

网络技术论文汇编

周国民 周义桃
丘 耘 樊景超 编

国家农业科学数据共享中心
中国农学会计算机农业应用分会

网格技术论文汇编

周国民 周义桃
丘 耘 樊景超 编

国家农业科学数据共享中心
中国农学会计算机农业应用分会
2007. 3

目 录

网格技术发展现状和趋势	国家“十一五”发展规划特聘专家: 苗再良	(1)
农业网格建设势在必行	周国民, 周义桃, 丘 耘等	(5)

网格系统

织女星网格的体系结构研究	徐志伟, 李 伟	(13)
支持网格的机群操作系统的设计	孙凝晖, 刘陶英	(21)
基于 LDAP 的网格监控系统	查 礼, 徐志伟, 林国璋等	(26)
基于有向无环图的两层网格监测系统	刘东华, 徐志伟, 李 伟	(32)
吴宇网络计算平台体系结构研究	章 勤, 廖 娟, 金 海等	(37)
Dagger: 一种散耦合的网格计算机体系结构	孙凝晖, 樊建平	(43)
网格化的动态自组织高性能计算机体系结构 DSAG	樊建平, 陈明宇	(49)
基于资源目录技术的网格系统	肖 依, 任 浩, 徐志伟等	(54)
网格环境中资源发现机制的研究	董方鹏, 龚奕利, 李伟等	(58)
网格环境下数据库系统的整合研究	周 维, 阎保平	(65)
数据网格及其关键技术研究	王意洁, 肖 依, 任 浩等	(69)
基于对等网络的全文信息检索	程学旗, 吕建明, 周昭涛	(74)
基于动态服务合成的应用型服务网格框架研究	倪晚成, 刘连臣, 吴澄等	(82)
网格用户管理体系结构的研究与设计	刘利民, 查 礼	(88)
面向网格的可视化系统研究	石敦英, 赵友兵, 仇应俊	(94)
分布式计算安全专家们谈论网络安全的新动向	Travis Van	(99)

环境工具

网格使用开发环境的研究现状和发展探讨	徐生亮, 刘东华	(107)
GSML 网格编程语言的一种实现方法	李丙辰, 徐志伟	(112)
K/G: 一种网格的使用模式体系结构及应用	杨 宁, 李晓林, 周浩杰等	(117)
一种支持业务用户编程的服务虚拟化技术--VINCA 聚合服务机制	赵卓峰, 韩燕波, 喻 坚等	(122)
网格计算环境 Globus 介绍	陈 萍, 余华山, 王 彬等	(129)
开放网格服务架构之旅--剖析 OGSA 的组件结构	Jay Unger, Matt Haynos	(133)
Globus Toolkit 3.0 和 OGSi 体系结构 - 概述	Joshy Joseph	(140)
基于 OGSA / OGSi 和 Eclipse 的网格计算平台设计	季一木, 王汝传	(152)
基于 SOA 构建网格应用程序	Martin C. Brown	(157)
将 GridFTP 集成到 Firefox 中	Karan Bhatia, Michela Taufer, Brent Stearn	(162)
Globus Toolkit 编写安全网格服务	Lavanya Ramakrishnan	(179)

应用实例

基于网格的中尺度数值天气预报系统设计与实现	宋君强, 张卫民, 朱小谦等	(189)
数据网格技术在淮河水质预报系统中的应用	王宝义, 陈华平, 田智慧等	(194)
生物数据网格的应用研究	张广治, 何洁月	(198)
GridGIS--基于网格计算的地理信息系统	王 铮, 吴 兵	(202)
数字灌区网格 GIS 的系统框架	方金云, 张汉松, 陈春仔等	(206)
中国地震减灾仿真网格试验床的系统结构	强 宇, 刘宗田, 陈吉荣	(212)
负载均衡的多级并行对等计算在新药研发网格中的实现	张 军, 常 艳, 谢向辉等	(217)
交通网格的研究与应用	谭国真, 李程旭, 刘浩等	(224)
网格技术在航空制造业的应用研究	修英妹, 崔德刚	(231)
基于网格的物流公共信息平台设计及实现	蒋 蕊, 王德占, 姜丽莉等	(236)
林业信息服务网格建设初探	赵秀芳	(240)
中国国家网格 (China National Grid)	中国国家网格运行管理中心	(242)

网格技术发展现状和趋势

国家“十一五”发展规划特聘专家 苗再良

1 综述

网格(Grid), 又称为虚拟计算环境, 是近年来兴起的一种重要的网络信息技术。网格利用计算机和网络把地理上广泛分布的计算资源、存储资源、网络资源、软件资源、信息资源、知识资源等连成一个逻辑整体, 然后像一台超级计算机一样为用户提供一体化的信息应用服务, 实现互联网上所有资源的全面连通、全面共享, 消除信息孤岛和资源孤岛。被称为是继传统因特网、Web 之后的第三个大浪潮, 可称为第三代因特网应用。它将带来因特网的新生, 极大改变我们的工作和生活。美国《福布斯》杂志预测, 网格技术将在今后几年出现一个高峰, 预计在 2020 年产生一个年产值为 20 万亿美元的大工业。

网格计算的核心构件是网格中间件(网格操作系统), 通过聚合计算设备、高性能存储器、数据库, 甚至包括科学仪器, 来为用户提供对计算力(Computing Power)随时随地的、透明的、远程的、安全的、可靠的访问。就象电网把电力提供给墙上的每个插座, 网格聚合可用的计算资源然后把计算力提供给网格环境中的每个用户, 使人们可以轻而易举地对于一些科研工作创建和使用大规模、多学科、动态的、分布式的、高性能的应用环境, 如高能物理数据分析、基因信息处理、气候建模、宇宙观测、实时遥感数据分析和数据同化、大型数据集交互分析和虚拟现实可视化等。

根据计算角度和应用特性, 网格可大体分为: 计算网格、存储网格、数据网格、信息网格、知识网格和各类应用网格等。各种网格在体系结构上基本一样, 都采用国际上比较公认的标准和规范。

国际上主要网格标准化组织包括全球网格论坛 GGF、对象管理组织 OMG、W3C、GridForum 以及 Globus.org 等。目前网格计算还没有世界统一的标准, 但在核心技术上, 相关机构与企业已达成共识: Globus Toolkit 已成为网格计算事实上的标准, Globus Toolkit 基于 Globus 和 IBM 共同倡议的网格标准 OGSA(Open GridServices Architecture), 把 Globus 标准与以商用为主的 Web Services 的标准结合起来, 网格服务统一以 Services 的方式对外界提供。

2003 年上半年符合 OGSA 规范的 Globus Toolkit 3.0 正式发布。这标志着 OGSA 已经从一种理念、一种体系结构, 走到了付诸实践的阶段。包括 IBM、Entropia、Micro soft、HP/Compaq、Cray、SGI、SUN、Veridian、富士通、日立、NEC 在内的 12 家计算机和软件厂商已宣布将采用 Globus Toolkit。作为一种开放架构和开放标准基础设施, Globus Toolkit 提供了构建网络应用所需的很多基本服务, 如安全、资源发现、资源管理、数据访问等。目前所有重大的网格项目都是基于 Globus Toolkit 提供的协议与服务建设的。

2 技术现状

面对即将到来的第三代互联网应用—网格技术及其应用, 世界上许多国家对网格技术非常重视, 很多发达国家都投入了大量研究资金, 希望能抓住机遇, 掌握未来的命运。并且把网格计算技术应用在各个领域, 如高能物理、全球变化、地震监测、商业计算、虚拟现实等。美国、欧洲、日本、印度等国都启动了大型网格研究计划, 并得到了产业界的大力支持。这方面的代表性研究工作包括美国国家科学基金资助的 NPACI、“国家技术网络”(NTG)、分布式万亿次级计算设施(DTF)、美国能源部的 ASCI Grid, 以及欧盟的 DataGrid、UNICORE 等。

美国网格计算研究主要由美国能源部(DOE)、美国国家科学基金会(NSF)、美国国防部(DOD)资助, 各个国家实验室、高等院校承担, 各个大型公司积极参与。欧洲的网格计算研究项目主要由欧盟和欧洲委员会资助。其中网格计算研究比较活跃的国家是英国、德国和意大利。

在美国, 无论是政府部门还是商业机构, 对网格的应用和发展都表现了极大的兴趣。美国军方正规划实施一个宏大的网格计划, 叫做“全球信息网格 GIG(Global Information Grid)”, 预计在 2020 年完成。作为这个计划的一部分, 美国海军和海军陆战队已启动了一个耗资 160 亿美元、历时 8 年的项目, 包括系统的研制、建设、维护和升级。美国能源部的山地国家实验室的“先进战略计算创新计划网格(ASCI Grid)”主要用于核武器研究。美国国防部和欧洲能源机构等在两三年前先后采用了网格技术。美国能源部下属的国家能源研究科学计算中心已经与美国国际商用机器公司(IBM)达成协议, 正式建造了美国能源部内部使用的计算机网格。美国的许多国家实验室的超级计算机和存储资源, 都将陆续与该网格相连。

英国建立了全球最大的网格计算机系统。该网络系统由 6000 多台分散在 78 个国家的计算机

组成,而英国就拥有其中的 1000 多台,它是“大型量子对撞机计算网格”(LCG)的一部分。LCG 是全球目前最大的网格系统,而且也是第一个永久性网格系统。

同时,随着网格研究在学术界的加速,信息产业界的大公司也相继公布了与网格目标一致的研究开发计划。惠普、IBM、微软、Sun 等公司最近取得共识,支持 XML、SOAP、UDDI 等网格和 Web 服务等标准,从而更有利于开发新一代的网格应用。近几年,IBM、Sun、Oracle、Platform、惠普等公司推出了一系列网格平台软件和网格实施方案,其目的是将因特网上的资源和信息汇聚在一起,组合成企业和消费者所需要的服务。世界上许多顶级企业集团都宣布了建设内部网格的计划。IBM 最近宣布,将投资数亿美元,部署内部网格研究机构,启动一个全公司的“网格计算创新计划”,以便于分散在世界各地的 IBM 研究所人员共享计算资源。医药、化工、通信、电子、汽车等领域的一些大公司,如辉瑞、爱立信、日立、宝马、联合利华、葛兰素威廉、史可必成等,都已经开始构造使用内部网格。

总的来看,在涉及企业应用时,网格计算目前仍处于早期阶段。在公用事业计算中,网格应用刚刚起步,IBM、HP、Oracle 和 Sun 等厂商目前都还没有将网格技术视为公用事业计算的基础,而是把它当做公用事业计算的关键组成部分。

网格技术引起我国高度重视。网格的研究和应用已被列为“863 计划”的一个专项,通过“国家 863 计划高性能计算专项”的形式在十五期间支持网格计算的技术研究和应用工作,早期的研究项目主要有中科院计算技术研究所承担的国家高性能计算环境项目和织女星计划((Vega 计划)、清华大学计算机技术系承担的“先进计算基础设施北京—上海试点工程”,这些项目起到了培养人才、跟踪世界先进水平的较好效果。目前国内正在进行的网格研究项目主要有:

- “国家网格 CNGrid”(科技部),有多家单位参加;
- “中国教育科研网格计划 China Grid”(教育部),有多家单位参加;
- “E-Science 网格研究计划”(国家基金委);
- “中国空间信息网格”;
- “上海教育科研网格”,多所上海的大学参加;
- “仿真网格”的研究,由航天二院和清华大学共同开展;
- “织女星网格”,由中科院计算所领衔开发;

另外,全国还有几十所大学和研究机构已经开展各种网格研究,一些计算密集、海量数据的专门机构也在建设专用网格。

可以看出,网格研究正在迅速展开。有关网格的其它一些重大战略决策也正在积极酝酿之中。许多大学和研究机构都投入了很大资源进行研究开发。国民经济和社会发展的很多领域已对网格技术提出了实际需求,比如生物科学中的“基因信息处理”、运输业的“综合信息平台”、制造业的“网络化虚拟制造”、电子政务中的“资源共享”与“协同工作”、能源业中的石油勘探和油藏模拟、资源环境中的海量数据处理与共享、科学研究中网络化虚拟实验室和虚拟天文台等等。

与网格研究的相关课题主要有:应如何利用科技界、产业界、应用界的资源,中国应如何加入国际网格技术研究开发,如何参与国际标准制定等问题。

3 发展趋势

3.1 标准趋向开放、统一

国际上网格技术的研究基本上采用了因特网的开放标准路线,以层次化开放式为基础,在多个层次上建立横向技术标准和平台,以满足不同层次资源和应用集成的需要。这些标准中最主要的是 Globus 和 OGSA。平台化也是网格的一个重要特征。

虽然 GGF、OMG、W3C 以及 Globus 等诸多团体都试图争夺网格标准的制定权。但目前大多数网格项目都是基于 Globus Toolkit 所提供的协议及服务建设的,例如美国的物理网格 GridPhyN、欧洲的数据网格 DataGrid、荷兰的集群计算机网格 DAS-2、美国能源部的科学网格和 DISCOM 网格、美国学术界的 TeraGrid,等等。迄今为止,基于 OGSA 的 Globus Toolkit 已经成为事实上的网格标准。

OGSA 是在网格计算技术的基础上,结合最新的 Web Service 技术提出来的。以服务为中心是 OGSA 的基本思想,在 OGSA 中一切都是服务,在原来 Web Service 服务概念的基础上,提出了“网格服务(Grid Service)”的概念,用于解决服务发现、动态服务创建、服务生命周期管理与临时服务有关的问题。

OGSA 是 Globus 和 IBM 共同倡议的网格标准,并得到业界的广泛支持,像微软、Platform Computing、AVAKI、Entropia 等从一开始就宣传支持 OGSA。到目前为止,OGSA 已经广为接

受, 几乎所有的业界同仁都认为它就是网格的未来。由于 OGSA 是在 GGF 会议上提出来的, GGF 也就顺理成章地成为 OGSA 标准化进程的领头羊。目前, GGF 有 4 个研究组负责与 OGSA 相关的标准制定工作: 开放网络服务体系结构工作组 OGSA-WG、开放网络服务基础设施工作组 OGSI-WG、开放网络服务体系结构安全工作组 OGSA-SEC-WG 和数据库访问和集成服务工作组 DAIS-WG。

3.2 多技术融合趋势

许多网格应用和系统是综合网格技术和其它多种主流技术建立而成的, 这些技术包括中间件技术、P2P、Web Service 等等。事实上, 网格技术解决了网格应用面临的一些问题, 而现行的许多技术很好地解决了系统集成的各种问题。同时, 网格作为一种信息基础设施, 其必然要综合运用各种技术来建立这种设施及其各种应用。

在 OGSA 出现之前, 已经出现很多种用于分布式计算的技术和产品。例如, SUN 公司的 ONE、OSF 的 DCE 和对象管理集团 OMG 的 CORBA、微软的 DCOM 等。这些机制互不兼容, 严重到了同一家公司的产品都不兼容的程度。其实, 在 OGSA 出现之前, 各种以填补异构平台之间的差异为己任的网格平台, 如 Condor、Legion、Ninf、Globus 等, 也都是各行其道、互不兼容的。

自 90 年代末以来, 基于 XML 的 Web Services 技术开始盛行。Web Services 在各种异构平台之上构筑了一层通用的、与平台无关的信息和服务交换设施, 从而屏蔽了互联网中千差万别的差异, 使信息和服务畅通无阻地在计算机之间流动。Web Services 得到了各大公司的支持, 解决方案精彩纷呈, 包括 IBM 的 WebSphere、微软的 .Net、SUX 的 SunOne、Oracle 的 Oracle9i、惠普的 eSpeak 等等。Globus 项目组看到了 Web Services 的巨大潜力, 在 2002 年迅速将 Globus Toolkit 的开发转向了 Web Services 平台, 试图用 OGSA 在网格世界一统天下。基于 OGSA 之后, 网格的一切对外功能都以网格服务(Grid Service)来体现, 并借助一些现成的、与平台无关的技术, 如 XML、SOAP、WSDL、UDDI、WSFL、WSFL 等, 来实现这些服务的描述、查找、访问和信息传输等功能。这样, 一切平台及所使用技术的异构性都被屏蔽。用户访问网格服务时, 根本无需关心该服务是 CORBA 提供的, 还是 .Net 提供的。

3.3 网格进入商用化

2004 年 1 月 20 日, Globus 联盟、IBM 和 HP 等联合发布了新的网格标准草案 WSRF (Web Services Resource Framework), 把 OCSI 转换成了 6 个用于扩展 Web Services 的规范, 使网格服务已经与 Web 服务彻底融为一体, 这标志着网格已经调整好方向, 将信息集成作为第一目标, 标志着网格商用化的时代已经来临。

此外, 数据网格也将会产生成功的应用模式。Globus 目前已经提供或即将提供一些基本的网格功能, 如数据高速传输、数据复制、数据复制的选择、元数据管理等。实现多种异构存储系统的统一接口, 屏蔽了存储系统的异构特性, 支持广域网络环境下多种数据源的访问, 让海量数据能够在网络化存储、高性能计算机、海量数据库之间快速传递。可以预见, Internet 已经变成了一个数据海洋, 对这种能将分布、异构、海量的数据集成起来的数据网格会有大量需求。

目前, 基于新标准 WSRF 的网格平台 Globus Toolkit 4.0 已经面世。此后几年, 会出现大量融 Web Services 和网格服务为一体的新型服务出现, 让计算机跨越时间和空间的限制, 自动获取和处理信息成为现实。

4 下一步的重点课题

4.1 网格计算效率的控制

除了标准需要进一步完善统一外, 目前网格存在的主要问题是: 计算效率的控制和网络安全。为了取得最高的效率, 每个网格节点用在计算上的时间必须比用在通信上的时间多。如果网格任务必须携带大量的数据, 这项任务将不能轻松地从一个节点转移到另一个节点。如果网格任务排队等待某些宝贵的资源, 网格则不具有优于普通技术的优势。实际上, 构建通用的计算网格是比较困难的。因为很多应用是通信密集型的, 子任务之间有频繁的通信, 而远程的网络带宽远不及系统区域网和局域网内的带宽, 很可能导致大规模网格计算的效果还不及使用规模小得多的集群计算。所以, 网格计算效率的保障与控制是下一步研究的重点课题之一。这方面的课题主要有:

网格资源动态分配; 任务分割与分配;
通信延迟控制; 互联网数据传输能力等。

如何在动态、异构虚拟组织间实现协同的资源共享以及协同计算是网格中非常重要的问题。已有的一些并行和分布计算系统的资源分配技术, 并不能很好地适应计算网格资源分配问题的特点。与此相关的是任务分割与分配问题, 而且其算法和策略的最佳化也是动态的, 是网格状态和

任务状态的函数。如果网络任务分割分配不合理,或者排队等待某些宝贵的资源,网络则不具有优于普通技术的优势。另外,数据通过因特网时形成的延迟问题,智能化的网络软件应确保按时传送数据,否则,网络处理问题的手段将被限定在“并行运算”。并行计算在不同的机器上完成,某一台计算机不需要等待另一台计算机的处理结果。

还有互联网的数据传输能力不足问题。为此,发展网格要和建设下一代宽带互联网结合起来。另一方面,采用无线移动和卫星通信,也是一种现实的途径。

4.2 安全问题

正如电网的首要问题是安全一样,网络计算的安全也是最为重要的。需要重视的安全问题主要来自两个方面,一是身份验证的问题,另一个是网络计算必须要受到有序的控制和管理,尤其是网络计算的有序管理问题。网络计算本身不能区分客户是一个善良的用户还是一个黑客,也不能区分所服务项目的内容属性,黑客们完全可通过付费租用 PC 的空闲处理能力的方式,或是发起一场什么冠冕堂皇的活动,从而吸纳大量 PC 的处理能力,从某些非法活动,由此引发的后果将是十分危险的,甚至是无法想象的。从某种角度上说网络计算,甚至是一种“无政府”状态的计算模式,因而如何进行有序的控制和管理就显得十分重要。

4.3 网络服务和运营问题

网络作为一种 IT 基础设施已经得到众多公司和学术界的一致认同。如何构建、管理、维护和运营这种能够按需提供服务的基础设施,如何在这个平台上开发应用,成为网络服务和产业发展的关键。在这样一个环境下,结合软件和计算机硬件的系统综合服务提供商将出现,并在网络产业的发展上起到关键作用。

要尽早研究通过法律的手段解决网上资源共享中的智能产权、相互信任和报酬等问题以及如何保障网络计算的安全性、认证和可靠性等。

4.4 开发商用的网络操作系统

网络操作系统或网络中间件,是网络系统软件的主体,是网络和多种技术的结合为利用网络资源提供的一个网络虚拟操作系统,应具有开放的体系结构、标准和协议,以形成信息获取、传输、访问、共享和处理的单一开放的信息处理基础设施平台,需要克服虚拟组织的管理和协同工作问题,需要解决网格的可用性和可开发性障碍,并解决系统安全问题。这个平台屏蔽了异构资源的差异和地理上的分布性,管理分布的计算、储存、网络等资源,提供实现应用之间的可交互性和互操作性的工具,提供用户开发网格应用的环境和网格应用的基本功能构件,支持各种应用的开发和按需服务的能力。

网络操作系统的开发必须要走实用化、市场化的道路,尽量避免像过去我们国家搞的一些操作系统和数据库项目,只能停留在实验室里,离实际使用相距太远,丧失产业化机会。

4.5 服务网格是重要产品方向

除了网络操作系统,各种应用网格的构建将成为今后几年的商业化重点,包括科学计算、教育、智能交通、医疗、生物信息、天气预报、资源探测、电子娱乐(e-Entertainment)、电子政务、电子商务和大型企业网络等。

科学计算、教育、生物信息、天气预报、资源探测等网格已经在建设当中。大型企业网络是电子商务的基础,其建设进程可能在一二年后兴起,但会成为网格产业化的最大空间。企业计算是高度分布和异构的,需要大量的计算和数据服务。通过应用网络技术和 Web service,遵循 OGSA、OCSI 的 Grid Service 的标准,将分散的功能整合起来,面向企业应用系统集成、支持服务连接、管理、集成优化和运行的单一的服务网格将成为商业网络系统的一个重要发展方向。企业网格的发展,最终将实现跨企业、行业或工作组的虚拟共享、管理和设备访问,动态地管理网络上的资源、服务和应用,在不同的组织中集成应用和交换数据,使企业能够共享资源,从而消除企业信息化中存在的“信息孤岛”。为实现多企业或部门之间广域分布业务应用的集成和协同提供按需服务、系统互操作和可监控等方面的有力支持。

4.6 应用系统向网格平台的平滑移植

这是网络计算走向主流应用的最大障碍。网络技术要求用户将原有的计算密集型应用标准化,并平移到新的系统之中。实际上,很多机构、企业的现有应用已经开发多年并运行良好,如果将其推向网格环境,必将面临重新编写应用代码的问题。虽然一些相关的工具套件已经推出,还需要深入研究开发相应的支撑技术,也需要厂商的进一步支持和相互之间合作的广泛开展。

来源:信息技术与信息化,2005年第2期 A17~A20

农业网格建设势在必行

周国民, 周义桃, 丘 耘, 樊景超

(中国农业科学院农业信息研究所, 北京 100081)

摘 要: 网格是因特网的发展和升级, 其实质就是实现真正的资源共享。网格将把人类带入理想的“三元世界”, 就像拿起电话能通话、插上电源能用电一样, 用户一旦接入网格就能根据需要取用各种 IT 资源。本文介绍网格的特性、国内外应用状况、农业上的应用及可能的应用, 指出建设农业网格的迫切性, 并提出以农业数据网格建设为突破点等几点建议。

关键词: 网格; 农业网格; 网格建设

The construction of agricultural grid is urgent

ZHOU Guo-min, ZHOU Yi-tao, QIU Yun, FAN Jing-chao

(Institute of Agricultural Information, CAAS, Beijing 100081, China)

Abstract: Grid is the development and upgrade of internet. The essential of grid is to realize the share of the resources. It will take human into the ideal “triplet world”. Just like picking up the phone you can talk, inserting into electrical outlet you own the power once the person access the grid he or she can get various IT resources as he or she need. This thesis introduces the features of grid, application at home and abroad, its usage on agriculture and possible application. Meanwhile this thesis also points out the construction of agricultural grid are urgent. Then the paper give several suggests about building the agricultural data grid to break out and so on.

Keyword: Grid; Agricultural grid; Grid construction

1 网格概述

1.1 网格特性

网格这一名字, 第一次出现在 90 年代中期, “网格之父”美国 Ian Foster 这样描述网格, 网格是构筑在互联网上的一组新兴技术, 它将高速互联网、高性能计算机、大型数据库、传感器、远程设备等等融为一体, 为科技人员和普通老百姓提供更多的资源、功能和交互性。互联网主要为人们提供电子邮件、网页浏览等通信功能, 而网格功能则更多更强, 能让人们透明地使用计算、存储等其他资源。这就是说网格包括全部计算资源。

网格与因特网相比具有高性能、一体化、知识生产、资源共享等技术优点。

网格的高性能是指网格比因特网具有更强的功能和性能, 表现在: ①具有更大的带宽; ②具有更高的速度; ③具有更多有效资源; ④具有更多、更大的网上社区, 这些网格社区, 把地球上所有公民联为一体。

网格的一体化是指网格不应该像因特网那样, 提供几百万个网站, 让用户去费尽心思寻找合适的地方。在逻辑上, 网格应该就像一台机器, 用户可随时随地获取信息、知识, 利用资源完成计算。

网格的知识生产特性反映了网格和因特网质的区别。因特网本身不创造或生产知识, 先要把信息或者知识生产出来, 放到网上供用户查找。网格则能根据用户的要求自动地生产知识。在知识生产的过程中, 高性能计算机将起到关键的作用。它把从数据源(传感器、贵重设备、数据库、信息库等等)得到的原始数据, 运行特定的程序加工成信息和知识。

网格的根本特征是资源共享而不是它的规模。实际上传统互联网实现了计算机硬件的连通, Web 实现了网页的连通, 而网格试图实现互联网上所有资源的全面连通, 消除信息孤岛, 协同工作。

网格将成为 21 世纪人类社会的信息处理基础设施, 包括网格节点、资料库、贵重仪器、视

频设备、宽带骨干网及网络软件等部分。网络是因特网的发展和升级,其实质就是将万维网(World Wide Web)升华为网络(Great Global Grid),把人类带入理想的全球大网络“三元世界”(物理世界、梳妆世界、人类社会),就像拿起电话能通话、插上电源能用电……一样,用户一旦接入网络就能根据需要取用各种IT资源。

1.2 网络应用

为什么需要网络,是因为网络有非常广泛的应用领域。就像网络从开始时主要为科研、军方等所用,而现在已为各个领域普遍应用一样,网络技术的应用也将有一个广阔的新天地。

1.2.1 在国外

国外网络技术发展态势值得关注。迄今为止,网络界的领头团队仍然是Globus团队。这个团队原来是做高性能计算应用的,对计算机系统经验较少,而今逐步转到面向服务的分布式计算软件体系结构。他们领导产生的OGSA框架和OGSI协议可以认为是计算网络技术的典型代表和事实上的规范,其核心部分元计算工具包,定义了构建计算网络最基础的服务。

网络在美国引起革命。加利福尼亚州最近投资3亿多美元在圣地亚哥附近成立了“加州通信与信息技术研究所”,研究所的所长是“国家高性能计算应用中心”(NCSA)的前主任,元计算的发明者Larry Smarr。该研究所的主要研究方向之一是开发新型的信息网络技术来构造一个“智能交通基础设施”。通过因特网将很多服务器与公路上的传感器和汽车中的电脑无线地联为一体,从而更有效地控制交通。

美国的民用计算网络已含有一百余个结点,每个结点有几十亿到几千亿次的计算能力。

美国国家航空航天局构造了一个国家技术网络(National Technology Grid),后来又出现了STAR TAP (Science Technology And Research Transit Access Point),将分布在各地的资源通过网络(包括无线通信手段)连接起来,解决目前无法解决的科学与工程计算与数据管理等问题。STAR TAP计划,试图将网络扩展到全世界,成为世界网格(International Grid)。目前加入STAR TAP计划的有加拿大、新加坡、日本和中国台湾;即将加入的包括巴西、俄罗斯、北欧、法国和亚太一些国家。

英国启动了著名的e-Science计划。英国政府非常重视网络技术的研究,认为网络是WWW的必然后继者,投资2.4亿英镑左右支持网络研究项目,其应用领域包括粒子物理、生物信息学、气候环境变化、工程系统设计等多学科领域。

日本启动了“国家研究网格”和“商务网格”计划;韩国的网格计划之一是N*Grid,五年计划,正在进行。政府对这些计划的投入数千万、上亿美元。

网络软件的研究开发早已不是停留在概念阶段,在部分领域已经进入了开发实用生产型系统的阶段。比如,美国的大银行JP Morgan-Chase已与IBM、Platform签约,开发实用的、具有业务整合、资源共享、协同工作特征的银行网络应用系统,该系统的一部分已经投入生产型运行。

1.2.1 在国内

国内网格也正处于快速发展时期。中科院牵头的“国家高性能计算环境(National High Performance Computing Environment)”项目,长期目标是提高计算网格系统的性能、可扩展性及可用性。2002年863信息技术领域“高性能计算机及其核心软件”专项启动,本专项所有工作都围绕着研制“国家高性能计算环境”这一大目标进行,以实现高性能计算机及其核心软件跨越为目标,确立“战略与系统综合研究”、“高性能计算机”、“网络结点”、“网络软件”和“应用网络”五个方面课题。具体目标是建设“中国国家网格”。

2002年4月,科技部决定建设“中国国家网格”,总投资高达3个亿。中国国家网格是聚合高性能计算和事务处理能力的新一代信息基础设施的试验床。通过资源共享、协同工作和服务机制,有效支持科学研究、资源环境、先进制造和信息服务等应用,推动国家信息化建设及相关产业的发展。

目前中国国家网格有包括香港在内的8个结点,装备了自主研制的面向网格的高性能计算机(联想深腾6800、曙光4000A),构成了开放的网格环境,通过自主开发的网络软件,支撑网格环境的运行和应用网格的开发建设。目前8个成员结点是中国科学技术大学、中科院网络中心、香港大学、上海超算中心、清华大学、西安交通大学、国防科技大学、北京应用物理与计算数学研究所等。

中国国家网格有包括教育网格、气象网格、资源环境网格、航空制造网格、科学数据网格、新药研发网格、生物信息网格、仿真应用网格、城市交通信息服务网格、油气地震勘探应用网格、森林资源与林业生态工程网格等应用网格。

另外，台湾也启动了“知识创新网格”三年科研计划。

这些项目都带有政府背景，说明我国对网格研究与开发的重视，相信经过努力，我国的网格基础设施就会发挥作用，为科研、教育、生产服务。网格对国民经济和社会发展的作用不亚于电网、交通网，对国家安全的影响不亚于两弹一星。网格将成为中国 21 世纪的战略基础性设施。

2 网格在农业上应用

2.1 应用成果

网格在农业上初见端倪，下面举几个应用事例。

①水稻基因组工作框架图应用

2002 年 4 月 5 日，中国在美国出版的权威杂志《科学》上向世界免费公布了有助于解决饥饿的重大科研成果——水稻基因组工作框架图，使人类朝远离饥饿的目标迈出了重要一步。该项成果估计水稻基因组中基因总数在 46022 至 55615 之间，每个基因还分为无数个序列，数据浩如烟海。研究能迅速、圆满完成，除了使用了国内最高端的曙光服务器外，还因为一种重要技术——网格在海量计算、数据处理、资源共享方面发挥了重要作用。

②农业气象研究应用

中尺度强对流天气是影响我国的最主要的灾害性天气，每年都给国民经济和人民生活财产造成了极大的损失。由于尺度小、时空分布不均匀、具有突发性和多发性的特点，使预报工作有较大难度。在中国国家网格节点上，实现了高解析度的中尺度数值天气预报软件研究环境，该研究环境具有：实现与美国全球数值预报系统输出产品的嵌套；引入本地区的高密度观测资料，采用每天 4 次 6 小时时间滚动同化方案；实现 48 小时中期天气预报；实现本地的气象可视化系统等特点。它提高了我国中尺度天气预报水平，对于防灾减灾、实现我国经济与社会的快速、持续、稳定发展具有重要的意义。

③卫星遥感图像处理应用

遥感，作为一门新兴的综合性技术学科，在农业领域有广泛应用，尤其是精确农业。随着传感技术的日新月异，遥感信息和数据数量巨大、处理复杂，单个超级计算机已经不能满足计算的需求，遥感图像处理与网格计算技术的结合变得自然而迫切。中国国家网格已为人们提供数字图像处理应用平台。通过网格分布处理技术，使数字图像处理变得容易。

④水资源管理应用

“淮河水资源保护与管理”应用网格已经问世。该应用网格以“淮河流域防洪防污智能调度”为主要应用背景，在网格计算环境下对系统进行了模块化分解和服务化封装，并充分利用工作流技术对原有系统进行业务流程的并行化和分布式配置，以获得较高的性能及良好的应用效果。目前该应用网格充分利用了数据网格平台的技术优势，整合了大量异地异构的数据，解决了淮河流域长期以来电算化数据环境混乱的问题，并把一些计算量大的水文学、水动力学计算模型部署到网络节点的高性能计算机系统上，有效解决了淮河水资源保护与管理业务应用中的超级计算问题。

2.2 发展前景

目前，一些发达国家的城市、地区，已酝酿将网格计算用于智能交通、城市应急系统、远程医疗、企业管理等众多方面。网格在农业上应用必将得到快速发展。网格在农业上可能应用，包括农业科学研究、农业教育及远程教育、农业技术推广、农业电子政务、农业电子商务等。

①农业科学研究

现代科学研究的问题空前复杂化，产生了一种崭新的科研协作模式和大科学工程，就是以新一代互联网技术和网格计算技术为基础，通过因特网联合组成一个共同的虚拟研究团队，共享资源和成果，协同工作共同完成大型现代科学研究。例如，最佳农业生产方案研究，广泛收集产、供、销等农业经济数据和土、肥、水、气象以及遥感图像资料等农业生产数据，利用网格分布计算，获得地、县最佳方案的基础上，可以制定省、乃至全国的最佳方案。

②农业教育及远程教育

将分布在农业高等院校中的图书馆数字资料、论文、多媒体课件和数字教学视频等各种海量信息资源整合起来，建立一个农业教育网格，提供统一的、高效的信息服务，提供主动学习方法和交互式的教学方式，使在家学习和交流成为现实。同时，这样的网格还可以为人们提供在网络上做模拟实验。

③农业远程咨询

将分散在农业科研院所、大专院校以及管理部门中的农业新品种、新技术、新知识以及农业专家信息集成起来,建立一个农业推广网格,提供咨询服务平台,农民可以足不出户学习新技术、了解种养新品种,并可以面对面向专家请教,解决种养中遇到的问题。该网格还可以收集病、虫、害资料,建立专家系统,提供远程诊断服务。

④农业电子商务

农业电子商务网格将为农业、农村、农民提供一个非常有用的、高效的服务平台。农、牧、渔产品产量信息、市场价格信息、供求信息、加工信息等等,及时收集并建立网格,对产销具有重要意义。例如欲进行某产品生产需做可行性分析。企业将自身的历史数据交付共享,财务顾问公司将自己的财务分析/预测系统交付共享,决策咨询公司也将自己的假设性分析系统交付共享。如此一来,参加可行性分析的各方形成了一个虚拟组织,所有资源由各方共享,大家可在不同的地点共同工作,最后得出可行报告。

⑤农业电子政务

国家电子政务应用的目标是建立跨领域、跨区域和跨企业的一体化信息网络。电子政务需要整合的资源和服务,种类繁多,数量巨大,地理上广泛分布,管理上属于不同的管理域,有着不同的管理策略和安全需求,网格技术能够很好地适应电子政务的这些需求,为解决这些问题提供了技术基础和标准平台。将分布的各个局部自治系统统一成为逻辑上的单一全局系统,从而以有限的资源,达到电子政务的更好效果。

3、农业网格建立建议

农业领域是一个多行业、多部门、分布广泛的大社会。农业网格可以分为农业信息网格、农业科技教育网格、农业技术推广网格等。农业信息网格把国家农业信息网、各省(直辖市、自治区)农业信息网、地区县市农业信息网整合起来,农业科技教育网格把国家农业科技网、国家农业教育网、各省(直辖市、自治区)农业科学院网、各农业院校网、地区县市农业科教网整合起来,农业技术推广网格把国家农业推广技术网、各省(直辖市、自治区)农业推广技术网、地区县市农业推广技术网整合起来,实现资源共享、信息互联互通、协同工作,为农业、农村、农民提供实实在在的服务。

建造农业网格意义深远、工程浩大。农业网格是一个有重要影响的基础设施,包括硬件、软件和网格组织。关于如何建造,必须遵照一定的原则来进行,否则就很难把网格建好。下面提出几点建议:

3.1 选准突破点

农业网格建设起步阶段,建议以农业科学数据网格构建为突破点,理由是:

首先参考、学习“中国科学院科学数据网格”构建的经验。该网格在“中国科学院科学数据库”数据资源的基础上,连接中国科学院分布在全国的四十多个研究所,通过网格计算的相关技术,特别是数据网格技术,对科学数据库中大量分布式异构数据资源实现有效共享,并在此基础上开发对科学研究有直接实用意义的应用系统。农业科学数据网格构建与中国科学院科学数据网格构建应该有许多相似的特点。

其次农业科学数据经几年的运作,已见规模。农业科学数据中心建设由中国农业科学院农业信息研究所主持,现有作物科学、动物科学与动物医学、农业资源与环境科学、农业微生物科学、热作科学、水产科学、草地与草业科学等12类数据160多个数据库。中国农业科学院部分专业研究所、中国水产科学研究院、中国热带农业科学研究院等单位的分中心正在建设中。

第三,农业科学数据网格构建内容有:构造农业科学数据网格的系统平台;开发所必需的中间件软件;开发示范应用系统。农业科学数据网格是一个信息网格,要解决的中心问题是如何描述信息、存储信息、发布信息和查找信息,为用户提供“信息在你指尖”式的服务。信息网格转变为知识网格是人们关心的一个重要研究课题。知识网格是由“知识结构”、“知识单元”、“知识元”采用关联技术组成的层次知识链网状知识关系图。“知识元”构成了知识网格的最小单元,是求解问题的证据,是知识网格的核心。分析“知识链”的构成规则,由知识链构成一个文本的“知识单元”。利用知识单元和知识元的对象关联关系,建立检索系统的“知识结构”框架(知识检索)。

农业科学数据网格构建研究,将为农业网格构建提供经验,并培养、造就一批农业网格研究、开发人才。

3.2 争取国家支持

农业网络建设必须有政府的大力支持，有如下几个方面的原因：

首先农业网络是一个涉及到整个国家的基础性设施，覆盖到整个国家的版图，只有政府出面才能够协调好各方面的关系。网络建设一定会有工程施工，这需要工程建设部门和相应的规划部门的介入。网络必然需要和传统的通信部门连接，因此也需要邮政、电信等提供积极的配合并且加入到网络的庞大网络体系中来。除了中央政府之外，还需要地方各级政府的配合。

其次就是农业网络这一基础设施的前期投入十分庞大、风险大、回报周期长，公司或者企业一般不愿先去投资的。但是，网络像以前的公路系统、电话与电报系统、电力网络、银行系统等传统的基础设施一样，其对整个社会产生的巨大影响，影响的深度和广度，决不逊色于传统的基础设施。因此，需要政府首先大力支持，使网络能够运转起来，然后逐步走向成熟和完善。

第三，网络的建设和使用会涉及到国家政策、法令和法规，以及相应的规章制度，只有国家和政府才有资格制定，网络从一开始的规划到建设直至完成和使用，都需要国家和政府进行全盘规划，并制订相应的政策和规定对网络进行规范化，这样才能够为网络这一基础设施的健康发展创造条件，提供保证。

3.3 利用市场与经济杠杆

在网络建设的初期，需要政府的大力投入，使网络能够初步发挥作用。当显示出市场的潜力时，大量的投资者会自觉地投入到网络的建设和完善过程中，使网络建设不断扩大和健全。举个例子来说，美国政府投资建成 NSFNET (NSFnet 是 1990 年 6 月彻底取代了 ARPAnet 而成为 Internet 的主干网) 后，社会各界后来在其上投资的总和达到了美国政府一开始投资的一百多倍，从此可以看出：政府开始的投入会带动以后的迅速发展，基础设施项目需要社会各界的共同介入才会产生重大的效果。

网络建成并不意味着大家愿意使用，如果没有人愿意使用它，那么网络就是一个失败。人们之所以会使用网络，是因为网络能够为他们的社会经济活动带来便利，增加收益，降低开销，利用市场这个推进器，会对网络的发展起到重要的推进作用。

3.4 培养和造就人才

农业网络是一个技术性很强的基础设施，网络的建设没有科技部门的投入和参与是不可能建好的，要组织一大批的科技工作者投入到这一工程中来，包括规划设计、可行性分析、工程施工等环节。

网络建设与网络建设一样，有路无车、有车无货是不行的，设施建设的同时要进行应用系统的研究和开发，才能使网络发挥作用。研究、开发人才是不可缺少的，而且网络还需要一批维护人员。目前农口这方面的人才十分缺乏，引进、培养、实践造就就是建设人才队伍的必由之路。

农业网络建设任务任重道远。在中国国家网络结点和应用网络建设中，至今未能将农业网络构建列入其中。我们认为，作为农业大国，农业网络是不可缺少的一块。我们呼吁，国家政府、科技部应该关注，农业部门领导、科技人员更要关注。让农业网络构建早日上马，让农业网络为社会主义新农村建设服务。

主要参考文献：

- [1] 周义桃，周国民，樊景超. 网络——下一代互联网络. 农业网络信息, 2005, (11): 67~71.
- [2] Albert Bunshaft. 网络：从理论走向应用. 软件世界, 2003, 12: 03
- [3] 都志辉, 陈渝, 刘鹏. 网络计算. 北京: 清华大学出版社, 2002
- [4] 中国国家网络 <http://www.cngid.org/>

作者简介

周国民(1969-), 男, 博士, 中国农业科学院农业信息研究所研究员、博士生导师, 研究方向: 农业信息技术。

周义桃, 中国农业科学院农业信息研究所研究员。

丘 耘, 中国农业科学院农业信息研究所副研究员。

樊景超, 中国农业科学院农业信息研究所硕士、实实。

来源：农业网络信息，2006.7, (121): 4~7, 19

网 格 系 统

织女星网络的体系结构研究

徐志伟 李伟

(中国科学院计算技术研究所, 北京 100080)

({zxt, liwei}@ict.ac.cn)

摘 要: 网格技术是高性能计算机研究的热点。网格的体系结构是构建网络系统的基础。在分析比较现有网格技术的基础上, 提出了织女星网络的体系结构设计, 并对其中的关键问题进行了研究。织女星网络体系结构的基本思路是实现一台虚拟的、单一的网格超级计算机, 其核心是 Vega 设计思想。织女星网络体系结构的关键组成包括网格硬件、网格互联系统以及网格操作系统, 同时, 提出了资源路由器、网格计算协议、网格浏览器等重要概念。

关键词: 织女星网络, 资源路由器, 网格计算协议, 网格浏览器

RESEARCH ON VEGA GRID ARCHITECTURE

XU Zhi-Wei and LI Wei

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

Abstract: The grid technology is a hotspot in HPC research. The architecture design is the basis for constructing the grid system. Based on the existing grid technologies, the design of Vega Grid is Proposed and make research on some important issues are studied. The fundamental idea of Vega Grid is to build a virtual, single grid super Computer. The kernel of Vega Grid is Vega Philosophy. The key components of the Vega Grid include the grid hardware, grid interconnection system and grid Operating system. Also Proposed are the important conceptions such as resource router, grid computing Protocol and grid browser etc.

Key words: Vega Grid, Resource router, Grid computing protocol, Grid browser

1 引言

随着高性能应用需求的迅猛发展, 单台高性能计算机已经不能胜任一些超大规模应用问题的解决。这就需要地理上分布、系统异构的多种计算资源通过高速网络连接起来, 共同解决大型应用问题, 即广域高性能的元计算技术(metacomputing)^[1], 也称为网格(grid)^[2]技术。

国外已有很多项目在进行网格体系结构的研究工作, 比较重要的有 Globus^[3]和 Logion^[4]。Globus 提出的网格体系结构是一种“沙漏”模型, “沙漏”的颈部定义了一组核心抽象和协议的集合, 位于顶部的各种高级操作可以映射到这个集合中, 同时这个集合可以映射到底部的各种不同的实现技术^[5]。Legion 系统则采用对象技术定义了一种支持元计算的软件体系结构^[6]。

网络技术的产生主要有以下几个原因:

- ①计算资源的广域分布。
- ②广域互联技术的发展。
- ③广域范围内资源共享的需求。

网络计算技术在产生的初期主要集中在高性能科学计算领域中, 现在网格计算不局限于此, 出现了使用于不同应用领域的网格技术。图 1 中根据不同类型的网络给出了对现有网络技术的分类。

中国科学院计算所正在进行织女星网络的研究工作。织女星网络体系结构的基本思想是把网络看成一台虚拟的超级计算机系统。从微机到小型机到大型机, 一个完整的计算机系统一般都由计算机硬件、操作系统和应用系统 3 部分组成。织女星网络体系结构设计继承了已有计算机系统的设计方法, 即将网络看成是一台虚拟的、具有单一系统映像的计算机系统。与现有的计算机系统类似, 织女星网络也将包括硬件、系统软件和应用 3 个组成部分。图 2 中给出了微机与网络、

织女星网格之间的对比。

在应用层,我们在研究信息网格^[1]和知识网格,在系统层,我们主要研究织女星系统软件即网格操作系统,而在硬件层,我们主要研究面向网格的曙光超级服务器^[8]。本文将重点讨论网格系统中的关键技术。



图1 网络技术的分类



图2 织女星网络与传统计算机系统的比较

本文的组织方式如下:第2节给出织女星网络体系结构概述;第3节给出了织女星网络硬件的组成和功能;第4节介绍了网格操作系统;第5节给出了织女星网络的使用模式;第6节给出了原型实现。

2 织女星网络体系结构概述

2.1 织女星网络的设计思想—VEGA

织女星网络总的特色是服务网络的思想,其概念及创新点和研究工作所遵循的指导思想是VEGA,即:①通用服务(Versatile services);②辅助智能(Enabling intelligence);③全局一体(Global uniformity);④自下控制(Autonomous control)。

(1) 通用服务

织女星(英文刚好是Vega)中的“v”包含3方面的含义:

第1个方面是“通用”,包含:支持多种应用类型;支持多种应用模式;

第2个方面是“服务”,其含义包括:应用和交换采用服务机制;网络及其应用的开发、构造、部署、运行、使用、管理都采用有一种称为“服务”的单一抽象接口;网络的功能和性能评价将不只是采用传统计算机的评价标准(如速度、加速比、性能价格比等),而应该以用户满意度为目标,用类似SLA(Service level agreement)的服务质量尺度来衡量。

第3个方面是“资源”。织女星网络提供的服务必须使用资源,而通用服务意味着广泛的、各种各样的资源。因此,“通用服务”的更深层次的含义是“通用服务与广泛资源”(universal services and resources)。

(2) 辅助智能

辅助智能(Vega)中的“e”有3层含义:

1)一个实用中的织女星网络应该比现在的Internet/web更具有智能性,包括:

①自动.信息、知识、服务的自动获取和自动生成;设备的自动接入。

②自我.自我描述、自我路由、自知(self aware),自在(context sensitive)等。