

# 基于 XML 的多级异构空间数据库的同步架构

解吉波<sup>1</sup> 吴华意<sup>1</sup> 龚健雅<sup>1</sup>

(1 武汉大学测绘遥感信息工程国家重点实验室, 武汉市珞喻路 129 号, 430079)

**摘要:**分析了多级异构空间数据库同步的关键问题和难点, 提出了一种基于 XML 的多级异构空间数据库的同步架构, 该架构在“全国国土资源数据交换中心试点建设项目”中得到成功应用。

**关键词:**多级异构空间数据库; 同步; 网络要素服务  
**中图分类号:** P208

随着测图技术的发展, 空间数据库可以在更短的时间内获取更新数据, 美国 USGS 提出要在 2010 年实现“七天内更新地图”<sup>[1]</sup>。我国的土地利用数据目前可以实现每年更新一次, 在一级空间数据库发生变更以后, 以怎样的机制保持各级异构空间数据库的同步和一致性成为一个亟待解决的问题<sup>[2,3]</sup>。

目前, 基于 Web 的 XML 技术正在成为构建跨平台异构应用系统的主流技术, XML 技术以其松散、灵活、易于跨软硬件平台等优点, 已成为 Web 服务的技术基础, 也已成为地理信息系统领域构建跨部门、跨行业、跨地区异构性地理信息系统以及开展地理信息 Web 服务的主要技术基础<sup>[4]</sup>。本文提出了一种基于 XML 的多级异构空间数据库的同步构架, 来解决多级异构空间数据库和异构 GIS 系统之间的数据交换与同步问题。

## 1 多级异构空间数据库的同步更新问题

所谓多级异构空间数据库是指某一专业领域不同级别部门建立的具有多分支、多层次的多级空间数据库, 这些数据库具有基本一致的内容模型, 空间数据一般为二维矢量形式, 地理位置分散在全国各地, 各级空间数据库所采用的数据库软件和 GIS 软件往往不一致, 具有异构的特点。本文研究的多级异构空间数据库, 上下级存在依存关系(上级部门的空间数据库包含其所辖的所有

下属部门空间数据库的核心内容), 数据库按一定的次序进行更新(更新首先发生在下级部门, 然后逐级将更新传递给上级部门, 而不能反方向)。国土部门的空间数据库具有典型的多级异构特点, 其构成如图 1 所示, 通常按行政级别分为国家、省(自治区、直辖市、特别行政区)、市、县(区)四级, 其组织形式为树状结构。

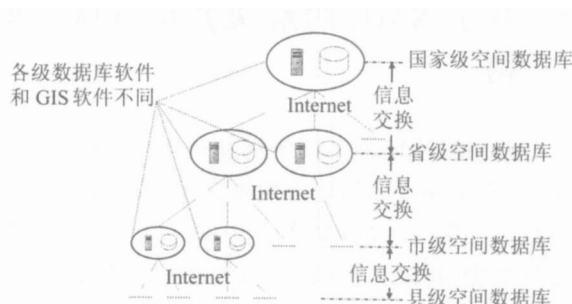


图 1 某一专题的多级异构空间数据库示意图

Fig. 1 Illustration of Multilevel and Heterogeneous Spatial Databases

数据库同步是指具有多个副本的数据库, 当其中一个副本发生变更后, 按某种机制将变更传递到其他副本, 使得多副本之间保持一致。因为多级异构空间数据库存在上下级之间的依存关系, 所以在发生变更后也面临同步问题。

我国政务网络的建设为空间信息的快速传输、交换和共享提供了网络基础设施, 然而目前的工作方式和现有的系统架构不能满足多级数据库网络同步的要求, 问题的妥善解决需要一种基于开放标准、松散耦合、可扩展性好的解决方案, 概

括起来有以下关键问题和难点需要解决。

1) 数据更新方式。实现异地空间数据库的同步,一种方案是将变更后的各下级空间数据全部传输到上级,但由于空间数据是海量的,在目前的网络条件下是不可行的,采用增量更新是一种可行的方案。

2) 空间数据网络交换格式。由于各部门异构 GIS 系统的数据模型和数据内容有所不同,因而需要规定统一的空间数据交换格式,且这种格式必须具有适合网络传输的特点。

3) 异构 GIS 系统之间的远程数据访问。下级节点为上级节点提供一种跨软硬件平台的网络数据访问接口,以实现异构空间数据库和 GIS 系统的集成与互操作。

4) 更新流程管理。通过网络实现多级空间数据库的数据交换,异常情况是难免的,为此必须考虑更新流程及控制机制。

5) 标准及规范。对于多级空间数据库业务系统,由不同的软件厂商开发实现,如果没有统一的标准,不同部门的系统建设将各行其道,不能实现互通。

## 2 基于 XML 的解决方案及体系架构

在网络环境下实现多级异构空间数据库的同步,需要在各级部门建立数据交换中心,在两级部门间建立交换系统,如图 1 中的结构要建立三级交换系统(县级-市级,市级-省级,省级-国家级),通过建设交换系统实现上下两级空间数据库之间的数据交换和更新。XML 具有结构化、自描述、可扩展、适合网络传输和可验证性等优点,可用其作为两级节点(数据中心)之间的信息交换格式。为了保证交换系统的开放性与可扩展性,交换体系不对具体开发软件和平台进行规定,而是对两级节点之间的增量更新机制、网络空间数据交换格式、数据远程访问以及消息通信机制进行了定义,统一了交换接口和开放的标准、规范。

交换体系中两级节点间信息交换如图 2 所示,通过消息通道和数据通道进行信息交换,消息通道是双向的,数据通道是单向的。消息通道用于传递更新控制消息,控制和管理更新行为。数据通道为更新数据获取的接口,空间数据交换格式采用 GML<sup>[5]</sup> 格式,通信采用 HTTP/POST 协议。

### 2.1 增量更新机制

实现多级数据库增量更新必须解决更新操作

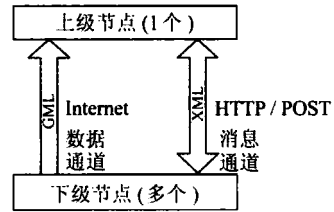


图 2 两级节点的信息交换

Fig. 2 Information Exchange Between Two Nodes

定义和地物唯一标识问题。本文定义了 3 个更新操作: Create(新增)、Destroy(删除)和 Update(变更)。这 3 个操作是原子级的,空间数据的变更往往比较复杂,一个地物的变更通常要牵涉周围的地物,对于复杂的更新过程最终可以分解为这 3 个最基本的操作。同一个地物在各级空间数据库中有唯一标识,这是保证各级空间数据库进行增量更新的重要条件。如国土资源数据由各个县市采集,县市内部可以保持唯一性,但在导入到省级库时,不同的县市的标识码可能会重复,解决办法是在地物的标识码前面加县级的 6 位行政代码(全国唯一)。

### 2.2 网络空间数据交换格式

在同步架构中,采用 GML 作为统一的空间数据交换格式,GML 适合作为空间数据网络交换的格式,各级交换中心只需开发相应的 GML 生成和解析模块,即可实现其异构空间数据的共享与互操作。

### 2.3 数据远程访问机制

数据交换在上下两级行政单位的空间数据库间进行。下级节点为上级节点提供数据的网络访问接口,为保证数据访问的灵活性,应采用静态和动态两种方式。动态方式采用 OGC 网络要素服务(WFS)<sup>[6]</sup>的方式,下级节点提供网络要素服务,上级节点可以动态地从下级空间数据库中查询和提取满足条件的地物要素。WFS 基于 Web Services 技术,具有松散耦合、跨软硬件平台等优点,能够实现网络异构 GIS 数据的共享和集成。静态方式是将更新数据文件静态存放于指定的网络路径,上级通过数据的 URL 地址下载数据。

### 2.4 基于消息驱动的更新机制

交换节点包含数据管理、消息通信和数据服务 3 个基本功能模块。数据管理模块的主要功能是从下级节点提取更新数据,解析并更新到本地数据库;消息通信模块的主要功能是处理两级节点之间的消息通信;数据服务模块是下级节点为上级节点提供数据访问的网络接口。

交换控制消息用于传递交换控制指令及更新数据内容描述信息,以对两级节点系统的数据交换进行协同控制,并为空间数据库的更新提供必要的更新数据相关信息,消息格式用 XML 定义。更新控制消息包含 4 个子节点: ExchangeSequenceID 是本次更新的标识; NodeID 是更新消息发送的源节点; TargetNodeID 是更新消息发送的目标节点; 最后一个节点有 3 个互斥的可选项,包含 DataExchangeRequest 的消息是数据交换请求消息,包含 ExchangeStatusReport 的消息为数据交换状态报告消息,包含 Acknowledge 的消息为响应消息,用来响应接受到的任何消息。

基于消息驱动的交流机制实现信息交换自动化,两级节点间的交换机制如图 3 所示。

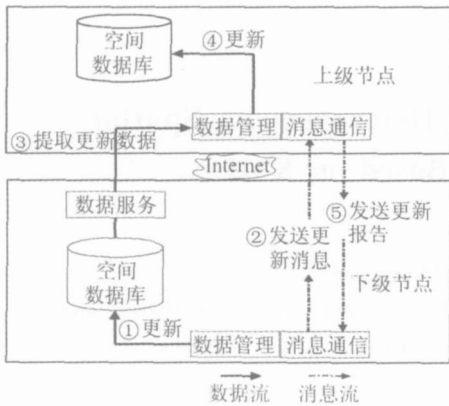


图 3 数据交换机制图

Fig. 3 Mechanism of Data Exchange

### 2.5 标准和规范

标准和规范是多级异构空间数据库同步架构的重要部分,主要包括《建库规程》、《数据更新管理规范》、《网络空间数据交换规范》和《网络要素服务规范》等。《建库规程》针对建库的数据内容,规范和指导建库的过程;《数据更新管理规范》定义了更新消息,并对更新行为及一些细节问题进行规范;《网络空间数据交换规范》规定了空间数据在网络上的传输格式,在 GML<sup>[3]</sup> 格式的基础上,定义适合行业和应用的模式;《网络要素服务规范》规定了空间数据的网络共享交换接口标准,其定义参照 OGC 的网络要素服务规范。

### 3 应用实例

全国国土资源数据交换中心建设试点项目采用了本文提出的同步架构,试点在国土资源部与

浙江省国土资源厅和广东省国土资源厅之间建立了数据交换系统,以此为依托,实现了国土资源信息的交换,省级数据中心汇总的国土资源数据通过网络自动更新到国土资源部数据中心。国土资源数据交换中心建设的示意图如图 4 所示,目前建立了国土资源部与浙江省和广东省三个数据交换中心,分别开发了数据交换软件,并制定了相关的建设规范,成功进行了系统的运行和测试。广东省国土资源数据交换中心软件系统的构成如图 5 所示,由“省级建库与管理”、“数据交换”、“网络要素服务(WFS)”、“在线共享”等子系统构成。系统功能包括土地利用数据的建库、管理、分析、更新和历史数据与管理等。网络要素服务是国家从省级国土资源数据库中提取数据的网络接口,可实现两级系统消息通信和数据交换业务,并可实现空间数据的发布与共享。

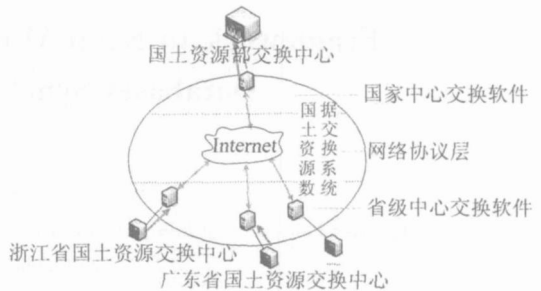


图 4 国土资源数据交换中心建设示意图

Fig. 4 Illustration of Land and Resources Exchange System

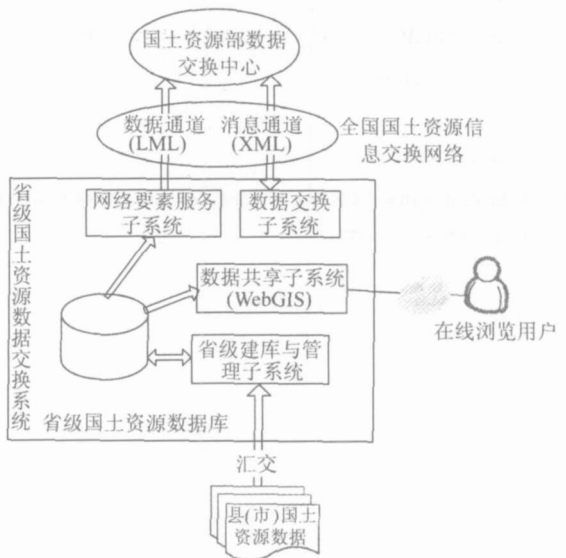


图 5 省级国土资源数据交换中心系统构成

Fig. 5 Architecture of Land and Resources Exchange Center

## 4 结 语

基于XML的多级异构空间数据库同步架构,采用XML和GIS互操作等相关技术,能够解决多级异构空间数据库同步的问题,确保多级空间数据库的整体一致性。由于该架构通过网络进行空间信息的网络传输与远程更新,数据安全性还有待进一步研究。

### 参 考 文 献

- [1] 陈军. 多维动态地理空间框架数据的构建[J]. 地球信息科学, 2002, 3(1): 8-13
- [2] 郭伦, 张毅. 分布式多空间数据库系统的集成技术[J]. 地理学与国土研究, 2002, 18(1): 6-10

- [3] 杜小平, 周顺平, 万波. 基于数据复制的多级空间数据库服务器体系结构设计及实现[J]. 计算机工程与应用, 2004, 21: 172-173
- [4] 龚健雅, 贾文珏, 陈玉敏, 等. 从平台GIS到跨平台互操作GIS的发展[J]. 武汉大学学报·信息科学版, 2004, 29(11): 985-989
- [5] Open GIS Consortium Inc.. OpenGIS Geography Markup Language(GML) Implementation Specification[OL]. <http://www.opengis.org/>, 2005
- [6] Open GIS Consortium Inc.. Web Feature Service Implementation Specification [OL]. <http://www.opengis.org/>, 2005

第一作者简介: 解吉波, 博士生。现主要从事地理信息系统、空间数据共享与互操作、分布式计算等研究。

E-mail: jiboxie@163.com

## Framework to Keep Multilevel and Heterogeneous Spatial Databases Synchronization Based on XML

XIE Jibo<sup>1</sup> WU Huayi<sup>1</sup> GONG Jianya<sup>1</sup>

(1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

**Abstract:** This paper analyses the key problems and difficulties to keep the remote and heterogeneous databases synchronization, promotes a framework to solve the problems using XML. The framework is used in the National Land and Resources Data Exchange Center Experimental Project successfully and is proved to be feasible.

**Key words:** multilevel and heterogeneous spatial Databases; synchronization; WFS(Web feature service)

**About the first author:** XIE Jibo, Ph. D candidate, majors in GIS, spatial data sharing and interoperability and distributed computing.  
E-mail: jiboxie@163.com