

网 格 计 算

将全世界的计算机

联合起来

1. 从一个试验项目开始

1999年5月,由美国加州大学伯克利分校发起的名为SETI@home的项目启动了。

SETI@home是Search for Extra Terrestrial Intelligence at Home的缩写,意为:在家里寻找外星文明。其核心思想是利用联接到Internet上PC的闲置能力分析世界上最大的射电望远镜获得的数据,以帮助科学家探索外星生物。该项目启动以来,现已有300万志愿者参加了这个项目,他们从指定的站点下载射电望远镜收集的信息的片断,用自己的计算机运行分析,从中寻找宇宙中生命的迹象,总处理数据量已经达到了15T,平均每位参与者让自己的电脑为SETI@home工作了17个半小时,这相当于一台PC机工作482,023年,相当于世界上最快的超级计算机工作48年。这个项目充

分利用了分布在世界各地计算机的力量。

SETI@home项目的大致流程是:研究部门将一项需要巨大运算量的任务以程序和数据的形式提交给服务器;服务器将数据和程序代码分成更小的部分,也称“子任务”;PC机安装一种特殊的客户程序,它们能自动同服务器联络,自动下载和处理子任务;子任务处理完后的结果被送回服务器,然后,客户程序下载新的子任务;一旦所有的子任务处理完毕,服务器就将各种结果汇总,生成最后的报告,并把最终结果发回提交人。

这是一种造价低廉而数据处理能力超强的计算模式,计算机科学家将这种全新的计算模式称为网格计算(Grid Computing)。

网格计算是伴随着互联网而迅速发展起来的,专门针对复杂科学计算的新型计算模式。这种计算模式是利用互联网把分散在不同地理位置的电脑组织成一个“虚拟的超级计算机”,其中每一台参与计算的计算机都是一个“节点”,而整个计算是由成千上万个“节点”组成的“一张网格”,所以这种计算方式叫网格计算。这样组织起来的“虚拟的超级计算机”有两个优势:一个是数据处理能力超强;另一个是能充分利用网上的闲置处理能力。

随着超级计算机的不断发展,它已经成为复杂科学计算领域的主宰。但超级计算机造价极高,通常只有一些国家级的部门,如航天、气象等部门才有能力配置这样的设备。而随着人们日常工作遇到的商业计算越来越复杂,人

■ 笛子

们越来越需要数据处理能力更强大的计算机,而超级计算机的昂贵的价格使得它很难进入普通人的工作领域。

实际上网格计算是分布式计算(Distributed Computing)的一种。网格计算模式首先把要计算的数据分割成若干“小片”,而计算这些“小片”的软件通常是一个预先编制好的程序,然后不同节点的计算机可以根据自己的处理能力下载一个或多个数据片断和程序。只要节点的计算机的用户不使用计算机时,程序就会工作。

网格计算资源都存在于Internet骨干网上,包括几十亿台计算机。Internet重组为混合虚拟计算平台是一个逻辑步骤。在新的Internet环境下,用户只需向网格提交“工作”,网格就会分配相应资源来完成工作,资源是指处理器周期、应用、存储或数据。

网格计算和以往标准的网络计算是有一定区别的。目前网络上的服务都是基于客户机/服务器模式来进行的。这种模式通过确认客户,检查他们的授权级别以决定客户可以在服务器上做什么来确保安全。而对于网格计算,客户机和服务器的功能划分没有这么明确。一台计算机可以在要求另一台计算机去完成一项任务的同时,还为其它的计算机进行着一项任务。用户无需确定这些互动是通过怎样的途径来实现的,也不必要在几百台完成你的工作可能用到的计算机上都去登录,或者去鉴别每一个可能利用你的计算机来完成他们工作的人。甚至可以在其它人的机器上创建一个任务,再由这台机器来向其它的机器进行进一步的任务分配。在开放标准上运行的内容不同,它们对

安全性的要求也是不一样的。在这种情况下,不可能把它作为一个单独的整体来管理:因此网格管理依靠的是社区授权。用户只有成为一个社区的一部分,在这个社区里可以做什么取决于这个社区被允许做什么。

现在大部分网格技术都把应用对象集中在科学和工程研究团体上。IBM已经与Department of Energy合作建立了一个互联的超级计算网络,其中拥有3000多个处理器和巨大的存储容量。英国也有一个类似的工程,它通过各个大学把超级计算机联接成共享资源。

美国国家科学基金会正在建立一台名为“Tera Grid”的超级计算机,它主要应用在研究领域,例如气象模型和生命科学,每秒13.5万亿次的计算速度会有效地推动这类研究。

网格计算不仅受到需要大型科学计算的国家级部门,如航天、气象部门的关注,目前很多大公司也开始推崇这种计算模式。

如美国IBM公司正在构筑一项名为“Grid Computing”的计划,旨在通过因特网,向每一台个人电脑提供超级的处理能力。IBM公司副总裁、也是这项计划的总设计师欧文·伯杰说,网格计算是一种整合电脑资源的新手段,它通过因特网把分散在各地的个人电脑连接起来,不仅可以使每台个人电脑通过充分利用相互间闲置的电脑能源,来提升各自的电脑处理能力,还可以使成千上万的用户在大范围的网络上共享电脑处理功能、文件以及应用软件。正如网络技术总是从科学开发领域转向企业商务领域一样,我们也希望看到网格计算能取得这样的进展。

另一个业界巨人SUN也推出新软件促进网络计算的发展。2001年11月,Sun推出了Sun Grid Engine企业版5.3版软件的β版,继续提升它的网络技术计算水平。至今全球有118000多颗CPU都是采用Sun Grid Engine软件管理的。

除此之外,一批围绕网格计算的软件公司也逐渐壮大和为人所知并成为受到关注的新商机。

有专家预测,网格计算将成为网络市场发展的热点。据《ForbesASAP》预测,网格技术将在2005年达到高峰,并带来Internet的新生。如果网格技术能促使市场按预期的17%年增长率持续成长的话,那么在2020年将会形成一个年产值20万亿美元的大产业。

不过,要使网格计算完全应用到企业或家庭中,仍存在着许多挑战,它包含了许多新的概念。全球化的分布式计算系统的发展还需要一段时间,还有许多的工作要做,但是它将会成为未来的主流。网格可以屏蔽掉建立虚拟资源的复杂性,但它模拟了人类复杂的神经系统,各组件层次之间不同的交互行为必须能自我管理和自我调整。IBM、HP以及其它公司已经开始着手解决计算的自我调整问题,例如自己检测拒绝服务攻击,自动产生适当回复等,但这些问题目前仍处于理论研究阶段。此外,远程操作请求、交换数据,任务管理和安全的标准协议是问题的关键,同时标准的应用程序界面(API)构建、代码库、可重复使用的组件和调试方法也是需要研究的问题。

网格计算技术是一个计算革命,它将全世界计算机联合起来协同工作。网格计算已经被人们视为21世纪的新型网络基础架构。