

计算网络的资源管理方法研究^{*}

武秀川^{1,2} 鞠九滨¹

(吉林大学计算机学院 长春130012)¹ (烟台大学计算机学院 烟台264005)²

Research of Methods for Computational Grid Resource Management

WU Xiou-Chuan^{1,2} JU Jiu-Bin¹

(School of Computer, Jilin University, Changchun 130012)¹ (School of Computer, Yantai University, Yantai 264005)²

Abstract From the perspective of resource sharing, computational grid is a system ranging from small kind of network system for home using to large-scale network computing systems even to the Internet. The management of resources in the grid environment becomes very complex as these resources are distributed geographically, heterogeneous in nature, owned by different organizations and managing fields, and each having their own resource management policies and different access as well as cost models. In this paper computational grid resource management policies were analysed and discussed.

Keywords Grid, Computational grid, Resource management

1 引言

计算网络^[1,2]被定义为一个通用的、大规模的计算处理虚拟系统,它由若干个分布式资源所组成。它能够由分布式资源所提供的不同的服务。计算网络的资源包括存储资源(如内存、磁盘、磁带等,也包括存储带宽和存储空间)、计算资源(如CPU、GPU)、网络资源(HPCswitch, LAN和WAN)以及其他资源(如客户账号、进程)等。资源管理系统从逻辑上把这些资源耦合起来作为一个单独的集成资源提供给用户,而用户与资源代理进行交互,从而该代理向用户隐藏了网格资源使用和网格计算的复杂性。

从资源共享的角度来看,计算网络可以是小型的家庭用的网络系统,也可以是大规模的网络计算系统,从理论上讲也可以扩大到因特网。由于计算网络的资源在地理上是分布的,分布于多个组织和管理域,本质上是异构的,并且各个组织和管理域有各自的资源管理策略和不同的访问代价模型,因此计算网络的资源管理必须要处理系统的边界问题,需要有助于安全和容错的特殊机制。大规模的计算网络系统的资源管理极其复杂。

本文对计算网络的资源管理策略进行深入地分析和研究。

2 计算网络资源管理系统的功能

计算网络资源管理系统的基本功能是接受来自计算网络范围内的机器的资源请求,并且把特定的资源分配给资源请求者。合理地调度相应的资源,使请求资源的作业得以运行。为实现上述功能,一般而言,计算网络的资源管理系统应提供三个基本的服务^[7,9]:即资源分发、资源发现和资源的调度。资源发现和资源分发提供相互补充的功能。资源分发由资源启动且提供有关机器资源的信息或一个源信息资源的指针并试图去发现能够利用该资源的合适的应用。而资源发现由网络

应用启动并在网格范围内发现适于本应用的资源。通常通过使用一个资源信息协议来发布计算网络环境中的资源信息,而资源信息协议由资源和分发功能所实现。资源分发和资源发现是资源调度的前提条件,资源调度实施把所需资源分配到相应的请求上去,包括通过不同结点资源的协作分配。

3 计算网络资源管理所涉及的若干问题

计算网络的资源管理方法涉及到很多方面^[9],如计算网络内部的机器组织、资源模型、分发协议、名字空间组织、数据存储组织、资源发现、调度组织、调度策略、状态评估以及对资源的再调度等,也包括QoS支持^[8,11]。

3.1 计算网络的资源模型和机器组织问题

资源模型决定了应用和资源管理系统如何描述和管理网络资源。不同的资源模型实现了对资源的数据和对资源操作的不同描述及表示方法。

计算网络的机器组织结构影响资源管理系统的通讯模式,因此也决定了计算网络综合体系结构的可扩充性问题。在集中式的机器组织结构中,单个的或一组控制者为所有的机器完成调度,因而这种方法在计算网络系统中是不适用的,因为集中式的这种机制面临可扩充性的问题。如果采用分布式的组织结构,则调度可被分布到计算网络系统中的机器上。

文[9]把计算网络系统环境中的机器组织划分为三类:平面(flat)、单元格(cell)(又划分为平面单元格和层次单元格)和层次结构等三类。机器组织描述了包含在资源管理系统中的机器如何做出调度决策及这些机器之间的通讯结构和机器在调度决策中所起的作用。在平面组织结构中所有的机器能够直接与其它机器通讯而不必要通过一个中介物。层次体系结构中,其结构对于在计算网络中的所有元素来说都是可见的,在同层的机器能够与其直接的父结点机和直接的子结点机或者同级进行直接的通讯,因此具有较好的可扩充性。在单元格的体系结构中,单元格内的机器通过使用平面结

^{*} 本文受到国家自然科学基金项目(69873018)的资助。武秀川 博士研究生,副教授,研究领域为分布式系统;鞠九滨 教授,博士生导师,研究领域为分布式系统和计算机网络。

构进行彼此之间的通讯,与层次结构之间的主要不同是单元格结构具有一个指定的边界,并且该边界具有一个隐藏的内部结构。

3.2 计算网格的资源名字空间的组织问题

资源名字空间的组织影响资源管理协议的设计和资源发现的方法。如在一个名字空间中,为了减少系统信息的冗余搜索,发现资源的代理要求对某类全局搜索空间进行划分。

名字空间组织可以划分为关系式的、层次结构式的和图解式的三种^[6,9]。关系式的名字空间把资源划分为关系并且使用来自于关系数据库的概念以指出在不同关系中的元组(tuple)之间的关系。层次式的名字空间把计算网格中的资源划分为层次结构,围绕着计算网格的物理或逻辑网络结构层次进行组织。关系/层次名字空间的综合组成了关系,在此关系的内容被分解为层次结构,目的是为了把他们分布在跨越计算网格的机器上。大多数基于名字空间的网络目录使用综合的名字空间组织结构。而基于图表式的名字空间使用结点和指针,以指向所要访问的计算资源^[9]。

3.3 计算网格资源信息的存储组织问题

资源描述、资源的状态及资源的预留数据存储组织为资源管理的总的性能赋予特征。该资源信息存储组织决定实现资源管理协议的代价,因为资源分发和资源的发现可能由数据存储实现来提供。资源信息存储组织可以划分为两大类即网络目录和分布式对象方式。它们的重要区别^[9]是在网络目录中模式和操作被数据存储模式所进行的外部定义所割裂开。而在面向对象的方法中,是由模式定义数据和操作。如 Globus 中的 MDS 组件基于 LDAP 提供对系统状态和资源信息的分布式访问。而在 Leigon 中每个资源用一个对象来表示,所有这些 Legion 对象导出一组对象代理成员函数,由这些函数实现关键的 Legion 服务如信息提供服务。

3.4 计算网格资源管理的 QoS 支持问题

QoS 表示为获得应用的要求功能所必需的一组定量和定性的特征。在网络中,QoS 是一组工具和标准,它赋予网络管理者控制带宽、延迟及其抖动和包丢失的综合能力^[3]。在通常情况下 QoS 只考虑网络的情况,而在计算网格系统中,QoS 不能仅被限制在网络的带宽还应扩展到计算网格中节点的处理能力和存储能力上。因此,计算网格的资源管理系统也能够提供端一端的 QoS 并能保证在网络连接上的 cpu 处理能力^[11]。

3.5 计算网格资源管理的体系结构模型问题

在大型计算网格系统中通常采用层次结构模型。文[6]把主要的部件划分为被动式部件和主动式部件。主动式部件是资源。任务是资源的消耗者,包括传统的计算任务和非传统的计算任务(如文件的分离和通讯)。作业是层次结构的实体,作业由于作业和任务组成。最简单的作业形式是包含一个单个的任务。调度是把任务映像到资源而不是把作业映像到资源,因为作业是任务的容器^[6,9,11],任务是实际的资源消耗者。被动式部件是调度程序,为作业的输入列表计算一个或多个调度。调度者的单元是作业,意味着调度程序试图同时映像在一个作业中的所有任务,作业而不是任务被提交到调度程序。信息服务如 LDAP 和一个商业数据库等都可以是被动部件。

3.6 计算网格资源管理中的资源调度问题

资源管理系统的有效性和可接受性在很大程度上依赖于所实现的资源调度系统。也与机器的组织结构密切相关。计算网格系统的资源是地理上分布的并属于几个不同的组织和拥

有者,而大多数的本地机器都有不同的调度策略,并且在大的计算网格系统中常遭受任意时间资源加入或退出的频繁的变化。计算网格的动态性和异构性决定了资源调度系统的复杂性。

通常采用集中式、分布式以及层次式的和计算经济的调度方法^[10]。在集中式的环境中,所有的资源由一个中央调度程序调度,所有可用系统的有关的状态信息被聚集在该中心机上。从概念上讲这种调度模式是非常有用的,例如在计算中心,因为所有的资源在相同的目标之下被使用。在这种情况下,即使缺乏网络带宽也可以被忽略。但它形成了一个瓶颈,使得当计算网格的大小规模扩充时是很困难的。在分布式的系统中,分布式的调度程序交互作用并且提交作业到远程系统中,没有一个中央调度程序负责作业的调度,所有系统的状态信息也没有被集中在单个的结点上,因此消除了集中式调度的通讯瓶。同样单个组件的失效将不会影响整个的计算网格系统,容错和可靠性更高。但由于一个并行程序的所有部分可能被分配在不同域的资源上,因此在相同的时间必须是活跃的,所以不同的调度程序必须同步作业并且保证同时运行,这使得调度系统的优化相当困难。在层次式的调度模型中,有一个集中式的调度程序,作业被提交到该集中的调度程序,而每一个机器资源使用一个独立的调度程序用于本地的调度。这种结构的主要优点是不同的策略能够被使用来用于本地和全局的作业调度。如 Globus 中的资源管理子系统 GRAM 根据用户的请求在网格范围内寻找最符合用户要求的一个或一组机器,然后实现资源的定位及分配,最后由本地资源的调度系统如 Loadleveler 等与用户进行交互。基于计算经济的调度模型的优越性是可调节供需矛盾^[4],因为做出资源调度策略决定的过程被分布在所有的用户和资源拥有者身上,因此使得调度由以系统为中心转向了以用户为中心,用户可以自己做出决定以最小的代价获取最好的性能。另外也可以帮助开发者开发调度策略,进而建立一个高度可扩充的系统。该模型最适合于计算网格的动态和异构特征,但首先必须寻找一个资源供应和要求函数,但是独立的供应和要求函数即使是在最好的情况下也很难度量。

结论 资源管理是计算网格系统的核心问题,GGF 调度工作组的网格资源调度协议提出了进行资源管理的10条要求。不失一般性,它对于计算网格也是完全适用的^[12]。文[6]提出了另外一种计算网格资源管理的体系结构模型即资源抽象拥有者 AO(Abstract Owner)模型。该模型主要集中于提供一种更为先进的作业提交和结果搜集的新方法,最终目的是实现更高层次上的资源共享方法。文[7]提出了基于服务的计算网格资源管理方法,该方法类似于计算门户的概念。该方法中的计算资源都不能被直接访问,用户仅被允许提交服务请求和输入数据,间接地通过一组预先安装的应用来访问计算资源,其结果是恶意用户不再可能把恶意的执行程序提交到计算网格系统中,从而确保了系统的安全。文[12]的资源管理协议基于数据报,该数据报在计算网格的结点之间被交换,其实类似于因特网的路由器—分配器,并不为通过它们的任一资源请求维持状态信息。资源状态和调度信息被维持在终端系统而不是维持在路由器—分配器中。为使不同类型的资源管理具有很强的可适应性还需要解决资源的再调度问题^[9]。从资源管理的角度看,资源预留也被考虑作为计算网格系统基本的 QoS 属性^[3],可通过或者使用应用过去行为或者有

(下转第104页)

有如下优点:①该算法所得出的评估结果与人的正常推理和经验值都较为吻合;②该算法并不是基于网络中某一条入侵路径进行量化评估,而是基于起始点和目标点中的整个网络可达路径集进行量化评估;③该算法的评估结果可以实时反映当前网络的安全状态值。无论是路径或结点的增加或删除,还是结点配置的改变,网络的安全状况都会通过该算法有所体现。

该算法的缺点是:当网络中主机数目达到上百台时,在生成网络入侵状态图时可能会产生状态爆炸问题,影响了网络安全评估的正确性及有效性,目前正在寻找状态爆炸问题的新的解决方案。

7 系统的进一步研究方向

该系统得到了充足的预研经费支持,目前在实验室环境下能够初步进行网络安全信息的收集、分析及评估,基本上取得了预期的结果。它大大节省网络安全管理员的时间与精力,对于发现潜在的网络脆弱性也是十分有益的。本系统的各个子模块以及相关的数学算法还在不断的完善中,一些功能实现还存在不足之处,系统的智能功能、自动化功能以及显示功能也需要进一步加强,为了进一步推动该课题的进展,以下几个问题将是我们下一步急需研究和解决的:①进一步研究分析如何提取网络安全量化评估中的各种量化指标要素,使其具有科学性和有效性;②进一步研究网络安全脆弱性之间的依赖关系,从条件概率的角度结合证据推理理论建立数

学模型,更加准确地评估漏洞存在的可能性;③结合人工智能的相关技术,进一步提高系统的智能性;④加快研发系统软件,增强系统的自动化和图形显示功能。

参考文献

- 1 Thorhuus R. Software fault injection testing:[master thesis]. Ericsson Telecom, Stockholm, Sweden, 2000
- 2 Sample C, et al. Quantifying Vulnerabilities In The Networked Environment: Methods and Uses. TISC, 2000
- 3 Shake T, Hazzard B, Marquis D. Assessing Network Infrastructure Vulnerabilities to Physical Layer Attacks. MIT Lincoln Laboratory. In: 22nd National Information Systems Security Conf. 1999
- 4 Ortalo R, et al. Experimenting with Quantitative Evaluation Tools for Monitoring Operational Security. LAAS Report 96369, 1997
- 5 Dacier M, et al. Quantitative Assessment of Operational Security: Models and Tools. LAAS Research Report 96493, 1996
- 6 Shafer L R. Implementing Dempster's rule for hierarchical evidence. Artificial Intelligence, 1987, 33: 271~298
- 7 Saaty T L. How to make a decision: the analytic hierarchy process. European Journal of Operational Research, 1990, 1 (48): 9~26
- 8 Barker R, Kelly G. The Vulnerability Instantiation Methodology prototype. Information Security technical Report, 1998, 13: 79~86

(上接第88页)

关类似的输入值或者通过使用启发式的方法实现其预测。但由于在计算网格这样一个高度动态的环境中计算网格资源信息提供资源管理系统的资源信息可能已经是过时的信息,因此资源管理系统根据此信息做出的资源预留和 QoS 性能严重下降,并且很有可能产生结点的拥塞。主动式网络中的 QoS 监控和 QoS 适应性能够预料到过载的情况并及时做出分流从而导致全局 QoS 和资源管理系统性能的进一步改善。

参考文献

- 1 Baker M, Buyya R, Laforenza D. The Grid: International Efforts in Global Computing. In: Intl. Conf on Advances in Infrastructure for Electronic Business, Science, and Education on the Internet (SSGRR 2000), l'Aquila, Rome, Italy, July 31 - August 6. 2000
- 2 Foster I, Kesselman C, eds. The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufman, 1999
- 3 Foster I, Roy A, Sander V. A Quality of Service Architecture that Combines Resource Reservation and Application Adaptation. In: 8th Intl. Workshop on Quality of Service (IWQOS 2000), June 2000
- 4 Buyya R, Abramson D, Giddy J. An Economy Grid Architecture for Service-Oriented Grid Computing. In: 10th Intl. Heterogeneous Computing Workshop (HCW 2001) (In conjunction with IPDPS 2001), San Francisco, California, USA
- 5 Kon F, Campbell R, Mickunas M, Nahrstedt K. 2K: A Distributed Operating System for Dynamic Heterogeneous Environments. In: 9th IEEE Intl. Symposium on High Performance Distributed computing (HPDC'9) August 2000
- 6 Buyya R, Chapin S, DiNucci D. Architectural Models for Resource Management in the Grid. In: First IEEE/ACM International Workshop on Grid Computing (GRID 2000), Springer Verlag LNCS Series, Germany, Dec. 17, 2000, Bangalore, India
- 7 Gehring J, Streit A. Robust Resource Management for Metacomputers. In: 9th IEEE Intl Symposium on High Performance Distributed Computing, 2000
- 8 Maheswaran M, Krauter K. A Parameter-based Approach to Resource Discovery in Grid Computing Systems. In: 1st IEEE/ACM Int'l Workshop on Grid Computing (Grid'00), Dec. 2000
- 9 Krauter K, Buyya R, Maheswaran M. A Taxonomy and Survey of Grid Resource Management Systems. Software Practice and Experience, 2002, 32(2): 135~164
- 10 Krauter L, Maheswaran M. Architecture for a Grid Operating System. <http://www.cs.umanitoba.ca/~anrl/PUBS/grid2000>
- 11 Krauter K, Maheswaran M. Towards a High Performance Extensible Grid Architecture. <http://citeseer.nj.nec.com/410299.html>
- 12 Czajkowski K, Sander V. GGF Scheduling Working Group Sched-WD 12. 1. Grid Resource Management Protocol: Requirements. Sept. 11, 2001