

基于网格计算的远程教育系统模型的研究

麦红¹ 李吉桂²

(广州大学数学与信息科学学院 广州 510405)¹ (华南师范大学计算机学院 广州 510630)²

摘要 网络远程教育是一种全新的知识传播方式和实现全民教育的有效手段。现在网络上的教育资源越来越多,且在地理上是分布的,运行平台上是异构的,因此彼此间不能有效地进行共享。利用网格技术,就可以实现远程教育资源的一体化集成、共享和管理,提高远程教育资源的使用效率。本文将网格技术与远程教育结合起来,提出了基于网格计算的远程教育系统模型,并针对该模型设计了在网格环境下对广域异构的远程教育信息资源进行统一访问的解决方案。

关键词 网格,远程教育,资源共享,资源管理

Research on the Model of Distant Education System Based on Grid

MAI Hong¹ LI Ji-Gui²

(School of Mathematics and Information Sciences, Guangzhou University, Guangzhou 510405)¹

(School of Computer, South China Normal University, Guangzhou 510630)²

Abstract Distant Education based on Web is a new method of transmitting knowledge and an efficient means of education. In Internet environment, there are more and more education resources, distributed and heterogeneous, which can not be sharing. Using Grid technology, we can achieve integrating, sharing and managing of distant education resource, and use the resource more efficient. In this paper, we put forward the model of distant education system based on Grid, then present the mechanism of accessing distributed and heterogeneous data by uniform way.

Keywords Grid, Distant education, Resource sharing, Resource management

随着网络技术特别是 Internet 技术的迅速发展,以网络技术为基础的远程教育也在全球范围内迅速发展起来。在基于网络的学习环境下,应该包含两个重要的组成部分:学习资源及其传输方法;教师与学生、学生与学生交流的工具。现今网络上的教育资源在不断地增加,但各教育机构的服务器却好象一个个网络世界上孤立的“小岛”,使得远程教育环境下的“学生”只能使用某教育机构提供的远程教学服务,已不能满足网络远程教育发展的需要。近年来,在国际上兴起了一种重要的信息技术——网格,被称之为是继传统互联网、Web 之后的第三代互联网。它可以对用户屏蔽物理上存在的单个信息服务器,并将这些服务器通过高速网络互联成为一个大的信息服务网络平台,实现广域、异构信息的共享、互连和互操作。利用网格技术,我们可以把跨地域的异构的各种教育资源(包括计算机、数据库、教学软件、教学仪器设备等)整合成一个大型的虚拟网络教育机构,消除“信息孤岛”,使得“学生”可以通过统一的网格门户透明地使用网络上不同地域的教育资源,并和其他“学生”进行有效的协同学习,这样就可以提高网络上教育资源的使用效率,扩大“学生”的学习资源和交流范围。在我国,地域广、人口众多、高等教育资源相对稀缺,发展基于网格技术的远程教育具有重大的现实意义。

1 系统体系结构

基于网格计算的远程教育系统模型的主要目标是实现教育资源(在这里主要是指信息资源,本文的以下部分如不特别指出,都是指信息资源)和远程教育系统平台的一体化共享、管理以及服务,为用户提供单一系统映像环境。设计系统的体系结构如图 1 所示。

在系统的体系结构中,采用分布式的层次结构和以资源

所有者为中心的策略,各个教学区的资源以及用户数据库都分布于教学区的管理范围内。“管理员”和“用户”通过 Web 服务器上的 API 进入系统,Web 服务器为教学区“管理员”和“用户”提供不同的统一访问界面,它还与用户管理一起提供单一登录的功能,“用户”在不同的教学区之间切换不需要多次输入用户名、密码等用户信息。“用户”进入系统即可得到全部共享信息的视图。访问控制主要用于资源访问控制、服务间的相互认证等。各个“教学区”的资源管理器利用对基于 LDAP 的信息存储服务器实现资源信息的存储、访问和管理,建立各个“教学区”的信息中心,负责对本“教学区”内的所有资源进行收集、组织、更新,并提供资源的发布与查看。存储资源代理用于将存储、检索数据集等高层用户的抽象请求映射为异构的、分布式存储中的底层操作,并负责收集系统中的动态资源信息报告给资源管理器。元数据服务器是一个层次式的分布服务器结构,由局部元信息服务器和中央全局元信息服务器组成,每个局部元数据服务器负责对应一组教学区的基于 LDAP 的信息存储服务器,并为每个服务器建立元信息的索引和数据缓冲,中央元信息服务器建立各个局部元信息的索引和数据缓冲。

在这里多次提到了“教学区”,这是参考人类社会中教育资源以学校为单位进行管理而提出的概念,目的是便于管理基于网格计算的远程教育系统中的各种教育资源,快速实时地把资源请求者和资源提供者联系起来。教学区是由某个教育机构的用户、资源以及管理策略组成的集合,它管理着系统中的部分资源、部分用户,在教学区中还可以定义子教学区。由于用户多数时候使用的资源都是属于本教学区的,仅在少数时候使用其他教学区的资源,这种分布式的层次存储结构能够提高资源检索速度,不会因为大量用户同时检索其他教

学区的资源而造成速度下降,而且,教学区的动态加入或退出

也因此而变得更加方便。

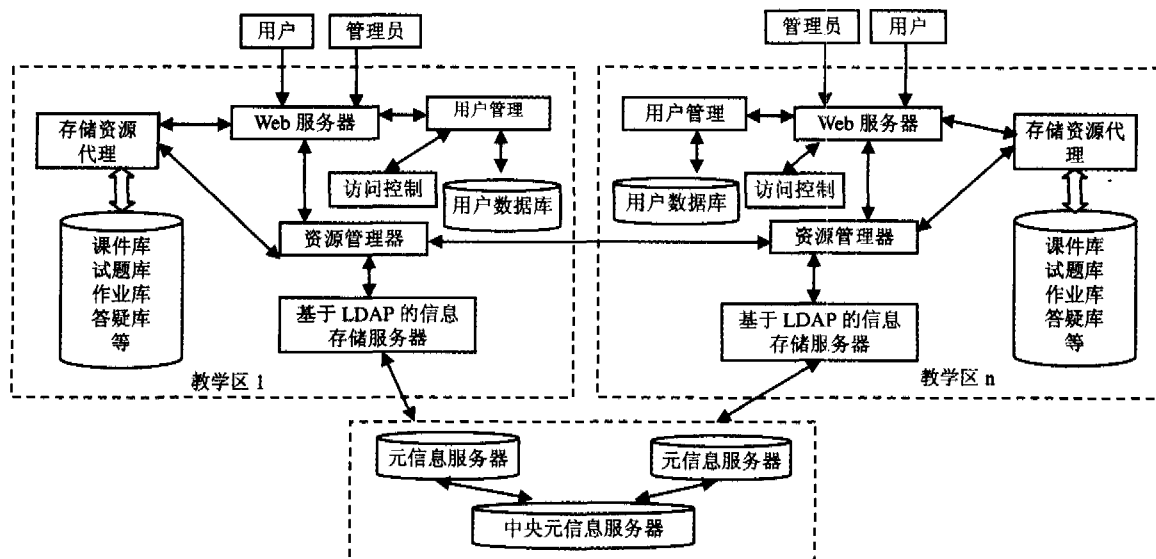


图1 系统的体系结构

2 系统实现的关键技术

实现基于网格计算的远程教育系统模型的关键技术是如何对广域异构的教育资源进行有效的组织管理、访问和权限控制。

2.1 资源的统一命名方式

为了实现网格环境下广域分布的信息资源的全局统一视图,需要有一个全局的命名空间,对系统中的信息资源(包括数据库、文件系统等)进行统一命名。为了将信息资源的物理特征与逻辑视图独立开,使物理层次上的变化不会对逻辑层次产生影响,逻辑层上的变化不会影响物理层次的变更,系统采用了分层次的三种命名空间:用户资源名、系统资源名和物理资源名。

用户资源名是指向用户、在用户逻辑视图中所使用的资源名称,它可以按用户的习惯自由拟定。它和系统资源名的关系是多对一的关系。

系统资源名是在系统内部使用的资源名,具有全局唯一性。每个系统资源名对应一个物理资源名,由它代表该资源在系统内部被系统进行管理、操作、访问、属性建立和查询等活动。

物理资源名是该资源在存储系统中的实际唯一标识符。在某个独立文件系统中,它是一个文件的完整路径,在网络文件系统中,它是可对该文件进行访问的系统路径。

2.2 资源的统一访问方式

资源的访问就是用户通过访问资源的特定接口以获取数据,在基于网格计算的远程教育系统中,主要利用存储资源代理来实现对广域异构资源的统一访问。存储资源代理为高层应用访问分布异构的存储资源提供了统一的接口,当它接收到用户或应用访问资源的请求后,首先获取相关资源的访问权限,然后根据请求的参数和LDAP服务器中的元信息重新转换请求的格式,分析资源种类,调用相应的适配器完成对资源的访问。由于每种资源都有自己的特性,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适配器,并以服务的方式封装起来提供给存储资源代理使用。在服务的内部,有数据适配器和服务适配器。服务适配器是以服务的方式集成遗留系统,而数据适配器则实现了

对数据库资源、文件系统以及URL资源的访问。对于数据库资源,我们在数据适配器中定义一个抽象的数据库基类,在这个基类中定义一些在概念上适用于所有不同数据库的方法(如添加、修改、删除等),然后为不同数据库管理系统构造的数据库,如SQL Server、Oracle等分别设计具体的数据库类,它们在继承基类的基础上添加自己对这些操作的实现。这样,当增加新的数据库时,仅仅需要增加一个具体的数据库类,对于访问接口的其他部分没有任何影响。而对于文件系统类和URL类资源,它们涉及各种复杂类型的文件,这些文件是非结构化的,难以设计统一的文件格式。为了达到统一访问的目的,我们将整个文件内容作为一个对象保存在消息的附件中,并通过消息头部的href属性引用附件。这样,不但可以不限文件传输的范围和大小,还可以保证文件的完整性。

2.3 统一的访问控制策略

在基于网格计算的远程教育系统中,具有多个教学区、多种安全策略、大量的用户和信息服务请求以及异构的信息资源,因此,系统的安全控制机制应该是全局统一的,而且全局的安全策略需要与本地的安全策略协调和交互,既要满足全局控制的需要,又要满足用户自主控制的需求。同时,系统又是一个开放、动态的系统,所以,应该使用一种易于管理的访问控制机制,以降低访问控制管理的复杂度。而且,系统的访问控制机制还应该能够发现跨教学区授权可能带来的安全漏洞。为解决上述问题,在下文中提出了一个以所有者为中心的基于角色的层次访问控制策略。它是在基于角色的访问控制的基础上,以保持资源所有者对资源的最终控制权为前提的分阶段的访问控制策略。它的抽象授权过程如下:

第一阶段:资源所在教学区(以下简称主体教学区)按某种角色将资源的权限授予其他教学区(以下简称客体教学区)。这时,客体教学区相当于主体教学区的一个特殊的用户,主体教学区用本地的方式对客体教学区进行授权,客体教学区获授权的全部共享资源是它所管理的用户能够使用的共享资源的最大集合。

第二阶段:客体教学区根据本教学区的授权方式,按共享资源访问权限所对应的角色将它授权给本教学区内具有相应角色的用户,并提供本地的调用方式给用户使用。

资源-客体教学区, 客体教学区-用户的分阶段授权方式, 利用客体教学区这一中间层就可以屏蔽访问控制机制的异构。主体教学区可以用本地的方式决定将资源共享给哪些客体教学区并拥有最终控制权, 它不必关心客体教学区中用户的具体信息, 所使用的访问控制策略以及组织结构, 客体教学区的有关变动对资源共享不会带来任何影响。同样, 客体教学区对所获得的共享资源在本教学区内使用, 可以依据自己的访问控制策略、组织结构、用户信息对资源进行分配, 主体教学区的内部变化对资源共享也不会带来任何影响。

在现实生活中, 各个教学区中的教育资源一般都能够为它的所有者带来一定的经济效益, 教学区之间的资源共享必然是以对经济效益的分配约定为前提的, 当这种约定改变时, 所有者为中心的基于角色的层次访问控制策略就能方便主体教学区将资源的访问权限收回或加以进一步的限制, 这就同时保证了访问控制策略的灵活性、可扩展性、通用性和自主性。

2.4 统一的身份认证机制

统一的身份认证机制是指用户可以在任何地点、通过任何接入设备登录基于网格计算的远程教育系统, 但用户不会因为地理位置的差异和接入设备的区别而看到不同的内容。由于用户的登录行为可能随时发生在网络的任一节点, 最简单的方法是在每一个节点都记录所有的用户信息, 但用户数量十分巨大, 这个方法显然是不可行的。另外一种方法是所有的用户信息都分别存储在每个教学区的用户数据库中, 这样虽然提高了效率, 但当某个教学区增加用户时就需要将用户的信息通过网络通知其他所有的用户数据库, 这是十分危险的。

通过观察分析, 在某个教学区中的用户, 一般情况下只在该教学区中的某一节点执行登录操作。所以基于这一出发点, 考虑只把用户注册在本地教学区中, 这样在绝大多数情况下都可以直接在本地进行登录处理, 减少了在网络中传播用户信息的危险。

但在极少数情况下, 用户会在其他教学区中登录系统, 这时, 登录教学区就找不到用户信息, 需要执行远程登录的步骤。首先, 要求用户输入所属教学区的名字, 然后由登录教学

区将用户名和密码传送到目标教学区中进行身份认证并返回结果。如果认证通过, 目标教学区会将用户的权限等信息传给登录教学区, 用户就可以在登录教学区中进行登录。当用户退出远程教育系统时, 登录教学区就将相关信息删除。

3 模型评价

① 以教学区为单位划分教育资源的归属, 符合现实社会中资源都属于某些人或某些组织的情况, 方便资源所有者对资源进行管理。

② 分层次的分布式元信息数据库结构, 避免了数据备份引起的不一致问题以及数据在网络中传输的安全问题。

③ 资源三重命名空间使资源的所有者和使用者可以按各自的习惯对资源命名, 任何一方对资源名称的修改都不会影响另一方对资源的使用。

④ 以服务方式封装了对用户和应用访问数据资源请求的响应过程, 可以方便地构造统一的数据访问接口, 还降低了集成遗留系统的复杂度。

⑤ 以所有者为中心的基于角色的层次访问控制策略, 将授权的过程分成两个阶段进行, 这样, 不同教学区之间异构的安全机制就不需要转换, 还降低了授权的工作量。

⑥ 用户管理采取分教学区分布式管理的模式, 将登录分为本地登录和远地登录两种, 降低了构造统一身份认证机制的难度, 避免了建立全局统一用户数据库的麻烦。

结束语 基于网格计算的远程教育系统模型具有较高的研究价值, 随着网格技术的发展及其应用的逐渐深入, 我们深信, 在不久的将来它必将获得飞跃的发展, 并为网络远程教育注入新的活力。

参考文献

- 1 徐志伟, 冯百明, 李伟. 网格计算技术[M]. 北京: 电子工业出版社, 2004
- 2 Foster I, Kesselman C. 金海, 等译. 网格计算(第二版)[M]. 北京: 电子工业出版社, 2004
- 3 Welch V, et al. Security for Grid Services. <http://xxx.lanl.gov>
- 4 Baker R, et al. A Model for Grid User Management [J]. Computer in High Energy and Nuclear Physics, California, March, 2003
- 5 McNab A. Grid based Access Control for Unix environments, Filesystem and Web Sites. <http://.xxx.lanl.gov>

(上接第 292 页)

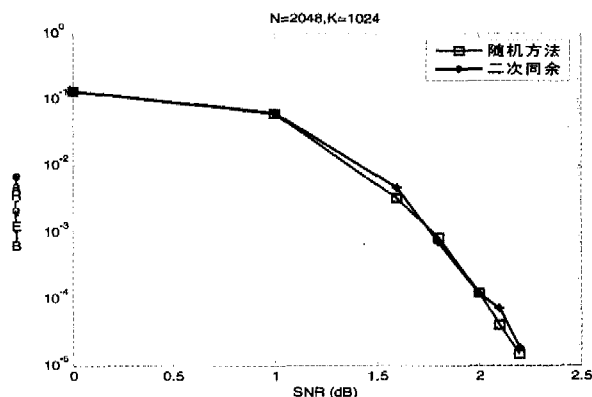


图3 长度 $N=2048, K=1024$ 的性能曲线(30次迭代)

参考文献

- 1 Gallager R G. Low Density Parity Check Codes [J]. IRE Trans on

Information Theory, 1962, 8(3):208~220

- 2 Gallager R G. Low Density Parity Check Codes [M]. Cambridge, Mass: MIT Press, 1963
- 3 Berrou C, Blavieux A, Thitimajshima P. Near Shannon Limit Error Correcting Coding and Decoding: Turbo Codes [A]. In: Proc. 1993 IEEE International Conference on Communications [C]. Geneva, Switzerland, 1993. 1064~1070
- 4 Mackay D J C. Good Error Correcting Codes Based on Very Sparse Matrices [J]. IEEE Trans on Information Theory, 1999, 45(2):399~431
- 5 Richardson T, Urbanke R. Capacity of low-density parity-check codes under message passing decoding. IEEE Trans. Inform. Theory, Feb. 2001, 47:599~618
- 6 Prabhakar A, Narayanan K. Pseudorandom Construction of Low-Density parity-check codes using Linear Congruential sequences. IEEE. Trans on commu. Sep. 2002, 50(9)
- 7 Teng H, Xu J, Lin S, Abdel-Ghaffar A S. Codes on Finite Geometries. IEEE Transaction Theory, Feb. 2005, 51(2):572~596