

网格 GIS 数据传输机制与策略

孙庆辉^{1,2}, 骆剑承³, 赵军喜²

(1 中国科学院遥感应用研究所, 北京 100101; 2 解放军信息工程大学测绘学院, 郑州 450052;

3 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, 北京 100101)

摘要: WebGIS 实现了数据的分布性, 而基于中间件的网格 GIS 的目标是实现 GIS 数据和 GIS 功能服务的分布性。本文介绍了网格 GIS 的信息处理流程, 并根据网格 GIS 对数据传输功能的要求, 提出了网格 GIS 数据传输的实现策略。在网格 GIS 数据传输过程中, 对地理空间数据信息的描述, 是通过 GML 语言来完成的, 并利用 Web services、消息中间件等构造网格 GIS 数据传输服务, 同时对于网格 GIS 数据传输中的安全问题作了阐述, 利用代理机制来完成数据传输过程中的安全控制。

关键词: 网格计算; 中间件; GML; Web services; 网格 GIS; 数据传输; 数据传输安全

中图分类号: P208

1 引言

WebGIS 实现了地理空间数据资源的共享和分布。典型的 WebGIS 主要通过集中模式来实现, 即将地理空间数据集合存储到一台或几台数据服务器(计算机)上, 远程用户(客户计算机)通过网络连接来实现对这些数据资源的远程调用, 对数据资源的处理对于 B/S 结构的 WebGIS 主要集中到地理信息功能服务器上, 对于 C/S 结构来说, 服务器和客户机分担不同的计算操作。这种实现方式造成了服务器的负担加重, 随着计算机技术、网络技术的发展, 用户对地理信息的需求将更大, 而服务器的处理能力是有限的; 同时, 客户对于数据资源的频繁调用是通过功能服务器来向数据服务器发送数据请求, 由数据服务器来实现对数据的操作, 这样就大大加重了网络的利用, 耗费大量的网络资源^[1]。

传统的 WebGIS 只实现了数据的共享和分布, 随着 GIS 用户对 GIS 处理能力和信息交流速度的日益增加, WebGIS 就显得力不从心。当前的 GIS 软件例如 ArcInfo、MapInfo 等都是高度集成化的软件, 用户在应用的时候必须学习这个软件的功能结构, 并且要求用户具有相当水平 GIS 专业知识, 同时用户在购买这些软件时必须购买这个系统, 并不能购买某个功能组件。这些问题是 WebGIS 所不能

解决的, 基于中间件的网格 GIS 不仅能提供用户高效的计算处理功能, 同时中间件的网格 GIS 可以发布 GIS 服务, 用户可以针对自己的需要从网络上查找自己需要的功能服务, 提交处理命令获得所需要的处理结果。

2 基于中间件的网格 GIS 分析

中间件是介于应用系统和系统软件之间的一类软件, 它使用系统软件所提供的基础服务(功能), 衔接网络上应用系统的各个部分或不同的应用, 能够达到资源共享、功能共享的目的。中间件具有以下特点^[2]:

- (1) 标准的协议和接口;
- (2) 分布计算, 提供网络、硬件、操作系统透明性能; 并可满足大量应用的需要; 运行于多种硬件和操作系统平台。

中间件带给 GIS 应用系统的, 不只是开发的简便、开发周期的缩短, 同时也减少了 GIS 系统的维护、运行和管理的工作量, 还减少了计算机总体费用的投入, 中间件作为新层次的基础软件, 其重要作用是将不同时期在不同操作系统上开发的 GIS 应用软件集成起来, 这是操作系统、GIS 数据库管理系统本身做不了的。中间件根据功能分为^[3-6]: 数据

收稿日期: 2004-04-19.

资助项目: 国家“863”项目(2002AA135230); 国家自然科学基金项目(40101021).

作者简介: 孙庆辉(1974-), 男, 博士生, 主要从事地理信息系统及地学计算等研究。E-mail: irisqh@163.com; sunqh@reis.ac.cn

访问中间件(Data Access Middleware)、远程过程调用中间件(RPC Middleware)、面向消息中间件(MO Middleware)和事务处理中间件(TP Middleware)等。

2.1 网络 GIS 的特点

基于中间件的网格 GIS 和传统 GIS 软件尤其是 WebGIS 相比,具有明显的优势和特点,主要表现在以下几个方面:

- (1) 实现了空间数据与 GIS 服务共享。
- (2) 采用高性能的网格计算模型进行 GIS 的操作计算。而中间件技术简化了 GIS 的开发。
- (3) 在网格 GIS 中,所有的用户都是对等的,不存在一个固定的服务器,一个客户即可以是客户也可以是服务器;在网格 GIS 中,用户对空间数据的操作是通过数据库管理中间件来实现的,中间件将

用户与数据库隔离起来,用户无权直接访问数据库,有利于安全管理,可有效防止恶意攻击。

(4) 面向应用的策略不同。

在网格地理信息系统中,由于网格计算以及基于 GIS 服务的计算和操作涉及大量空间数据和非空间数据的移动、传输和复制,这就需要一种高效的数据传输机制的支持,这种传输机制要保证在网络的环境下可靠地传输数据。本文结合网格技术、软件中间件技术等,着重探讨了网格 GIS 的数据传输策略。

2.2 网络 GIS 的信息处理流程

在实现网格 GIS 的过程中,利用中间件技术进行功能组件的构造,整个网格 GIS 的功能结构和信息处理流程如图 1 所示:

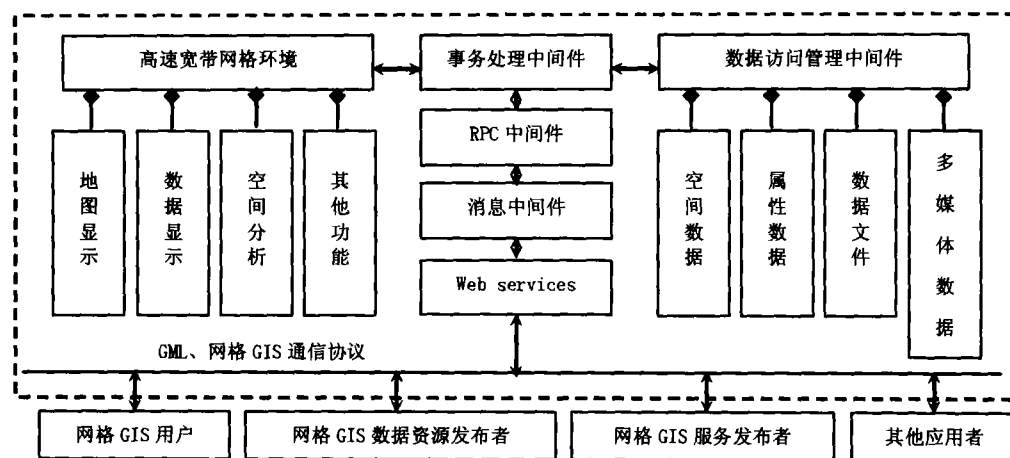


图 1 网络 GIS 信息流程结构图

Fig.1 Functional architecture of Grid-GIS

3 网络 GIS 的数据传输策略

3.1 网络 GIS 数据传输内容

数据传输就是在接收机与计算机之间进行数据信息的交换,数据传输主要包括 5 方面的内容:数据来源(data source)、数据发送者(transmitter)、数据通道(communication channel)、数据接收者(receiver)和数据目的(data sink)。在网格 GIS 中数据传输的内容如表 1 所示。

3.2 网络 GIS 数据传输的功能机制

为了保证网格 GIS 的高性能的计算能力,对于

网络 GIS 的数据传输应达到如下功能指标^[7,8]:

(1) 高速数据传输:要支持广泛接受的协议和广域网络上的数据传输,可以采用并行数据传输机制等。

(2) 分块数据传输:数据网格支持各种数据存储形式,一些数据集的数据本身是广泛分布的,因此,需要支持多个分数据块的并发数据传输,汇总后形成一个完整的数据集。

(3) 部分数据的传输:用户经常需要数据集中的一部分数据,而不是整个数据集,例如一个文件中的一段数据,因此支持这种数据的传输方式是必要的。

Web 服务链^[7],形成如图 2 的传输服务结构。

其中,UDDI(Universal Description, Discovery & Integration)用于 Web 服务注册和查找,WSDL(Web Service Description Language)用来描述 Web 服务的接口和功能,SOAP (Simple Object Access Protocol)用来构建 Web 服务和请求之间的数据传输。

在网格 GIS 环境中,利用 Web services 来实现异构的计算机软件和硬件之间的连接,同时进行异构系统用户之间的数据传输。

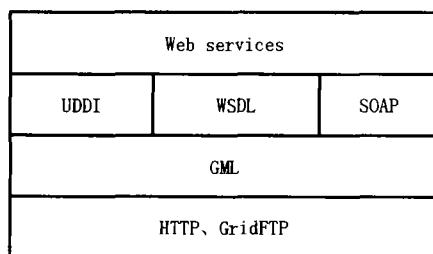


图 2 数据传输服务结构

Fig.2 Architecture of data transmission service

在 WebGIS 发展过程中,使用传统的 B/S 和 C/S 结构开发,造成了 GIS 系统结构和功能过于复杂、开发周期过长,而且还会带来系统效率严重降低,应用开发无法满足实际需要,数据传输不可靠,数据加密无法完成等问题。伴随分布式应用的迅猛发展,使用或开发必要的中间件产品显然是最好的解决方法。消息传输中间件为应用系统提供了可靠的消息通信手段,能够实现不同操作系统平台、数据库和硬件系统平台的数据通信。

对于消息传输中间件而言,最重要是及时提供可靠的网格 GIS 数据传输手段。为了能够完成数据的可靠传输,一般情况下,使用队列的方式进行消

息管理,也就是说,通常在进行数据传输时,将数据按照用户定义的大小,拆分成若干消息放入消息队列,中间件可以按照同步或异步的通信方式发送或者接收消息。在实际的操作过程中,为了保障消息可靠传输,经常使用诸如消息优先级、断点续传、可靠消息队列、内存队列等技术。

为满足异构平台的数据传输要求,消息传输均使用标准的 IP 包封装。此外,消息传输中间件通常都提供丰富的 API 函数接口,可以支持标准 C/C++、Microsoft Visual C++、Microsoft Visual Basic、Borland C、PowerBuilder、Java 等开发语言,特别是对标准的 C/C++和 Java 的支持,方便在不同应用平台的应用开发,同时实际上也增强了消息传输中间件的跨平台能力。

在网格 GIS 数据传输中,对于消息传输中间件来说,应该具有基本的安全措施,包括完善的认证机制、数据加密和对异常情况的处理等。如图 3 是消息传输中间件的功能结构。

3.4 网格 GIS 数据传输的安全

网格 GIS 数据传输的安全服务包括以下几个方面的内容:

(1) 访问控制服务,用来保护各种资源不被非法使用;

(2) 数据传输过程中的安全服务,用来提供认证、数据保密性与完整性和各种通信端的不可否认性服务。

(3) 支持跨虚拟组织的安全,由于网格 GIS 的用户和数据资源以及服务资源是属于多个组织,因此必须支持跨虚拟组织的安全。

可见,网格 GIS 的数据传输安全问题主要集中

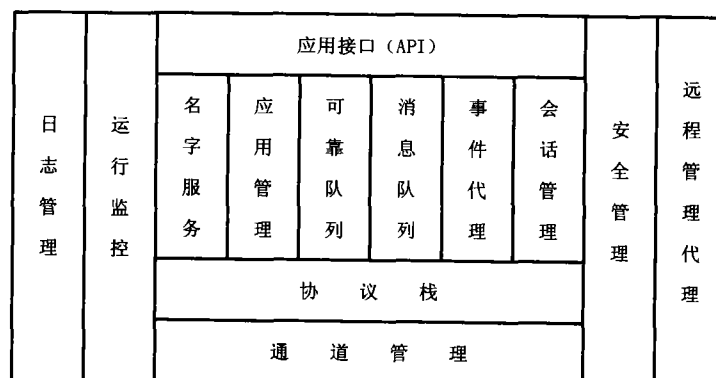


图 3 消息传输中间件的功能结构

Fig.3 Functional architecture of MO Middleware

在网络的传输层和应用层上。

在网络的传输层和应用层上,当前流行的安全技术主要有:

- 安全认证,每个用户和服务都需要通过认证证书来验明正身;

- 安全身份相互鉴别,数据传输和应用双方都具有认证证书,而且彼此都信任对方的认证中心,即相互鉴别问题。

- 通信加密,数据在传输过程中进行加密处理,数据的接收者知道数据的解密方法。

- 私钥保护,用户在使用数据资源或者服务资源时,必须提供私钥,来获取系统的信任。

在网格 GIS 项目实施过程中,我们采用基于 Web services 的安全服务利用代理机制来实现数据在传输过程中的安全性。代理对象是 Web 服务调用的一个基本概念。有了代理对象,就可以在各种平台上通过代理对象传输数据。其实现机制如图 4 所示。

在网格 GIS 数据传输中,所有的用户都通过代理对象和网格 GIS 环境以及其他用户进行数据的传输,避免把网格环境上的敏感的业务逻辑暴露给非法使用者或恶意的黑客,从而起到了一定的保护作用,同时代理对象的引入,简化了认证过程的复杂性,使数据传输操作简单。

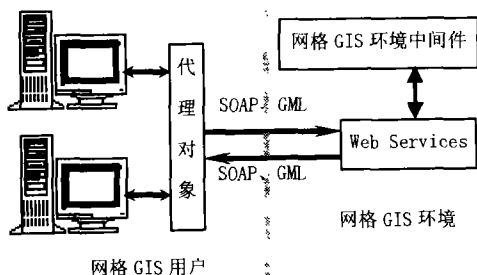


图 4 网络 GIS 的数据传输代理

Fig.4 Data transmission proxy in Grid GIS

4 结论

对于网格 GIS 环境,由于是在网络的架构上进行的,因此数据传输问题是网格 GIS 研究中的关键环节之一。为了保证网格 GIS 的数据传输的高效和可靠性能,传统的网络传输协议不能满足要求,必须对现有的网络传输协议进行扩展,在网格 GIS 数

据传输中,利用 GridFTP 协议来完成大数据量的数据传输,对于小数据量和日常的事务处理采用传统的 HTTP 协议。在数据传输构成中,利用 GML 语言对地理数据进行描述,利用 Web services 构造数据传输服务,同时在数据传输构成中,安全问题是影响网格 GIS 的关键问题之一。对于安全问题的解决,除了采用常规的安全技术外,还利用代理机制来实现网格 GIS 数据传输过程中的安全。在网格 GIS 的研究中,还有不少技术问题,有待深入研究。

参考文献

- [1] Apirak Panatkool, Sithichai Laoveerakul. Decentralized GIS Web Services on Grid, GIS - GRASS users conference, 2002, Trento, Italy, 11~13 September, 2002.
- [2] 徐志伟. 网络的称谓. <http://www.ccw.com.cn>, 2001.
- [3] 周园春, 李 森 等. 中间件技术综述. 计算机工程与应用, 2002, 15: 80~82.
- [4] <http://www.webopedia.com/TERM/M/middleware.html>
- [5] <http://www.eecs.wsu.edu/~bakken/middleware-article-bakken.pdf>
- [6] <http://www.middleware.de/>
- [7] 肖 依. 编织“数据网格”. 计算机世界报, 2002, 40.
- [8] 都志辉, 陈 渝, 刘 鹏. 网格计算. 北京: 清华大学出版社, 2002.
- [9] 李 伟, 谈恩华. Globus 项目进展和技术水平分析. 北京: 中国科学院计算所资料.
- [10] 齐 锐, 张大力, 黄 磊, 李 琦. 网络化地理信息系统中数据传输技术的探讨. 计算机研究与发展, 1999.
- [11] 刘荣高, 庄大方, 刘纪远. WEB 环境下实现空间数据表达的框架研究, 测绘学报, 2001, 3, 276~280.
- [12] <http://www.zdnet.com.cn>.
- [13] Open GIS Consortium, Inc. OpenGIS Geography Markup Language (GML) Implementation Specification, version 3.0.0.2003.
- [14] <http://www.opengis.org>.
- [15] 王春樵. 面向服务架构—分布式网络应用的方向, 广东通信技术, 2002, 1, 40~44.
- [16] 谢 澍, 张洪伟, 赵 立. Web services 基本特征及其在构造供应链管理系统中的应用. 成都信息工程学院学报, 2002, 2: 76~82.
- [17] 闫新庆, 李文锋, 陈定方. Web 服务的体系结构和应用. 武汉理工大学学报·信息与管理工程版, 2002, 24(5): 28~31.
- [18] 郑 江. 网络地理信息系统中数据传输安全的探讨. 地球信息科学, 2002, 4(2): 23~27.
- [19] 杨 强, 卢建军, 唐善成. 一种基于 .net 的安全服务决策平台的具体实现. 计算机应用, 2002, 22(9): 83~84.

Strategy of Data Transmission in Grid-GIS

SUN Qinghui^{1,2}, LUO Jiancheng³, ZHAO Junxi²

(1 Institute of Remote Sensing Applications, CAS, Beijing 100101, China;

2 Institute of Surveying and Mapping of Information Engineering University, Zhengzhou, Henan 450052, China;

3 State Key Lab of Resources and Environment Information System, IGSNRR, CAS, Beijing 100101, China)

Abstract: Data distribution of GIS has been made realism through WebGIS. With the increasing need of data sharing and GIS, WebGIS can no longer meet this requirement. The application of Grid Computing and Middleware technology in GIS make a promise for the development of Middleware Grid-GIS. Grid-GIS not only makes spatial data distributed, but also makes the functional services of GIS distributed. Grid-GIS has the following characteristics:

(1) Sharing of spatial data and GIS function; (2) highly performance of computation; (3) easiness for GIS software development; (4) equity in architecture; (5) high security; and (6) easy for GIS application.

In this paper the procession of data transmission was recommended, and the strategy of data transmission in Grid-GIS was put forward in order to meet the functional requirement for data transmission.

In data transmission of Grid-GIS, GML was used as the description of geographic data. GML is the extending use of Extreme Markup Language (XML) in the area of Geography. Using GML user can store and transmit the attribute and location information of geographical feature. GML can describe geographical feature into text of XML format, and supply the spatial information transmission between GIS components with cogent technology.

The data transmission service of Grid-GIS was constructed by taking the advantage of Web services and MO Middleware technology. Web services are a new breed of web applications. They are self-contained, self-described and the modular application can be published, located and invoked across the web. Web services perform functions that can be anything from simple requests to complicated business process. Once a web service is deployed, other applications (and other web services) can discover a deployed service. Security of data transmission in Grid-GIS was also expounded. It was accomplished by using the proxy method.

Key words: grid computing; middleware; GML; Web services; Grid-GIS; data transmission; security of data transmission