

面向网格环境的 Java跨平台 GIS系统 原型实现的关键技术

高 昂^{1,2}, 陈荣国¹, 张明波¹, 李飞^{1,2}, 赵斯思^{1,2}

(1 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;

2 中国科学院研究生院, 北京 100039)

摘要: 针对传统地理信息系统在空间数据处理中无法有效进行数据访问与集成的问题, 提出了面向网格计算环境下空间数据访问与集成的地理信息系统体系架构, 并给出了面向网格环境的 Java跨平台 GIS系统原型的设计方案, 以及地理信息网格服务化提升的关键技术。进一步整合 Java拓扑模型、Eclipse富客户端平台与面向海量数据访问的数据网格组件, 以标准规范的开发流程, 给出构建面向网格环境具备可拓展结构的跨平台地理信息系统原型的实现过程。

关键词: 地理信息系统; 数据网格; 拓扑模型; Eclipse; 富客户端平台

1 引言

网格是面向下一代网络的新兴技术, 通过网格平台的资源整合能力, 可以实现多种资源的整合, 包括计算资源、存储资源、数据资源、信息资源、通信资源、软件资源、知识资源的全面共享与连通。通过共享与集成有效消除资源孤岛, 促进资源的充分利用。

地理信息系统面向大规模空间数据进行管理、分析、计算与存储。地理空间数据信息量大、分布式存储覆盖面广, 分析和计算具有数据密集和计算密集的显著特点。随着地理信息数据的不断膨胀, 要求数据管理者必须采取更有效的数据和计算共享方式来促进资源的充分集成与利用。网格技术的核心是解决广域网络环境下各种计算资源的共享和协同工作, 因此网格技术与空间信息技术的结合, 将可以使地理信息系统的发展跨上一个新的台阶。

随着网格GIS的体系结构模型的完善与发展, 与之相适应的地理信息系统平台设计与开发也成为当务之急。在面向空间数据访问与集成的网格GIS系统设计基础上, 研究地理信息网格服务化提

升的关键技术, 并基于Eclipse富客户端平台与Java拓扑模型实现UStudio地理系统原型, 以服务方式整合空间数据访问与集成的模型组件, 为进一步完善与实现网格GIS体系做出基础性的实践工作, 据此可以构建一个拓展性强, 跨越各种操作系统平台, 支持多种语言并且面向网格环境的GIS应用系统平台。

2 系统原型结构及关键技术的实现

2.1 网格GIS数据访问与集成体系架构

网格GIS软件体系平台是汇集和共享空间信息资源一体化组织与处理, 具有按需服务能力的空间信息基础设施。平台的目标是提供一体化的空间信息获取、处理与应用服务的基本解决方案, 以及智能化的空间信息处理平台和基本应用环境。网格GIS软件平台的研究为协调基础地理空间数据的收集、管理、分发和共享的基础设施提供技术支撑, 保证空间信息基础设施所涵盖的各系统之间能有效地实现数据集成和功能集成, 提升空间信息资源的整体应用水平^[1]。

网格GIS软件体系平台通过由下而上的层次,

收稿日期: 2006-02-23 修回日期: 2006-12-26

资助项目: 国家高技术研究发展计划 863信息获取与处理技术主题项目 (No. 2004AA132020)。

作者简介: 高昂 (1982-), 男, 博士生, IEEE-CS、CCF学生会员。主要研究方向: 空间数据库、地理信息系统、网格计算。E-mail: gaoang@lreis.ac.cn

使得分布的计算和数据资源在网格环境下可以充分地访问和集成,并将面向地理信息的 Web 服务向网格服务提升,整合各方面资源,形成数据、计算共享的研究环境^[2]。同时根据异构系统间信息共享和服务的需要,通过各种协议及标准将空间信息系统在分布式环境下互连,在网格 GIS 体系框架下将空间信息资源、地理信息服务及空间数据服务进行汇集、共享和协同处理。

针对空间信息处理技术的发展,网格 GIS 软件平台体系结构^[3]的设计,将集中在现有的空间信息资源的有效组织与管理,以及对空间信息资源的快速处理方面。其中,网格 GIS 系统框架下的空间数据服务以及数据资源的访问与集成在网格 GIS 体系中占据重要位置,图 1 所示是空间数据访问与集成及数据服务模块在网格 GIS 体系框架结构中的组织结构。

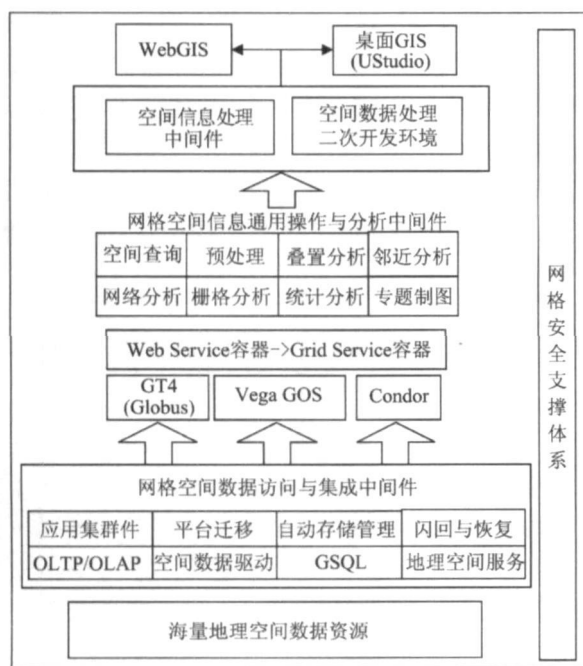


图 1 空间数据访问集成及服务在网格 GIS 体系框架中的组织结构

Fig. 1 Spatial Data Service and Integration in the Grid GIS Architecture

由图 1 可以看出,地理空间数据访问与集成以及数据服务与网格 GIS 体系结构的组织方式相互对应,层与层之间有着严格的关联依赖关系。海量地理空间数据资源是整个体系结构的基础,作为空间信息处理与共享的基础,涉及已有空间数据及数据存储系统,基础空间数据通过数据采

集设备可进行数据的在线实时更新。网格空间数据访问与集成中间件为系统服务层提供数据提取的环境,将分布式数据集成后提交给服务层对外提供访问接口。网格系统服务层借助 Globus Toolkit 或 Vega GOS 等网格环境支持对数据服务的访问进行控制与授权管理、全局用户管理、统一资源管理,并且提供元信息服务、文件管理服务、数据库服务、日志和审计服务等基本面向网格的系统服务。领域支撑层为空间数据提供 GIS 通用分析与操作的接口中间件,包括空间查询、统计分析、专题制图等通用操作,并将现有空间专题信息处理功能进行网格化提升,提高软件复用程度。位于组织结构顶层的应用层为用户提供网格 GIS 前端,作为应用层模块,USudio 系统原型可以完成通用制图及发布功能并且可以加载网格环境下的 Web 地图服务,集成与处理分布式的地理数据资源。

针对地理空间数据库的访问和集成方式,整个体系嵌入符合开放网格服务架构 OGSA 标准^[4]的数据访问和集成中间件 DAI 存取来自网格节点上各个分散数据源中的数据^[5]。数据访问与集成中间件可以将异构数据源提供的空间数据在网格数据服务注册表上登记,调用网格数据服务工厂 GDSF (Grid Data Service Factory) 来创建数据服务 GDS (Grid Data Service),并以 Web 服务的方式为数据访问提供统一的调用接口,利用 SOAP 消息实现客户机和服务器之间的通信。针对不同数据源的集成与分布式查询等操作中,DAI 的内部机制还屏蔽了如数据库驱动、数据格式以及客户端传输机制等技术细节,以透明访问的方式提供用户功能调用接口^[6]。

2.2 地理信息网格服务实现的关键技术

网格计算将分布式的计算机资源通过合理调度,组成共享的网络资源集合。在以空间数据访问集成为核心的网格 GIS 架构设计中,将异构资源通过服务化的方式整合在网格 GIS 体系中,利用空间数据访问与集成中间件提供的基础设施,实现 GIS 空间分析和处理服务的网格调度、负载均衡和空间数据的服务化发布。

实现网格计算技术与网络服务结合的 WSRF 网络服务资源框架是将现有 GIS 数据和处理功能进行网格服务提升的核心与关键。WSRF 在已有的

OGSA 规范上进行网络服务的扩展, 加入服务状态描述并制定描述规范, 以确保网络计算和网络服务在相同的基础上向前推进。网络服务资源框架 WSRF 融合了网格与网络服务两者的兼容性^[7], 并充分利用已有 GIS 空间分析处理算法通过网格服务拓展到分布式的环境之下。

网络服务资源框架 WSRF 通过隐含的资源模式为在网络服务间创造有状态的资源定义, 并且指定使用网络服务来访问有状态资源的一系列规范。包括网络服务资源特性 (Resource Properties)、网络服务资源生命周期 (Resource Lifetime)、网络服务基本故障 (Base Faults) 和网络

服务组 (Service Group) 等作用于整个服务周期的规范。网络服务在实现交互过程中有规则的处理操作状态信息, 在网络服务资源框架中, 状态作为有状态资源建模并且通过一个隐含的资源模式使网络服务间的关系模式化。

网络服务资源特性使用网络服务技术, 查询和更改有状态资源相关的地理空间数据服务, 并提供用于数据关联和协调由客户端访问的协议标准。WSRF 通过网络服务接口定义访问资源特性的描述, 网络服务资源生命周期包含服务资源回收与清除的方法。网络服务资源框架对于资源的存取描述可以由图 2 表示。

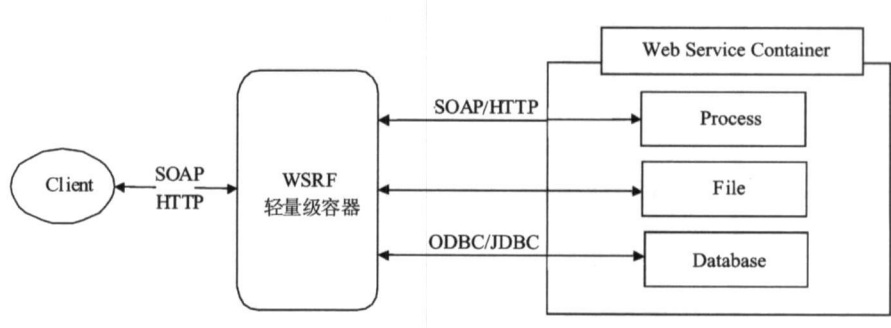


图 2 网络服务资源框架对于资源的存取

Fig.2 Access and integration of Web service resources

地理信息网格服务客户端可以通过服务发布的接口定义调用服务处理过程, 遵循 WSRF 规则发布的服务在轻量级网络服务容器中为客户端提供服务调用的 WSDL 描述, 网络服务容器中的处理过程、文件资源或数据库交互通过 WSRF 规范相应协议实现。网络服务组将网络服务和网络服务资源绑定在一起, 使请求者能够根据服务组的内容通过特定接口查询, 并对客户端调用资格通过分类机制进行限制与管理, 同时要求各个级别的成员必须通过统一的访问接口进行服务查询或服务过程调用。

2.3 Java 拓扑模型中间件与 RCP 富客户端插件体系

网格环境下空间数据存储于复杂的分布式环境, 存在异源、异构等情况, 采用传统的数据库访问方法很难满足应用要求, 需要有面向服务的跨平台桌面应用环境作为支持。在面向网络服务的 UStudio 原型实现过程中, 集成以 Web 服务模式进行数据访问与集成的中间件支持和 Java 拓扑模

型, 并使用 Eclipse 富客户端平台将各个功能模块以插件形式封装在系统内部。

Java 拓扑模型 (Java Topology Suite) 是具有完善的空间拓扑操作的地理信息通用模型, 通过调用其内置的空间分析算法^[18]与空间数据处理接口, 集成在基于 Java 语言的跨平台桌面系统中来实现地理数据处理功能。模型还提供了实现空间分析功能的 Java 应用程序接口, 并且符合 OpenGIS 标准定义的简单特征规范, 屏蔽底层数据环境的复杂操作, 实现对不同种类及格式地理数据和处理方法的透明访问。

作为实现地理信息通用功能的中间件, Java 拓扑模型提供基于内存的空间索引及健壮的空间索引算法, 系统通过 Java 拓扑模型功能调用接口, 屏蔽信息访问的底层细节, 实现空间数据的互操作, 保持应用的相对独立性。UStudio 系统在拓扑模型的基础上进一步拓展, 定义属性数据、要素类、空间参照等数据模型, 以及 Web 地图服务、地理标记语言、符号化、和 XML 文件读取的支

持，封装地理信息基础算法和空间数据通用处理操作，架构设计结构一致、层次清晰、易于拓展升级的系统平台。

在 UStudio 系统原型的功能集成上，基于 Eclipse 的 RCP (Rich Client Platform) 富客户端平台框架^[10]进行模块组织，最大限度地利用已有基础平台，实现资源的整合。Eclipse 是基于 Java 跨平台集成开发环境，Eclipse 的富客户端平台 RCP 应用重组 Eclipse 的功能模块^[11]，规范各部分组织结构，保证程序健壮性和拓展性，使软件设计者集中精力在业务逻辑侧面，简化软件的设计与实现过程。

富客户端平台基于 OSGi (Open Services Gateway Initiative) 标准以插件方式建立可扩展体系结构，OSGi 体系^[12]也是构建 Java 桌面插件体系最为高效可靠的方式。在 UStudio 中，除了集成 Eclipse 平台最小核心模块，其他的各项功能都作为插件加载到富客户端平台当中，并遵循统一的 OSGi 规范来拓展通用、安全、可配置的功能模块。OSGi 规范中的 Bundle 配置模块对各部分功能插件的生命周期作为对象进行管理，并且描述插件包的对外引用，提供的服务类型、依赖项、动态引用类库等。开发者可以将 Eclipse 提供的各种资源包集成到符合 OSGi 规范的插件中，针对不同目标操作系统发布 Java 富客户端软件产品。

3 面向网格环境的跨平台 GIS 系统原型实现

3.1 网格数据访问与集成原理

在 UStudio 系统的功能架构实现过程中，以服务为中心的统一 Web Service 调用接口为系统服务的主体，使符合 OGSA 标准的网格服务继承 Web Service 标准的多协议绑定、本地与远程接口的通用性。在 UStudio 系统中，服务发现与集成可以将高级网格服务功能向已有平台设施映射与应用。

面向服务的特点允许用户在不同层面上虚拟化资源，将异构关系数据库有机集成到网格环境架构中，并利用 SOAP 消息传递机制服务进行通信，实现跨越多个主机结点的分布式网格支持与协同工作。整个 UStudio 体系以 Java 运行时库和富客户端创建体系为基础，在其上以模块创建的方式插入各个功能部件，现实由底层运行环境到服务发布体系的各个层次之间的有机结合，UStudio 系统原型的内部协同工作模型如图 3 所示。

如图 3 所示，Eclipse 的富客户端平台以及 Java 拓扑模型作为系统架构的基础，在其上利用数据访问与集成中间件以 Web Service 的方式与异构的数据源进行交互与协同工作。运行在 Web 服务容器中的数据访问与集成中间件通过网格数据服务功能建立与异构数据源之间的数据传输通道，

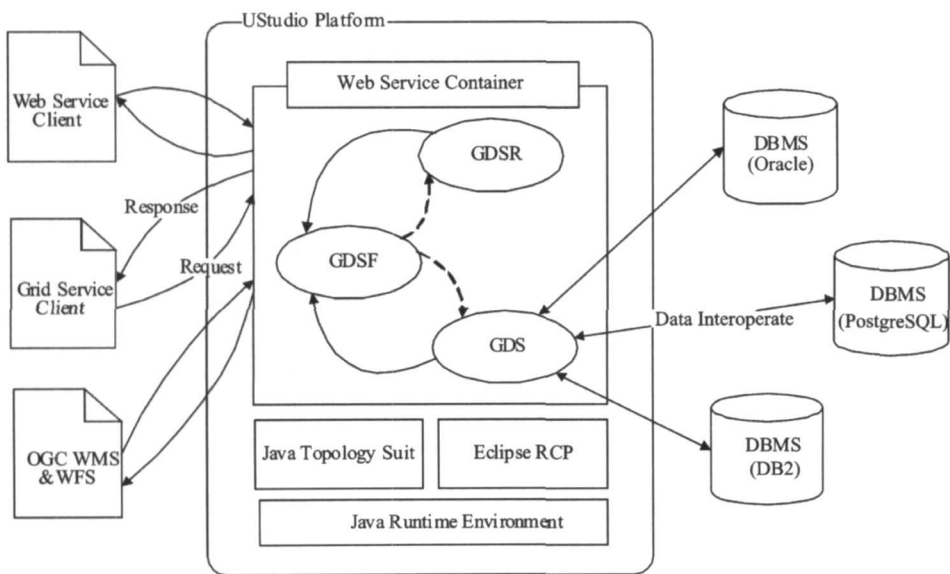


图 3 UStudio 系统原型的内部协同工作过程

Fig 3 Collaboration process in UStudio Prototype System

在接收到 Web 服务发出的数据请求之后，转向网格数据服务注册表（GDSR）查询已存在的服务或进行通信服务的注册，同时按照服务注册信息调用网格数据服务工厂 GDSF（Grid Data Service Factory）创建数据服务 GDS（Grid Data Service）。新创建的数据服务与异构的数据源建立通信链接，对数据库进行查询检索并向客户机返回数据处理请求结果。

在 UStudio 系统原型架构中，服务请求的生命周期由服务注册表进行统一的管理和控制，整个数据处理过程基于数据访问服务进行有机组织，针对存储于不同数据源的地理空间数据，实现交互访问与集成。

3.2 基于富客户端平台的功能模块拓展

从功能模块组织管理角度看，各部分功能结构插件的有机结合需要有统一的动态模块管理体系作为支撑，符合 OSG i 标准规范的 Eclipse 富客户端插件结构保证系统平稳的运行体系，所有功能扩展都通过富客户端插件进行配置与管理。并且整个系统基于松散耦合结构组成，保证各插件部分具有各自的相对独立性，系统中的插件通过功能接口来提供外部调用插件所实现的各种功能。

在 Eclipse 富客户端平台中配置插件，需要在功能部件清单 `feature.xml` 部分定义系统的各部分组成模块。在功能部件清单文件中，需要指定使用统一的 Unicode 编码方式。遵循 Java 包命名规范定义系统基础运行插件的名称，同时定义插件依赖项包含基础平台插件、功能拓展插件以及多国语言支持插等各个部分。

针对各个插件部分，在插件配置文件 `plugin.xml` 中定义相对应的拓展信息和拓展点标记。标准的 `plugin.xml` 插件内容定义包含插件名、ID 号、版本号等信息，以及插件的依赖关系、插件的运行支持库和插件各部分扩展点标记。由于遵循 OSG i 插件体系，插件基本信息和依赖关系在专门的配置文件 `MANIFEST.MF` 中定义，所以只需将插件的扩展点标记定义在配置文件 `plugin.xml` 中。配置文件中，所有的拓展都是定义在闭合的 `extension` 标签之间。基于服务的 OSG i 插件体系平台模式支持面向服务的交互，其交互模式如图 4 所示。

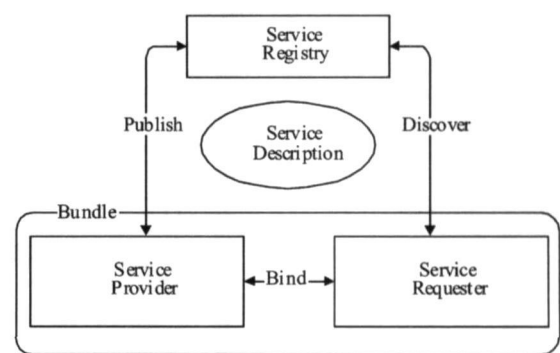


图 4 UStudio 系统原型中符合 OSG i 标准的面向服务交互模式

Fig.4 OSG i Interoperation Mode in UStudio Prototype

如图 3 所示，基于服务的 OSG i 平台的服务交互过程包含服务提供者、服务请求者和服务注册。服务提供者发布服务的描述信息，服务请求者根据描述信息发现已发布的服务并且与服务提供者进行绑定。在 OSG i 体系中，服务被表示为具有不同属性的 Java 类或接口，服务的属性允许不同的服务提供者使用相同的服务接口，同时服务注册表允许使用符合 LDAP（Lightweight Directory Access Protocol）轻型目录访问协议规则的查询语言进行服务的检索，此外，OSG i 的消息机制可以使服务请求者随时监控并获得服务注册表中的最新变动。

OSG i 为 Web 服务提供了一套标准的，面向组件的规范，可以在网络的任何地方管理组件的生命周期，并且组件可以动态的安装、加载、升级和卸载。软件组件可以动态的发现和使用其他类库或者应用程序接口。在 OSG i 体系中，服务的提供者和请求者都是 Bundle 包中的有机组成部分，服务接口由 Bundle 包创建的对象实现，同时 Bundle 负责应用运行时的服务依赖管理，包括服务的发布、发现以及动态绑定等内容。

3.3 UStudio 系统原型模块配置

UStudio 系统原型中的地理信息系统基本功能以模块化的插件形式整合在系统原型中，包括添加地图绘制、浏览放缩等标准功能，并且加入对地理标记语言 GML（Geography Markup Language）的支持，用来读取符合 GML 规范的数据，实现系统对符合 GML 规范的空间数据支持。在 Eclipse 富

客户端产品的插件导出为 JAR 文件时, 要求 MANIFEST.MF 文件与其相应的 JAR 包关联在一起, 插件各部分遵循 OSGi 规范的定义。位于 META-INF 目录下的 MANIFEST.MF 文件对 JAR 中各个依赖项进行配置, 每个插件包 (Bundle) 都需要通过配置文件描述其自身信息。

UStudio 系统原型发布时, 基于 Eclipse 提供的本地语言支持 NLS (National Language Support) 实现 RCP 富客户端国际化语言支持, 使其可以处理本地语言的数据格式文件而不出现乱码。在 UStudio 系统中, 针对每个插件包都建立带有 “nls” 名称的插件片段 Fragments, 在插件片段包中编写中文以及其他多国语言支持的配置文件, 在 Eclipse 的富客户端平台支持下, 不必修改程序的源代码, 只需将代码中的常量字符串外部化到遵循统一规格的 Properties 属性文件中, 再针对每种语言编写不同版本的 Properties 属性文件, 然后分布时部署到 JAR 打包文件中即可。在 Properties 属性文件中需要定义程序中常量与中文 Unicode 编码文字的对照映射关系, 当 UStudio 系统原型启动时, 加载中文界面只需在启动参数中加上相应的编码标志符号 “ZH_CN”, 即可实现软件产品的国际化支持。

遵循富客户端平台规范的 Eclipse 应用具有统一的界面风格, UStudio 系统原型的菜单、编辑窗口以及透视图都符合 Eclipse 默认的风格模式。基于 Eclipse 的 RCP 富客户端规范发布的产品可以运行在不同的操作系统平台之上, 由于纯 Java 语言程序运行与操作系统平台无关, 只要目标操作系统有 Java 运行时环境的底层支持, 系统即可正常启动运行。

在网络服务支持方面, 在获得 WSDL 描述信息之后, 在 UStudio 环境下可以直接加载网络过程处理服务的描述信息, 并将服务处理参数与地理空间数据提交到服务端处理后得到返回结果。系统原型以基于网络处理服务方式的访问与集成, 为地理信息的网络服务化提升提供了可行的实践基础。

4 结语

在面向网络 GIS 的地理数据资源组织与管理体系设计中, 利用数据网络的解决方案将分布的空间信息资源进行访问与集成, 实现空间信息资

源的共享和协同工作。体系框架基于面向服务的思想, 将开放网格服务结构下的数据访问与集成技术引入到空间数据库的访问中, 以网格服务的形式为用户提供地理空间数据的全面共享与互操作。

在系统原型的实现过程中, 将 Java 拓扑模型成熟的空间分析算法与空间数据处理功能集成到地理信息系统应用中, 遵循可拓展的标准规范并增加了程序的健壮性和适应性。Eclipse 富客户端平台的使用, 在重用标准软件界面组件功能的同时, 规范开发的流程与组织结构, 使产品具有与 Eclipse 相同的可拓展插件结构, 并利用优秀的开源框架将各组件优势充分整合。通过框架间的结合, 充分发挥了已有框架的优点, 使资源得到最大限度的节省和利用的同时, 也使得项目开发简洁、结构清晰, 并且具备了更好的可扩展性和可维护性。

参考文献

- [1] 沈占锋, 骆剑承, 盛昊. 基于 Web Service 的分布式遥感影像数据库设计与实现. 测绘通报, 2006 (2): 14~17.
- [2] 孙庆辉, 骆剑承, 周成虎. 网格 GIS 数据信息发布的关键技术. 地球信息科学, 2004, 6 (1): 22~26.
- [3] 方金云, 何建邦. 网格 GIS 体系结构及其实现技术. 地球信息科学, 2002 (4): 36~42.
- [4] 都志辉, 陈渝, 刘鹏. 以服务为中心的网格体系结构 OGSA. 计算机科学, 2003 (7): 26~29.
- [5] 刘建新, 阎保平. OGSA-DAI 体系结构及其关键技术研究. 计算机应用, 2004 (11): 81~87.
- [6] 刘雪梅, 柳永坡, 顾国昌. OGSA-DAI 框架模型的应用. 大庆石油学院学报, 2005 (6): 112~114.
- [7] 刘会斌, 都志辉. 网格与 Web 服务的融合-WSRF 和 WS-Notification. 计算机科学, 2005 (2): 76~79.
- [8] 王劲峰, 孙英君, 韩卫国. 空间分析引论. 地理信息世界, 2004 (5): 6~10.
- [9] 张晓东, 陈华斌, 池天河. 应用 Web 服务技术实现 GIS 互操作. 林业科学, 2006 (S1): 133~139.
- [10] 高昂, 陈荣国, 卫文学. 基于 Java 拓扑模型和 RCP 的 GIS 平台研究. 计算机工程与应用, 2007 (5): 106~209.
- [11] 高恒国, 罗克露. Eclipse 平台架构技术分析及其基于 RCP 的应用研究. 计算机与信息技术, 2006 (5): 1~3.
- [12] 熊江. OSGi 的分析和实现及其改进思路. 计算机科学, 2004 (3): 192~194.

A Grid Oriented Java GIS Architecture and Its Prototype Application

GAO Ang^{1,2}, CHEN Rongguo¹, ZHANG Mingbo¹, LI Fei^{1,2}, ZHAO Si^{1,2}

(1 *LERE, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101;*

2 *Graduate School of the Chinese Academy of Sciences, Beijing 100039, China*)

Abstract In order to deal with the complicated operation and poor expansibility in integrating spatial data, we design a grid oriented Java GIS framework for the spatial data access and integration as well as the key technology of GIS grid service upgrade. Also we put forward a cross platform Java GIS prototype to combine the topological model, Java rich client framework and data grid components together with the uniform standard to build high performance grid computing oriented Geographic Information System as well as to provide a best practice for the realization of prototype development efficiency towards grid environment.

Key words GIS; data grid; JTS; Eclipse; rich client platform