

文章编号: 1007-4619 (2004)01-0089-08

数字地球网格计算雏议

薛勇, 王剑秦, 郭华东

(中国科学院 遥感应用研究所 遥感信息科学重点实验室, 北京 100101)

摘要: 数字地球将为人提供关于我们地球的海量自然和人文数据与信息, 是我们生活的行星的一个多分辨率, 四维虚拟表达。网格计算被认为是解决数字地球问题的最好方法。数字地球问题的解决必须通过异构的计算资源, 信息系统, 设备, 人之间的相互协作, 而这些都是地理位置或组织结构分散的。本文介绍我们在生成用于解决数字地球问题的核心中间件的研究工作以及结果。因为网格计算本身是一门比较新的领域, 网格计算与数字地球的有机结合将为数字地球提供一个全新的计算工具。

关键词: 数字地球; 网格计算; 虚拟组织; 空间信息网格; 远程通讯地学处理

中图分类号: TP393 **文献标识码:** A

1 数字地球

地球模型包括地图和描述地表整治过程的数学模型。在所有以往的工作中, 包括精确地模拟和粗略地模拟地表或近地表特征, 数字地球是一个新的描述地球的方向。它不同于以往的方法主要表现在以下两个方面:^[1]

- 它是全球性的, 将地球的所有信息归于一个单一的计算机环境中;
- 它是数字化的, 所有的信息均以数字形式表示。

以下是构筑解决数字地球问题的基础的几个技术和想法^[1]。首先, 沉浸式环境提供了一个信息存储和学习者之间进行交流的非常丰富的形式, 其没有因为简单媒介的约束而受到阻碍, 而且没有因为可视化渠道或是狭义的地图概念的限制而受到制约。其次, 视觉将多种类型的数据合成在一起, 通过将数据信息译成它们真正所代表的外形形式以便于交流, 例如地形图和地面覆盖物数据, 而别的数据不得不用符号来表示, 例如人口数量, 健康, 环境质量的信息。地图制作者熟悉混合这两种类型的问题, 凭他们的经验用符号来强化正射投影图片。数

字地球更为根本的体现了一个新颖的数字信息组织形式和用户界面的构建方法。数字地球通过用户界面将整个地球呈现在办公室中。从这种意义上讲, 它通过地理位置或地理图书馆而有力地推动了信息查询。

数字地球提出的研究问题涉及许多学科。Goodchild^[1]概括如下:

- 暗含的规模范围至少超过四个等级的序列, 从恰当地表现整个地球的 10km 的分辨率到用来表现邻近事物所需的 1m 分辨率。
- 将来数字地球将是用户为中心的统一的系统, 而以往传统的几乎所有的地图制作者都持有独立于用户的观点。
- 数字地球中的模糊信息的问题, 由于在支持协同工作服务中重要的作用, 必须被解决。
- 如以上所提到的, 数字地球必须应用符号和图标来表示。
- 数字地球环境应该了解和具有访问给定地方的信息的权限和入口。这就引发了在数字图书馆、票据交换所、和万维网方面的关于信息查询和发现的许多有趣的技术问题; 关于质量保险和信誉的社会公共机构的问题; 知识产权和保密性的社会问题。
- 虽然数字地球环境是一个沉浸式的虚拟的环

收稿日期: 2002-09-16; 修订日期: 2003-02-26

基金项目: 中国科学院“百人计划”项目资助, 中国科学院知识创新工程“数字地球基础理论问题研究”项目资助(KZCX2-312)。

作者简介: 薛勇(1965—), 男, 北京大学学士、硕士, 英国 Dundee 大学博士, 现在中科院遥感所任研究员, 博士生导师, 英国皇家特许物理学家, 英国皇家物理学会专业会员, 英国遥感与摄影测量协会专业成员, IEEE 会员。主要研究方向为远程通讯地学处理(Telegeoprocessing), 数字地球、遥感、GIS 及图像处理。已发表论文 40 余篇。Email : y. xue @unl. ac uk

境,但是数字地球的原理可以同样很好地用在传统的结构上,即由用户,键盘,屏幕显示器,和按键设备所构成的结构。

- 数字地球指的不仅仅是以往的数据。
- 一定的标准(包括表示全球地理数据的一般形式)的发展有益于数字地球的发展。

• 数字地球需要新的技术来克服通讯带宽的限制。数字地球需要一个一致的数据结构和检索方案以支持超过四个量级的缩放。最理想的演示方案不可能在模拟地球表面时也是最理想的,也就需要一定的处理和折衷。

• 数字地球所需的制图技术尚未可得和实用。虽然卫星影像覆盖了大面积的地球,但是地形的信息在数量等级和可用性上变化很大。这样就需要强大的全球空间数据基础设施的发展和需要合适的组织机构来协调它。

2 网格计算

网格计算作为一个重要的新兴领域出现,不同于以往传统的分布式计算,其重点在于大规模的资源共享、创造性的应用和在某些情况下高性能计算的应用^[2], Foster 和 Kesselman 将网格计算问题定义为在动态个人集合、机构,和资源(他们指虚拟的组织)之间的灵活的安全的协作的资源共享。“网格计算”的术语是在 20 世纪 90 年代的中期提出的用于高级科学和工程方面来表示分布式计算基础结构的^[3]。在基础结构的构造方面已经取得了比较大的发展(例如文献[4]—[7]),但是网格的概念已经被合并了,至少以一种流行的说法来说,包含了从先进的网络到人工智能方面的所有东西。

网格计算所涉及到的用户群体包括从事实时再现和快速仿真赋值的计算机科学家和工程师,监测臭氧损失、气候变化、污染以及偶合的模型,知识数据库的环境学家,还包括致力于远程设备+超级计算机先进的可视化方面的实验科学家。同时包括分布式资源(CPU,数据,人)的协作,全球性的企业,虚拟环境,计算机辅助设计公司和虚拟授课空间一分布式的用于培训和教育的课堂^[8]。

网格计算通过认证,授权,协商和制定安全协议能提供一种访问全球性的分布式计算环境途径。网格计算通过资源发现来审查信息服务和通过监测工具,能够查询目前的状态以便能够鉴别和分配信息源,并进行战略决策,例如通过流程传送数据或应用

程序以及相互分配资源。而且,网格计算进度任务和问题的分解使得在远端的机器上可以执行应用程序代码,可以传递数据或复制目录和更新目录,一旦问题出现可以解决问题,可以监测问题的执行和解决,找回和分析结果等^[9]。

现在有许多可用的工具和工具包,例如 Globus, Condor, Legion, SRB, LDAP, OOFs 等等^[10]。Globus 提供资源分配和管理(GRAM)、信息存取(MDS)和鉴定(CSI)等功能。Condor 通过具有检查站的分布式网络工作站提供巨大的计算吞吐量。Legion 是一个基于对象的大规模的分布式计算环境,它的设计是用来处理百亿数量级的对象的。资源存储代理(Storage Resource Broker)提供了管理一个包括多重拷贝的分布式仓库的便利工具。

网格计算的应用领域从环境科学(有关的高分辨率的大气和海洋模拟),生物学,天文学-虚拟的天文台到材料科学。这些应用涉及到计算机建模,分散的群体和多学科仿真的数据分析,在线设备访问和实时分析,共享数据档案(EO 数据和染色体组数据)和协作性的可视化和分析。这就引发了下述问题:网格计算是协作性,但是又面临着区域分散的世界范围的研究群体、从桌上型电脑到超级计算机的异构的计算机资源,多学科、有关的仿真模型/代码、数据文档,对大型数据源远程存储方法和对在线设备结果的数据返回的问题^[8]。

许多学科和工程分析需要网格计算^[6]。许多异构的计算和数据资源需要统一的管理。现有的各种仿真和数据分析组件需要互联在一起通过分布式资源同步协同地工作。计算机和数据分析工具的界面必须为不同学科问题的解决者提供一个应用标准,在不同的地理位置,涉及到许多计算和密集数据型步骤的工艺流程必须在分布式资源间进行安全和可靠的管理。自动纠错和恢复管理工具在应用程序和基础结构方面都必不可少。大的复杂的数据文档——例如几何学(结构)和飞机和涡轮机的运转——都需要有专门的学科专家在不同的地点进行维护,而且必须由合作的分析家进行存取和更新。庞大的数据设备是分布式的,必须进行有效的管理,这样在世界各地的科学家就可以进行共享和访问。从工具系统所得到的数据流必须通过计算机的数据分析系统进行实时的分析和管理的,科学家和工程师必须能够安全地和有选择性地共享他们工作过程中所需的任何数据。应当有一个单一的安全机制,一旦被用户所激发,他就可获得访问解决一个问题所需

资源的所有权限。

对于工具开发者—应用领域的计算机科学家—来说提出了另外的要求,他们起初只是进行代码的编写和访问以及管理执行环境。例如:

- 跨越所有系统的统一的大批量的队列管理。
- 使队列等待时间最小的工具。
- 计算资源共享管理工具,包括通过共享分配管理。
- 确保大量的有关的或相互依赖的工作以一个整体可靠地运行的工具。
- 使等待时间最少的文件定位管理工具。
- 描述特征和存储的编目机制。
- 独立于物理存储位置的数据体系数据设备。

技术方面的挑战包括跨越多国家和多管理领域方面的,例如安全性,访问权限政策,费用的付还和无中心控制;资源的稳定性,如特征的时空变化;复杂的分布式应用程序,例如相互分配任务,高级预定和优化;确保从终端到终端的执行,如异构性、容错性、隐藏的复杂性的问题。

网格计算技术是现实的而且在目前是非常有用的。基础的网格计算服务用来支持为计算、数据、和工具系统提供统一和安全的入口,对于地理位置和组织形式分散的计算、数据源、工作管理以及安全性提供一个统一的权限,安全性包括单一的签约(用户一旦签订则具有访问所有授权资源的权限)和安全的内部通信和网格计算管理。网格计算的执行管理工具(例如 Condor-G)目前正在被使用。提供对三重存储系统和全球元数据目录(例如 GridFTP 和 SRB/MCAT)统一的访问入口的数据服务目前也被投入应用。支持应用框架和科学入口的网络服务也被做成原型。着眼于长远的基础设施也正在建设。基于计算和数据系统以网格计算的原型产品为形式的网格计算服务也得到了维持和继续。PCI(公钥基础设施)提供支持简单签约的加密证明。资源发现服务正在维持和继续(网格计算信息服务——分布式目录服务)。这些都已经出现了,例如美国宇航局(NASA)的信息能源网格计算(IPG),能源部(DOE)的科学网格计算,欧共体(EU)的数据网格计算,英国的 eScience 网格计算,国家科学基金会(NSF)的全国工业生产咨询委员会和美国超级计算应用研究中心(NASA)。

2.1 信息能源网格计算

NASA 所涉及的大型科学和工程问题需要利用

许多计算机和数据资源来解决,包括大型计算机和大型数据存贮系统,这些都必须将那些由不同部门的研究人员开发的应用程序和从不同设备获取的位于不同的地理位置上的数据集成到一起。信息能源网格计算(IPG)是 NASA 全力推出的一种持久、安全、鲁棒性强的网格计算^[8]。该网格提供一种统一的用于动态地、大规模地解决分布式和异构资源问题的环境。

2.2 数据网格计算

数据网格(Data Grid)项目将开发、实现和利用一个基于大型数据和 CPU 导向的计算机网络计算。这将允许那些由三个学科得到的分布式数据和 CPU 集约型科学计算模型在一个地理上分散的试验台进行验证。DataGrid 项目中包括四个主要部分^[9],分别是:利用基于 GRID 的基本资源(计算结构,大容量存贮,网络),开发通用中间件(安全性,信息服务,资源分配,文件复制),建立应用服务(工作安排,资源管理,过程控制),以及对 3 类科学应用的测试(粒子物理学(LHC),地球观测和生物科学)。定义 Grid 中间件的各部分为 WP1—WP11 工作包^[11]。

3 空间信息网格计算

许多人造卫星正在日夜不停地收集着我们这个星系信息,特别是地球的有关数据。卫星数据有许多用途,例如用于无线电通讯、航空和航海系统以及环境监测等。然而,即使应用功能最强大的计算机来处理这些数据,仍是相当耗时的,代价昂贵。如果把这些数据处理任务分配到许多低成本的由网络连接的处理平台上,那么将会以较低的代价为许多空间应用提供巨大潜力。空间信息网格计算(SIG)是一种基础服务设施,用于空间信息收集以及为信息综合管理和处理提供信息共享。SIG 提供一种技术体系,用于综合的空间数据收集、数据处理和应用,以及提供一种智能空间数据处理平台和基本环境。

网格计算在地理学方面的应用指的是结合多重的数据源(例如空间,大气,和地面)、模型(例如预报和评估)……,从中提取有用的信息和知识(科学工作者),以便于为环境监测和自然资源管理(评价增加的合理的用户)提供服务。在欧共体数据网格计算项目中地球观测科学应用程序组件 WP9,主要用于地球观测的这个工作包定义和发展了地球观测(EO)专用的组件来与数据网格计算平台接口,同时

将网格感知应用的概念引入地学环境中。这就为需要大的计算机资源和需要对地理上分散的大型数据文件和文档访问的地球观测科学应用提供了一个很好的机遇。本文在下面的内容里,设计了可用于支持数字地球的空间信息网格的原型,基于这个空间信息网格,使数字地球实现成为可能,对空间信息网格支撑下的数字地球作了详细的叙述。

4 远程通讯地学处理

远程通讯地学处理是一门建立在实时空间数据库基础之上的新兴学科,为进行决策支持和实时控制,空间数据库必须通过远程通讯系统定期更新^[12]。远程通信需求的用于远程通信地学处理的多媒体系统主要依据所从事的远程通信地学处理的类型。远程通信地学处理的类型包括环境的远程诊断和环境远程咨询。环境远程咨询可以相互共享图像、图形和地理信息,通过这些信息科学家对现场可作出初步的诊断。远程咨询的目的主要是利用远程专家提供的一个第二手的参考,帮助证实一下现场专家的诊断结果或帮助现场专家达成一个正确的共识。在这种情况下,可以进行有声或无声面对面的会议,包括同步全双工视频和音频的视频会议就显得尤为重要。既然视频仅用于开会,那样的话么其图像的质量就要求不太严格;然而声音必须清晰,没有间断而且没有时间延迟。虽然带有某些瑕疵的图形图像也可以在远程咨询中应用,但还是需要质量比较好的图形和图像以便作出精确的判断。

环境远程诊断可以共享图像、图形、和地理信息,通过这些信息距离现场较远的科学家可以得出一个初步的诊断结果。

远程通信地学处理是应用网格计算技术的最好的例子。远程通信地学处理所涉及到的技术和工作包括:遥感数据的实时处理和成像,实时空间数据结构,实时 GIS 检索,通用性,地理数据交换标准,平行式和分布式 GIS, GIS 和 Internet 工作,组件和 GIS 计算机支持协同工作(CSCW), GIS 的客户-服务器结构,超图 and 多媒体通讯及 GIS 应用,无线 GIS 映射(制图),在任何时间和地点地球观测数据和 GIS 的实时接口。

5 数字地球网格计算

5.1 网格技术是数字地球的支撑

数字地球是 21 世纪的信息领域和地球科学的

重要发展领域。当 Al Gore 强调数字地球能够成为开放的“无墙实验室”时,我们应将其概念扩展为用于理解人类同我们星球上周边环境间的复杂关系^[13]。数字地球必须通过人、异构计算机资源以及仪器间的相互联系才能够得以实现,而这些资源在地理上、组织上都是分散的。数字地球包括多源多比例尺多分辨率的、历史和现时的、矢量和栅格格式的数据。数据的组织是无缝的。不同的用户对不同的数据和信息具有不同的使用权限。“数字地球”采用开放平台、构件技术、动态互操作等最先进的技术方案。

数字地球实现资源的高度共享,这种共享不是简单的文件交换,而是大范围的协作性地资源的直接访问。其具有易于高度控制的特点,也就是说应该清晰地划分资源的提供者和消费者,而且要明确共享的内容、允许共享的用户、在哪种情况下可以进行资源的共享。数字地球所涉及的设备和参与者具有这些特点,决定其形成一个典型的虚拟组织(Virtual Organizations-VOs)。

目前的分布式技术不能满足数字地球技术的要求。例如目前对于用于计算的多个站点,Internet 技术没有提供一个完整的平等的使用方法^[3]。商业到商业的技术只能通过中心服务器实现信息的共享。企业分布式计算技术(例如 CORBA)只能在简单的小组织中实现共享。开放组的分布式环境虽然可以实现数据的安全共享,但其显得累赘不够灵活。存储服务提供者(SSP)和应用服务提供者(ASP)只能通过有限的方式进行数据的存储和计算。总而言之,以上这些技术既不能提供大范围的资源,又不能对数字地球所要求的共享关系进行灵活地控制。为了构建“数字地球”,需要技术支持,如计算科学、海量数据存贮技术、高分辨率的卫星图像、宽带网络、互操作技术、元数据标准等,保证建立在不同计算机平台的、分布式的、异构的、不同数据源的数据能在网络环境中共享。而这些特点只有通过网格技术来实现。

网格技术通过协议、服务和必要的工具使得数字地球这种大范围的、灵活的可控的资源共享成为可能。

5.2 基于网格技术的数字地球原型

“数字地球网格计算”的主要目的是使这些常规的资源形成一个无缝集成的协同式计算环境。针对于数字地球技术的要求,互操作性当然成为网络要

处理的中心问题。在网络环境下,互操作性意味着通用的协议。通过协议可以定义基本的机制,参与数字地球的用户和资源可以以此基本的机制进行协商、建立、管理和利用共享的关系。标准的协议使得定义功能强大的服务功能变得比较容易。协议之上

需要开发一系列服务。APIs 和 SDK 可为数字地球应用程序的开发者提供便利的条件,加速了代码的开发。

空间信息网格支撑的数字地球原型具有以下层次特点(图 1)。各层的结构功能如下(参考 Ian Foster)

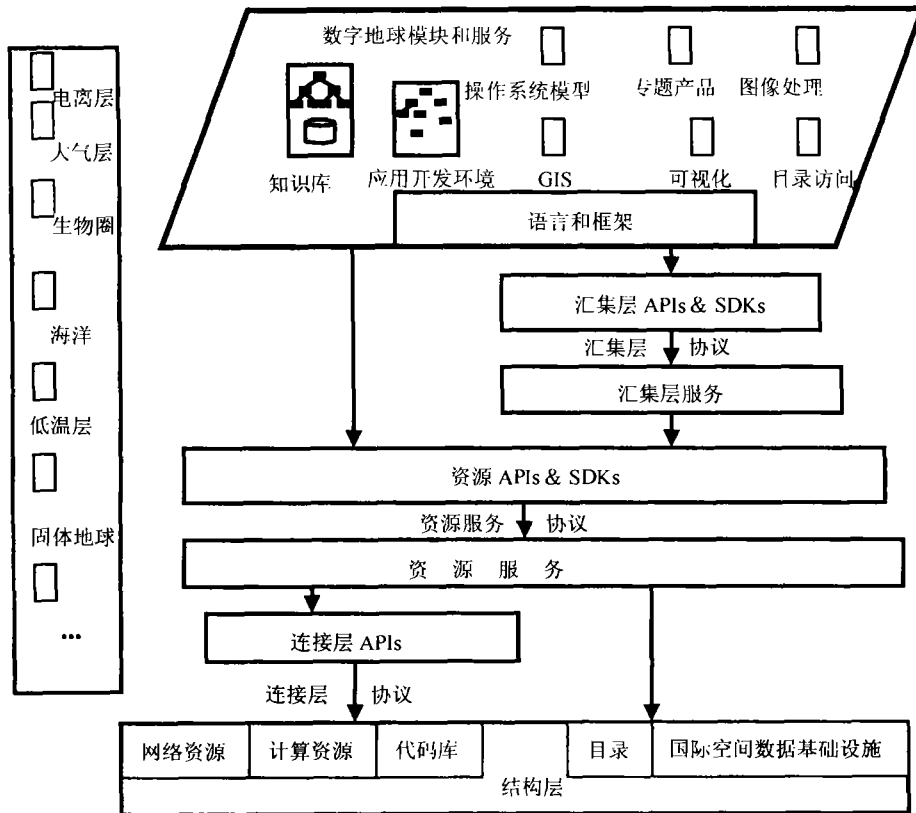


图 1 空间信息网格支撑的数字地球应用框架(由[15]改写)

Fig.1 Spatial information grid supported digital earth application architecture (revised from[15])

(1) 构造层(Fabric):它的功能是向上提供网格中可供共享的资源,它们是物理或逻辑实体。常用的资源包括计算资源、目录、网络资源、代码库和国际空间数据基础设施。比较典型的空间数据基础设施(SDI)内涵包括四部分^[14]:1)数据交互网络体系,包括作为网络节点的各基础地理信息系统和空间数据通讯网络;2)基础数据集,包括空间定位控制数据、地形框架数据、土地利用覆盖数据、地籍测绘数据、航摄与遥感影像数据以及其他与空间位置相关的基本自然人文信息;3)法规与标准,包括有关信息共享机制的法规和政策,以及地理信息技术标准;4)机构体系,包括权威的空间数据和地理信息协调领导机构、从事数据获取和加工的骨干生产组织、专门从事空间数据维护更新的单位。在本文中结构层空间基础设施指的是:在一定的机构体系的支持下,依

据法规和标准通过数据交互网络体系连接的基础数据集。分布式存储海量资源和技术是实现空间数据共享的基础,而基础数据的标准化是构建数字地球的关键。在结构层中,开发相应的组件负责侦测可用的软硬件资源的特性、当前负荷、状态等信息,并将其打包供上层协议调用。

(2) 连接层(Connectivity):定义了网格中处理通信与授权控制的核心协议。构造层提交的各种资源间的数据交换都在这一层的控制下实现。各资源间的授权认证、安全控制也在这里实现。目前的通信协议我们可以使用已经存在的网络通信协议。在本层中应该开发基于公钥协议的组件,提供通信协议,服务发现(DNS),授权、认证等功能。

(3) 资源层(Resource):本层的的功能主要是对单个资源实施控制,与可用资源进行安全握手、对资

源做初始化、监测资源运行状况、统计与付费有关的资源使用数据。资源层应用这些协议调用结构层功能来对本地资源进行访问和控制。在本层中用来实现资源注册、资源分配和资源监视组件应该遵守资源信息协议、基于 HTTP 的网格存取和管理协议和 GridFTP 协议。

(4) 汇集层 (Collective): 这层的作用是将资源层提交的受控资源汇集在一起, 供数字地球虚拟组织的应用程序共享、调用。为了对来自应用的共享进行管理, 汇集层提供目录服务、资源分配、日程安排、资源代理、资源监测诊断、网格启动、负荷控制、账户管理等多种功能。目录服务允许数字地球虚拟组织的参与者能够利用姓名、特征和路径查询数字地球虚拟组织和其他通过空间信息网格连接的虚拟组织的资源。当然这个功能还需调用资源层相应的相应功能的 APIs 和 SDKs。相互定位、日程安排、代理服务允许数字地球的参与者查询资源的位置, 依据合适的资源对任务进行安排日程。监测诊断服务用于监测数字地球虚拟组织中资源受到敌意攻击过载。数据复制服务负责管理数字地球虚拟组织存储、网络和计算资源, 以使其稳定地、快速地、最大限度地为数字地球虚拟组织的参与者服务。

(5) 应用层 (Applications): 数字地球虚拟组织提供的中间件和服务在应用层, 包括知识库、应用开发环境、GIS 服务、可视化、主题产品、目录访问、图像处理 and 操作系统模型。其必须以下述技术作为支撑:

1) 高性能的科学计算-数字地球的海量数据计算功能突破了基础研究和理论研究的限制。本文提出的空间信息网格在结构层提供了海量可共享的计算资源, 而且其他相应的层提供了资源的查询和定位、任务的分配功能, 数字地球成员通过授权和认证, 确认用户身份后进行按照网格协议安全地使用计算资源, 以为数字地球提供高性能的科学计算。

2) 海量数据存储和更新技术, 由于空间信息网格通过协议很好解决了互操作性, 接入空间信息网格的空间数据服务基础设施为数字地球海量数据的需求奠定了比较好的基础。使得数字地球成员可以对等地、安全地、对海量数据进行存取。

3) 网络与互操作技术, 数字地球需要通过高速网络连接和浏览数据库群中的数据库, 互操作性是实现多智能体计算的核心技术。这与本文所提出的空间信息网格的基本特点相吻合。作为数字地球虚拟组织成员应用程序通过各层的 API 调用相应的服

务, 再通过服务调用网格上的资源来完成。应用程序的开发涉及大量库函数。为便于网格应用程序的开发, 需要构建支持网格计算的库函数。

基于空间信息网格技术的数字地球构建的中间件主要集中在以上应用层的三个关键支撑技术方面。GIS 既服务于数字地球, 又是中间件的工具。应用其开发环境, 用户可以发展自己的界面、处理模型等。图 2 是知识仓库的框图。数字地球参考模型^[16]是基于 ISO/IEC10746 的开放分布式处理参考模型。数字地球参考模型提供的技术框架可以引导数字地球的开发、信息的获取、标准议程的制定, 交换数字地球设计信息, 定义数字地球元系统 (meta-system) 的内容和功能, 详述关键数据及标准, 并总结了工程和技术方面的选择方案。

数字地球参考模型 (DERM) 不同于全球空间数

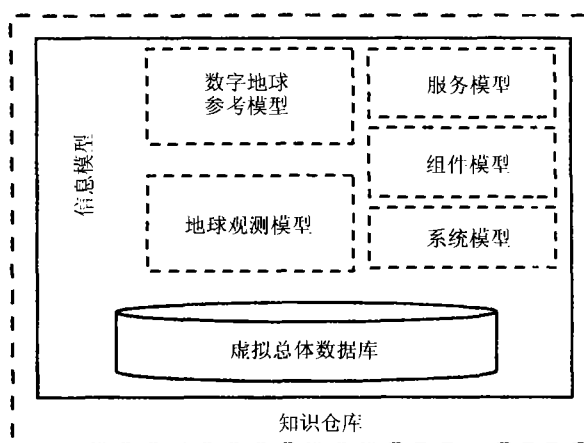


图 2 数字地球知识仓库系统

Fig.2 Digital earth repository system

据基础设施技术委员会 (GSDI) 的技术参考模型。二者相近但并不相同。开放式地理信息系统协会 DERM 委员会主要致力于无缝联络、交叉和会聚。DERM 被更狭义地定义为支持早期数字地球示范。其尽可能地应用 FGDC、OGC、ISO 等来研究数字地球的一个方面。此项技术将成为一全球性标准。如果正常的话, 全球各个国家将应用此模型, 用以理解需要什么标准, 它们为什么重要, 怎样应用。太空对地观测技术能够提供对地球表面的全球性的、重复性的和连续的表面覆盖, 所获得的数据用于对地理解。地球是一涉及地学各个分支学科的复杂系统, 空间对地观测与复杂的计算机模拟在对地观测中起着不可替代的作用, 是建立数字地球的关键性技术。郭华东^[17]列举了在没有空间信息网格技术支撑下的数字

地球知识仓库,全球对地观测系统应由三部分组成:信息获取、信息处理与信息应用。信息获取是指利用各种航天、航空观测系统和地面信息获取测量系统的三层对地观测系统(陆地-空中-太空)对陆地、海洋及气象进行观测以获取信息。信息处理指快速处理各种传感器的数据,并能对信息应用进行快速传输;数据应用的目的是寻求可持续发展的战略、人类资源与环境,发展具有自动识别和智能分析的能力应用技术,并希望此应用领域能够工业化。

在空间信息网格技术支持下数字地球信息获取功能在结构层中为空间数据基础设施提供丰富的空间数据集。信息处理指应用开发的中间件所提供的数字地球与空间信息网格基础平台的接口,利用分布式大规模的安全的计算资源对来源于国内航空航天传感器的数据、国外先进的对地观测卫星的数据进行快速处理。

此系统是模块化的,其中一些个别的组件允许被其它组件所简单替换。

6 结论与讨论

信息社会建立在快速地聚敛计算机资源、海量媒体数据、远程通信信息的基础之上。移动通信,互联网以及它们的相互互联构成了影响将来信息服务和应用的全球性主动力。数字地球概念成为一个全球性的相互协作的一个项目,而不仅仅是一个国家的事情。网格计算技术的目的是使常规所有资源的相互联系与共享更加方便,这样可以支持广泛的分布式的各种特点的科学和工程。但是目前网格计算技术还不够成熟和可靠,还达不到作为一个普遍的解决问题的方法来轻松易懂地使用,也没有经过长时间的运作和经济上的验证^[10]。目前还有许多方面缺陷需要完善:知识框架,科学入口/解决问题的环境, workflow 管理,协作框架,全球文件系统,应用的组合和动态执行,监测/全球事件,授权,支持复杂的分布式的应用软件的可执行环境的动态构造,网格计算同现有的商业框架(例如微软的 DCOM 或 IBM 的 MQ)的接口界面^[11],除此之外,网格计算并不是基于网格计算技术免费的大型资源站。网格计算对于联系紧密的计算(例如 CFD 计算)并不见得有效^[8]。

迄今为止,数字地球问题最好的解决方法是利用网格技术。数字地球必须进行人、异构的计算资源、信息系统、设备的参与,而这些都是地理上或组

织上分散的。目前我们借助于 GIS 技术开发空间信息处理和应用的中间件,这些中间件将直接用于数字地球的地理计算,结果将于近斯完成。

参考文献(References)

- [1] Goodchild M F. Implementing Digital Earth: a research agenda, In Proceedings of the International Symposium on Digital Earth-Towards Digital Earth[C], 1999, Science Press, Beijing.
- [2] Foster I, Kesselman C and Tuecke S. The anatomy of the Grid[J]. *Intl J. Supercomputer Applications*, 2001
- [3] Foster I and Kesselman C (eds.). The Grid: Blueprint for a New Computing Infrastructure[M]. Morgan Kaufmann, 1999.
- [4] Beiriger J, Johnson W, Bivens H, Humphreys S and Rhea R. Constructing the ASCI Grid[C]. In Proc. 9th IEEE Symposium on High Performance Distributed Computing, 2000, IEEE press.
- [5] Brunett S, Czajkowski K, Fitzgerald S, Foster I, Johnson A, Kesselman C, Leigh J and Tuecke S. Application Experiences with the Globus Toolkit[C]. In Proc. 7th IEEE Symp. on High Performance Distributed Computing, 1998, IEEE Press, 81-89.
- [6] Johnston W E, Vaziri A, Hinke T, Lissota T. The Information Power Grid-The Information Power Grid[R], <http://www.ipg.nasa.gov>.
- [7] Stevens R, Woodward P, DeFanti T and Catlett C. From the I-WAY to the National Technology Grid[J]. *Communications of the ACM*, 1997, 40 (11): 50-61.
- [8] Jeffreys P. Why we need the 'Grid'-the successor to the WWW[R]. <http://www.escience.clrc.ac.uk/>
- [9] Boyd D. The Grid -an overview and some current activities[R]. <http://www.escience.clrc.ac.uk/>
- [10] Boyd D. Introduction to Grids [R], <http://www.escience.clrc.ac.uk/>
- [11] Information society technologies (IST) programme-DATAGRID [R]. <http://www.datagrid.cnr.it/>.
- [12] Xue Y, Cracknell A P, and Guo H D. Telegeoprocessing: The Integration of Remote Sensing, Geographic Information System (GIS), Global Positioning System (GPS) and Telecommunication[J]. *International Journal of Remote Sensing*, 2002, 23: 1851-1893.
- [13] van Gendern J L. The potential of virtual reality imagery to the Digital Earth concept[C]. In Proceedings of the International Symposium on Digital Earth-Towards Digital Earth, 1999, Science Press, Beijing.
- [14] Yang Kai, Accelerating the construction of Chinese national spatial data infrastructure[R], <http://www.nsd.gov.cn/asa/YK.html>. [杨凯谈如何迎接“数字地球”挑战.]
- [15] Fusco L. GRID: Earth and Space Science Applications Perspective [C]. In Proceedings of Space GRID KO meeting to be held in ESTEC on 26 September 2001, <http://www.esa.int/>.
- [16] McKee L. Digital Earth Reference Model[R], www.cacr.caltech.edu/Publications/DVC/
- [17] Guo H D. Building up an Earth Observing system for Digital Earth [C]. In Proceedings of the International Symposium on Digital Earth-Towards Digital Earth, 1999, Science Press, Beijing.

Digital Earth GRID

XUE Yong, WANG Jian-qin, GUO Hua-dong

(*Laboratory of Remote Sensing Information Sciences, Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101, China*)

Abstract: Digital Earth will be a multi-resolution, four-dimensional virtual representation of our planet that enables a person to explore and interact with the vast amounts of natural and cultural information gathered about the Earth. It needs Interactive 3-D visualisation, display and navigation through immersed and non-immersed environments; it also needs high-performance computation to create derived information and model simulations on demand, storage and real-time access to very large, multi-resolution datasets, fusion of satellite imagery and other geo-referenced data sources of diverse content, satellite and terrestrial broadband networks for high data rate transmission, interaction, and collaboration, standards and metadata for interoperability among and access to differing geo-spatial databases are inevitable. Digital Earth concept becomes an international cooperative program instead of merely one country's initiative.

So far, Grid technique is the best solution for Digital Earth. Digital Earth can only be done through the interaction of people, heterogeneous computing resources, information systems, and instruments, all of which are geographically and organizationally dispersed. Earth observation includes information acquisition, processing and applications. Information acquisition provides a vast amount of spatial data for building the fabric resource infrastructure. Information processing means that spatial information processing middleware is used with distributed large, secure grid computing resources for real time processing of all kinds of spatial data.

With the help of GIS, we are currently working on the development of core-middleware for Earth observation data processing and applications. The results will be available soon.

Key words: digital earth; grid; virtual organization; spatial information grid; telegeoprocessing