

支持e-Science的协同工作环境

南凯, 董科军, 马永征, 杨德婷, 李华飏
中国科学院计算机网络信息中心 北京 100190

摘要: e-Science将为科研手段和方法带来新的革命性的进步。协同工作环境是e-Science环境的重要组成部分。本文阐述协同工作环境对实现e-Science的重要意义和作用, 简介国内外相关的主要工作及研究热点, 并介绍我们正在开发的一个支持e-Science的协同工作环境套件 (Duckling) 的进展情况。

关键词: 协同工作环境 科研信息化 协同实验室 社会计算

Collaboration Environment for e-Science

Kai Nan, Kejun Dong, Yongzheng Ma, Deting Yang, Huabiao Li
Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190 China

Abstract: e-Science is bringing us new revolutionary progress on the way how we do research. Collaboration Environment is a critical part of e-Science environment. This paper states the importance and functions of Collaboration Environment towards implementing e-Science, surveys related works and main research topics, and introduces a software suite for Collaboration Environment for e-Science, namely Duckling.

Keywords: Collaboration Environment, e-Science, Collaboratory, Social Computing

1. 引言

以协同的方式进行工作自古就有，随着生产力和技术的进步，协同的范围不断扩大、水平逐渐提高。计算机的出现使得协同工作环境的方式和手段有了第一次飞跃，这次飞跃的典型标志是上世纪80年代计算机支持的协同工作（CSCW）。随着Internet的飞速发展以及信息便携设备（PDA）的大量普及，协同工作发生了以网络协同为显著标志的第二次飞跃，进入了一个新的阶段。

在实现e-Science的过程中，往往通过高速的网络联接高性能计算资源、海量科学数据库、科学数字图书馆、野外观测台站、大型科学装置、计算模拟的软件工具等等组成e-Science的基础设施与资源，实现资源集成共享，形成支撑科研人员大范围交流协作的协同工作环境，支撑科研人员组成跨学科、跨组织、跨地域的虚拟研究团队或者组织，通过远程实验观测、计算模拟、协同工作等新的科研方法和手段开展科学研究活动。

本文将阐述协同工作环境对实现e-Science的重要意义和作用，简介国内外相关的主要工作及研究热点，并对我们正在开发的一个支持e-Science的协同工作环境套件（Duckling）的进展情况进行介绍，最后是总结和展望。

2. 协同工作环境在e-Science活动中的意义和作用

数字化技术与计算机网络的发展，使人、科学仪器与装置、计算工具和信息联接在一起，提供了协同工作环境，消除了地域、时间、机构之间的边界，为科研活动提供了革命性的新模式，从而大大促进了科研活动中的信息共享、合作与交流，促进了学科的交叉，提高了工作效率和创新发现能力。虚拟组织、虚拟实验室或虚拟研究团队成为信息化条件下的新型研究组织形式。

支持e-Science的协同工作环境是一种为科研人员服务的虚拟实验室。它的主要用户是科学家，它建立在互连网络环境、超级计算环境和数据应用环境等信息化基础设施之上，利用和集成它们提供的基础性服务，构建支持网络化科研模式的综合性服务，支持面向具体科研活动的特色协同环境。具体

而言，这种虚拟实验室是由支持资源共享与协同工作的硬件、软件、数据、信息等资源和人员组织共同构成的有机的整体，是e-Science环境中面向用户（科研人员）的一个基础使用界面^[1]。

进一步来讲，支持e-Science的协同工作环境或虚拟实验室在技术之外，还可以有组织上的涵义。e-Science的三大资源要素——高性能计算、科学数据和协同工作的科研人员——通过虚拟实验室有机地结合起来。在e-Science示范应用的建设和实施过程中，根据各个应用的特点和需求，虚拟实验室提供的综合性服务和工具与各个e-Science应用的具体科研流程无缝地融合为一体，形成面向应用领域科研活动特点的网络协同工作环境。在e-Science发展的过程中，虚拟实验室在粘合信息化基础设施、催化新的科研模式和促推科研机制变革方面具有重要的战略意义。

协同工作环境在支持科研人员开展e-Science活动方面可提供以下主要功能和服务：

(1) 科研项目可以作为“虚拟组织”的形式成为虚拟实验室的用户。科研人员可参加项目相应的虚拟组织，使用虚拟实验室中相应的服务；也可以在虚拟实验室中与同行交流或寻找合作伙伴。重大的科研项目可以建立专用的虚拟实验室运行系统。

(2) 科研人员在虚拟实验室中享有个性化的使用环境，可以享受灵活定制的服务，可根据自身需要获取、组织和管理各类信息资源。

(3) 科研人员在虚拟实验室中可以集成使用院信息化建设的计算、存储、数据库、文献等多种资源，集成使用一些工具和服务，包括领域特定的应用。

(4) 科研人员可以与项目/虚拟组织内的合作者或其他人员高效、安全、方便地共享各种电子文档，基于智能搜索的方式使用和管理文档，并可进行协作式写作。

(5) 科研人员可以更高效地参与或组织一些学术活动和项目工作，如会议、论文评审、检查验收、考察活动等。

3. 相关研究工作及关键问题

协同工作环境相关技术受到国内外学术界的普遍关注和重视。国际上，荷兰的Virtual Laboratory

for e-Science (VL-e)^[2]项目、英国JISC的Virtual Research Environments (VRE)^[3]项目、GGF的Open Grid Computing Environment (OGCE)^[4]项目和SAKAI^[5]等项目在这方面进行了大量的研究和探索,取得了丰富的成果。在国内,国家自然科学基金委员会制定了“以网络为基础的科学活动环境研究”重大研究计划;科技部在国家科技基础条件平台的建设中启动了“网络协同研究与工作环境建设”项目。这些项目探索了如何将各种信息化资源集成为一个综合协作环境。

下面重点对VL-e, VRE, OGCE, SAKAI等进行介绍。

(1) VL-e

荷兰的VL-e (Virtual Laboratory for e-Science)项目的目标是将高性能网络和网格技术发展带来的推动力与更广泛的科学实验应用的需求结合起来。为科学研究中的分布协作以及实验提供一个虚拟的工作空间,使用分布的信息和通信技术来产生和递送结果。提供通用的功能以支持广泛和特定的e-Science应用环境。通过创建e-Science环境推动e-Science应用,并且开展方法学上的研究。VL-e项目的策略是沿着从应用到网络整个完整的e-Science技术链开展具体研究,并且关注于新的方法学和可重用组件的研究。

VL-e项目中包括四个任务:创建特定学科领域e-Science环境的科学原型;为可重用的信息技术组件开发应用方法学;逐步扩展并开展真实的科研实验应用;在有效实现的学科领域特定的e-Science环境中构建和传输知识。

(2) VRE

VRE(Virtual Research Environments)被定义为包含一组在线工具和网络资源,通过互操作技术支持或改善研究机构内或跨研究机构间广泛的科研人员开展科学研究活动。VRE的一个关键特征是能够促进研究人员和研究团队间的协作,为科研人员提供更有效的方式来协同地采集、处理和管理数据,以及协同的知识创建。VRE项目的目标并不是产生一个单一而且完整的VRE,而是定义和开发一个通用的架构,推进满足要求的组件的开发,并且将这些组件与满足科研人员需求的应用、服务和资

源组装为一个虚拟研究环境。VRE项目2006年得到英国JISC两百万英镑的滚动支持经费。

(3) OGCE

Global Grid Forum (GGF)的Open Grid Services Architecture (OGSA)计划为基于Web服务的、面向标准服务的网格框架定义了一个基础。OGSA允许不同供应商使用各种技术提供网格服务实现,不过这些技术能够通过GGF定义的OGSA标准兼容来实现互操作。GGF中的一个研究组Grid Computing Environment (GCE)重点研究了与客户端的网格开发工具有关的问题。它提供了一组工具集和用户界面来访问网格提供的服务和功能。为了解决并满足日益增长的网格用户社区的需求,GCE计划构建一个框架为客户端的网格开发提供一组标准的接口,这个接口类似于OGSA为服务器端的网格开发提供的接口。这种框架称为OGCE。

OGCE允许网格社区进行以下操作:开发客户端应用程序,这些程序可以跨多个网格后端实现进行互操作。提供可重用的代码,以支持基本网格访问模式的快速原型设计。提供一种开源的、可扩展的架构,可以根据社区的反馈,整体地、递增地构建该架构。访问使用异构技术实现的相同接口集。

(4) SAKAI

SAKAI是一个开源社区项目,聚焦于构建核心协作功能的资源,开发一个协同工具箱,用于教学以及即时协作。该项目提供一个框架,一方面“开箱即用”;另一方面,通过增加一些特别的工具组件,使之也能应用于特定领域。

SAKAI核心功能包括:公告、聊天室、Threaded Discussion、Drop Box、邮件归档、今日消息、新闻/RSS、偏好、讲演工具、资源、日程、web Content、Worksite Setup、WebDAV。SAKAI社区正在积极地开发新的工具扩展核心协同工具集。这些工具集没有包含在Sakai2.0发布中,但是正在由Lancaster和Cambridge大学的研究者进行积极地开发。这部分内容包括:Wiki based on Radeox、博客、共享显示、共享白板、音频多播和视频多播。

对于构建e-Science应用,SAKAI能够作为一个单独的系统提供人对人的交流技术,协调和社会化项目组成员的责任。通过这种方式,SAKAI提供了项目



图1: 协同工作环境套件 (Duckling) 组成结构图

或项目组管理工具的功能。当e-Science应用的其他元素跟上，SAKAI可以被集成到整个项目的门户中。

支持e-Science的协同工作环境的研究和发展是一个十分具有挑战性的问题，不仅仅是软件工具的研发，也不仅仅是若干计算机技术的应用，将涉及到信息技术和一些社会科学的交叉，这当中存在一些需要从更广阔的视角来看待的问题。在我们正在开展的一些e-Science应用的尝试中，已经切实感受到了这方面的问题和困难：如何让科研人员接受新的技术和工具，科研人员怎样才能积极有效地适应和参与新的科研环境和流程，e-Science活动中的各类角色应如何发挥作用，e-Science虚拟实验室该如何建、如何用、如何管？

网格计算^[6]、协同工作、Web服务和SOA^[7]等领域的前沿技术，无疑是协同工作环境研究开发中的一些关键问题。此外，“社会计算”是一个新的研究主题，有可能为e-Science协同工作环境的开发、应用和服务从一些新的角度提供理论指导。社会计算可以从两个方面或角度来看：一个是从计算机或更广义的信息技术在社会活动中的应用；另一个是从社会知识，或更具体的人文知识在计算机或信息技术中的使用和嵌入，反过来提高社会活动的效益和水平^[8]。应用社会学的知识指导e-Science协同工作环境的研究和开发，开展信息技术与社会学的交

叉研究，这也是协同工作环境研究发展中的一个重要的关键问题。

4. Duckling研究进展

协同工作环境套件 (Duckling) 是中科院计算机网络信息中心面向e-Science及中科院“十一·五”信息化发展的需求，研究开发的一套支持e-Science的协同工作环境的软件。这套软件包括虚拟工作台、核心工具集、资源与服务池、安全基础设施等主要部件，组成结构图如图1。

虚拟工作台是用户在虚拟实验室中主要的工作环境和界面，其主要功能是为用户提供统一的入口，支持用户以个性化的方式组织管理资源、使用资源，提供灵活的工具集成，提供方便、易用的使用界面。虚拟工作台以基于Web的服务门户的方式来实现。其目标是实现一个开放、灵活和可扩展的集成平台。虚拟工作台必须保证其开放性和可扩展性，支持虚拟实验室的持续开发并满足科研人员灵活多变的需求；支持构件的可重用性，使其能够有效地使用长期积累的科研资源；同时确保虚拟实验室的可用性和可访问性，使得它能获得科研人员的广泛接受。

虚拟实验室的核心工具集包括支持科研人员在e-Science环境下开展协同工作所需要的基础性、关

键性的工具。这个核心工具集会随着e-Science环境的发展，科研人员工作内容、方式、习惯的改变而逐步扩展或变化。核心工具集的功能设计应和现阶段科研人员从事科研活动的特点和水平相适应，优先提供需求迫切、条件成熟的工具。核心工具集的系统架构应具有良好的可扩充性。经过大量的调研和分析，我们认为现阶段具有普遍性和典型性的协同科研活动是科研人员在完成合作科研项目过程中的相关工作，这类活动对协同工具的需求也最具有现实的迫切性。

基于上述思路，当前的核心工具集包括虚拟组织管理工具、文档协同工具和活动组织工具等三大工具。

虚拟组织管理工具是一种面向群组（虚拟组织）的用户管理系统。管理工具中，个体科研人员以虚拟组织为单元进行灵活的组合。虚拟组织管理工具提供虚拟组织从创建、运行到撤销整个生命周期的各项管理功能，特别是面向虚拟组织的权限管理，包括解决用户在多个虚拟实验室应用中实现单一登录的问题。它的另一项重要功能是帮助科研人员与合作伙伴进行联系和交流，或寻找潜在的合作者。虚拟组织管理工具为整个虚拟实验室的应用提供了一个统一的用户管理的解决方案。

文档协同工具是一种面向组用户的协作式写作、文档共享和管理工具，支持一个虚拟组织当中的成员便捷、高效地共享和协同开发各类数字化文档，其重要特点是基于组的协作式写作和文档管理、基于搜索的文档定位，可以方便地实现一键上载、全文检索、智能标签、各处引用等功能。文档协同工具可在项目申请、项目执行等各个环节的大量文档工作中发挥重要作用。

活动组织工具的作用是协助科研人员组织开展各类科研活动。一个典型的例子是学术会议的组

当前科研活动中有大量的会议，从项目组内的讨论到大型的国际学术交流，在e-Science环境中这类活动更加频繁和日常化（如视频会议）。安排会议日程、共享会议材料等往往涉及到参会人员之间的大量沟通和协作，活动组织工具可帮助虚拟实验室的用户在这些工作中提高效率。除会议外，其他的科研活动如野外考察、联合观测等也有类似的需要组织安排的工作，结合e-Science示范应用在这些方面的具体需求，在会议类活动的基础上进行扩展或定制，逐步支持更多类型的科研活动。

e-Science环境中包括高性能计算、科学数据、科技文献、软件/算法、科学仪器等大量资源，协同工作环境中的资源与服务池部件负责实现将这些资源根据用户需要集成到虚拟实验室中供用户方便地使用。资源集成的方案不仅要在技术上可行，在管理上、机制上也要可行，这是相当有挑战性的。虚拟工作台插件技术将科研应用中的特定功能和用户接口连接起来，使得它们构成了虚拟工作台的一个可插拔的组件，保证虚拟工作台的定制性和可扩展性，简化虚拟工作台的构建，使得它们在长时间的运行过程中更加具有可扩展性、可重用性、可管理性和可维护性。

workflow支撑平台采用基于服务的工作流来构建科研业务流程，通过科研业务工作流的自动或半自动执行，提高科研的效率和产出。

安全基础设施是实现e-Science协同工作的一项重要前提，必须在协同工作环境下建设一套可靠、方便、高效的安全基础设施，实现安全环境下的虚拟组织管理、数据传输及协同协作。协同工作环境中采用数字证书作为基本的用户身份标识，建立了基于PKI的安全基础设施，由自建的CA系统向协同工作环境用户提供证书服务。



图2: 协同工作环境套件 (Duckling) 应用案例

到目前为止，协同工作环境套件 (Duckling) 的原型版本已在中科院和国家的一些重大科研项目中得

到应用（如图2）。

在国家科技基础条件平台项目“天文学网络虚拟实验室示范”中，与大型天文望远镜的申请审批、联合观测以及观测数据管理等科研流程结合在一起，使用Duckling建设了提供包括大型望远镜观测时间在线申请、多人同时在线的网络远程协同观测、观测数据管理等学科特点鲜明的网络协同工作环境，该虚拟实验室已逐步成为国内天文领域研究和相关活动的一个基础性平台，发挥了重要的信息化支撑作用。

在中国科学院知识创新工程项目“面向蛋白质科学的高性能计算研究”和“北京及周边地区奥运大气环境监测和预警联合行动计划”中，使用Duckling为项目组中多单位不同地区不同学科的科研人员提供了网上的协同工作环境。

Duckling也应用在“青海湖联合科研基地”、“计算化学虚拟实验室”等跨学科、多个研究所参

与的科研活动中。此外，Duckling开发团队自身也在使用这样的协同工作环境。

5. 总结与展望

e-Science将为科研手段和方法带来新的革命性的进步，进而极大地推动科学技术的发展^[9]。协同工作环境是e-Science环境的重要组成部分。协同工作环境套件Duckling是这个方向上的一个探索，经过近一年的开发已经形成了原型版本并在少数项目中试用，计划将在2008年底之前发布Duckling的第一个正式版本。同时我们看到，再先进的软件也只是帮助和促进科研协同的工具，e-Science的发展需要多方面因素的共同推进，本文第3节中讨论的社会学方面的问题也将是支持e-Science的协同工作环境的一个有待进一步研究的关键问题。🌀



参考文献

- [1] 中国科学院信息化工作领导小组办公室. 中国科学院“十一·五”信息化发展规划. 2006年4月.
- [2] Virtual Laboratory for e-Science. 2008. <http://www.vl-e.nl>.
- [3] Virtual Research Environment programme. 2008. <http://www.jisc.ac.uk/whatwedo/programmes/vre2.aspx>.
- [4] Open Grid Computing Environment. 2008. <http://www.collab-ogce.org/>.
- [5] Sakai Project. 2008. <http://sakaiproject.org/>.
- [6] I. Foster, C. Kesselman and S. Tueckle. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. in Lecture Notes in Computer Science, 2001, Vol. 2150.
- [7] I. Foster, H. Kishimoto, A. Savva, D. Berry, A. Djaoui, A. Grimshaw, B. Horn, F. Maciel, F. Siebenlist, R. Subramaniam, J. Treadwell, and J. Von Reich. The Open Grid Services Architecture, Version 1.0. <http://www.ggf.org/documents/GFD.30.pdf>, 2005.
- [8] 王飞跃. 社会计算——科学、技术与人文的数字化动态交融. 中国基础科学, 2005, 7(5):5-12.
- [9] US National Science Foundation. Cyberinfrastructure Vision for 21st Century Discovery. March 2007.

作者信息



南 凯

中国科学院计算机网络信息中心，博士，副研究员。主要研究方向为网格与协同计算、数据网格等。