

# 网格计算基本原理及其系统管理关键技术问题

林晓鹏<sup>①②</sup>, 李晓潮<sup>③</sup>, 郭东辉<sup>②③</sup>

①厦门海洋职业技术学院, 福建 厦门 361012; ②厦门大学 物理系, 福建 厦门 361005;

③厦门大学 电子工程系, 福建 厦门 361005)

**【摘要】** 网格计算可以充分利用网络资源, 为海量计算用户提供集成方便的数据处理服务, 文章从网格计算的基本概念和工作原理介绍入手, 分析了资源管理、任务调度、服务质量和安全性等网格计算的关键技术问题, 并对其技术应用与发展现状进行总结。

**【关键词】** 网格计算; 关键技术; 资源管理

**【中图分类号】** TP309

**【文献标识码】** A

**【文章编号】** 1002-0802(2009)02-0218-03

## Principles and Key Technologies of Grid Computing

LIN Xiao-Peng<sup>①②</sup>, LI Xiao-Chao<sup>③</sup>, GUO Dong-Hui<sup>②③</sup>

①Xiamen Ocean Vocational College, Xiamen Fujian 361012, China;

②Department of Physics, Xiamen University, Xiamen Fujian 361005, China;

③Department of Electronic Engineering, Xiamen University, Xiamen Fujian 361005, China)

**【Abstract】** Grid computing can improve the utilization of resource, provide integrated services for customer. Based on description of the concepts and principles of grid computing, the key technologies and progresses of grid computing are analyzed, including resource management, job scheduling, quality of service, and security.

**【Key words】** grid computing; key technologies; resources management

### 0 引言

随着计算机网络技术的应用和普及, 网络设备的数量和种类越来越多, 网络的规模不断扩大, 但许多网络资源(如计算能力、存储空间等)日的利用率很低<sup>[1]</sup>, 然而在一些应用领域如气象模拟、设计仿真等, 由于要处理海量数据, 往往对计算能力和存储空间等又有很大的需求。因此, 对现有资源进行整合和管理, 为用户提供聚合、安全和可靠的服务是目前信息行业的一个重要研究课题<sup>[2]</sup>, 网格计算技术发展和推广正是符合了解决这方面技术问题的需求。

网格计算可以说是一种为用户提供统一、标准、可靠的网络计算服务, 其目的是充分挖掘现有网络资源, 即希望将数据、计算能力和仪器设备等各种资源整合成一个与地理位置的集成环境。它突破了传统网络的限制, 实现了网络资源的全面共享, 不仅在科学计算领域得到应用, 而且在社会经济领域也有广阔的应用前景。但是目前网格计算还有大量的技术问题有待于解决, 如: 如何有效地管理数量多、种类繁

杂的各种资源, 实现异构网络的资源共享; 如何实现任务的分解与调度, 并为各个子任务选择最适合的资源<sup>[1]</sup>; 以及网格计算的安全机制、服务质量(QoS)控制<sup>[3]</sup>等技术细节问题。

为此, 本文介绍网格计算的基本概念和工作原理入手, 具体分析网格计算的资源管理、任务调度、QoS控制和用户安全等关键技术, 总结其技术应用与发展现状, 为下一步基于网格计算的应用研究提供理论基础和研究思路。

### 1 网格计算基本原理

网格计算研究的先驱者 Ian Foster 认为网格计算必须同时满足以下三个条件<sup>[4]</sup>: 在非集中控制的环境中协同使用资源; 使用标准的、开放、通用的协议和接口; 提供非平凡的服务, 随后在网格计算中引入 VO<sup>[5]</sup>(Virtual Organization)的思想, 认为网格计算是由网络资源组成的一个虚拟的整体, 在此基础上为用户提供服务, 协同解决问题, 是一种在动态多机构虚拟组织内实现资源协调共享和解决问题的过程。

收稿日期: 2008-11-10。

作者简介: 林晓鹏(1972-), 讲师, 主要研究方向为网络资源管理。

提供资源的网格计算节点、网格用户以及用户与节点之间的网格中间件构成了网格计算的基本组件，其工作原理如图 1 所示。网格计算节点可以是高性能的服务器、集群、多个机器组成的网络，也可以是工作站或其它仪器设备。各计算节点向网格信息服务 (GIS: Grid Information Server) 注册，发布各自节点的资源信息。各网格计算节点上的本地资源管理器负责管理本地资源，并与网格中间件中的资源管理器协商资源的使用，本地任务管理器负责管理子任务在本地的执行。

网格用户通过网格计算中间件提交任务，获得网格计算服务。网格计算中间件由任务控制管理、资源索引、资源管理、任务调度器、任务执行器等组成。任务控制管理组件负责网格计算与用户之间的交互，将用户任务提交给任务调度器；任务调度器将用户任务分解成一系列的子任务，并将任务描述转换成资源需求信息，并通过资源定位从网格信息服务 GIS 中获得合适资源的具体信息，并通过资源管理组件与资源节点进行协商获得资源，然后通过任务管理将各子任务及对应的资源信息交给任务执行器；任务执行器根据资源协商的情况将各个子任务部署到相应的资源上执行并对管理任务的执行，收集任务执行的结果并通过任务控制管理传递给用户。

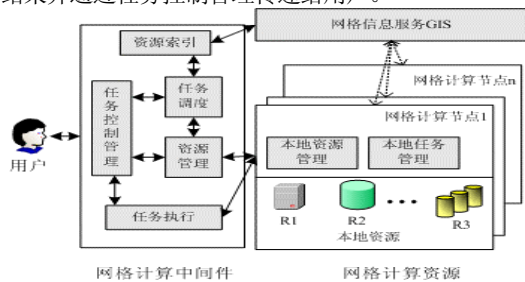


图 1 网格计算的工作原理

## 1.2 网格计算的特点

从工作原理可知，网格计算是将网络虚拟成一个“超级计算机”为用户提供服务，需要对大量的、异构的、广域分布的网络资源进行整合，实现资源共享和任务协作。虽然网格计算中的资源情况很复杂，具有异构、动态、自治等特点。但从逻辑上看，网格计算应当形成一个统一的整体，用户可以方便地使用网格计算服务而不需考虑资源的具体位置和网格计算服务具体的实现方式，需要对多样、异构的资源进行抽象，形成统一、用户可见的服务能力，在资源管理与分配、任务调度与管理、安全机制以及 QoS 等方面需要解决大量的技术问题。

## 2 网格计算的关键技术

为了实现虚拟组织内部的资源共享和任务的协同解决，为用户提供集成的服务平台的目标，在网格计算环境中，首先必须考虑资源的有效管理，并在此基础上，实现用户任务的管理，因此资源管理和任务管理是网格计算的关键技术。此外，服务质量控制机制和安全机制是网格计算服务的基础和有效保证，也是需要引起关注的重要问题。

### 2.1 资源管理

资源是网格计算服务的基础，是实现网格计算服务极为

关键的部分，因此资源是网格计算研究的主要对象。资源管理系统包含的基本操作有：资源信息的收集、更新和发布；根据资源信息和用户的资源需求描述，为用户分配合适的资源或预留资源；协调资源的共享使用；在任务运行时监测和收集资源使用情况等动态信息，监控资源的使用情况等。

目前资源管理模型主要有分层式模型、抽象所有者的模型和基于经济理论的模型等。分层式模型是借鉴现有网络管理的方式对资源进行整合，可以充分应用现在的网络技术，但要实现跨区域资源共享协作仍存在许多困难；抽象所有者的模型是以“订购”的方式提供服务，是一种理想化的模型，还缺乏具体的实现方式；基于经济市场的模型在资源使用和任务调度中引入对服务成本和代价的约束控制，能较好地符合网格商业化运作的要求，正成为网格资源管理的研究热点。基于经济理论的资源管理模型主要有基于一般均衡理论的资源管理模型和基于纳什均衡<sup>[6]</sup>的资源管理模型。前者如商品市场模型、拍卖模型等<sup>[7]</sup>，后者运用博弈理论<sup>[8]</sup>分析用户与用户之间，以及资源与资源之间的合作与竞争关系，更符合网格计算的实现情况，但其研究还需要进一步深入。

### 2.2 任务调度

网格计算服务是通过用户对任务的执行来实现，任务调度是网格计算的另一个关键问题。首先，任务调度的基本功能包括任务提交、评估完成任务所需要的资源、任务分配、任务修订、任务查询、任务运行监测、任务迁移和任务删除等。其次，在网格计算中通常将用户的任务分解成一系列的子任务后分配到适合的资源并行执行，因此任务调度需要考虑子任务的粒度大小和相互之间的约束关系。第三，各个子任务之间要处理大量的静态或实时数据，需要解决数据的存储、传输和管理等问题，在 Globus 中用网格文件传输协议 GridFTP、GASS 等子模块来实现数据的传输、访问和副本的管理等功能。此外，还需要考虑子任务之间的通信等因素。

任务调度系统主要有集中式任务管理、分布式任务管理和混合式任务管理这三种模型。集中式任务管理由一个全局任务管理器进行任务的统一调度，如 LSF (Load Sharing Facility)、PBS (Portable Batch System) 等，这类模型中全局任务管理器往往成为性能瓶颈，不利于系统的扩展。分布式任务管理系统采用多个分布的局域任务管理器进行任务的调度，但由于局域任务管理器只有部分资源的信息，可能导致负载的不均衡。混合式模型结合全局任务管理器和局部任务管理器完成任务的调度，可以较好地应用于网格计算中，但难以满足快速反应的要求。

### 2.3 QoS 控制

在网格服务过程中，不同的用户对服务的功能、性能、成本有着不同的考虑，对网格计算服务有不同的需求，而用户的需求往往必须通过 OoS 参数来描述，但不同的网格应用有不同的性能评价标准，因此网格环境中的服务质量控制和传统网络有很大的差异。此外，由于网格计算中的资源是动态变化的，在任务调度时或任务执行过程中，资源状态可能

随时发生变化,而资源状态的变化必然对任务的执行产生影响,从而最终影响到服务质量。可见 QoS 控制的相关技术在网络计算中也是发挥着基础性的作用。目前针对 QoS 的层次结构、描述、类型、预留和实现等也已有了相应的技术研究,但也存在层次比较模糊、QoS 参数之间缺乏合理的映射方案等问题。

## 2.4 安全机制

网络计算的安全性主要指访问控制和通信安全两个方面。由于网络计算环境比传统的计算机网络更为复杂,如网络计算中资源的多样性、异构性和动态性,存在着多种通信机制;资源可能属于不同组织或个人,因此有不同的访问控制机制。网络计算中既要保持资源的自治性,又要进行统一管理、协作组合而不产生安全漏洞,因此网络计算的安全性比传统网络面临更大的挑战。网络计算的安全主要考虑的问题有:认证要求,如用户认证、资源认证、协同认证等;授权要求,如授权、代理等;通信安全要求,如信息加密等。常用的网络计算安全技术有 Kerberos、PKI (Public Key Infrastructure) 和 SSL/TLS (Secure Socket Layer/Transport Layer Security)。安全性虽然可以保障计算和数据节点的安全,但也会制约网络和用户规模的扩展,增加网络使用成本,因此研究高效的安全机制是网络系统的难点之一。

除了以上的几个方面的关键技术问题有待进一步研究外,网络计算还须解决网络计算应用的开发工具,以及需要考虑应用工具图形化等有利于简化用户操作的技术问题。

## 3 网络计算的应用与发展现状

网络计算的思想自 20 世纪 60 年代就已经提出,但到 90 年代以后才开始真正的发展。上世纪 90 年代初,针对网络上主机数量增加但利用率不高的情况,美国国家科学基金会 (NFS) 将 4 个超级计算中心构建成一个能够进行元计算 (Meta Computing) 的整体,为重大科研领域的研究提供计算资源,解决科学与工程计算问题,此后相关的项目有 I-WAY<sup>1</sup>和 FAFNER<sup>1</sup>。90 年代中期,网络计算研究内容主要关注网络计算的中间件研究与开发,以解决不同节点和资源有效共享和协同工作的问题,如 Globus 和 Legion 项目基础探讨了基本系统级网络基础设施的开发,Condor 项目研究了高吞吐率调度问题,其他一些项目研究了异构数据资源的统一访问、远程计算等问题。2000 年后,网络计算关注的是在动态、多机构的虚拟组织中协调资源共享和协同解决问题,提出了开放网格服务结构 (OGSA: Open Grid Services Architecture),并逐步与 Web 服务相结合,支持 XML、SOAP、UDDI 等标准,形成 Web 服务资源框架 WSRF (Web Services Resource Framework)。

当前网络计算已不仅仅局限于高性能科学计算,出现了许多适用于不同应用的网络计算研究领域,如:信息网络、数据网格、制造网格、教育网格等。网络计算已得到学术界和产业界的广泛关注,有多个国家的机构与组织参与到该领

域中,并产生了全球性网络论坛 GGF。美国的网络计算研究处于世界前列,其多家机构参与网络计算的研究, Globus、Legion 和 Condor 等具有影响性的网络计算软件和工具就来自于此。欧洲的网络计算研究也非常活跃,已启动了多项研究项目,其中欧洲数据网格 (European Data Grid) 是由欧盟支持的一个项目,目的是将其建设成下一代的计算基础设施。日本是亚洲中较早开展网络计算研究的国家,已进行了 Ninf 等多个项目的研究。印度、韩国等国家也相继开展了网络计算项目的研究。

近几年来,网络计算研究在我国得到迅速发展,设立了网络计算研究的重大专项,如:用于高性能计算的国家高性能计算环境 (NHPCE)、CNGrid、中国教育科研网格 ChinaGrid、E-Science 网络研究计划、上海网格和中国空间信息网格等。

## 4 结语

网络计算的目的是通过共享网络中的资源,协同解决问题,向用户提供服务,并产生了多个相应的应用领域,虽然在资源管理、任务管理等方面还面临技术挑战,但网络计算的应用必将对人们生活和工作方式产生巨大的影响,并带来信息产业的新市场。本文分析了网络计算的工作原理,指出网络计算的异构、动态、自治等系统特点,说明网络计算在资源管理、任务管理、安全机制、QoS 控制等技术方面的关键问题,并介绍了网络计算的关键技术进展对推动网络计算的应用领域拓展所起的作用,同时概括当前国内外网络计算应用发展现状,为基于网络计算的应用研究提供方向性指导和相应参考文献。

## 参考文献

- [1] Foster I, Kesselman C. The Grid: Blueprint for a new Computing Infrastructure[M]. Morgan Kaufmann, San Fransisco, CA, 1999.
- [2] 郁志辉,陈渝,刘鹏. 网络计算[M]. 北京:清华大学出版社,2002.
- [3] Bart Jacob, Michael Brown, Kentaro Fukui, Nihar Trivedi. Introduction to Grid Computing[C]. <http://www.redbooks.ibm.com/redbooks/pdfs/sg246778.pdf>.
- [4] Ian Foster. What is the Grid? A Three Point Checklist[J]. GRIDToday, July 20, 2002.
- [5] Ian Foster, C Kesslman, J M Nick, S Tuecke. The Physiology of the Grid—An Open Grid Services Architecture for Distributed System Integration[J]. In Processings: Open Grid Service Infrastructure WG, Global Grid Forum, 2002. <http://www.globus.org/alliance/publications/papers/ogsa.pdf>.
- [6] Nash JF. Non-Cooperative games. Annals of Mathematics[J], 1951, 54(2):286-295.
- [7] Buyya R, Abramson D, Giddy J, et al. Economic Models for Resource Management and Scheduling in Grid Computing[J]. The Journal of Concurrency and Computation: Practice and Experience (CCPE), Wiley Press, May 2002.
- [8] Osborne M J. An Introduction to Game Theory[M]. Oxford, 2004.