下数据可以穿透防火墙或代理服务器而正常传输. 分发代理与接收代理不仅可以使更新的数据在广域网上有效的传输,还能使发布和接收地理数据库不必直接暴露于广域网环境下,增加了地理数据库的安全性.

基于分布式数据层提供数据并进行数据管理,其完成数据的定义、存储、检索、完整性约束以及有关的空间数据库管理工作,其处理服务组件或 Web Service 应用逻辑层的数据请求,并将处理结果交送服务组件或 Web Service 应用逻辑层.

Web 应用服务层和 Web 数据服务层直接和数据管理层打交道;当客户端访问 Web 服务器时,Web 服务器调用 Web 数据服务层或者 Web 应用服务层,并将最终处理的结果通过 Web 服务器传递给客户端.

在空间数据库引擎和地理数据库管理的支持下,数据可以转换为 GML 发布给用户,同样也可以将 GML 数据转换为内部格式的数据. 通过 GML 和 Web 服务,可以构建异构的空间数据管理. 通过其他 GIS 系统提供的符合 OGC 标准(OGC, 2003)、支持 GML 的 Web 服务,提供的 Web 服务也可以成为其他 GIS 平台的数据源,实现真正的跨平台信息共享.

1.2　元数据管理

空间元数据是地理空间数据的描述信息,主要用于海量、分布式信息数据的检索. 空间元数据管理模块主要由元数据采集器和元数据服务器组成,其设计基于 XML 技术构建,实现元数据的建库及管理、网络发布及分布式检索,具有分布式、跨平台、能兼容异构数据源等特点,主要表现在以下几个方面:

(1)元数据模式管理:元数据的标准由元数据模式来描述;系统提供元数据模式的注册、注销及根据模式校验元数据有效性等功能;

(2)元数据建库:元数据库的基本数据组织单位是符合某一元数据模式的元数据集合;系统提供元数据库的创建、删除、备份及导出等功能;

(3)元数据录入编辑:通过元数据采集器可对元数据进行编辑,根据元数据模式动态调整录入界面,确保数据与模式的一致性.

服务元数据库包含 WFS(web feature service)元数据、WMS(web map service)元数据、WCS(web coverage service)元数据、WCTS(web coordinate transformation service)元数据等. 它描述了数据和服务的各种信息(如分布位置、接口格式、数据范围、规模等). 用户通过查询服务的元数据来确定服务的位置和获取服务的接口,如图 2 所示.

1.3　基于消息的分布式计算和负载均衡

通常的 Web Service 调用是一种静态引用的调用过程,在跨越网络防火墙时,由于大量的数据传输和分布式数据处理工作效率不是很理想,本模型应用了消息传输机制,可以根据客户端的动态请求来动态决定调用 Web Service,实现了跨网络防火墙的快速服务,同时能进行高效的负载均衡. 消息处理和服务调用流程如图 3 所示.

Web 站点服务器提供 Web 功能服务和数据服务,每个不同的服务可能分布在不同的服务器上执行. 为了安全,这些服务器都隐藏在防火墙后面. 不同的用户调用的数据不同,需要的功能也不同,但浏览器端的用户并不需要知道他使用的数据服务和功能服务实际所在的位置信息,因为他并不直接调用这些服务器上的服务. 用户通过 Web 站点调用相应的服务,而 Web 站点动态地调用相应的服务器上的 Web Service 来完成用户的服务请求.

系统采用消息队列的方式建立 Web 站点与每个 Web Service 服务器的通信. 在每个 Web Service 服务器上都有一个消息服务,负责从消息队列中接受 Web 站点服务器的 Web Service 调用请求,并将 Web Service 处理的结果传输给 Web 站点端的消息服务.

Web 站点可以根据 Web Service 服务器状态在多个服务器间进行任务分配和调度,实现负载均衡.

1.4　SIG 与分布式空间信息系统集成

空间信息栅格(spatial information grid, SIG)是一种汇集和共享地理上分布的海量空间信息资源,对其进行一体化组织与协同处理,从而具有按需

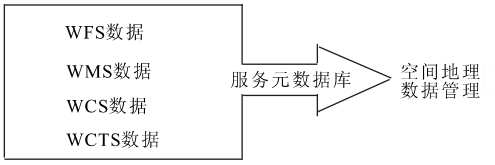

图 2　服务元数据与空间数据管理的关系

Fig. 2　Relation of metadata and spatial data

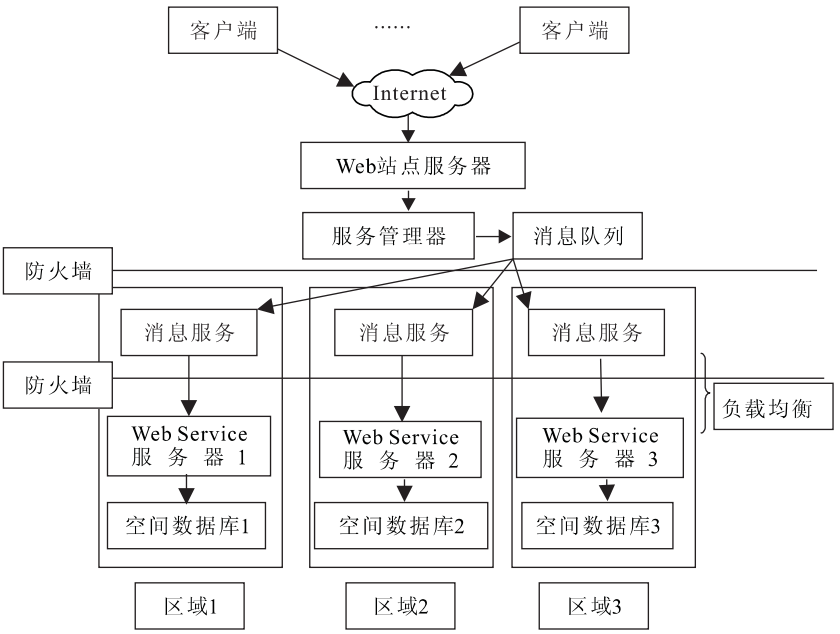

图 3 基于消息的分布式计算和负载均衡

Fig. 3 Distributed computing and load balance based on message

服务能力的空间信息基础设施.

SIG 应用先进的网格技术和综合集成技术,通过高速网络连接并集成地理上分布的、异构的各种高性能计算机系统、处理工具和软件系统、大型数据存储系统、数字化仪器设备和控制系统等各种资源为一体,实现跨地域的、分布的高性能联合、协同计算,为用户提供一体化高性能计算服务、信息处理服务和决策支持服务,发挥网络上资源的综合效能. SIG 建立空间信息标准和规范体系,提供统一的空间参考手段,能够集成和协同各种空间信息资源,针对多层次空间信息用户提供一站式无障碍的按需服务(图 4).

空间信息服务与集成是解决分布式空间信息处理功能共享、应用协同与综合集成的重要技术手段,是实现汇集和共享分布的空间信息资源、进行协同信息处理、提供按需服务的主要技术途径. 本模型已应用在采用"面向服务"的设计思想研发的第四代大型 GIS 平台——MAPGIS7. 0 上,该平台提供的海量空间信息管理、多层体系结构、完备的空间分析组件、全面的空间信息 Web 服务等特征能很好的支撑 SIG 的建立,并在 SIG 节点中作为支撑平台得到了很好的应用.

在 MAPGIS7. 0 的 Web 服务支撑下,可以实现空间数据的地域分布,空间信息应用的行业应用分布. 每个地方的每个行业应用服务都可以是网格上的一个节点. 用户通过网格提出他的空间信息应用要求,在网格调度和网格管理系统的调度下,网格中的众多共同节点完成运算任务,从而实现系统的集成.

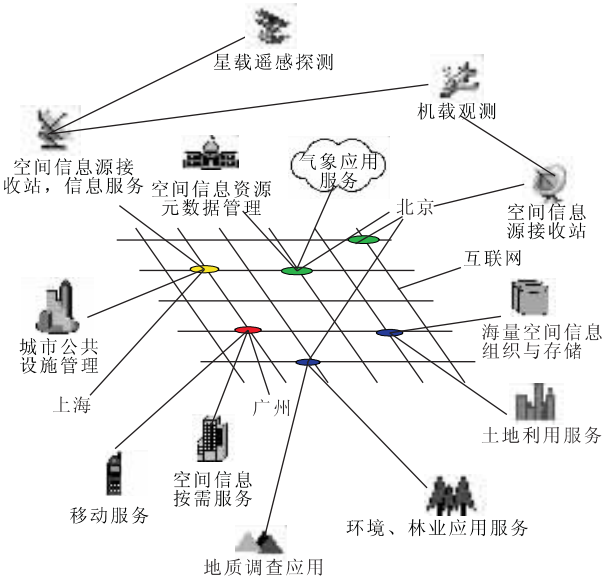

图 4 SIG 体系结构

Fig. 4 Architecture of SIG

2 结论

本模型已成功运用到大型 GIS(MAPGIS7. 0)和 SIG 的空间服务计算中:采用面向服务的地理实

体模型,基于对象、关系和规则定义各种行为(操作),能更适应长事务、分布式空间数据库的一体化描述和组织,能较好地解决面向空间实体及其关系的数据组织、高效海量空间数据的存储与索引、大尺度多维动态空间信息数据的管理及实体建模和分析;采用面向服务,基于 Web 服务技术构建 Internet 范围内空间信息资源共享与协同操作的模式,有助于提高网格系统的可扩展性和互操作能力,是实现多层体系结构,处理 TB 级空间数据,支持网络环境下空间数据的分布式计算、分布式空间信息分发与共享、空间信息服务的一条较为理想的技术路线.

本文的研究工作仍需深化和延续,网格系统作为一个大规模的分布式系统,对资源的自管理、自组织、自适应、自修复等提出了新的需求,因此,如何构造一个自主、可信的服务计算环境是我们今后研究和应用的重点.

References

Erl, T., 2004. Service-oriented architecture: A field guide to integrating XML and Web Service. Prentice Hall PTR, NJ, USA.

Foster, I., Kesselman, C., 1998. The globus project: A status report. In: Proc. of the IPPS/SPDP'98 heterogeneous computing workshop. IEEE Computer Society Press, Orlando. 4—18.

Foster, I., Kesselman, C., 1997. Globus: A metacomputing infrastructure toolkit international. *Journal of Supercomputer Applications*, 11(2): 115—129.

Li, D. R., Yi, H. R., Jiang, Z. J., 2003. Introduction and analysis of grid technology. *Geomatics and Information Science of Wuhan University*, 30(9): 757—761 (in Chinese with English abstract).

Mock, S., Mueller, K., Youn, C., et al., 2002. A batch script generator Web Service for computational portals. In: Proc. of the 11th IEEE international symp. on high performance distributed computing (HPDC-11). IEEE Computer Society Press, Edinburgh. 422.

OGC, 2003. OpenGIS geography markup language (GML) implementation specification 3. 0. http://www.opengis.org.

Papazoglou, M. P., Georgakopoulos, D., 2003. Service-oriented computing: Introduction. *Communications of the ACM*, 46(10): 24—28.

Sarmenta, L., Chua, S., Echevarria, P., et al., 2002. Bayanihan computing . NET: Grid computing with XML Web Services. In: Proc. of the 2nd IEEE/ACM international conf. on cluster computing and the grid (CCGRID). IEEE Computer Society Press, Berlin. 404—405.

附中文参考文献

李德仁,易华蓉,江志军,2003. 论网格技术及其与空间信息技术的集成. 武汉大学学报(信息科学),30(9):757—761.