

网格中的信息计算技术^{*})

王意洁 卢锡城

(国防科技大学并行与分布处理国家重点实验室 长沙 410073)

Information-Based Computing Technology in Grid

WANG Yi-Jie LU Xi-Cheng

(National Key Laboratory For Parallel & Distributed Processing, NUDT, Changsha 410073)

Abstract In the view of application requirement, the concept and framework of information-based computing technology in grid are discussed. The key technologies--metadata catalog and storage resource broker are studied, and the application of information-based computing technology to biology is analyzed.

Keywords Grid, Information-based computing, Mass information, metadata, Storage resource broker

1. 引言

目前,我们正处于一个信息爆炸的时代,信息访问随处可见。核模拟、气象、地球科学、高能物理学、医学、生物学等学科的数据量已经达到TB级(10^{12} 字节),很快将达到PB级(10^{15} 字节),各学科的研究团体都是由分布在各个地方的许多科学家组成,同样,科学家们用来存储和分析数据的计算资源和存储系统在地理上也是分布的^[1]。这些学科应用的显著特点是:数据量大,用户和资源在地理上分布,数据处理的计算量大、性能要求高。

网格(Grid)^[2]是国家级高性能计算和信息服务的战略性基础设施,它的目标是将地理上分布、异构的各种高性能计算机、数据服务器、大型检索存储系统和可视化、虚拟现实系统等,通过高速互连网络连接并集成起来,共同完成一些缺乏有效研究办法的重大应用研究问题。它不仅实现了对各种计算资源的访问,而且实现了对所有数据资源的统一访问。

一方面,目前网格的研究侧重于实现对计算资源的合理高效使用,并满足用户一般的信息访问需求。另一方面,数据密集型计算的飞速发展对信息处理能力的要求越来越高,从科学研究到军事决策、战场仿真,从分子生物学到核模拟,从数字地球计划到海湾战争,人们迫切需要缩短从原始数据提取高层信息,并进一步完成高层信息处理的时间,这就需要扩充网格的数据管理能力,建立一种有效的数据管理体系结构,将超级计算机的数值计算能力和数字图书馆的数据处理及发布能力有机结合起来,为科学应用在网格中实施资源发现和发现,并进行高效的数据访问提供有力支持,使网格具有较强的信息计算(Information-based Computing)能力^[3]。

当前,在网格研究的基础上,信息计算技术的研究工作也在世界各地逐步开展起来,并得到了各国政府的高度重视。由美国DARPA投资的海量信息分析系统(MDAS)计划^[4]旨在研究如何对异构环境中的海量信息进行有效存储管理和数据处理,以及为信息的拥有者和使用者提供各种高质量的服务。其研究内容包括分布并行计算技术,信息发现技术,数字图书馆技术,数据管理技术,存储资源代理技术,数据库技术,以及

档案存储技术等。欧洲原子核研究中心(CERN)正在开展面向高能物理学、地球观测、生物信息学的信息计算技术的研究工作,推出了国际网格计划^[5],研究内容主要包括:数据访问、数据副本管理、元数据管理、数据安全、查询优化等。美国的Argonne国家实验室和南加州大学在研制开发网格系统工具Globus的基础上,积极探讨信息计算的系统框架和关键技术^[6],重点研究数据存储、元数据管理和数据副本管理等方面。

2. 信息计算的基本软件框架

从应用需求出发,网格面向广域的异构环境,因此,信息计算的设计目标^[3]归纳为:

- 命名透明性——网格中的数据集成千上万,信息计算技术应该支持基于属性的访问,允许用户根据属性查找所需的数据集。

- 定位透明性——用户给出某个特定数据集的标识,信息计算技术应该负责在分布式的异构环境中精确定位所需的数据集,并为后续处理提供支持。

- 协议透明性——网格应该支持对异构数据资源的访问,网格为用户提供统一的访问接口,选择适当的访问协议来实现用户提出的数据访问请求。

- 时间透明性——信息计算技术应该尽量提高网格中数据访问的效率,减少数据访问的时间。

信息计算技术将超级计算机的数值计算和分析能力与数字图书馆技术有机结合起来,从而为科学应用在分布式异构计算环境中实施资源发现和发现提供支持。具体而言,信息计算技术通过提供一组服务来支持资源和发现,通过存储资源代理使计算可以在异构的存储资源上进行,它的基本软件框架如图1所示。

从图1可以看出,信息计算技术提供了目录服务、注册与发布、信息发现、存储资源代理(Storage Resource Broker)、身份认证与访问控制、调度、方法执行等服务,其核心是元数据目录,它负责维护异构环境中各种系统实体的信息。目录服务提供元数据目录的统一逻辑视图,而元数据目录可以采用分

^{*} 本文得到高等学校全国优秀博士学位论文作者专项(200141)资金项目和国家自然科学基金项目(69903011,69933030)的资助。王意洁 博士,副教授,研究领域为网络计算和数据库技术。卢锡城 博士生导师,工程院院士,研究方向为MPP实现技术、分布处理技术和网络技术。

布式的实现方式。信息计算的软件框架由一组服务构成,所有的服务都使用元数据目录中存储的信息,并通过目录服务的

API 访问元数据。应用程序也可以通过目录服务的 API 访问元数据。各服务之间可以相互调用。

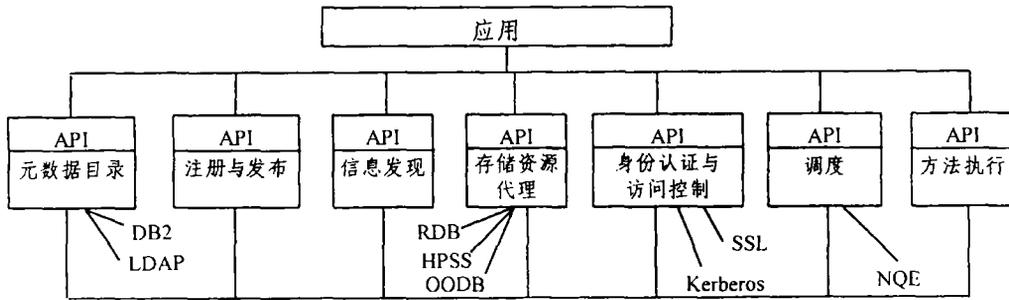


图1 网格中信息计算的基本软件框架

3. 信息计算的关键技术

信息计算在网格基本功能的基础上扩充数据管理功能,提供与信息有关的各种服务,其最核心的关键技术是元数据目录和存储资源代理。

3.1 元数据目录

为了实现命名的透明性,网格需要有效管理数量繁多的名字和属性,以及它们之间的关系;为了实现定位的透明性,网格需要有效管理数据集的定位信息;为了实现协议的透明性,网格需要有效管理数据资源的有关信息。实际上,这些信息就是用于描述资源、方法、数据集和用户的元数据。概括地说,元数据目录为用户身份认证、数据定位、访问控制、数据复制等提供支持。

元数据可以分为系统元数据、副本元数据和应用元数据。系统元数据主要是关于网格自身结构的信息,比如,网络互联情况、存储系统的容量等。副本元数据主要是关于数据副本的信息,比如,文件与具体存储系统之间的映射信息。应用元数据主要是关于与具体应用相关的文件的逻辑结构或语义的信息,比如,数据的内容和结构、获取数据的必要条件等。

元数据的管理包括元数据的命名、发布和访问,并为用户提供统一的访问接口。网格中的所有元数据构成元数据目录。目前,元数据目录的实现方式有两种:对象关系数据库(如:DB2)和轻权目录访问协议(Lightweight Directory Access Protocol,简称LDAP)。采用数据库建立元数据目录可以较好地利用数据库完善的管理机制(比如,事务管理机制)和较为强大的管理能力(比如,维护数据的一致性),之所以采用对象关系数据库是因为对象数据模型易于建模等特点,特别是元数据需要对各种实体(资源、方法、数据集和用户)之间的关系进行描述,而对象数据模型擅长描述各种关系。LDAP采用层次数据模型,支持层次式的命名结构,同时具有对象类描述能力,能够提供分布式的目录服务,但是需要先将要描述的信息进行层次映射,而且许多管理机制有待元数据目录设计者建立。目前,网格已经开始采用LDAP提供目录服务,比如,Globus系统^[7]。

随着应用的不断发展,网格也在不断发展,元数据在不断增多,其结构也日趋复杂。为了保证在网格规模不断扩大的情况下,仍然提供高效的元数据服务,元数据目录应该采用具有良好可扩展性的层次式分布式结构,这需要一套合理的管理机制。

3.2 存储资源代理

众所周知,网格中的数据位于广域范围内分布的异构存

储设备(数据库、文件系统或档案库)上,为了提高数据访问的效率,在网格范围内建立多个数据副本(Replica),从而实现数据的就近访问。存储资源代理是网格中的数据管理核心。存储资源代理(如图2)是为高层应用访问分布的异构存储资源和数据副本提供统一接口的中间件。存储资源代理^[3]利用元数据目录为用户提供面向集合(Collection-oriented)的数据视图,在物理上,同一集合中的数据可能分布在异构的存储系统中。用户利用存储资源代理提供的API提出数据访问请求,存储资源代理利用元数据目录中的信息进行协议转换,并将转换后的数据访问请求发向不同的存储系统(数据库、文件系统或档案库),从而实现对异构存储资源的统一访问。

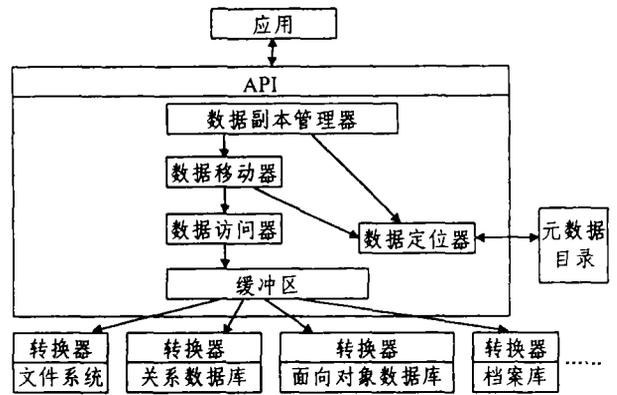


图2 存储资源代理的软件结构

在系统功能方面,存储资源代理可以划分为:数据副本管理器、数据移动器、数据访问器、数据定位器和缓冲区等几个部分。数据副本管理器主要负责管理数据副本和维护数据副本的一致性,它通过数据定位器确定数据或其副本的物理位置,通过数据移动器实现对数据及其副本的访问和一致性维护。数据移动器主要负责数据在不同存储资源之间的移动,通过数据定位器确定数据的物理位置,通过数据访问器实现对数据的访问。数据访问器主要负责与具体的存储设备和其它的存储资源代理联系,实现数据访问。数据定位器利用元数据目录获取有关数据的各种信息,从而实现数据定位。缓冲区的主要目标是提高数据访问的效率。为了实现存储资源代理对各种存储资源的数据访问,需要为不同的存储资源设计转换器,实现数据访问接口的转换。

美国的San Diego超级计算中心(简称SDSC)参加了NAPCI计划的研制工作,已经推出了一套存储资源代理软件

(下转第51页)

MapObjects、ArcObjects 和 SuperMap)、Web GIS (Map Extreme, Arc IMS, SuperMap IS)、采用关系数据库和对象关系数据库管理空间数据 (Oracle Spatial 和 ESRI 的 Geo-Database^[1]) 等,它是组件技术、Internet 技术和数据库技术发展的结果。而 GIS 技术的发展又促进了 IT 技术的发展, GIS 对 IT 技术最主要的贡献在于解决了 IT 对现实世界建模过程中涉及的空间问题,促进了统一对象模型的发展,从而促进了信息系统的空间化 (Spatially Enable),空间化只是在信息系统原有对象或实体中增加空间属性,并通过 GIS 软件加以处理的过程,“空间对象 (Feature) 在关系表中是一个具有几何属性 (Geometry) 的行 (Row),空间对象类 (Feature Class) 是具有几何属性列 (Geometry Column) 的表”^[1]。GIS 使 IT 所处理信息能够纳入到一个统一空间框架体系中,发现其存在的空间关系和模式,这种空间框架不局限于现实地理空间,也可以是一种虚拟的空间,从而促进了基于空间关系的知识挖掘技术的发展。

现有采用关系数据库或对象关系数据库管理空间数据的 GIS 技术适合于解决宏观的空间问题,难于解决微观的空间问题,即反映微观空间对象 (Micro Spatial Objects)。在现实世界中,微观空间对象往往是一个复杂空间对象,如一个房屋的内部结构,并且往往基于独立、局部的坐标系。采用现有的方法,一个微观空间对象用一个表来表示,将大大增加数据库中表的数量,在实现上是不可行的;用一个表来存贮多个微观空间对象,由于一个表对应一个统一的坐标系,难于解决微观对象独立的、局部坐标系问题,而且实现效率较低。因此,需要发展一种微观 GIS 技术,采用表中的一个列 (Column) 来存贮微观空间对象,并记录微观空间对象的独立坐标系及其与统

一坐标系之间的关系,可以微观 GIS 视图 (Micro GIS View) 的方式对一个微观 GIS 对象进行操作,也可以在宏观 GIS 视图中通过坐标转换进行操作,微观 GIS 将成为现有 GIS 的一个扩展。

随着时态数据库 (TDB)^[9] 和时态 GIS (Temporal GIS) 技术的发展与成熟,基于 IT 技术,人们将进一步从时空的角度认识现实世界,最终实现人类对时空认知的统一。GIS 与 IT 技术的发展将使人类从制造和使用信息工具阶段,进入到制造和使用知识工具的时代。

参考文献

- 1 Zeler M. Modeling Our World-The ESRI Guide to Geodatabase Design. Environment Systems Research Institute, Inc. 1999
- 2 白光润. 地理学的哲学贫困. 地理学报, 1995, 50(3)
- 3 Amler S W. The Object Prime. Second Edition. Cambridge University Press, 2001
- 4 Bézin J. Who's Afraid of Ontologies?. OOPSLA'98 Workshop #25, Model Engineering, Methods and Tools Integration with CDIF
- 5 Architecture Board MDA Drafting Team. Model Driven Architecture--A Technical Perspective. Review Draft 14 February 2001, Document Number ab/2001-02-01
- 6 Nicola Guarino Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press
- 7 Object Management Group (OMG). OMG Unified Modeling Language Specification Version 1. 3, June 1999
- 8 Rumbaugh J, Jacobson I, Booch G 著. UML 参考手册. 姚淑珍, 唐发银等译. 机械工业出版社, 2001. 1
- 9 何新贵, 唐常杰, 李霖, 刘云生著. 特种数据库技术. 科学出版社, 2000

(上接第39页)

SRB V1. 2^[8], 并在全国14个站点上安装了 SRB 软件。

4. 应用

涉及核模拟、气象、地球科学、高能物理学、医学、生物学等领域的科学计算是网格信息计算技术的典型应用。其中,生物学的发展关系到一个国家国民经济的发展,甚至整个人类社会的进步。

生物学的发展与精确的数据表示和高效的数据处理息息相关。目前,生物数据的表示和查询技术的发展已经能够适应大规模的软件分布。为此,美国的 NPACI 计划试图利用信息计算技术为生物学的研究提供支持,推动生物学研究工作的广泛开展。SDSC、斯坦福大学、德克萨斯大学联合推出了“生物数据表示与查询”(Biological Data Representation and Query, 简称 BDRQ) 计划。BDRQ 主要研究分子生物序列结构数据库的高效查询方法。为了支持对序列结构数据库的分析,需要从核酸数据库和蛋白质数据库中提取三维结构信息。因此,需要联合数据库进行统一查询,并进一步研究新的查询算法,以允许生物学家通过类比已知的分子来从序列和结构信息中推断出功能。

斯坦福大学已经开发了一种自动注释蛋白质三维结构中功能部位和结构部位的方法,并将通过 NPACI 向外提供 WEB 服务。德克萨斯大学正在研究解决物种分析中的复杂问题的工具,编程实现时将充分利用 NPACI 的并行计算资源。目前,已经利用 SDSC 的 SRB 对 SDSC 和德克萨斯大学等站点的数据库进行管理,并向有关生物学的研究机构提供信息服务。除了 BDRQ 计划外, NPACI 还推出了一系列相关的研

究计划,如,“增强型生物成象”(Enhanced Biological Imaging) 计划等。

结论 从社会发展到国防建设,从地球科学、高能物理学、分子科学、神经系统科学、医学、生物学等学科到作战指挥、战场仿真,各领域明确提出了分布式异构环境下的海量信息高效复杂处理的应用需求,迫切 need 要加强网格的信息计算能力,因此,针对应用需求开展信息计算关键技术的研究迫在眉睫,这对各领域的发展将起到极大的推动作用。

参考文献

- 1 Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. International J. Supercomputer Applications, 2001, 15(3)
- 2 Foster I, Kesselman C. The Grid: Blueprint for a Future Computing Infrastructure. Morgan Kaufmann Publishers, 1999
- 3 Stockinger H, Rana O, Moore R, Merzky A. Data Management for Grid Environments. HPCN Europe 2001. 151~160
- 4 Moore R. Massive Data Analysis Systems: [Technical Report]. San Diego Supercomputer Center, 1997
- 5 Hoschek W, Jaen-Martinez J. Data Management in an International Data Grid Project, ACM International Workshop on Grid Computing (Grid'2000), 17-20, Dec. 2000, Bangalore, India
- 6 Chervenak A, Foster I. The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets, Journal of network and computer applications, 2000
- 7 Foster I, Kesselman C. The Globus Project: A Status Report. Proc. IPPS/SPDP '98 Heterogeneous Computing Workshop, 1998. 4~18
- 8 Rajasekar A K, Wan M. SRB & SRBRack - Components of a Virtual Data Grid Architecture, Advanced Simulation Technologies Conference (ASTC02) San Diego, 2002. 15~17