

水利信息网格体系结构设计及关键技术研究

谢禹, 王志坚, 许峰

(河海大学 计算机及信息工程学院, 江苏 南京 210098)

摘要: 网格技术为解决数据共享问题提供了解决方案。从网格技术在领域中的应用出发, 结合水利数据分布异构性、计算密集、资源分布等问题分析了水利行业对信息网格的需求, 设计了一个层次式的、以“服务”为中心的五层水利信息网格体系结构。重点研究了水利元数据管理、元数据目录服务、资源交易服务等关键技术。

关键词: 水利信息网格; 数据共享; 元数据; 层次式体系结构

中图分类号: TP393

文献标识码: A

文章编号: 1673-629X(2006)06-0020-03

Design on Water - Resources - Information - Grid Architecture and Research on Relevant Key Technologies

XIE Yu, WANG Zhi-jian, XU Feng

(College of Computers & Information Engineering of Hohai University, Nanjing 210098, China)

Abstract: Grid technology provides a solution for data sharing. Based on the application of grid in relevant area and water resources survey which is characterized by data transparent, computing - intensiveness, distribution of resources, requirement analysis of grid is discussed. A new service - oriented and five layered system architecture of information grid in water resources are designed to solve application problems of water resources. Meanwhile, some crucial technologies are discussed in detail, such as metadata management, metadata directory service, resource trading service, etc.

Key words: information grid of water resources; data sharing; metadata; layered architecture

0 引言

网格(grid)是近年来兴起的一种重要信息技术,它构筑在互联网上,将互联网、计算机、数据库、传感器、远程设备等融为一体,实现计算资源、存储资源、信息资源、知识资源等的全面共享,消除信息孤岛和资源孤岛。它是划时代的技术,传统因特网实现了计算机硬件的连通,Web实现了网页的连通,而网格试图实现互联网上所有资源的全面连通,在动态变化的多个虚拟机构间共享资源和协同解决问题。

正是由于网格共享、服务、标准的技术特点,为解决各个行业资源共享难题提供了解决方案。

1 相关工作

美国、欧洲、日本和印度等都启动了大型网格研究计划,并得到了产业界的大力支持。Globus 是美国 Argonne

国家实验室的网格计算研发项目,目前 Globus 已经推出 Globus Toolkit Version 3 正式版,并提出了与 Web Service 兼容的 Web Service Resource Framework(WSRF)规范,已在 NASA 网格(NASA IPG)、欧洲数据网格(data grid)、美国国家技术网格(NTG)等8个项目得到应用。目前,国内在网格技术方面取得了一定进展,2003年底中国国家网格推出了网格操作系统1.1版并初步建立了实验床,行业应用部门也重点开展了网格应用的部署和开发工作。2004年中国网格推出了相应的网格系统软件和试验床。

网格技术应用于各个领域方面相关的研究与工作如:郭权等人^[1]介绍了网格技术在药物开发领域药物分子对接中的研究和进展;谭国真等人^[2]介绍了网格技术在交通信息方面的研究,交通信息网格通过异构计算机的整合实现更高层次的资源共享,提供更强计算能力;张英朝等人^[3]针对于目前电子政务系统建设中的问题,提出了一种基于网格技术的电子政务平台体系结构;唐宇等人^[4]介绍了国家地质调查应用网格体系及关键技术研究。网格技术在行业中的应用为水利信息网格的建立奠定了良好的基础。

2 水利行业对网格技术的需求分析

近年来,水利信息化建设取得了显著成效,但由于水

收稿日期:2005-11-08

基金项目:国家“九七三”重点基础研究发展计划(2002CB312002);
国家自然科学基金(60573098)

作者简介:谢禹(1981-),男,江苏南京人,硕士研究生,助教,研究方向为信息网格;王志坚,教授,博士生导师,研究方向为基于网络的计算机应用技术、软件复用、基于网络的软件系统集成技术。

利数据存储的异构性、分布性、动态性,信息类型纷杂,缺乏统一规范的存储模式和标准,决定了当前技术不能满足跨行政区域、跨学科的数据共享和协作,成为一直困扰着水利业务发展的瓶颈,具体表现在以下几个方面:

(1)分布于不同地域、分属于不同业务的水利部门信息资源缺乏有效的共享与利用,当前各流域机构、省(市)建设了基于各业务部门的系统,呈现条块分割的特征,形成以专业、部门等为边界的信息孤岛。各数据库之间缺乏信息共享机制与手段,有些内容还相互重复。

(2)水利系统的分布性与异构性,导致了不能方便地实现跨流域、跨部门、跨系统的互连互通互操作,各部门、各子系统“各施其职”的状况难以满足水利综合业务开展的需要,尤其是在水资源保护等方面进行多层次、全方位的协同操作,需要通过行政手段协调,极大地降低了水利行业的工作效率。

(3)对水利系统的应用系统集成缺乏统一的、开放、标准的技术支持。在水利信息化的发展过程中,已对数据存储格式、功能模块划分、系统总体框架等方面的标准逐步完善,但在跨流域、跨部门、跨系统的系统集成实施过程中,更需要一种开放的系统集成技术,以提供全面的服务质量支持,降低系统动态集成的复杂性,并最终实现资源的全面共享和协同操作。

信息网格技术是网格技术的一个组成部分,它的提出与发展为解决与上述类似的系统集成、资源共享与协同操作问题提供了一种有效的手段。信息网格及其基础设施通过动态、分布的虚拟组织来集成地理上分散的高性能计算资源、信息资源、存储资源、控制资源、传感资源、软件和应用系统以及人力等各种资源,以支持在其之上的共享和协作。

文中将信息网格技术应用于解决水利业务的信息资源共享问题,结合水利行业的数据分散、专业跨度大的特点设计出五层开放式的水利信息网格体系结构,在已有的织女星网格的基础上,构建自己行业的水利信息网格。

3 水利信息网格总体框架

水利信息网格包括了一组相应的软件系统,向水利管理者、决策者、使用者、研究者提供透明、实时、准确、高效的信息服务,通过借鉴开放网格服务体系结构^[5](OGSA),设计了水利信息网格的总体框架,将各类计算

资源、存储资源、控制资源、传感资源、仪器设备、网络资源、程序资源等以服务的形式提供统一的表示,服务间通过消息交换实现广泛的共享和协作,图1给出了一个分层的、面向服务的水利信息网格体系结构,描述了水利信息网格软件的构成及其相互关系,这种层次化、模块化的设计思想大大地提高了系统的可理解性、可维护性、可扩展性和可复用性。

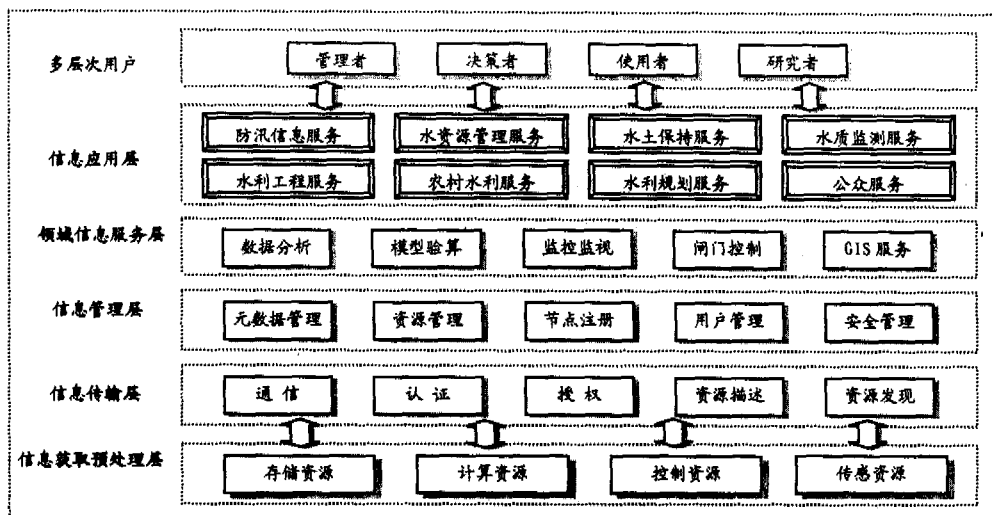


图1 水利信息网格总体框架

3.1 信息获取、预处理层

信息获取、预处理层将各类基础资源以服务的形式提供给上层系统使用,主要功能包括提供远程控制各种闸门、泵站等水利设施的能力;远程获取遥测传感设备产生的数据的能力;远程使用计算资源的能力;远程访问存储设备上数据的能力,由信息管理层对各类资源进行统一的管理、注册和发布。

3.2 信息传输层

信息传输层实现了传统网格体系结构的连接层(Connectivity)的主要功能,经资源描述后,通过核心通信及认证过程,实现资源的相互通信,向上层提供基本的信息资源发现通信及传输服务,用于网格中的通信处理。

3.3 信息管理层

基础服务层提供构建网格的一系列共性服务,例如认证和授权、资源发现、数据管理、元数据生成和管理等等。信息管理层独立于具体的领域信息服务和应用领域,为不同的网格应用提供共同的基础支撑服务,可以充分利用当前已有的一些网格中间件系统和软件开发包,比如织女星网格、Globus 工具包^[6]等。

3.4 领域信息服务层

领域服务层提供一组面向水利领域的共性服务,这些共性服务来源于各专业应用系统的核心功能,如模型演算、闸门监控、水文遥测等,更重要的是基于跨区域、地区的联机数据的数据分析、GIS服务等。领域服务层构建于信息管理层之上,将水利领域重要的业务逻辑通过封装成网格服务的形式供上层系统使用。

3.5 信息应用层

信息应用层面向众多的用户,通过网格门户(portal)向管理者、决策者、使用者、研究者等提供多方面的服务。

4 关键技术研究

4.1 水利信息网格与织女星网格的关系

织女星网格(Vega)^[7]是国家在网格领域的研究重点,国内各行业在网格应用方面基本上是基于织女星建立的。例如:国家地质调查网格(NNG)^[4]就是基于织女星建立起来的全国性地质调查系统,目前取得了初步应用成果。水利信息网格参照现有行业网格的形式,作为织女星网格的有益补充,通过对水利领域信息服务的任务化、服务组合形成综合服务流,为用户提供服务。所有的服务分布式地部署到水利信息网格中,同时在织女星网格中注册,为其他用户提供一致性服务。

4.2 元数据管理

(1)水利元数据。

元数据(Meta Data)是“关于数据的数据”,即关于数据的内容、质量、状况和其它特性的信息。水利科学数据^[8]按其数据属性分为:工程地物属性数据和水利物理量数据。地物属性数据主要包括:河流、测站、水利工程、行蓄洪区等相关物体的工程特性数据,这类数据属相对静态数据,通常不随时间变化;水利物理量数据主要包括:水文水资源、水环境、泥沙,以及科学研究、管理过程中采集或产生的数据,这类数据是自然界对水利现象作用而产生的结果,数据随时间的变化而变化,属时变数据。

为表示上述数据,解决水利元数据位置透明、命名透明、并行透明、异构透明和模式变更透明等问题,在参考关系数据库数据抽象层次、织女星网格资源空间模型的基础上,构建出水利信息网格元数据层次化结构模型,如图 2 所示。

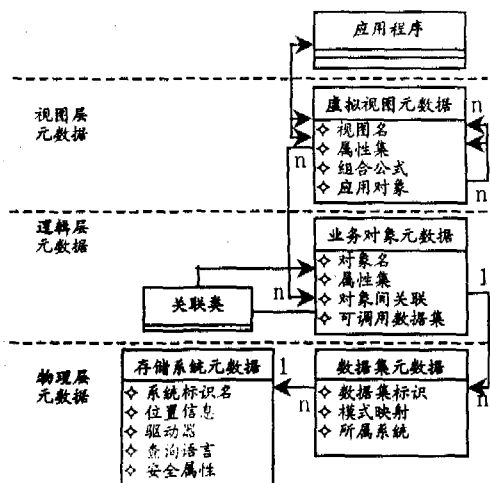


图 2 水利信息网格元数据层次化结构模型

物理层描述信息网格中存储系统和数据集;逻辑层对业务对象进行描述;视图层的元数据描述虚拟视图外部特征(如该虚拟视图的使用者、存在领域等信息)和组合公式

(与之有引用关系的业务对象或其他虚拟视图及其之间的关系表达式)的视图元数据。其中逻辑层和视图元数据是与特定应用领域相关的^[9]。

(2)元数据目录服务。

水利元数据目录服务 MDS^[9](Metadata Directory Service)主要使用 LDAP 作为元数据信息存储和访问的统一界面,包括元数据生成、分布、存储、搜索、查询和表示等功能,如图 3 所示。其主要模块包括本地元数据目录服务 Local MDS(Metadata Directory Service)、全局元数据目录缓存 Global MDC(Metadata Directory Cache)和底层目录管理协议 LDAP(Lightweight Directory Access Protocol)。

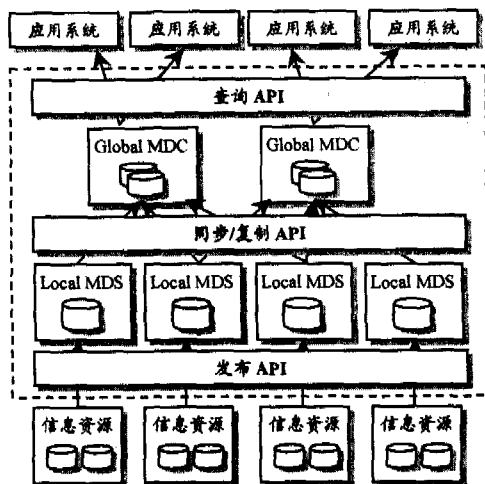


图 3 MDS 结构模型示意图

4.3 资源交易服务

在水利信息网格中,大量的原始数据都存放在各级不同的水利部门,跨流域、跨省市的数据交换是水利信息网格的重要保障机制。当用户频繁地使用异地水利元数据时,其中部分资源需要付费,需要建立资源交易管理机制,规范资源交易的行为,制定合理的价格,使数据使用频率较高的所有者得到相应的报酬。因此,水利信息网格资源交易服务必须建立资源交易服务流程^[4],及时、正确地记录用户和资源提供者之间的交易行为。资源交易服务流程如图 4 所示。

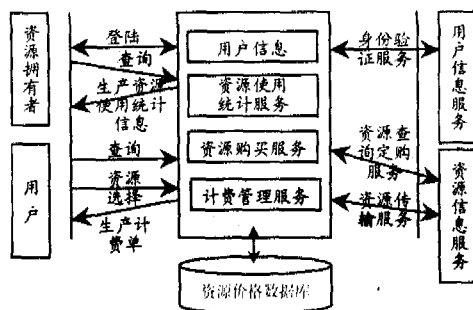


图 4 资源交易服务流程

5 结束语

目前的工作主要是在已有的织女星网格的基础上,设
(下转第 25 页)

是开放的程序系统,基本的存储结构是 Intel 的 IPL 库,用户界面层代码和算法应用代码为高层次部分,主要由 C++ 编写;底层的数据处理和计算函数,封装为库函数,主要由 C 语言编写。

系统要求处理具有一定的实时性,传统方法无法满足系统实时性的要求。OpenCV 提供了新型的图像数据结构和处理函数接口,具有很高的处理速度和效率。

实验硬件部分采用 10MOONS SDK-2000 采集卡,CCD 摄像头。整个系统在 WindowsXP 及 Microsoft Visual C++ 6.0 环境下构建。

3 实验结果与分析

图 4 中图(a)是待分割的原图,图(b)是其直方图,图(c)是人为给定的一个阈值 130 进行分割的结果,图(d)是 Ostu 算法进行分割的结果,图(e)是改进算法进行分割的结果。

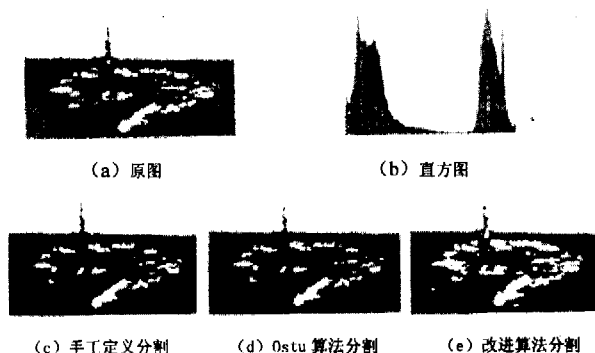


图 4 实验结果

由结果可以看出,改进算法在边缘的分割效果较好,

能较好地分割出目标物体。

4 结论

基于视频图像的监控系统是一个很热门的专题。这种技术已经广泛运用到了智能交通的各个领域,如视频车流检测、视频车牌识别等。文中介绍了该技术在桥梁安防方面的应用,并重点介绍了背景提取算法和图像分割的 Ostu 算法。在系统设计上采用了 Intel 公司提供的 OpenCV 开发包,保证了程序的高效和稳定性。这些计算方法具有明显的优势,能得到较合理的结果。但由于江面视场较广,环境复杂,加之受天气、光线(特别是夜间)等因素的干扰,故不能得到很好的效果。这也是目前这个领域的重点和难点,也是笔者目前研究的方向。

参考文献:

- [1] 吴祖林,沈庆宏,都思丹,等.背景提取基础上运动车辆视频检测[J].交通与计算机,2003(6):18-20.
- [2] 方敏,徐俊艳,王建平,等.一种新的文本图像二值化方法[J].合肥工业大学学报,2001(2):167-169.
- [3] Leung H, Young A. Small target detection in clutter using recursive nonlinear prediction[J]. IEEE Trans on Aerospace and Electronic Systems, 2000, 36(2): 713-718.
- [4] Meier T, Ngan K N. Automatic segmentation of moving objects for video object plane generation[J]. Circuits and Systems for Video Technology, 1998, 8(5): 525-538.
- [5] Intel Corporation. Open Source Computer Vision Reference Manual [EB/OL]. <http://www.intel.com/research/mrl/research/OpenCV/>, 2000-12-08.

(上接第 22 页)

计出五层开放式的水利信息网格体系结构,对水利信息网格与织女星网格的关系、元数据表示及管理、资源交易服务等关键技术进行重点研究,设计了元数据目录服务、资源交易服务模型,构建出水利行业信息网格框架。下一步研究的重点主要集中在对资源管理、安全管理等信息管理层的核心技术的深入研究,并对领域信息服务、综合信息服务等方面提出切实可行的集成方法。

参考文献:

- [1] 郭权,王希诚,李纯莲.药物分子对接中应用网格的研究与进展[J].计算机研究与发展,2004,41(12):2054-2059.
- [2] 谭国真,李程旭,刘浩,等.交通网格的研究与应用[J].计算机研究与发展,2004,41(12):2066-2072.
- [3] 张英朝,张维明,肖卫东,等.基于网格技术的电子政务平台体系结构[J].计算机应用,2002,22(12):28-30.
- [4] 唐宇,何凯涛,肖依,等.国家地质调查应用网格体系

及关键技术研究[J].计算机研究与发展,2003,40(12):1682-1688.

- [5] Foster I, Kesselman C. The physiology of the grid: An open grid services architecture for distributed systems integration, open grid service infrastructure WG, Global Grid Forum, Tech Rep: Draft 2.9, 2002 [EB/OL]. <http://www.gridforum.org/ogsi-wg/drafts/ogsa-draft2.9-2002-06-22.pdf>, 2002.
- [6] Raman V, Narang I, Crone C, et al. Services for data access and processing on grids Argonne: Global Grid Forum, 2003 [EB/OL]. <http://www.gridforum.org/>, 2003.
- [7] 徐志伟,李伟.织女星网格的体系结构研究[J].计算机研究与发展,2002,39(8):923-929.
- [8] 朱星明,张行南,白婧怡,等.水利科学数据共享元数据理论的应用探讨[J].水利学报,2005,36(8):946-949.
- [9] 廖华明,程伯羽,刘新周,等.信息网格中元数据层次化结构模型的研究和应用[J].计算机研究与发展,2003,40(12):1694-1699.