

引言：网格作为一种能带来巨大处理、存储能力和其他 IT 资源的新型网络，可以应付临时之用。网格计算通过共享网络将不同地点的大量计算机相联，从而形成虚拟的超级计算机，将各处计算机的多余处理器能力合在一起，可为研究和其他数据集中应用提供巨大的处理能力。有了网格计算，那些没有能力购买价值数百万美元的超级计算机的机构，也能利用其巨大的计算能力。

## 一. 网格技术

### 1. 计算的“乌托邦”？

Gartner 公司的 Rob Batchelder 认为，网格的构想一直是计算领域的“乌托邦”，在科技应用上虽有巨大前景，但最大的缺陷是缺乏明显的商业应用。自 20 世纪 90 年代在欧美出现以来，网格主要被用于帮助分散的大学研究人员分析粒子加速器和巨型望远镜的数据。但在过去的两年中，网格的概念和 Globus Toolkit 已在研究和教育领域得到广泛应用，数十项全球性的大项目采用这些技术，以挑战科学计算中的海量计算问题。

目前网格技术虽主要为学术机构所控制，但企业也在陆续跟进。事实上，全球网格论坛（GlobalGridForum）的主要赞助企业就包括 Unilever ——一家以经销肥皂、冰淇淋著称的企业。与许多正在研究和评估网格技术的企业一样，Unilever 自己对于如何利用此技术仍秘而不宣。而 Johnson&Johnson 与 Merck 等制药公司、BMW 与波音等制造企业却已利用这一技术的处理能力和存储空间进行仿真试验，例如药品能否保护细胞免受病毒侵袭？飞机机翼是否会在暴风雨中折断？

基因研究是网格技术的自然应用，这一领域所需的投资很难由一家企业来承担，生物科技企业可用网格技术来分析基因数据；医生可以用网格技术制作出病人器官的三维模型，作为诊断疾病的辅助手段；网格可以处理来自商店现金记录或金融市场的数据流。其他行业，如航空、保险、运输和国防，也会从中受益。如此看来，网格计算并非是可望不可及的乌托邦，其商业应用的广阔前景就在眼前。

## 2. 网格核心技术

为解决不同领域复杂科学计算与海量数据服务问题，人们以网络互连为基础构造了不同的网格，有代表性的如计算网格、拾遗网格、数据网格等，它们在体系结构和需要解决的问题类型等方面不尽相同，但都需要共同的关键技术，主要有如下几种：

**高性能调度技术** 在网格系统中，大量的应用共享网格的各种资源，如何使得这些应用获得最大的性能，这就是调度所要解决的问题。网格调度技术比传统高性能计算中的调度技术更复杂，这主要是因为网格具有一些独有的特征，例如，网格资源的动态变化性、资源的类型异构性和多样性、调度器的局部管理性等。所以网格的调度需要建立随时间变化的性能预测模型，充分利用网格的动态信息来表示网格性能的波动。在网格调度中，还需要考虑移植性、扩展性、效率、可重复性以及网格调度和本地调度的结合等一系列问题。

**资源管理技术** 资源管理的关键问题是为用户有效地分配资源。高效分配涉及到资源分配和调度两个问题，一般通过一个包含系统模型的调度模型来体现，而系统模型则是潜在资源的一个抽象，系统模型为分配器及时地提供所有节点上可见的资源信息，分配器获得信息后将资源合理地分配给任务，从而优化系统性能。

**网络安全技术** 网格计算环境对安全的要求比 Internet 的安全要求更为复杂。网格计算环境中的用户数量、资源数量都很大且动态可变，一个计算过程中的多个进程间存在不同的通信机制，资源支持不同的认证和授权机制且可以属于多个组织。正是由于这些网格独有的特征，使得它的安全要求性更高，具体包括支持在网格计算环境中主体之间的安全通信，防止主体假冒和数据泄密；支持跨虚拟组织的安全；支持网格计算环境中用户的单点登录，包括跨多个资源和地点的信任委托和信任转移等。

## 3. 网格的商业应用

**生物医学：**网格可提供药品开发人员所需的计算能力，用以研究

药物和蛋白质分子的形态与运动。

工程：波音、福特、bmw公司都在尝试用网格计算进行复杂的仿真与设计。

数据搜集 / 分析：制造、石油加工、货物运输、甚至零售企业都要维护昂贵的设备，时常会出现问题，造成不好的结果。同无线传感器一样，网格能够存储和处理所有交易。

娱乐产业：特殊效果设计。

网格作为一个集成的计算与资源环境，能够吸收各种计算资源，将它们转化成一种随处可得的、可靠的、标准的且相对经济的计算能力，其吸收的计算资源包括各种类型的计算机、网络通信能力、数据资料、仪器设备甚至有操作能力的人等各种相关资源。

网格是借鉴电力网的概念提出的，网格的最终目的是希望用户在使用网格计算能力解决问题时像使用电力一样方便，用户不用去考虑得到的服务来自于哪个地理位置，由什么样的计算设施提供。也就是说，网格给最终的使用者提供的是一种通用的计算能力。

电力网中需要有大量的变电站等设施对电网进行调控，相应的网格中也需要大量的管理站点来维护网格的正常运行。网格的结构及资源的调控将更复杂，需要解决的问题也更多。因为网格所关心的问题不再是文件交换，而是直接访问计算机、软件、数据和其他资源。这就要求网格具备解决资源与任务的分配和调度、安全传输与通信实时性保障、人与系统以及人与人之间的交互等能力。网格提供的资源是随时间动态变化的，原来拥有的资源或者功能，在下一时刻可能会出现故障或者拒绝被使用，而原来没有的资源，可能随着时间的进展会不断加入进来。

#### 4. 网格技术在科学计算领域的应用

分布式超级计算。网格计算可以把分布式的超级计算机集中起来，协同解决复杂的大规模的问题。使大量闲置的计算机资源得到有效的组织，提高了资源的利用效率，节省了大量的重复投资，使用户

的需求能够得到及时满足。

高吞吐率计算。网格技术能够十分有效地提高计算的吞吐率，它利用CPU的周期窃取技术，将大量空闲的计算机的计算资源集中起来，提供给对时间不太敏感的问题，作为计算资源的重要来源。

数据密集型计算。数据密集型的问题的求解往往同时产生很大的通讯和计算需求，需要网格能力才可以解决。网格可以药物分子设计、计算力学、计算材料、电子学、生物学、核物理反应、航空航天等众多的领域得到广泛的需求。

基于广泛信息共享的人与人交互。网格的出现更加突破了人与人之间地理界线的限制，使得科技工作者之间的交流更加的方便，从某种程度上可以说实现人与人之间的智慧共享。

更广泛的资源贸易。随着大型机的性能的提高和微机的更加普及，及其资源的闲置的问题也越来越突出，网格技术能够有效地组织这些闲置的资源，使得有大量的计算需求的用户能够获得这些资源，资源的提供者的应用也不会受到太大的干扰。需要计算能力的人可以不必购买大的计算机，只要根据自己的任务的需求，向网格购买计算能力就可以满足计算需求。

## 二. 网格计算

网格计算即分布式计算，是一门计算机科学。它研究如何把一个需要非常巨大的计算能力才能解决的问题分成许多小的部分，然后把这些部分分配给许多计算机进行处理，最后把这些计算结果综合起来得到最终结果。最近的分布式计算项目已经被用于使用世界各地成千上万志愿者的计算机的闲置计算能力，通过因特网，您可以分析来自外太空的电讯号，寻找隐蔽的黑洞，并探索可能存在的外星智慧生命；您可以寻找超过1000万位数字的梅森质数；您也可以寻找并发现对抗艾滋病毒更为有效的药物。用以完成需要惊人的计算量的庞大项目。

分布式计算是利用互联网上的计算机的CPU的闲置处理能力来解决大型计算问题的一种计算科学。随着计算机的普及，个人电脑开始进

入千家万户。与之伴随产生的是电脑的利用问题。越来越多的电脑处于闲置状态，即使在开机状态下 CPU 的潜力也远远不能被完全利用。我们可以想象，一台家用的计算机将大多数的时间花费在“等待”上面。即便是使用者实际使用他们的计算机时，处理器依然是寂静的消费，依然是不计其数的等待（等待输入，但实际上并没有做什么）。互联网的出现，使得连接调用所有这些拥有限制计算资源的计算机系统成为了现实。

那么，一些本身非常复杂的但是却很适合于划分为大量的更小的计算片断的问题被提出来，然后由某个研究机构通过大量艰辛的工作开发出计算用服务端和客户端。服务端负责将计算问题分成许多小的计算部分，然后把这部分分配给许多联网参与计算的计算机进行并行处理，最后将这些计算结果综合起来得到最终的结果。

当然，这看起来也似乎很原始、很困难，但是随着参与者和参与计算的计算机的数量的不断增加，计算计划变得非常迅速，而且被实践证明是的确可行的。目前一些较大的分布式计算项目的处理能力已经可以达到甚而超过目前世界上速度最快的巨型计算机。您也可以选择参加某些项目以捐赠的 CPU 内核处理时间，您将发现您所提供的 CPU 内核处理时间将出现在项目的贡献统计中。您可以和其他的参与者竞争贡献时间的排名，您也可以加入一个已经存在的计算团体或者自己组建一个计算小组。这种方法很利于调动参与者的热情。

随着民间的组队逐渐增多，许多大型组织（例如公司、学校和各种各样的网站）也开始了组建自己的战队。同时，也形成了大量的以分布式计算技术和项目讨论为主题的社区，这些社区多数是翻译制作分布式计算项目的使用教程及发布相关技术性文章，并提供必要的技术支持。那么谁可能加入到这些项目中来呢？

当然是任何人都可以！如果您已经加入了某个项目，而且曾经考虑加入计算小组，您将在中国分布式计算总站及论坛里找到您的家。任何人都能加入任何由我站的组建的分布式计算小组。希望您在中国分布式

总站及论坛里发现乐趣。

参与分布式计算——一种能充分发挥您的个人电脑的利用价值的最有意义的选择——只需要下载有关程序，然后这个程序会以最低的优先度在计算机上运行，这对平时正常使用计算机几乎没有影响。如果你想利用计算机的空余时间做点有益的事情，还犹豫什么？马上行动起来吧，你的微不足道的付出或许就能使你在人类科学的发展史上留下不小的一笔呢！

分布式计算是近年提出的一种新的计算方式。所谓分布式计算就是在两个或多个软件互相共享信息，这些软件既可以在同一台计算机上运行，也可以在通过网络连接起来的多台计算机上运行。

分布式计算比起其它算法具有以下几个优点：

- 1、稀有资源可以共享，
- 2、通过分布式计算可以在多台计算机上平衡计算负载，
- 3、可以把程序放在最适合运行它的计算机上，其中，共享稀有资源和平衡负载是计算机分布式计算的核心思想之一。实际上，网格计算就是分布式计算的一种。如果我们说某项工作是分布式的，那么，参与这项工作的一定不只是一台计算机，而是一个计算机网络，显然这种“蚂蚁搬山”的方式将具有很强的数据处理能力。

## 工作原理

下面，我们看看它是怎么工作的：

首先，要发现一个需要非常巨大的计算能力才能解决的问题。这类问题一般是跨学科的、极富挑战性的、人类急待解决的科研课题。其中较为著名的是：

1. 解决较为复杂的数学问题，例如：GIMPS（寻找最大的梅森素数）。

2. 研究寻找最为安全的密码系统，例如：RC-72（密码破解）。

3. 生物病理研究，例如：Folding@home（研究蛋白质折叠，误解，聚合及由此引起的相关疾病）。

4. 各种各样疾病的药物研究，例如：United Devices（寻找对抗癌症的有

效的药物)。 5.信号处理，例如： SETI@Home（在家寻找地外文明）。

从这些实际的例子可以看出，这些项目都很庞大，需要惊人的计算量，仅仅由单个的电脑或是个人在一个能让人接受的时间内计算完成是决不可能的。在以前，这些问题都应该由超级计算机来解决。但是，超级计算机的造价和维护非常的昂贵，这不是一个普通的科研组织所能承受的。随着科学的发展，一种廉价的、高效的、维护方便的计算方法应运而生——分布式计算！

### 网格计算优势

网格计算的目的是，通过任何一台计算机都可以提供无限的计算能力，可以接入浩如烟海的信息。这种环境将能够使各企业解决以前难以处理的问题，最有效地使用他们的系统，满足客户要求并降低他们计算机资源的拥有和管理总成本。网格计算的主要目的是设计一种能够提供以下功能的系统：

提高或拓展型企业内所有计算资源的效率和利用率，满足最终用户的需求，同时能够解决以前由于计算、数据或存储资源的短缺而无法解决的问题。建立虚拟组织，通过让他们共享应用和数据来对公共问题进行合作。整合计算能力、存储和其他资源，能使得需要大量计算资源的巨大问题求解成为可能。通过对这些资源进行共享、有效优化和整体管理，能够降低计算的总成本。

### 网格计算环境

现在，网格计算主要被各大学和研究实验室用于高性能计算的项目。这些项目要求巨大的计算能力，或需要接入大量数据。

网格计算的目的是支持所有行业的电子商务应用。例如，飞机和汽车等复杂产品的生产要求对产品设计、产品组装和产品生命周期管理进行计算密集型模拟。其他一些实例还有，通过 Monte Carlo 方法对复杂金融环境的模拟，以及生命科学领域的许多项目。

网格环境的最终目的是，从简单的资源集中发展到数据共享，最后

发展到协作处理和有质量的服务 (Quality of Service)。

资源集中 —— 使公司用户能够将公司的整个 IT 基础设施看作是一台计算机，能够根据他们的需要找到尚未被利用的资源。

数据共享 —— 使各公司接入远程数据。这对某些生命科学项目尤其有用，因为在这些项目中，各公司需要和其他公司共享人类基因数据。

通过网格计算来合作 —— 使广泛分散在各地的组织能够在一定的项目上进行合作，整合业务流程，共享从工程蓝图到软件应用程序等所有信息，协同处理项目中的问题。有质量的服务 (QoS)——是指能针对不同用户或者不同数据流采用相应不同的优先级，或者是根据应用程序的要求，保证数据流的性能达到一定的水准。为同一网络中的各结点提供质量有保障的服务。