

面向服务的语义网格框架及其关键技术

曹蓉¹, 褚伟¹, 朱颖²

(1. 合肥工业大学 计算机网络系统研究所, 安徽 合肥 230009;

2. 上海大学 计算机学院, 上海 200072)

摘要: 由于传统的网格存在信息格式异构等问题, 导致网格无法像原先设想的那般高效无缝自动化, 因此在网格中引入语义的概念, 希望通过语义网格来弥补这种缺陷。分析了语义网格与网格及语义 Web 的关系, 阐述了语义网格中与语义 Web 共享的关键技术, 并基于面向服务的思想, 提出了一个面向服务的语义网格框架, 并对它做了详细的描述。

关键词: 语义网格; 语义 Web; 元数据

中图分类号: TP393.02

文献标识码: A

文章编号: 1673-629X(2007)01-0071-03

Service-Oriented Semantic Grid Architecture and Its Key Technology

CAO Rong¹, CHU Wei¹, ZHU Ying²

(1. Institute of Computer Network Systems, Hefei University of Technology, Hefei 230009, China;

2. School of Computer Science, Shanghai University, Shanghai 200072, China)

Abstract: Due to the problems such as heterogeneous information format existing in the traditional grid, grid cannot be as highly efficient, seamless and automatic as it is originally expected. Therefore the conception of semantics is introduced into grid and it is expected to make up for the defects by semantic grid. Analyzes the relations among semantic grid, grid and semantic Web and illustrates the key technology shared with semantic Web in semantic grid. It also proposes a service-oriented framework of semantic grid on the basis of the service-oriented ideology and make a detailed description for it.

Key words: semantic grid; semantic Web; metadata

0 引言

随着网络技术的发展, 为了更加有效地利用网络上的各种资源, Ian Foster 和 Carl Kesselman 提出了网格^[1]的概念。简单地说, 网格是一种信息社会的网络基础设施, 是利用互联网把分散在不同地理位置上的多个资源, 包括计算资源、存储资源、通信资源、软件资源、信息资源、知识资源等全面连通和统一分配、管理及协调起来, 组成一台“虚拟的超级计算机”。但现有网格上存在信息格式异构、信息语义的多重性以及信息关系匮乏和非统一等问题, 无法满足原先设想的网格应具有高度的简单实用和无缝自动化的需求, 这是网格发展现状和应用需求之间目前存在的主要差距, 因此需要在传统的网格中引入语义来改变这种局面。

语义 Web 是互联网的创始人 Tim Berners-Lee 在 2000 年提出的。语义 Web 是对现有 Web 的扩展, 其目标是通过使用本体和标记语言, 如: XML, RDF, OWL 和 DAML+OIL, 对 Web 文档中的术语含义进行形式化描述, 使计算机能够理解 Web 资源的内容, 分辨和识别这些

语义信息, 并对其自动进行解释、交换和处理, 为用户提供智能索引、基于语义的内容检索和知识管理等智能服务。

语义网格是语义 Web 和网格相结合产生的新的研究领域。语义网格小组对其进行了定义: 语义网格就是对当前网格的一个扩展, 其中对信息和服务进行了很好的定义, 可以更好地让计算机和人们协同工作。将网格和语义 Web 技术组合在一起, 能够提高集成和数据计算的能力^[2], 如图 1 所示。

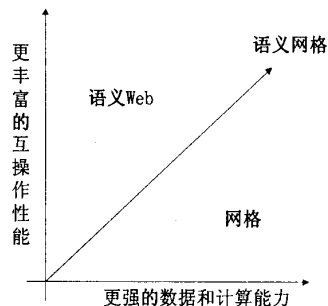


图1 提高集成和数据计算能力图示

目前国外一些围绕语义网格的项目已经展开, 如英国 e-Science^[2] 导航项目 mygrid, CoAKTinG (网格高级协同知识技术) 项目等; 中国也已经重视这方面的研究工作, 2004 年 9 月科技部批准了国家重点基础研究计划 (简称

收稿日期: 2006-04-13

作者简介: 曹蓉 (1979-) 女, 安徽合肥人, 硕士, 研究方向为计算机网络、网格计算。

973 计划):“语义网格的基础理论、模型与方法研究”。目前对语义网格的研究还处于起步阶段,主要集中在标准的制定和理论模型的建立等方面^[3],文中就是从面向服务的角度建立了一个语义网格框架。

1 语义网格相关技术

在语义网格中,要求网格能够运用和处理一些语义信息。这些信息包括服务的可用性,服务的目的以及服务连接、配置和替代的方式,还有服务如何被发现、调用和演化。语义网格使用元数据来描述网格中的信息,将信息转化为一些更有意义的东西,而不只是一个数据集,这意味着要正确理解数据的内容、格式和重要性。语义 Web 就遵循这种模型,即提供其他一些元数据来帮助描述在 Web 页面上显示的信息,这样浏览器、应用程序和用户就能够更好地决定如何处理数据。语义 Web 的核心技术和语义网格在分布式系统和信息管理中有着共同的基础,它们所需的核心技术有相似之处。

元数据(Metadata)简单地讲就是“关于数据的数据”,元数据可以用于创建文档,跟踪代码中的依赖性,甚至执行基本编译时检查。

(1)可扩展标记性语言(XML, Extensible Markup Language)。XML 可以让信息提供者根据需要,自行定义标记及属性名,从而使 XML 文件的结构可以复杂到任意程度。它具有良好的数据存储格式和可扩展性、高度结构化以及便于网络传输等优点,再加上其特有的 NS 机制及 XML Schema 所支持的多种数据类型与校验机制,使其在数据交换与信息共享方面及 Web 文档上得到广泛运用。

(2)资源描述框架(RDF, Resource Description Framework)。RDF 是 W3C 推荐的描述和交换元数据的框架,是处理元数据的基础。RDF 是一个由资源(Resource)、属性(Property)和声明(Statement)组成的模型。通过命名属性的值,这个模型可以描述资源之间的内在关系。RDF 把 XML 作为元数据交换和处理的公共语法。很多网格项目也正在开始采用 RDF 作为元数据的一个共同数据模型。

(3)本体^[4](Ontology)。文献[5]给出了本体的定义:“本体是用来帮助程序和人共享知识的概念的规范描述。”文献[5]还指出:“概念化是关于世界上的实体,如:事物、事物之间的关系和约束条件的知识表达,而规范一词是强调这种表达是用一种固定的形式来描述。”传统的本体表示方法是四元素表示法。四元素表示法的基本思想是:一个本体中的四个主要元素是:概念(concepts)、关系(relations)、实例(instances)和公理(axioms)。

DAML + OIL 是一种专门为 Web 设计的本体语言。它建立在已有 Web 标准,如 XML 和 RDF 的基础上:本体作为 XML 文档被存储,概念由 URI 来关联。它由一种有着很强表达能力的描述逻辑来支持,它的形式化语义使得机器能够解释并且推理。DAML + OIL 也是 W3C OWL

Web 本体语言的基础。

2 面向服务的语义网格框架

语义网格实际上有两种用途:用来发现处理数据的可用资源,以及对数据进行集成。目前已有的一种是文献[6]中提出的语义网格体系结构,这种结构是将语义网格看作一个网格中间件,但对于语义网格是如何实现语义转化,以及如何具体使用和管理语义信息没有详细描述。而文献[7,8]中的模型,无法体现目前网格将所有资源看作一种服务的概念,因此,结合这两种模型,给出了一个基于 WRSF 的面向服务的语义网格框架,如图 2 所示。将整个语义网格按照其所提供的服务分为三层,分别为基础服务层、语义服务层和应用层。

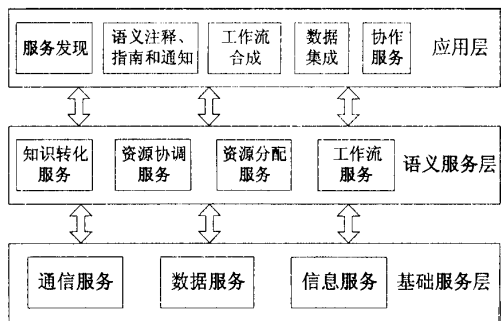


图 2 面向服务的语义网格框架图

(1)基础服务层包括数据和计算服务、通信服务和信息服务,例如网络访问、资源定位和调度、进程间数据迁移。信息服务响应众多计算进程的请求,这些进程请求若干数据资源和处理阶段来实现所期望的结果。这些服务包括分布式查询处理、工作流发布、事件通知和检测管理。基础服务使用网格服务和实体相关联的元数据描述网格服务和网格的处理能力,使得网格在 Internet 上更容易被发现。

(2)语义信息服务层是在网格服务的基础上应用本体技术对不同结构、不同领域的知识进行统一语义定义,并通过服务实现对分布式的知识与数据进行统一的组织、执行以及协调,以便有效地管理语义、数据元、数据挖掘工具以及可视化工具的语义描述。例如,一个为科学社区提供计算资源的网格可以由大量的用户和组织使用。这不用为网格环境开发一个单用户的应用程序,其他人都可以使用现有的网格基础设施和网格应用程序。复杂性在于如何描述网格服务和网格的处理能力。这是语义网格的特点所在,加上对能力和功能的详细分类和描述,使得判断特定网格可以提供哪些功能更加容易。

(3)应用层包括服务发现,语义注释、通知和指南、工作流合成、数据集成、协作服务等多种应用服务^[6]。例如服务发现是指在开放式网格服务体系(OGSA)中,每个服务都能够静态或动态地与上下文结合,上下文包括特定用户、执行代价、执行速度、可靠性和对用户可用的适当认证机制。能否发现正确的服务依赖于每个服务的语义信息。

服务描述的附加元数据使用 RDF 陈述来进行声明。服务提供者发布他们的服务,服务使用者通过多种机制,如名字、词语、特征、类型或本体描述来查找、匹配服务。DAML+OIL 提供了表达服务描述的词汇。在执行 workflow 前,对服务类和它们的实例进行发现、匹配和选择。匹配首先对域进行匹配,然后对操作的属性进行匹配。又如数据集成,与 Web 和语义 Web 中的信息一样,语义网格对网格中存储和可以使用的信息有进行关联和协调使用的的能力。对于资源网格来说(这种网格共享的是磁盘和存储空间,而不是提供 CPU 处理能力),使用网格技术在信息之间提供关联和连通性可以为存储和信息提供一种有效的方法。例如,使用存储照片的语义网格组件,以及存储影像资料的语义网格组件,就可以建立连接和关联。例如,您可以搜索某个主题的照片,比如大熊猫,不但能找到相关的照片,还可以找到相关的影像资料。

3 小 结

语义网格承诺将为网格带来语义上的可互操作性、智能自动化以及灵活重用的优势,使其日益成为网格计算发展过程中的一个重要阶段。文中描述了语义网格中发展过程及研究状况,介绍了它同语义 Web 共享的关键技术和标准,并针对现有语义网格模型的不足,提出了基于 WRSF 的面向服务的语义网格框架,下一步的研究工作将围绕实现该模型来进行。

(上接第 70 页)

测方法。



图 2 原始 SAR 图像



图 3 Canny 算子



图 4 Ratio 算子

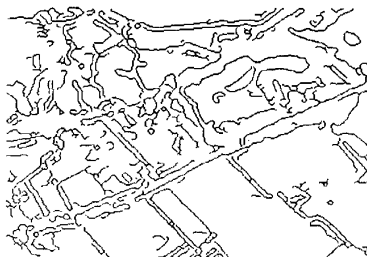


图 5 文中方法

实验结果证明,该方法在受乘性噪声影响较大的 SAR 图像中有着良好的边缘检测性能,提取的边缘图准确描述了 SAR 图像中的各种结构,能较好地避免边缘过

参考文献:

- [1] Foster I. The Grid: A New Infrastructure for 21st Century Science[J]. Physics Today, 2002,55(2):42-47.
- [2] De Roure D, Jennings N R, Shadbolt N R. Research agenda for the semantic grid: A future e-science infrastructure[EB/OL]. 2004-02. <http://www.semanticgrid.org/v1.9/semanticgrid.pdf>.
- [3] Hai Zhuge. China's E-science knowledge grid environment [J]. Intelligent Systems, IEEE, 2004, 19(1):13-17.
- [4] Gruber T R. Towards Principles for the Design of Ontologies Used for Knowledge Sharing[C]//Proceeding of the International Workshop on Formal Ontology. Padova, Italy: [s.n.], 1993.
- [5] De Roure D, Jennings N R, Shadbolt N R. The Semantic Grid: Past, Present and Future [EB/OL]. 2004/2006. <http://www.semanticgrid.org/documents/semgrid2004/semgrid2004.html>.
- [6] 李从东, 李晓宇. 面向分布式知识发现的知识网格框架及其实现[J]. 华北工业学院学报, 2005, 26(1): 27-28.
- [7] Smirnov A, Pashkin M, Chilov N. Knowledge logistics in information grid environment[J]. Future Generation Computer Systems, 2004, 20: 61-79.
- [8] 李 玲, 唐胜群. 知识网格中基于 RDF 的知识表示技术和应用[J]. 计算机应用研究, 2005, 22(12): 223-225.

检测和欠检测,可以用于遥感图像矢量化、自动目标识别等方面。

参考文献:

- [1] Canny J. A computational approach to edge detection [J]. IEEE Trans on Pattern Analysis Machine Intelligence, 1986, 8 (11): 679-698.
- [2] Touzi R, Lopes A, Bousquet P. A Statistical and Geometrical Edge Detector for SAR Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 1988, 26(6): 764-773.
- [3] Oliver C J, Blacknell D, White R G. Optimum Edge Detection in SAR[J]. IEEE Radar Sonar and Navigation, 1996, 143 (1): 31-40.
- [4] Bovik A C. On detecting edges in speckle imagery[J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1988, 36 (10): 1618-1627.
- [5] 王 程, 王润生. SAR 图像直线提取[J]. 电子学报, 2003, 31(6): 816-820.