

下一代的网格技术——知识网格^{*}

张 海¹, 关伟豪²

(1 第一军医大学网络中心, 广东广州510515; 2 中山大学信息与网络中心, 广东广州510275)

摘 要: 语义网和网格的出现是为了解决 Web 技术存在的各种缺陷, 更好的实现网络环境下的信息共享和协同工作, 它们代表互联网技术未来发展的两个方向。而知识网格是这两种新技术的结合, 被喻为是下一代网格技术。知识网格通过语义网技术来实现信息处理和知识获取, 采用层次性网格体系结构来构建资源共享、信息共享的平台, 为信息服务的提供更高层次的集成和共享服务。

关键词: 语义网; 网格; 知识网格

中图分类号: TP393 **文献标识码:** A **文章编号:** 0529-6579 (2005) S1-0171-03

一般认为, 第一代的互联网技术的代表是 TCP/IP 协议, 而第二代互联网技术则是从上个世纪 90 年代兴起的 web 技术。Web 技术采用 HTML 描述发布各种数据和信息, 通过友好的界面 (浏览器), 更加生动地发布与显示不同类型的数据以满足不同的需求。但是, HTML 仅仅由于表示数据显示的布局, 主要是面对用户直接阅读, 没有把信息的表现形式、内在结构和表达内容分离出来, 非常不利于计算机的直接处理。最近的几年来, 针对目前 web 技术在信息表达、检索、共享等方面存在的缺陷, 研究人员逐渐把目光投向了未来互联网发展的两个方向, 一个方向是改进现有的 web 技术来建立下一代的 WWW 技术——语义网 (Semantic Web), 另一个方向则是建立一套与现有 web 完全不同的应用平台——网格^[1-8]。

语义网的目标是, 为互联网上的信息提供具有计算机可以理解的语义, 从而满足智能主体 (Agent) 对 WWW 上异构、分布信息的有效检索和访问, 实现网上信息资源在语义层上的全方位互联, 并在此基础上, 实现更高层的、基于知识的智能应用^[8]。在语义网中, 我们不仅需要描述 web 内容的元数据来实现计算机之间的互操作, 还应提供为每个 web 资源自动进行语义描述的能力。语义网应用的层次模型包括以下几部分:

(1) 全局唯一标识的 web 资源, 采用元数据作形式化的描述, 以及作用在资源上的一些规则, 通过本体来推导新的元数据和知识。

(2) 基本的语义服务, 实现在元数据、本体、语义搜索引擎上的推理和查询。

(3) 使用基本语义服务的应用程序。

其中元数据是描述数据的数据, 而本体的定义是共享概念模型的明确的形式化规范说明。实现元数据、本体的语言有 XML、RDF、DAML+OIL、OWL。XML 是资源内容描述的标记语言, RDF 是语义描述和关系描述, DAML+OIL、OWL 是满足逻辑和证明要求的本体语言。支持语义网应用的开发工具有 Jena。

网格所倡导是分布式计算资源的共享、管理、协作和控制, 目标是使得应用程序在网格环境中, 就像电力网中的插座一样, 随时即插即用就可获得所需的服务。Web 技术刚开始出现时是为科学协作服务的, 到了今天大量应用于商业领域。同样, 网格刚出现的时候, 主要是应用在科学研究, 它的主要概念和技术是用于科学协作中的资源共享。现在, 人们发现网格对于商业计算也是特别的重要, 不仅可以作为一种增强企业计算能力的方式, 而且是一种构造可靠、可扩展、安全的分布式系统的重要解决方案。开放网格服务结构 OGSA (Open Grid Services Architecture) 正是在这种需求下产生的, 是目前最新的网络体系结构。OGSA 的重点是商业应用而不是象早期网格应用是偏向于科学计算, 但是其架构又可以同时使用于两种环境。OGSA 是以服务为中心的服务架构。这里的服务是指具有特定功能的网络化实体, 包括各种计算资源、存储资源、网络、程序、数据库等。与 web 中的 Web Service 相对应, OGSA 定义了网格服务 (Grid Service)。网格服务也是一种 Web Service, 该服务提供了一组接口, 这些接口的定义明确且遵守特定的惯例, 解决

* 收稿日期: 2005-02-02

作者简介: 张海 (1972 年生), 男, 讲师, E-mail: zhanghai@fimmu.com

服务发现、动态服务创建、生命周期管理、通知等问题。在 OGSA 中, 将一切都看作是网格服务, 因此网格就是可扩展的网络服务的集合, 即网格 = { 网格服务 }。

网格应用领域逐渐从科学计算转向工业和商业应用, 为了满足这些领域的需求, 网格需要支持在分布式数据敏感型的应用环境中大数据集的存储、移动、管理。语义网的技术优势在于信息处理和知识获取, 而网格则是提供资源共享、信息共享, 通过对两者的研究结合, 人们提出了下一代的网格——知识网格 (语义网格)。图 1 可以很好表示语义网、网格、知识网格之间的发展关系^[5]。

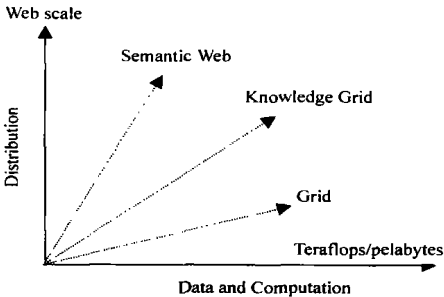


图 1 语义网、网格、知识网格的发展
Fig 1 The development of the semantic web, grid, and knowledge grid

1 知识网格

所谓的知识网格^[3]是指一个智能的、可支撑网络应用的环境, 使用户可以有效地获取、发布、共享、管理知识资源。知识网格提供按需服务 (On-demand Service) 来支持应用的创新、协作、决策和解决问题。实现知识网格需要采用许多现有的可能技术和应用平台, 网格技术并不是实现知识网格的唯一平台, 但是知识网格要吸取网格的理念, 如共享和协同。知识网格使用的关键技术有^[7]:

(1) Web 服务 (Web Services)。OGSA 以网格服务为中心的架构, 非常有利于我们构建知识网格环境。在 OGSA 网格体系结构中, 由于网格环境中所有的组件都是虚拟的 (指对相同的接口不同实现的封装), 因此通过提供一组相对统一的核心接口, 所有的网格服务都基于这些接口实现, 就可以很容易构造出具有层次结构的、更高级别的服务 (语义服务、知识服务)。

(2) 软件代理 (Software Agents)。多 Agent 系统研究提出的问题空间与知识网格是相一致的, (Aget) 中决策、分散、对等、自治的行为, 正是构建网格虚拟组织所必须的。基于 Agent 的计算模

式也是一种面向服务的模式, Agent 与面向服务的网格可以建立直接的映射, Agent 可以是服务的生产者、消费者和代理者。

(3) 元数据。元数据是关于数据的数据, 是对知识网格上信息的一种描述方式, 这种描述方式使得信息变成计算机可以理解的信息。元数据最基本的作用就是管理数据, 从而实现查询、阅读、交换和共享。

(4) 本体和推理。本体是共享概念模型的形式化规范说明, 目标是获取、描述和表示相关领域的知识, 提供对该领域知识的共同理解, 确定该领域内共同认可的词汇, 并从不同层次的形式化模式上给出这些词汇 (术语) 和词汇间相互关系的明确定义。

(5) 语义服务 (Semantic Web Services)。现有的 Web Services 围绕着 UDDI、WSDL 和 SOAP, 提供了有限的机制来实现服务的发现、配置、组合和自动协商, 对服务的调用来说抽象层次较低。语义服务的目标是通过提升服务的描述, 来指明服务的能力和任务完成的特性, 使得在同一语义空间下的服务更容易整合。

2 知识网格的体系结构

知识网格体系结构^[4,7]从低到高分别是: 网格中间件基础结构、信息网格服务层、语义网格服务层、知识网格服务层、基于知识网格的应用程序, 如图 2 所示。

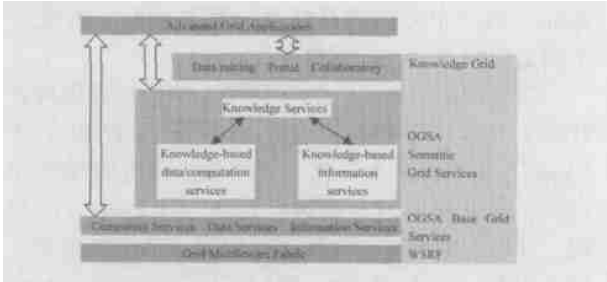


图 2 知识网格的体系结构
Fig. 2 Knowledge grid architecture

(1) 网格中间件基础结构由网格操作系统组成, 把网络上的信息资源连接在一起, 为上层的应用提供数据层面的连通和共享。网格操作系统一般采用网格工具包软件 (Globus Toolkit) 来构建, Globus 项目是目前世界上与网格相关的最有影响力的项目, 开发出能在多种平台上运行的网格工具包软件。

(2) 信息服务层实现信息的无缝共享, 包括信

息数据库的构建、信息的发现、连通和处理。

(3) 语义服务层按照计算机理解的格式来表示知识, 使用本体和元数据语言来描述信息。通过语义网, 为上层应用提供资源管理 另和开发复杂应用 (信息检索、数据挖掘) 的支持。

(4) 知识服务层为知识网格环境提供一个智能化访问入口, 同时为上层的应用程序提供发现、获取知识的接口。

(5) 基于知识网格的应用程序, 通过 UDDI 来发现服务, 通过 SOAP 来调用服务。应用程序可以根据需求, 通过 SOAP 直接访问信息服务层、语义服务层、知识服务层所提供的服务。

实际应用中这几个层次在软件体系结构上是彼此不可分割的软件实体, 而是互相渗透的。例如, 知识不仅驻留在网络应用中, 而且也渗透到网格应用和基础框架得全部有效范围内, 知识在信息服务层需要实现知识存储服务, 在语义服务层需要实现知识表示服务。具体的实现可以依据基于组件的网格体系结构, 每一层工作由各种各样的组件来完成。

3 结 论

知识网格是下一代的 Web 高级体系架构, 它通过各种知识资源、计算机可理解的语义和资源实体, 提供按需的知识服务。在网格应用从科学计算

转向商业应用, 从计算密集型应用转向数据密集型应用, 知识网格具有光明的应用前景。

参考文献:

- [1] FOSTER I, KESSELMAN C, TUECKE S The anatomy of the grid: enabling scalable virtual organizations[J] . International J. Supercomputer Applications, 2001, 15(3): 1.
- [2] FOSTER I, KESSELMAN C, JEFFREY M N et al. The physiology of the grid——an open grid services architecture for distributed systems integration, Globus Project[DB/OL] . <http://www.globus.org/research/papers/ogsa.pdf>. 2002—06—22.
- [3] HAI Zhuge, China's e-science knowledge grid environment [J] . IEEE Intelligent Systems, 2004, 19(1): 13—17.
- [4] CANNATARO M, TALIA D. Semantics and knowledge grids: building the next-generation grid [J] . IEEE Intelligent Systems, 2004, 19(1): 56—63.
- [5] DAVID De, JAMES A. Hendler, E-Science: The grid and the semantic web[J] . IEEE Intelligent Systems, 2004, 19(1): 65—71.
- [6] FOSTER I, KESSELMAN C. , 网格计算[M] . 北京: 电子工业出版社, 2004.
- [7] DAVID D R, NICHOLAS R J, NIGEL R S. The Semantic Grid: Past, Present and Future [DB/OL] . <http://www.semanticgrid.org/documents/semgrid2004/semgrid2004.html>.
- [8] 宋炜, 张铭. 语义网简明教程[M] . 北京: 高等教育出版社, 2004.

The Next-generation Grid — Knowledge Grid

ZHANG Hai¹ QUAN Wei-hao²

(1. The Network Center of The First Military Medical University, Guangzhou 510515, China;

2. The Information and Network Center of Sun Yat-sen University, Guangzhou 510275, China)

Abstract: To overcome the shortcomings in the Web technology and implement sharing information and cooperative teamwork, scientists are working toward the Semantic Web and the Grid which are two developmental trend in Internet. As the next-generation Grid, the Knowledge Grid is the combination on both. The Knowledge Grid, based on multi-layers application architecture, processing information and capturing knowledge through the Semantic Web technology, provides integrating and sharing information services in higher layer.

Key words: semantic web; grid; knowledge grid