

# 基于共相思想和 Java EE 的 Internet GIS 研究与开发

李玮<sup>3</sup>, 朱江<sup>1,2,3</sup>, 胡中南<sup>3</sup>, 苏博<sup>3</sup>(1. 中国科学院地理科学与资源研究所, 北京 100101; 2. 中国科学院研究生院, 北京 100039;  
3. 北京超图地理信息技术有限公司, 北京 100085)

**摘要:** Java EE 是构建分布式企业应用信息系统的计算平台, 本文提出基于共相思想和 Java EE 的新一代 Internet GIS 的技术框架, 针对 Internet 环境中多用户并发访问和海量空间数据难以有效发布的难题, 在共相思想指导下, 引入集群并行技术, 提出动态负载均衡器中基于区域分治的负载平衡算法, 另外一项关键技术是针对 Java EE 多层体系结构的智能缓存技术, 为面向互联网公众 Internet GIS 提供了解决方案。

**关键词:** 共相思想; Java EE; Internet GIS

## Research and Development of Internet GIS Based on Universal Idea and Java EE

LI Wei-gu<sup>3</sup>, ZHU Jiang<sup>1,2,3</sup>, HU Zhong-nan<sup>3</sup>, SU Bo<sup>3</sup>(1. Institute of Geographical Sciences and Resources Research, Chinese Academy of Sciences, Beijing 100101;  
2. Graduate School of the Chinese Academy of Sciences, Beijing, 100039;  
3. SuperMap GIS Technologies, Inc. Beijing 100085)

**Abstract:** This paper introduces a new architecture of Internet GIS based on universal idea and Java EE technique, the distributed enterprise computing platform. The Internet GIS can handle the incoming multi-access request concurrently from Internet users. Another big challenge is to disseminate the increasing massive spatial data via Internet. We address the cluster computing technology with the region-division loading balance algorithms as the core component of dynamic load balancer. Multi-tier intelligent caching for Java EE platform improves the performance of Internet GIS dramatically.

**Key words:** Universal Idea; Java EE; Internet GIS

## 0 引言

Internet GIS 是 Internet 技术和 GIS 结合的产物<sup>[1]</sup>。Internet GIS 具有数据统一管理、空间分析计算集中处理、空间数据实时发布、简单易用等特点, 目前已经广泛应用于城市规划、设施管理、智能交通等众多领域。

在 Internet 发展的不同阶段, 国内外学者分别在基于 Internet GIS 项目、Internet GIS 软件体系结构等方面进行了深入的研究, 并提出了胖客户端和瘦客户端模式, 陈能成等<sup>[2]</sup>对基于 Java EE 的分布式 GIS

进行了初步的研究, 陈静<sup>[3]</sup>等进一步研究了使用 Java EE 技术实现 Internet GIS。但是, 目前还没有相关研究对大用户量并发访问的 Internet GIS 进行探讨, 特别是针对海量空间数据的发布与共享。

同一时期, 业界也意识到 Internet GIS 相对于传统 GIS 的优势, 国内外主要 GIS 厂商已经着手研制 Internet GIS 产品, 并推出了面向 Internet 的 Internet GIS 开发平台, 如 ESRI 的 ArcIMS, 超图公司的 SuperMap IS, MapInfo 的 MapXtreme 等。

传统 Internet GIS 偏重于提供地图网络发布能力, 随着 Internet 技术、定位技术和通信技术的发展,

更加深入的工程应用例如物流配送、移动定位服务等地理信息服务需求纷纷出现,对 Internet GIS 平台软件提出了更高的要求,要求集成行业模型和业务处理流程,集成地理空间数据和业务数据,进行基于地理位置的空间信息挖掘。另一方面,GIS 已经融入 IT 主流,GIS 社会化、大众化的进程使得 Internet 用户可以随时随地享受空间信息服务(如 LBS)。因此,Internet GIS 的研究具有很重要的研究价值和现实意义。

## 1 Internet GIS 与共相思想

要实现 Internet GIS,目前比较可行的解决方案有微软的分布式网络应用体系框架(.NET),对象管理组织 OMA 的 CORBA 标准(CORBA)和 SUN 公司的 Java 企业级平台(Java EE)。其中,CORBA 的规范大而全,互操作性和开放性好,但是实现的中间件软件非常复杂,并且技术和标准的更新相对较慢;而 .NET 和 Java EE 由于具备完善的架构体系成为了目前两大主流开发平台。但是如果针对每个平台都从头设计和开发 Internet GIS,开发和维护升级的代价都很大,那么是否能够从基于不同平台的 Internet GIS 上提取出一种共同的设计思想和技术框架呢?

哲学中用“共相”和“殊相”来分别指代普遍性和个别性的概念与事物。比如,使用 Java RMI 或 .NET Remoting 技术实现的 Internet GIS 集群是一个殊相,而超脱特定实现语言的 Internet GIS 集群分布式架构技术则是一个共相。基于 .NET 或 Java EE 实现的 Internet GIS 可以看作一个“殊相”,而具备支持任意开发语言和平台能力的 Internet GIS 技术则渗透着“共相”的思想。

如何应用 .NET 开发 Internet GIS,在文献[4]文章中已经进行了充分的阐述和分析,这里就不再作更多论述。本文试图研究 Internet GIS 中针对不同分布式平台技术的“殊相”,重点介绍采用 Java EE 分布式计算技术作为实现空间应用服务平台,并以此阐述提炼出来的具有“共相”性质的关键技术。

## 2 基于 Java EE 的 Internet GIS 框架

Java EE 得到了业界广泛支持和参与平台规范,并且具有良好的跨平台能力,基于 Java EE 的应用程序可以运行于任何平台。业界已经推出了多种企业应用服务器,如 BEA 的 WebLogic,IBM 的 WebSphere 以及开源项目 JBoss,这些应用服务器为 EJB 提供了

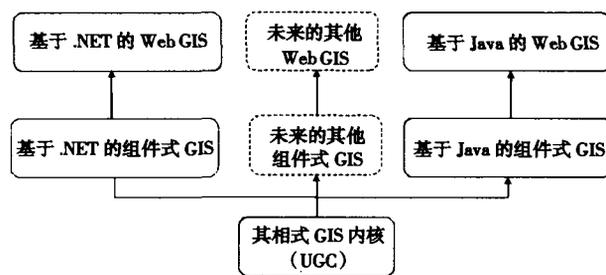


图 1 共相式 Web GIS 框架结构

Fig. 1 The architecture of Internet GIS based on universal idea

部署和运行环境,包括安全、事务处理、部署,以及并发和实例生命周期管理等服务。基于 Java EE 平台与企业应用服务器,GIS 软件厂商可以集中于网络地理应用的功能实现和 GIS 组件的封装。

基于 Java EE 规范,我们提出了 Internet GIS 的框架(如图 2 所示),在逻辑上将整个框架划分为客户端表现层、服务器端呈现层、服务器应用逻辑层和数据服务层。

客户端表现层是 Internet GIS 框架的前端,根据运行的实体不同,划分为浏览器客户端、Java 应用程序、支持 Java ME 的移动终端设备等。前端为浏览器客户端的网络应用,称为纯 B/S 的 Web 应用系统,这类客户端程序受到浏览器编程接口和安全性限制,UI 表现能力和交互能力有限。但是,系统部署和维护容易,是目前 Internet 使用最广泛的系统部署方式。桌面 Java 应用具有强大的图形能力和客户机访问能力,结合 JWS 部署技术适合构建胖客户端(Rich Client)模式的应用系统。支持 Java ME 的移动设备是 Java 的最新应用领域,无线网络技术的快速发展使得移动设备成为网络地理信息服务的重要客户端。客户端呈现程序与 Web 服务程序之间的通讯采用 Internet 的 HTTP/SOAP 协议或者无线网络的 WAP 协议。

服务器端系统逻辑上分为 3 部分:前端表现层、GIS 功能处理层(应用逻辑层)和数据服务层。服务器端系统的物理实现既可以在同一企业应用服务器容器中,也可以分别部署到 Web 服务器容器和 EJB 容器中。表现层的实现实体是 JSP、Servlet 等 Web 应用程序,使用 Session 管理客户端用户,用于响应客户端的请求并且输出处理结果。GIS 服务组件既可以作为 EJB 实体运行在应用服务器容器中,也可以作为独立的服务程序运行于后台,提供地图制图、

坐标转换、地理编码、路由导航等服务,Java EE 提供分布式计算服务,包括命名服务、注册服务、事件服务以及安全服务等,因此,GIS 应用服务通过 JNDI 进行执行名字和目录服务注册,向其他组件发布空间

信息处理功能。空间数据库扩展 Oracle 等商用 RDBMS 数据库,将地理实体的空间拓扑关系与地理实体集中统一管理,提供统一的访问接口(USDA, Universal Spatial Data Access)。

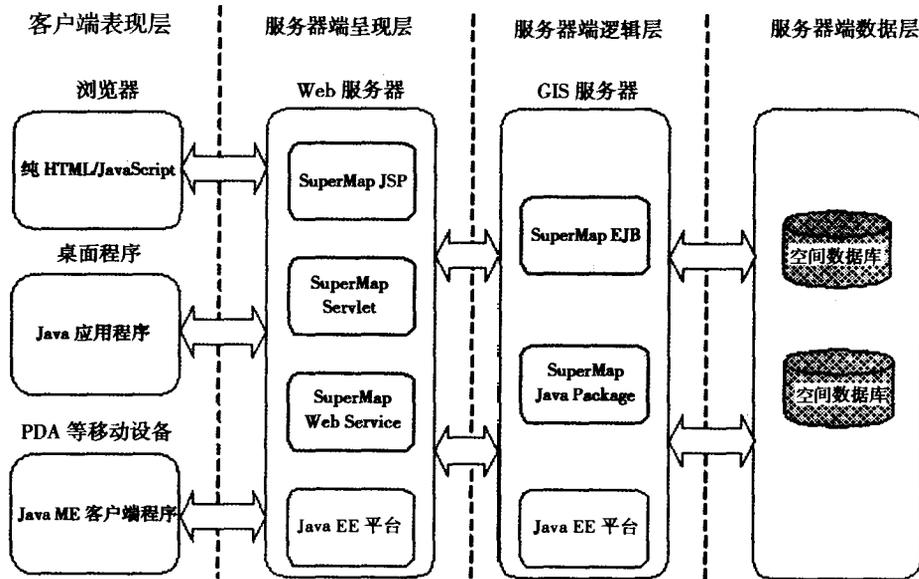


图 2 Internet GIS 框架

Fig. 2 The architecture of Internet GIS

### 3 关键技术研究

#### 3.1 集群技术

单应用服务器构建的系统能够适应中小网站的建设,随着访问用户的增加,容易达到硬件系统的极限。因此,我们引入集群技术,使用多台服务器协同服务。针对地理信息的空间分布特性,我们改进通用集群技术,提出面向地图内容的集群构建技术,其中动态负载器组件是其核心部件,负责分配服务器节点执行空间信息查询和空间分析。

在 Internet GIS 系统中,地理空间数据的吞吐量巨大,为了提高性能,必须采取必要的缓存策略,将经常访问的影像数据直接缓存在内存中,从而减少磁盘的 I/O 操作,实现了基于区域分治的负载平衡算法,提高了缓存的命中率。

对服务节点进行分组,一般采用静态分组方法(冗余格网法),即预先划分空间范围区域,相邻的空间范围区域具有一定的冗余重叠,对于每个请求,计算其对应划分空间范围区域的归属度  $P_i$ ,  $P_i = (A_u \cap A_w) / A_w$ , 其中  $A_u$  为用户请求空间范围的面积,  $A_w$  为

划分空间范围区域的面积,其最大归属度  $P_{\max} = \max(P_i)$ 。寻找  $P_{\max}$  对应服务节点中相对负载量最小的一个,如果当前节点的负载量  $L$  超过预先设置的预制阈值,那么该请求转入系统队列的尾部或者丢弃。实验证明,区域分治法简单高效,可以提高数倍的缓存命中率,优化系统性能。

#### 3.2 智能缓存技术

缓存技术是提供 Internet GIS 性能最有效的手段,Java EE 多层分布式计算平台将网络地理信息系统划分成多个独立物理层次,基于 Java EE 平台可以有效实现多层缓存机制,搭建高效的缓存体系框架。

地理信息数据预取机制。作为 Session Bean 存在的 GIS 服务引擎一旦被企业服务器装载,常驻内存直至企业服务器将主动进行系统资源回收。因此,预先从数据库服务中获取部分数据在内存中可以减少 GIS 服务组件访问空间数据库的次数,减少数据库系统资源的消耗。空间数据的分布具有明显的 2 维聚类分布的特点,相邻地理元素具有一定的相似性,Hilbert 曲线是一种空间填充曲线,它可以把  $n$  维欧氏空间划分成  $n-1$  维或更低维空间,利用它

可获得最佳的聚类效果。Hilbert 曲线能较好地体现空间对象的邻接关系,利用 Hilbert 曲线对地理元素进行编码,具有相同或者相近 Hilbert 码的地理元素其空间距离也相邻,因此,利用关系型数据库 SQL 查询操作,可以快速批量提取地理空间相近的地理元素,避免多次访问数据库。

**处理结果缓存。**根据 Internet GIS 用户访问统计规律,对地理空间按感兴趣度进行划分和建立等级制度,不同于传统 Web 服务器的 LRU 缓存算法,对于感兴趣度指标大的地理空间,我们缓存了查询条件和对应的查询结果;对于需要制图服务的请求,矢量数据栅格化的结果按照影像金字塔的数据组织方式进行管理,从影像金字塔提出的影像数据块的时间复杂度是  $O(\ln(n))$ 。

**网络前端缓存。**网络前端缓存是客户端程序的本地缓存方法,减少客户端程序连接服务器是提高 Internet GIS 系统最有效的方法。早期的 Internet GIS 系统主要手段是同时下载插件程序和矢量数据以减少后继的数据访问,这种方式带来了系统配置和维护上的难题,并且不能发布海量空间数据,不适于在 Internet 上部署应用。网络前端缓存采纳了这种系统的优点,采用了逐步获取一定比例尺度数据的方法,在客户端进行拼接,曾经访问过的区域数据缓存在客户端本地,减少了网络传输量和网络之间的交互次数。

根据 Java EE 的规范,我们设计了缓存测试环境,在 100M 局域网内使用 3 台 PC 服务器构建试验床,高速局域网消除了网络带宽的瓶颈因素,Web 服务器 Apache 作为系统的入口,应用服务器采用独立 GIS 应用服务器,数据库服务器采用 Oracle9i 独立运行后台服务器,测试数据是中国 1:1 000 000 全国基础地形图,测试程序使用 ACT 模拟多用户并发访问,基本操作包括地图操作、空间定位、空间查询等。

图 3 是性能测试的序列图,横轴记录时间,纵轴记录单位时间(s)处理的用户请求数(RPS),同样是 32 位服务器,启用 Cache 能够在单位时间内处理用户请求个数几乎是非缓存模式的百倍。在启用缓存模式的情况下,64 位服务器的处理能力略高于 32

位服务器,大约能够提高 25% 的用户请求量。性能测试表明 Cache 是否启用是提高单服务器网络服务能力的关键。

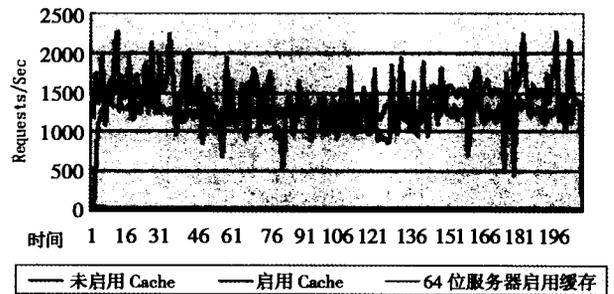


图 3 Cache 测试序列图

Fig. 3 The Diagram of Cache Testing

## 4 结 论

Java EE 是理想的分布式计算平台,也是实现 Internet GIS 的基础平台,本文提出基于共相思想和 Java EE 的 Internet GIS 框架,并且对其中的关键技术——集群技术和智能缓存技术进行深入剖析,同时进行了性能测试,验证该体系结构的可行性和高效性。

Java EE 在不断发展与改进,纳入新的技术和规范,在引入共相思想后,基于 Java EE 的 Internet GIS 也需要进行新的探索和技术创新,以适应面向服务的软件体系结构,满足各种企业级应用和门户网站的需求。

## 参考文献:

- [1] 宋关福,钟耳顺,王尔琪. Internet GIS——基于 Internet 的地理信息系统[J]. 中国图像图形学报,1998,(3).
- [2] 陈能成,龚健雅,朱欣焰,等. 基于 Java EE 的分布式 GIS 研究[J]. 测绘学报,Vol. 32, No. 2.
- [3] 陈静,龚健雅,朱欣焰,等. 基于 Java EE 的分布式 Web GIS [J]. 测绘通报,2004,(2).
- [4] 朱江,宋关福,钟耳顺,等. 基于 Web Services 和 .NET 技术的新一代 Web GIS 研究与开发[J]. 地理信息世界,2004,2(2):17-20.