

适合遥感影像处理的网格计算环境研究^{*})

沈占锋 骆剑承 郑江 陈秋晓 明冬萍

(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

中国科学院研究生院 北京 100101)

摘要 网格计算是近几年来发展迅速的一种网络资源共享模型,其目的是网络资源的完全共享。网格计算相对于传统的C/S及Web GIS来说,它在网络计算、数据处理、资源共享、任务协同等方面都有了进一步的发展。在遥感影像处理与理解应用方面,由于涉及大量的网络计算及网络传输等耗时操作,应用网格计算思想进行遥感图像处理就更有实际意义。结合网格计算的网格计算及资源共享的优势,设计并开发适合遥感影像处理的网格计算环境及软件,是解决当前遥感影像海量数据处理的有效途径之一,同时也是遥感图像处理软件的一个发展趋势之一。本文结合实际研究成果,首先给出了基于网格计算思想的遥感影像处理的设计流程,并对其具体实现技术进一步进行了讨论。在分析其在图像处理方面应用的基础上,对面向网络的智能化遥感图像网格处理系统进行了分析与设计,并给出了其实现的一些关键技术。

关键词 网格计算,遥感图像处理,网络服务,数据传输

Grid Computing Environment Study for Remote Sensing Image Processing

SHEN Zhan-Feng LUO Jian-Cheng ZHENG Jiang CHEN Qiu-Xiao MING Dong-Ping

(Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences,

Graduate School of the Chinese Academy of Sciences, Beijing 100101)

Abstract Grid computing is a resource sharing model presented by computer experts to solving the problem of network resources imbalance and lopsided usage. Compared with C/S model and Web GIS model, Grid computing develops well at network computing, data processing, and resource sharing and task cooperation and so on. Remote Sensing image processing involves many network computing and data transfer, it is more valuable to apply grid computing to this field. It is a good solution to solve the problem of large data processing to design and develop a series of software with the support of grid computing environment, which is a development direction of this field. This paper aims to build the grid computing environment for remote sensing image processing, firstly gives the system design work flow, and then talks about its implementation method, at last gives some of its key techniques.

Keywords Grid computing, Remote sensing image processing, Network services, Data transfer

1 引言

近几年来,随着全球信息化的要求及遥感卫星技术的发展及成熟,遥感作为一种信息快速获取的手段就变得尤为突出,相应地,遥感图像处理与理解技术也就变得更加重要。特别是随着高分辨率遥感技术及全球卫星定位技术(GPS)的发展,遥感图像处理所要解决的问题越来越复杂,所需处理的数据量也越来越大,同时对遥感图像处理的精度及速度要求也越来越严格,其中还涉及大量的网络数据传输及互操作等问题,这就给图像处理技术带来了很大程度的挑战。为适应这种挑战,人们正试图从两个主要方面进行改进。一方面就是针对不同类型的遥感图像的特点及图像处理算法的特点进行遥感图像处理算法上的改进,这种方法试图从遥感获取、处理等原理出发,从处理算法方面进行进一步的优化,以达到遥感图像处理的快速、精确化的要求。这种方法在遥感图像处理技术发展的初期发展得很快,到现在一些特定算法已趋于成熟。另一方面,人们正从计算机处理及网络计算的方面进行改进,以适

应现在的网络化大数据量处理的应用及多机分布式并行处理的要求,其表现形式主要有图像的并行、分布式处理以及集群式处理,应用方面表现为C/S及Web GIS应用等,本文要讨论的适合遥感图像处理的网格计算环境研究也属于这一范畴。

2 网格计算

网格计算是近几年来计算机研究人员为解决当前网络资源分配与利用不均而提出的一种先进的网络资源共享模型,这里的资源不仅包括传统意义上的数据资源,还包括网络计算资源、存储资源、信息资源、知识资源、软件资源等^[3,4]。在遥感图像处理领域,针对不同的任务而设计出能够同时实现计算资源与数据资源共享的模型是非常必要的。网格计算为有效解决遥感图像大容量空间数据共享提供的各种底层处理功能的支持,并为建立上层的图像处理及各种空间数据应用提供了环境支持。

Web GIS是近几年发展迅速的一种网络共享模型,它具

^{*})项目资助:国家“863”项目(2002AA135230),国家自然科学基金项目(40101021)。沈占锋 博士生,主要从事网格计算,网格GIS,遥感图像处理与理解等方面研究工作。

有多用户性、客户端平台独立性、操作简单、成本较低等优点^[5],可应用于 GIS 及遥感图像处理中的数据共享及网络计算部分。但 Web GIS 的局限是各个 GIS 站点是独立的,不能互相访问和调用,并存在信息孤岛^[6]。网络计算在解决网络资源共享及消除信息孤岛方面给出了很好的解决方案。

并行计算及集群计算也是解决大数据量计算的一个较好模型,它们也被广泛应用在遥感图像数字处理领域。相比网络计算模型来说,它们的资源共享范围更小,需要更高的配置,并且都需要专门的技术人员的大量编程才能较好地工作,而网络的思想则是在此基础上的更广泛的网络资源共享,因此能够更好地为遥感图像的大数据处理服务。

网络计算的思想是在网络计算方面兼具了并行计算与分布式计算的优点^[1,2],在数据管理方面提供了一种可供全球资源共享的机理。同时,网络计算还融入了“知识”方面的内容,进而使得网络处理方面更具智能化,网络计算的此特点对于改善现有图像处理特别是大数据量的高分辨率影像处理无疑都是非常必要的。

随着越来越多的专家学者致力于网络计算的研制与开发工作,网络计算已从其概念阶段发展为现在的设计与实现阶段。目前国际上已经出现了一些网络计算的雏形,其中一些已投入试运行阶段,比较著名的有:美国自 1997 年底的“分布式网络”研究项目开始,先后有 NASA 的 IPG 网络,美国能源部的 ASCII 网络,美国国防部的 GIG 网络试验运行;欧洲共同体的 Euro 网络和 Data 网络也投入了研究实验;我国科技部在“九五”开展了国家高性能计算环境(网络)的建设和关键技术的研究。“十五”期间科技部加大了对网络技术研究和推广的力度,目标是突破网络关键技术,建立网络计算技术标准,将网络计算技术应用到行业和企业应用中,建立行业和企业应用网络,进一步加强全社会共享的国家高性能网络计算环境的建设,推动我国网络产业的形成和发展^[14],而网络在各领域中的应用也进入了探索性研究阶段。

遥感数据处理是遥感领域及空间信息应用领域的重要基础,它涉及各种图像基本预处理算法、图像分类与分割、图像特征提取算法等。近年来,基于统计学习以及人工神经网络等方法的介绍极大地提高了计算机化遥感数据处理分析的效果,但一直很难克服其中的计算复杂性问题,特别是针对大数据量的海量遥感数据处理时,这个问题就显得尤为重要^[11,12]。因此,在保证处理效果的前提下,在遥感图像处理中引入网络计算的思想,用以提高图像的处理速度,是解决当前海量遥感图像数据处理的最佳方式^[13];同时,应用网络计算模型,也可解决海量遥感图像在广域范围内的数据共享,防止了资源的浪费。

本次设计的遥感图像处理网络计算环境主要是针对遥感图像处理中的图像分类与分割等比较耗时的工作,以建立适合遥感图像处理的网络计算环境为目的,在分析不同的遥感图像处理算法的基础上,根据适合不同算法的网络通信方式,应用 XML、Web Services、RPC 远程调用等技术,创建适合遥感图像处理的网格化软件。

3 针对遥感影像处理的网络计算环境设计

3.1 网络计算环境工作流程设计

考虑到实际遥感图像处理任务及环境的要求,我们在微软平台上采用 Web Services、ATL 及中间件等技术,在 Visual Studio .NET 环境下结合网络计算中计算资源共享的工作

原理,根据遥感图像处理的实际需要,对应用网络计算思想进行遥感图像处理的环境进行了系统的设计。在网络计算系统设计中,我们保留了 C/S 及 Web GIS 模型中逻辑上三层结构的优点,用户不能直接操作底层数据资源而只能通过特定的接口对其数据进行访问,这样有助于保证底层数据资源的安全性。同时,我们采用了微软的 ATL 类库对系统进行了实现,因为 ATL 相对于微软的另一个大的类库 MFC 来说,ATL 主要用于开发轻量级的适于网上传输下载的 COM 组件,它通过模板机制来实现代码的复用,体积小,效率高。在数据传输方面,我们采用 TOMCAT + Java(JSP)进行数据传输,实践证明传输速度较快。

网络架构体系是指系统的软件结构体系,它决定了整个系统的运行可行性与扩展性。网络架构中定义了网格内及网格结点的功能,以及其各种数据传输及通信协议,用以指导网络系统及相应应用程序间的各种操作^[10]。

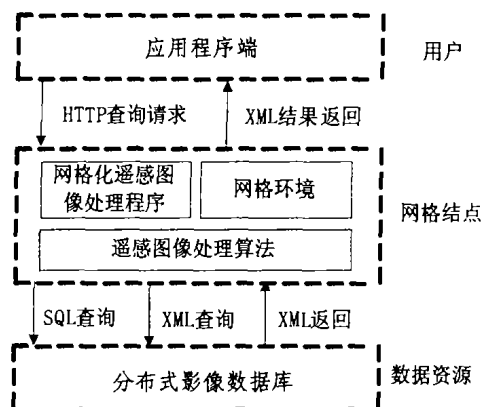


图1 应用网络计算环境进行遥感图像处理的体系结构

图1展示了本次应用网络计算思想进行遥感图像处理的体系结构。结合总体分为三层:用户应用层,网络结点层及数据资源层,这三层结构对应了实际网络物理结构的三种功能。数据资源层是指各分布式影像数据库,用以提供用户任务所需要的各种数据^[7,8]。网络结点层由各种图像处理算法及网络计算环境组成,它是系统的核心组成部分,主要负责完成用户提交的任务,包括数据资源的查找、处理及返回用户处理的结果,其具体实现可由各种网络中间件辅助完成。用户层指与用户交互的界面部分,负责用户的登录、任务提交及返回结果的显示。整个通信过程均采用 XML 形式进行通信,符合 Web Services 及网络计算标准,并具跨平台性、跨编程语言性,并可根据实际的需要进行相应的扩展。

图2是遥感图像处理的网络结点运行策略示意图。在网络环境中,各网络结点之间在功能上具有相同的地位,即任意两网络结点在 Internet 上地位是对等的。系统对遥感图像处理的流程大致如下:①用户从某一代理处登录网络环境(连接网络结点),并向网络系统提交图像处理任务(如遥感图像分割、分类、目标自动识别等);②网络结点对任务进行分析与分解,并建立与本地网络结点所拥有的资源服务器之间的连接;③通过查找本地资源服务器内的数据库查找任务所需要的资源;④如果所需的资源,则建立与所需资源间的连接;⑤如果本地没有所需的资源,则通过 HTTP 连接到其它网络结点并进行资源的查找^[9],各网络结点会有相应的服务处理其它网络结点的消息响应;⑥⑦步骤类似③④步,进行用户所需资源的远程查找;⑧是网络计算环境具体进行遥感图像处理的过程。默认情况下,系统将只使用用户代理端网络结点下属的代

理计算机进行计算,用户也可以根据实际需要规定具体参与

图像处理的计算机及个数。

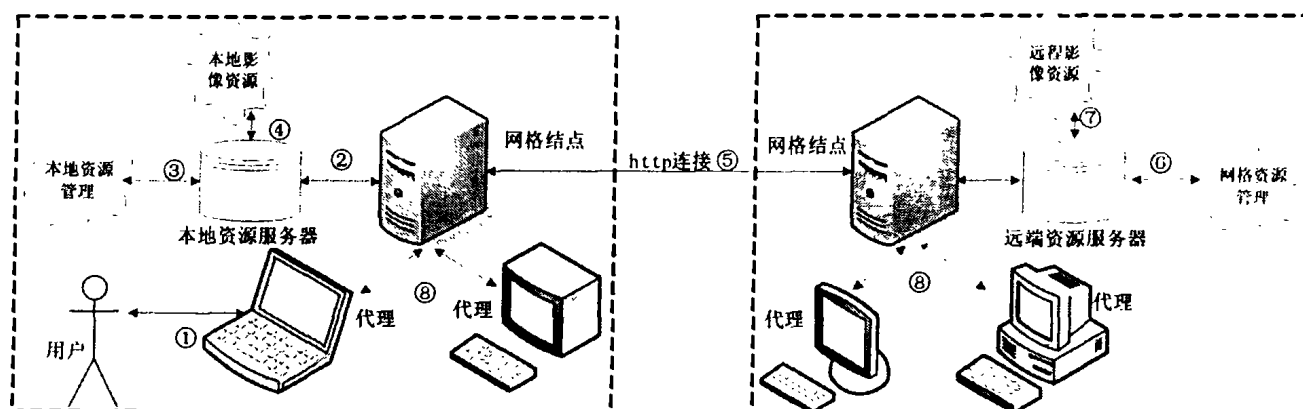


图2 遥感图像处理的网络结点间运行策略

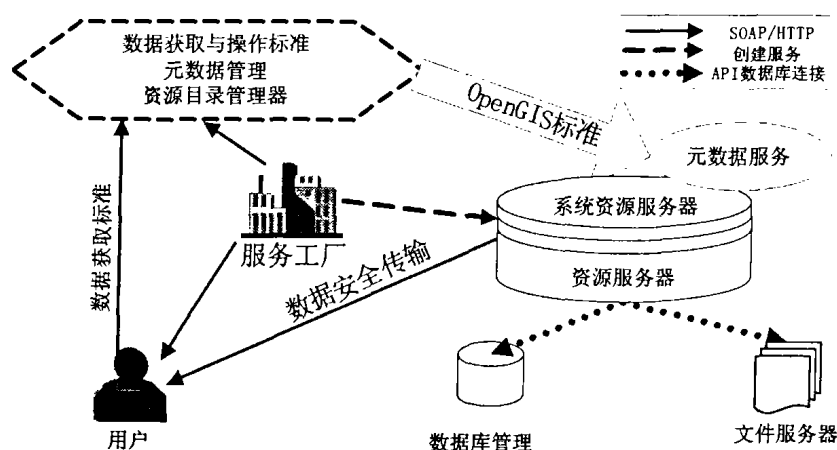


图3 单个网络结点内部处理遥感图像的资源管理模式

图3是一个网络结点内部的资源管理及产生模式过程的示意图。用户在登录系统后,可以将其某些资源声明为可共享资源,并进行资源注册。资源注册中心的资源目录是按层次结构目录组织的基于数据库与LDAP形式的资源管理机制与索引服务结合的形式,即二者相结合以管理分布式的任务、计算资源及其它的各种资源,这样可以方便对任务的查询和控制,方便不同应用的信息交流,同时为不同地理位置的计算机系统之间的任务迁移、任务调度、负载平衡提供技术支持。图中的服务工厂的作用是负责搜索合适的算法及计算机以供用户使用,服务工厂在适当时候会派出不同的服务以完成不同的任务,系统的资源管理符合OpenGIS原理。资源服务器以API数据连接的方式连接文件系统及数据库系统,在我们的系统中,数据库管理系统是应用Oracle的数据管理及Oracle Spatial的矢量数据管理的功能。

3.2 网络计算环境实现关键技术

3.2.1 资源注册与发现 资源注册与发现是网络系统的一个重要组成部分,它决定着系统对不同资源整合的速度及质量。表1列出了我们系统中各资源注册所需要的参数的列表,其中含*的为必须填写的项。

表1 网络系统资源注册所需服务参数

服务代号	服务所属节点*
服务名称*	服务具体描述
服务发布人*	服务计费说明
服务发布时间*	服务参数文件*
服务有效时间*	服务状态
服务位置*	服务所属类别*

通过填写服务参数,用户可以将其某些资源发布为共享,其它用户就可以根据此用户注册的内容发现服务并进行使用了,服务的使用过程即为网格内部实现的过程。

3.2.2 网格数据传输 数据传输是网格系统中不可避免的一个环节,数据传输的效率直接影响着系统的整体速度。本次应用中采用ASP+JSP方式共同组成,ASP可以方便地与系统核心实现部分的ATL技术进行消息通信,JSP由于采用Java语言实现,可以方便地使用网格上的一个可利用资源,服务器采用IIS-TOMCAT组成,即JDK开发语言环境+IIS-TOMCAT(Web Server)+JSP(开发语言·环境)。网格结点间及网格结点与代理计算机间通过http请求通信。表2规定了本网格结点服务的各参数。在我们的网格系统中,暂时规定了两种服务:图像处理服务与数据传输服务。图像处理服务包括图像分割与图像分类,其参数包括命令cmd、操作op、分割或分类方法method、处理文件名fileName、原始文件类型fileType、登录用户名(本网格内操作采用登录代理计算机的用户名,涉及网格结点间操作采用网格系统统一授权用户名);数据传输服务包括数据下载至本地计算机和本地数据上传至其它网格结点,其参数与图像处理基本相同,只是需要指定目标target,当需要进行这些服务时,网格结点将按表2所示的参数构造http请求,http://address/GIS/ImageProcessService? cmd = imageSegmentation&op = segmentation&method = otsu&fileName = lreispc-shenzf-20030915122105&fileType=BSQ&userName=shenzf 为图像分割的http请求,http://address/GIS/DataService? cmd = importData&

(下转第66页)

结束语 根据以上的技术分析, TCP 协议和自定义 IP 数据包存在着那些与楼宇自控网络和 BACnet 标准特点相违背的地方, 所以在 BACnet/IP 技术中没有采用这两种方式, 而是采用了 UDP 协议。BACnet 作为专门为楼宇自控领域开发的互操作标准, 正是由于在各个方面都充分考虑了楼宇自控网络自身的特点, 才使其成为楼宇自控领域中最具有竞争力和最能代表楼宇自控网络发展方向的标准。

参考文献

1 Newman M. BACnet - A Tutorial Overview. <http://www.bac->

[net.org/Tutorial/HMN-Overview/sld001.htm](http://www.bacnet.org/Tutorial/HMN-Overview/sld001.htm), 2000

- Bender J, Newman M. BACnet/IP. <http://www.bacnet.org/Tutorial/BACnetIP/default.html>, 2000
- Frequently Asked Questions (FAQ) about BACnet. <http://www.bacnet.org/Bibliography/HPAC-3-97.pdf>, 2000
- 深圳市鑫王科技发展有限公司译. BACnet 楼宇自动控制网络数据通讯协议. 广东: 广东经济出版社, 2000
- 董春桥编著. 智能楼宇 BACnet 原理与应用. 北京: 电子工业出版社, 2003

(上接第60页)

op = importImageFile&fileName = lreispc-shenzf-200309

28101311&fileType = BSQ&username = shenzf&target = Export 为下载影像文件时的 http 请求。

表2 网络结点服务种类参数说明

服务名称	Parameters							应用范围	数据传输方式
	cmd	op	method	fileName	fileType	userName	target		
ImageProc-ssService	imageSegmentation	segmentation	不同算法名称	域-用户时间	影像文件类型	Client 机登录用户名	—	局域或广域	共享或 http
	imageClassify	classify							
DataService	importData	importImageFile	—	影像文件名	影像文件类型	网格授权用户名	目标主机接收文件夹名或网络数据库	局域或广域	共享, http 或 ftp
	exportData	exportImageFile							

当请求端向服务端发送出 http 请求时, 服务端解析请求并执行相应操作, 以 XML 的形式返回相应的执行结果, 图像分割任务回信的 Instance 实例如下:

```
<?xml version="1.0" encoding="Shift-JIS"?>
<SegmentationResult cmd="imageSegmentation" code="101" op="segmentation" method="otsu"
time="2003/09/15 12:21:25" user="shenzf">
  <Abstract>
    <Text>...</Text>
  </Abstract>
  <Detail>
    <Text>...</Text>
    ....
  </Detail>
</SegmentationResult>
```

请求端通过解析此 XML 文档的各种参数以决定下一步的工作(XML 文档中 code="0" 时为程序执行正常), 如果正常则继续执行后面的操作, 发生异常则进程转至错误管理处进行进一步的处理。

4 实验效果分析

根据以上的设计, 我们对系统在 Windows 2000 平台上进行了实现, 实现以遥感图像分割为例, 系统含有3个网络结点, 并有一个结点负责服务及资源的注册管理(相当于 Web Services 技术的 UDDI 功能)。实现效果表明, 系统对文件传输部分效率较高, 系统能够按预定的程序流程运行, 在各不同网络结点上对遥感图像进行处理后返回给用户最终的处理结果。

结论及下步工作 网格计算是当前计算机领域的一个新的研究问题。在给遥感图像处理提供了很好的应用前景的同时, 网格计算也给我们提出了更多的挑战。针对遥感图像处理的特点及当前网格计算的发展, 我们提出了应用网格计算进行遥感图像处理的方法, 对其具体流程之构建方法进行了设计, 并应用 Web Services、XML 等技术对其进行了实现, 以服务于遥感图像网络计算。

网格计算是一项跨地区、跨平台、跨专业的研究项目, 需要付诸很大的投入。鉴于网格计算系统的复杂性, 基于现有网

络, 我们设计了网格计算环境的框架, 并为实际应用保留了接口。下步工作将是完善现有网格系统, 如考虑数据传输的断点续传, 并设法提高整个系统的进行效率等; 同时改善现有图像处理算法, 使其更适合于网格环境的分布式并行特点, 以更适合于大数据量遥感图像处理的速度要求。

参考文献

- XIAO Nong, FU Wei. SDPG: Spatial Data Processing Grid[J]. Journal of Computer Science & Technology, 2003, 18(4): 523~531
- CHEN Lin, WANG ChoLi, Lau F C M. A Grid Middleware for Distributed Java Computing with MPI Binding and Process Migration Supports[J]. Journal of Computer Science & Technology, 2003, 18(4): 505~514
- Foster I, Kesselman C. The Grid: Blueprint for a New Computing Infrastructure[J]. Morgan Kaufmann Publishers, Inc. San Francisco, CA, 1999
- Tuecke Steven, Czajkowski Karl, Foster Ian. Grid Service Specification [J/OL]. 2002. <http://www.gridforum.org/ogsi-wg>
- 于雪英, 江南. 基于 XML 三层 B/S 结构的 WebGIS 构建[J]. 地球信息科学, 2003, 6(2): 56~59
- 许春杰, 邹乐君, 沈晓华. 基于 Web Service 的 WebGIS[J]. 测绘通报, 2003, 1: 41~43
- 赵峰, 郭容赛, 蒋建国. 基于 GeoMediaPro 的遥感影像数据库研建[J]. 测绘通报, 2003, 4
- 王密, 龚健雅. 基于扩展关系数据库的遥感影像数据库管理系统的研究与实现[J]. 测绘信息与工程, 2002, 27
- 黄国满, 郭建峰. 分布式并行遥感图像处理中的数据划分[J]. 遥感信息, 2001, 2
- The Globus Project™ Argonne National Laboratory USC Information Sciences Institute. Grid Architecture [J/OL]. <http://www.globus.org>
- 刘纯波, 李琦, 承继成. 基于 XML-RPC 的分布式地理信息系统计算模型[J]. 地球信息科学, 2003, 1
- 郑江, 骆剑承, 陈秋晓, 等. 遥感影像理解智能化系统与模型集成方法[J]. 地球信息科学, 2003, 1
- 沈占锋, 骆剑承, 蔡少华, 等. 网格 GIS 的应用架构及关键技术[J]. 地球信息科学, 2003, 4
- 肖依, 卢锡城, 王怀民. 网络计算的四种形式[J-OL]. 计算机世界报, 2002 (40): B-B4. <http://www.yesky.com/20030313/1656808.shtml>