

一种高效的三维 GIS 数据库引擎设计与实现

朱 庆¹ 李晓明¹ 张叶廷¹ 刘 刚²

(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)

(2 中国地质大学(武汉)计算机学院,武汉市鲁磨路 388 号,430074)

摘 要:针对大规模三维城市建模与数据库协同应用,设计实现了一种高效的三维 GIS 数据库引擎,支持基于 Oracle 11g 的多模式数据库管理;提出了顾及语义的三维空间数据库模型,为地上下室内外三维空间数据的一体化组织管理奠定了基础。介绍了该引擎涉及的多层次三维空间索引、多级缓存、多线程调度以及异步通信传输等关键技术,并用武汉市三维城市模型数据进行了试验分析,验证了该引擎的有效性和可靠性。

关键词:三维 GIS;数据库引擎;一体化数据库管理;试验可视化应用

中图法分类号:P208

随着大范围三维城市模型的普及应用,地上下室内外三维空间数据的一体化管理、多用户协同操作、网络共享、数据的安全等需求日益迫切。而现在大多采用的文件系统管理方式要么将大范围三维空间数据分成多个规划单元分别管理,很难满足一体化管理和无缝漫游的要求;要么采用松散的大文件管理方式,难以保证数据的安全性和高效性,因而采用大型关系数据库管理系统一体化管理三维空间数据已逐渐成为必然的选择^[1,2]。

地上下室内外三维空间数据的一体化管理亟需真三维的空间数据模型,而已有的三维空间数据模型大都局限在特定的专业领域,如地质模型、矿山模型、地表景观模型等,其通用性和扩展性不强^[3],难以满足地上下室内外三维空间信息的语义表达、动态更新与一致性维护以及综合分析的需要。OGC 标准 CityGML 1.0 针对城市地表目标的统一表示,提出了一种抽象层次的语义信息模型^[4],但仍缺乏三维地质和地下设施的语义表达。虽然主流的商业关系数据库管理系统都已扩展了空间数据的管理,如 Oracle Spatial、DB2 Spatial Extender、SQL Server Express 等^[5],但只有 Oracle Spatial 11g 直接支持简单三维数据类型的表达^[6],如基于 CityGML 模式,利用 Oracle

Spatial 管理了柏林三维城市模型的几何数据^[7],但 Oracle Spatial 仅支持简单的 3D 几何,无法直接满足具有三维曲线或曲面等复杂三维空间实体的表达^[6]。现有的商业 GIS 平台的空间数据库引擎,如 ArcSDE、SuperMap SDX⁺ 等^[8],能有效支持二维空间数据在 Oracle SQL Server 等关系数据库管理系统中的存储与管理,并开始逐渐扩展了三维空间数据的管理,但是大都难以满足地上与地下、室内与室外三维空间数据的一体化管理,尤其是对三维空间数据丰富的语义信息都未完整地表达。因此,针对大范围多细节层次地上与地下三维空间数据高效的一体化组织管理的重大需求,本文介绍了一种高效的三维 GIS 数据库引擎的设计与实现。

1 顾及语义的三维空间数据库模型

1.1 概念模型

针对地上、地下和室内外三维空间实体的基本特点,首先从概念上建立其统一表达的通用三维空间数据模型。如图 1 所示,地上下三维空间

实体集成表示的统一数据模型通过几何层、多尺度表达层和专题语义层三个层次进行完整的表达。在几何层,通过点、线、面、体等基本元素实现地上三维空间实体几何表达的统一;在多尺度表达层,通过 LOD 对象实现地上三维空间对象的多尺度表达;在专题语义层,对建筑物模型、道路网络模型、地质体模型等专题进行面向专题语义的扩展。其中,在 DEM 表面层次上确保合理的空间划分与区域识别;在地上下立体层次,解决地理实体在二维抽象表示中产生的地上下交叠问题,满足三维立体空间层面的实体精确表达与分析需求;基于语义层次关系进一步详细描述相

关实体的三维内部结构或组成关系。

为了实现地上下三维空间实体几何表达的统一,提出了三维空间实体统一表达的几何对象模型,如图 2 所示。该模型首先定义了三维几何统一表达的基类 CGeometry,并直接派生三维几何基类 C3DGeometry,派生了点(C3DPoint)、线(C3DCurve)、面(C3DSurface)、体(C3DVolume) 4 种基本空间元素,实现了整个三维空间的抽象表达,每个基本几何元素都可以泛化为具体的几何类型表达,各类几何元素又可以组合成复杂对象,并引入组对象(C3DGroup)来表达混合类型的复杂几何对象。通过几何基类 CGeometry 与

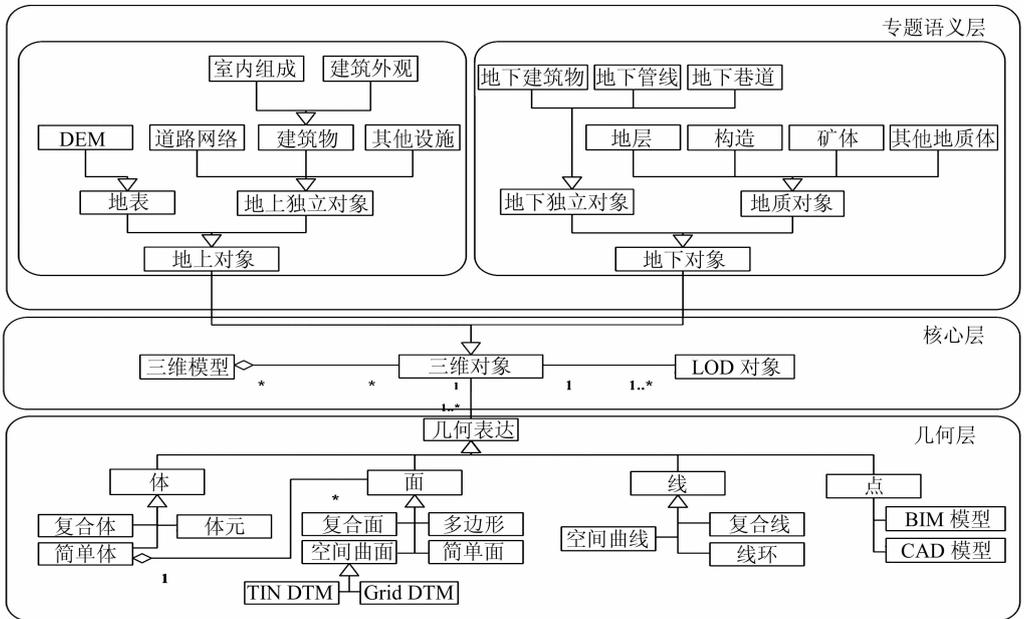


图 1 一体化的三维空间数据模型
Fig. 1 An Integrative 3D Spatial Data Model

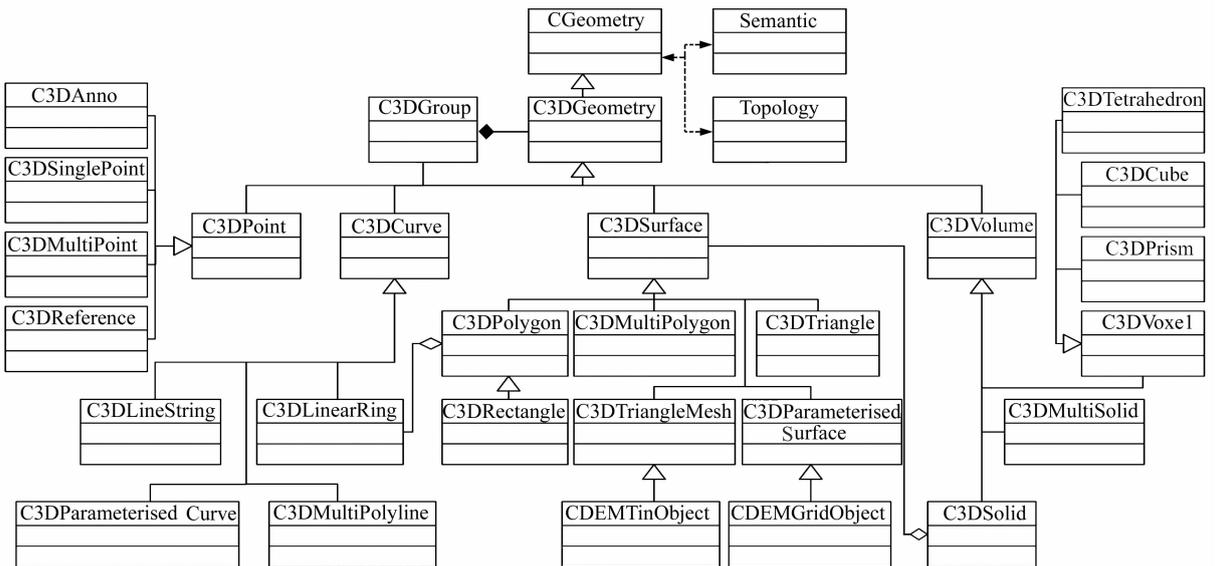


图 2 统一表示的三维几何模型
Fig. 2 Unified Representation of 3D Geometrical Model

语义类(Semantic)和拓扑类(Topology)之间建立的映射关系,实现了三维空间实体语义拓扑关系的统一表达。

1.2 数据库模型

基于上述概念模型,建立了地上下三维空间数据一体化管理的数据库模型,如图 3 所示,三维 GIS 数据库由一个或多个工程组成,每个工程由多个数据集组成,包括地形数据集、三维模型数据集、三维语义拓扑数据集等,每个三维模型数据集和地形数据集的数据都必须具有相同的空间参考。地形数据集由 DEM 数据集和 DOM 数据集组合而成,DEM 数据集和 DOM 数据集采用金字塔分层+均匀分块的技术管理多分辨的数字高程模型数据和数字正射影像数据,分别由多个 DEM 金字塔层数据集和多个 DOM 金字塔层数据集组成,并建立了明确的对应关系。三维模型数据集管理地上下室内外的三维空间实体数据,由具有相同空间参考的三维要素类以及对象类和关联类组成;三维要素类是具有相同属性结构的同类三维空间实体的集合;而对象类是没有空间特征、具有相同行为和属性的对象集合,三维要素类与对象类通过关联类建立关系。三维空间实体由具有

多细节层次(LOD)的三维几何模型及其属性数据组成,并通过 LOD 描述结构统一管理每个细节层次几何模型的包围盒、重要性因子、可见范围等信息。三维几何模型与共享模型数据、材质数据以及纹理数据之间通过 ID 号建立关联关系。三维语义拓扑集管理三维要素类之间以及三维空间实体之间的语义拓扑关系,包括位于、部分组成、连通等,如通过“位于”和“部分组成”语义关系表达建筑的内部逻辑构成以及建筑内外空间和实体的“所属关系”,通过建筑内部空间之间明确的连通关系建立建筑内部的三维网络,能够支持室内外一体化寻径等复杂空间决策应用。

地上下三维空间数据的一体化数据库模型具有如下特点:① 支持大规模的地上下室内外各种类型三维空间数据在关系数据库中的一体化管理;② 采用面向对象技术,通过具有行为和特征的三维要素类,使三维空间实体的表达更面向实际的应用;③ 支持复杂的语义关系管理,可以完整地表达三维要素类之间以及三维空间实体之间的语义拓扑关系;④ 基于关系数据库的优势,采用 workflow 机制和长事务机制支持多用户并发处理。

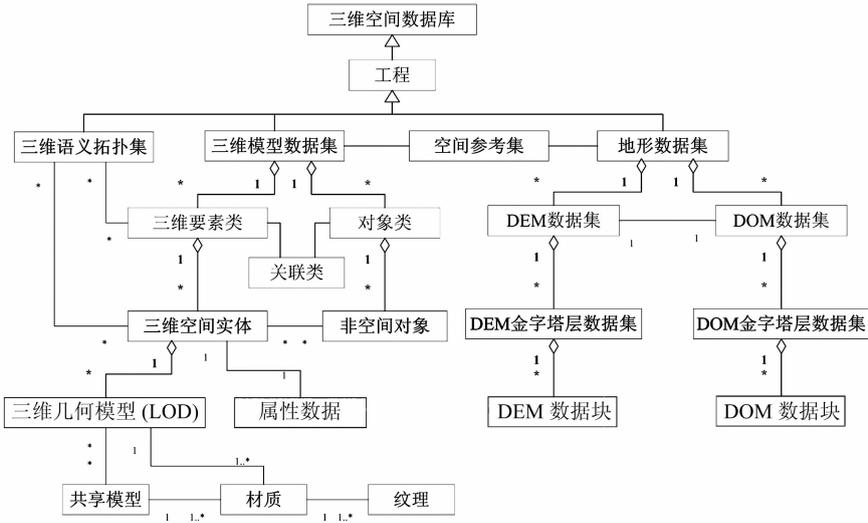


图 3 顾及语义的三维空间数据库模型

Fig. 3 A 3D Spatial Database Model Concerned with Semantics

2 多层分布式架构的三维 GIS 数据库引擎设计

为了在大型关系数据库中高效地管理和调度大规模的三维空间数据,设计了一个高效的三维 GIS 数据库引擎,该引擎是三维 GIS 平台与关系数据库数据交互的连接通道,能高效地存储、索引

和访问在关系数据库中的三维空间数据,并解决关系数据库管理系统的多样性和复杂性。三维 GIS 数据库引擎采用多层分布式的客户/服务器体系架构,如图 4 所示。在客户端层,数据库统一访问接口是三维空间数据建库和访问等操作的唯一通道,通过客户端调度管理模块实现数据请求在客户端的快速获取数据,并进行索引检索、缓存查找等操作,通过客户端通信模块与服务器建立

网络连接;在应用服务层,采用异步通信机制的服务器通信模块并发处理多客户端发送的数据请求,通过服务器调度管理模块将这些请求转换成多线程可以执行的调度任务,通过多线程调度模块在服务器缓存与三维空间数据库中快速获取数据;在数据库存储层,通过基于地上下一体化的三维空间数据库模型实现了大规模地上下三维空间数据在大型关系数据库管理系统中的高效组织与存储。

采用多层分布式架构的三维 GIS 数据库引擎在安全性、可维护性、可伸缩性、开放性上具有明显的优势,可以支持流行的关系数据库管理系统,并针对不同数据库管理系统的特性进行有针对性的性能优化。首先,本文基于 Oracle 11g 数据库管理系统搭建了三维 GIS 数据库引擎的原型,利用 Oracle 11g 的新特性,以充分发挥数据库管理的性能,包括新的 LOB 架构——Securefiles BLOB、增强的分区功能以及结果集缓存等^[9],并支持基于 Oracle RAC(真正应用集群)的高可用的并行数据库管理环境。为了满足不同的应用需求,三维 GIS 数据库引擎支持基于 Oracle 11g 的多模式管理,直接访问模式可以通过三维 GIS 数据库引擎客户端直接与数据库建立链接,不需要通过应用服务层,具有操作方便的特点;应用服务器访问模式需要通过三维 GIS 数据库引擎的应用服务器与数据库建立链接,对多用户并发的快速处理更有优势。

需求,三维 GIS 数据库引擎面临的关键技术问题有高效的数据库结构、三维空间索引和多用户并发访问等。

3 三维 GIS 数据库引擎的关键技术

3.1 多层次三维空间索引方法

高效的三维空间索引是实现大规模三维空间数据快速检索的关键技术^[10]。由于建筑、地质体、管线等各种专题模型在几何形状与纹理特点、空间分布特点等多方面的明显差异,往往难以采用一种通用的三维空间索引来满足所有类型三维模型数据的高效组织,本文基于“高效统一、合理区分”的原则,在统一的三维空间索引机制下,兼顾地上下各类三维空间实体,发展了高效的多层次、多类型混合的三维空间索引方法,采用由粗到精的多层索引和多级过滤,实现三维空间数据的快速、准确查询。多层次、多类型混合的三维空间索引方法首先通过格网索引作为一级索引实现快速定位;然后采用改进的 R 树索引、X-List 索引等作为二级索引实现地上下三维目标的精确查找,并有效顾及地上下三维空间目标的 LOD 特性。针对离散分布、目标形状各异、具有多细节层次的地上下建筑模型采用了结构平衡的三维 R 树索引扩展方法^[10],中间节点引入细节层次模型,实现目标查询和细节层次查询的双重任务,通过先自下而上、后自上而下全局搜索的节点选择算法和基于 *k*-medoids 聚类算法的节点分裂算法,保证节点尺寸均匀、形状规则以及重叠减少。针对地质模型分层连续分布、形态随机变化、拓扑关系复杂、属性丰富的特点,采用含拓扑关系的 X-List 三维空间索引方法和 CSR-Tree 数据预处理技术来提高地质数据的效率^[11]。

3.2 面向多层架构的自适应多级缓存管理方法

针对多用户并发引起的数据库服务器的磁盘 I/O 瓶颈和服务性能瓶颈,高效的缓存机制成为有效的解决方法之一^[12],其可以有效地减少网络的通信量,减轻服务器的负载,显著提高响应速度。本文设计了自适应的三维空间数据多级缓存结构,通过多级缓存的有效配合,提高三维空间数据的调度效率。三维 GIS 数据库引擎的多级缓存包括客户端内存缓存、客户端文件缓存、应用服务器内存缓存等三级缓存,客户端内存缓存采用与三维可视化缓存共用的思想,通过智能指针的引用计数机制进行自动维护,采用基于缓存池的思想实现各类三维空间数据的分池管理;客户端

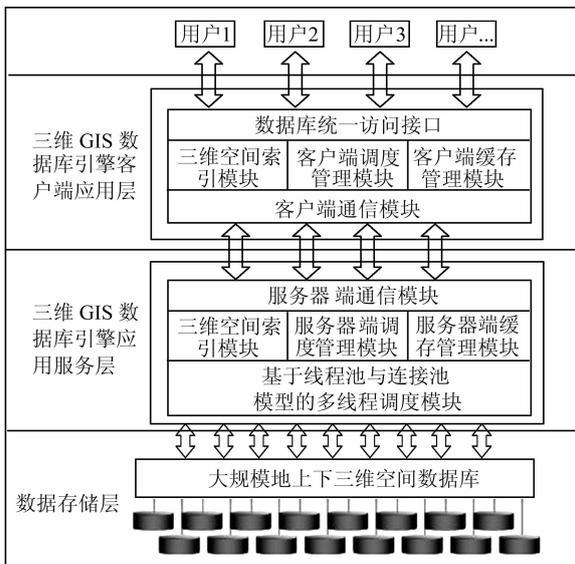


图4 三维 GIS 数据库引擎多层分布式体系结构
Fig. 4 Multi-level Distributed Architecture of 3D GIS Database Engine

为了满足三维实时可视化和高性能分析等的

文件缓存主要由三维空间数据索引文件和数据文件组成,采用基于空间索引节点统一组织、面向内存结构一致的三维空间数据文件缓存组织方式,有效提高文件缓存的数据访问效率;应用服务器内存缓存采用面向内存块的缓存组织方式,所有类型的数据都采用统一结构的内存块进行管理,直接应用于网络传输,无需进行序列化等额外操作。通过多级缓存自适应管理模块提供各级缓存统一的注册和查询接口,对各级缓存的命中率、空间利用率等信息进行统计,对各级缓存进行缓存空间、缓存粒度等方面的自适应配置。

3.3 基于线程池和链接池的多任务并行调度方法

针对多细节层次三维空间数据自适应可视化的特点,在三维实时可视化过程中通常会在瞬间产生大量的三维空间数据查询请求,尤其是多用户并发对三维 GIS 应用服务器造成的巨大负载,需要在三维 GIS 应用服务器上建立支持多用户并发的多线程模型,达到多用户请求并行处理的能力。基于此,本文搭建了基于线程池和数据库链接池的多线程调度模型,线程池模型实现了多线程并发控制,且动态自适应调整;数据库链接池模型实现了数据库链接的重用,提高了对数据库操作的性能。通过建立基于数据内容的多线程调度任务分配机制,使不同类型数据的调度任务能并行处理,并合理分配在多核处理器上的工作负载。同时建立了三维空间数据实时调度与预调度相结合的调度机制,通过实时调度和预调度分别建立不同的线程池和数据库链接池,达到实时调度和预调度并行处理的目的,并通过基于视点预测和数据内容相关性的三维空间数据预调度方法,将下一步可能实时调度的数据预先从三维空间数据库中读取出来,并在应用服务器缓存中进行管理,提高了应用服务器缓存的命中率,使实时调度不再频繁地进行数据库查询操作,显著提高了数据调度的效率。

3.4 面向多用户的异步通信传输方法

面向多层分布式架构的三维 GIS 数据库引擎需要客户端和应用服务器之间进行大量的网络通信,实现数据请求与大数据量三维空间数据的快速且稳定的网络传输。针对多用户并发的三维空间数据快速传输的特殊需求,三维 GIS 数据库引擎采用异步通信传输机制,在客户端,采用专门的线程处理数据请求的发送和返回三维空间数据的接收;在服务器端,采用半同步/半异步线程池模型支持多用户/多任务并发。针对三维空

间实体数据和纹理数据等大数据量的数据传输特点,采用了高效的分包和重组机制来达到数据传输的高效率。

4 试验与分析

为了验证三维 GIS 数据库引擎对于大规模三维空间数据管理与调度的能力,采用了武汉市现有的三维城市模型数据作为试验数据,包括武汉全市域 5 m 分辨率的 DEM 数据、1 m 分辨率的航空影像以及主城区 0.2 m 分辨率的航空影像,原数据量将近 100 GB。武汉市主城区的三维模型数据分布范围如图 5 所示,由建筑模型、交通设施模型、植被模型、城市小品模型、地下管线模型以及建筑室内模型等数据组成,原始数据量达到 200 GB 以上,三维目标数量达到 150 万个。通过三维 GIS 数据库引擎,将上述试验数据采用 Oracle 11g 数据库管理系统实现了一体化的建库和管理。

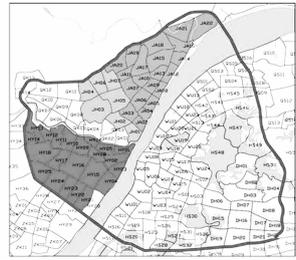


图 5 试验数据分布示意图

Fig. 5 Distribution of Experimental Data

试验环境包括 1 台数据库服务器、1 台应用服务器和 10 个客户端,数据库服务器与应用服务器通过千兆交换机相连,客户端通过千兆网络与服务器建立连接,数据服务器配置两路四核 CPU 2.53 Hz 16 GB 内存,安装 64 位 Oracle Linux 5.3 操作系统;应用服务器配置两路四核 CPU 2.53 Hz 16 GB 内存,安装 64 位 Windows Server 2003 操作系统;客户端为双核 CPU 2 G 内存,Windows XP/7 操作系统;每个客户端都基于全市三维空间数据库进行动态漫游,随着多用户并发访问的增加,对比客户端三维几何模型数据平均调度效率的变化。通过无客户端缓存的客户端直接访问模式(模式 1)、开启客户端缓存的客户端直接访问模式(模式 2)、多级缓存的应用服务器访问模式(模式 3)等三种模式进行试验分析,试验结果如图 6 所示。由图 6 可知,模式 2 的数据调度效率基本是模式 1 的 1 倍,验证了三维 GIS 数据库引擎客户端缓存的作用;与模式 1 和

模式2相比,模式3在多级缓存、多任务并行调度、异步通信传输等关键技术的有效协同下,数据调度效率进一步提升,尤其是对于多用户并发的数据调度效率更有优势。如图7所示,基于全市大规模三维空间数据库的多客户端动态漫游测试,每个客户端的三维实时可视化效率基本都稳定在25帧/s左右的流畅效果,与一般文件系统的效率已基本相当,从而验证了多层分布式架构的三维GIS数据库引擎在对大规模三维空间数据管理与调度方面的高效性。

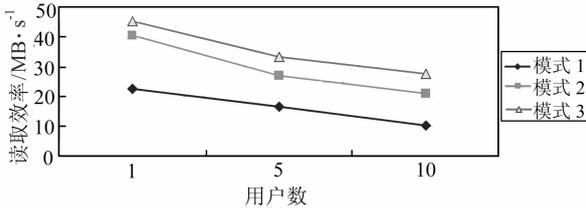


图6 多用户不同访问模式下的数据调度效率对比
Fig. 6 Comparison of Data Scheduling Efficiency in Different Multi-access Mode

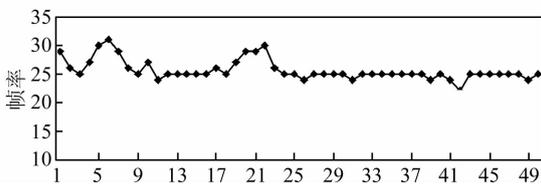


图7 基于三维空间数据库的实时可视化效率
Fig. 7 Real-time Visualization Efficiency Based on 3D Spatial Database

5 结语

本文针对地上下三维城市模型数据一体化的高效管理,提出了顾及语义的地上下三维空间数据一体化的数据库模型,并在此基础上设计实现了多层分布式架构的三维GIS数据库引擎,在多层次三维空间索引、多级缓存管理、多线程调度以及异步通信等关键技术方面取得了重要进展。通过武汉市全市范围三维空间数据库的实际应用分析,验证了三维GIS数据库引擎在大规模三维空间数据库高效管理和实时应用方面的有效性,突破了传统三维可视化系统将一个城市范围的三维模型数据分割为若干独立子系统的局限,有助于在全市统一的数据库基础上开展更广泛的应用。下一步将考虑增加服务器集群,并发展负载均衡技术支持更大规模的多用户协同应用。

致谢:感谢龚俊、杜志强、何珍文、周东波和许伟平等在三维GIS数据库引擎的设计与开发工

作中作出的贡献。

参考文献

- [1] 朱庆. 三维地理信息系统技术综述[J]. 地理信息世界, 2004, 2(3): 8-12
- [2] Zlatanova S, Stoter J E. The Role of DBMS in the New Generation GIS Architecture [C]/Rana S, Sharma J. Frontiers of Geographic Information Technology. Berlin: Springer, 2006
- [3] 郑坤, 负新莉, 刘修国, 等. 基于规则库的三维空间数据模型[J]. 地球科学——中国地质大学学报, 2010, 35(3): 369-374
- [4] OGC. OpenGIS City Geography Markup Language (CityGML) Encoding Standard 1.0 [OL]. <http://www.opengeospatial.org/standards/citygml>, 2008
- [5] Zlatanova S. 3D Geometries in DBMS [C]/Abdul-Rahman A, Zlatanova S, Coors V. Innovations in 3D Geoinformation Systems. Berlin: Springer, 2006
- [6] Kothuri R, Godfrind A, Beinat E. Pro Oracle Spatial for Oracle Database 11g [M]. New York: Apress, 2007
- [7] Stadler A, Nagel C, König G, et al. Making Interoperability Persistent: A 3D Geo Database Based on CityGML [C]. The 3rd International Workshop on 3D Geo-Information, Seoul, Korea, 2008
- [8] 汪明冲, 赵军, 李玉琳. 空间数据库引擎及其解决方案分析[J]. 地理信息世界, 2006, 8(4): 63-66
- [9] Bryla B, Loney K. Oracle Database 11g DBA Handbook [OL]. <http://thjfk.download.csdn.net/user/thjfk/Databasellg>, 2009
- [10] Zhu Qing, Gong Jun, Zhang Yeting. An Efficient 3D R-tree Spatial Index Method for Virtual Geographic Environments [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2007, 62(3): 217-224
- [11] He Zhenwen, Liu Gang, Weng Zhengping, et al. R-Lists: A Dynamic Spatial Index Structure Based on Generalized Lists [C]. International Conference on Tuture Computer& Communication, 2010
- [12] He Xubin, Ou Li, Kosa M, et al. A Unified Cache for High Performance Cluster Storage Systems [J]. International Journal of High Performance Computing and Networking, 2007, 5(1/2): 97-109

第一作者简介: 朱庆, 教授, 博士, 博士生导师, 长江学者。研究领域为虚拟地理环境、三维GIS。代表性成果有《数字高程模型》、《数码城市地理信息系统: 虚拟城市环境中的三维城市模型初探》。
E-mail: zhuq66@263.net