



本图片由作者提供

天文学作为一个古老而又神秘的学科，伴随最近十几年来天文望远镜和终端设备技术的发展、天文观测数据速度的成倍增长，科学家面对海量数据发现新天体新问题的传统研究方式也将面临着新变革。虚拟天文台的概念于2002年在美国被提出，该概念的诞生正是基于e-Science技术的应用。

天文学领域的e-Science应用 ——虚拟天文台

杨阳 崔辰州
中国科学院国家天文台, 北京 100012

摘要: 作为e-Science应用的天文学领域, 虚拟天文台是数据密集型的在线天文学研究和教育环境, 它利用先进的信息技术实现对天文信息的无缝和统一访问。上世纪九十年代末提出的虚拟天文台的概念满足了由于天文学领域数据雪崩所带来的挑战。本文简要地介绍了全世界VO项目的国际虚拟天文台联盟的状况, 着重说明了中国虚拟天文台的体系结构、数据访问系统以及各种各样的服务。

关键词: e-Science; 天文学; 网格; 虚拟天文台

Virtual Observatory: the e-Science Implementation in Astronomy

Yang Yang, Cui Chenzhou
National Astronomical Observatories, Chinese Academy of Sciences, Beijing
100012, China

Abstract: Virtual Observatory (VO), the e-Science implementation in Astronomy, is a data intensive online astronomical research and education environment, taking advantages of advanced information technologies to achieve seamless and uniform access to astronomical information. The concept of VO was introduced in late of 1990s to meet challenges brought up with data avalanche in astronomy. In the paper, current status of International Virtual Observatory Alliance from world wide VO projects are reviewed, a detail introduction of architecture, data access system and various services of Chinese Virtual Observatory is given.

Keywords: e-Science; Astronomy; Grid; Virtual Observatory

1. 引言

虚拟天文台的概念于2002年在美国被提出,该概念的诞生是基于e-Science技术的应用。e-Science由英国在2000年提出,是为了应对当时各学科研究领域所面临问题的空前复杂化,利用新一代网络技术(Internet)和广域分布式高性能计算环境(Grid)建立的一种全新科学研究模式,即在信息化基础设施支持下的科学研究活动。此后,该理念便席卷世界各科研领域,大家都试图用它来解决本学科在研究过程中遇到的种种难题。天文学作为一个古老而又神秘的学科,伴随最近十几年来天文望远镜和终端设备技术的发展、天文观测数据速度的成倍增长,科学家面对海量数据发现新天体新问题的传统研究方式也将面临着新变革。e-Science的产生对于天文学领域来说十分及时,也促进了虚拟天文台的诞生。2002年6月,国际虚拟天文台联盟(IVOA)^[1]正式成立,它的使命在于推进国际合作与协作,建设能够综合利用各国天文台数据的、完整的、能协同工作的虚拟天文台,开发、配置必要的工具、系统,建立相应的组织结构。

美国、欧洲、日本、加拿大、英国、印度、法国等纷纷提出了自己的虚拟天文台计划。中国于2002年成立了中国虚拟天文台(China Virtual Observatory, China-V0),并于同年加入IVOA。目前,IVOA的成员国已达16个国家。

2. 虚拟天文台的体系结构

中国虚拟天文台最大的特点是:它是伴随国家重大科学工程大天区面积多目标光纤光谱天文望远镜(LAMOST)^[2]项目成长起来的,LAMOST是目前光谱获取率最高的天文望远镜,届时将产生海量的数据。如何完成一个完整的“从数据到科学”的网络化协同工作平台是非常重要的一个课题,因此,LAMOST的这些现实需求刺激了China-V0的创立和发展。此外,China-V0的发展也处

于中国科研信息化的大潮之中。这样,China-V0就获得了两个外力的带动:一方面它的服务目标与LAMOST以及相关的科学目标紧密联系,为中国天文学家建立了一个便捷高效的天文资源访问平台,定位明确、清晰;另一方面它的发展也将得到国内e-Science相关项目的支持。

2003年,China-V0的系统设计第一次得到完整的阐述^[3],2004年,崔辰州和赵永恒提出了China-V0的体系结构^[4],至此,China-V0的概念框架已经形成。在这个概念框架中,China-V0类比于网格的层次结构设计了自己的层次结构(见图1)。图中,China-V0被定义了四层结构,从下至上依次为构造层、资源层、汇集层和用户层。构造层包括各种数据资源、计算资源、网络资源和存储资源等。各种数据资源在虚拟天文台这样一个数据密集



图1 中国虚拟天文台体系结构

型在线研究平台中占有非常关键的作用，是V0成功运作的基础和前提。它主要包括星表、星图、光谱、时序数据、计数测量数据、模拟数据、多媒体数据、天文文献等。资源层是以网格为代表的中间件为基础，配合其它系统服务工具，利用标准的数据模型和服务模型，通过抽象化实现统一的数据访问和计算访问等功能。汇集层提供具有天文特色的各种服务，如数据处理、数据挖掘、统计分析、可视化等应用服务。用户层是整个体系的最高层，包括V0客户端和V0门户，直接与虚拟天文台用户接触，用户层的基本职能是用户任务提交和处理结果返回，主要功能包括用户登录、身份认证、V0资源浏览、任务编制和提交、结果显示、数据下载等。物理上，整个系统是分布式的，在网络环境下实现；逻辑上，通过网格等中间件技术实现一个统一的整体。

由此我们可以很明显的看出网格技术将承载着中国虚拟天文台成长与发展，并且网格的研究目标从资源的角度上来看和虚拟天文台的目标是一致的，那就是通过Internet架构实现最大限度的网络资源共享。因此，它就自然而然地在虚拟天文台的研究中得到了广泛的应用。由于网格技术作为一种较新的技术，也同样处于一种探索阶段，其体系结构在短短几年时间内也有所变迁，从一开始的OGSA架构转向了WSRF架构。WSRF的全称是Web Service

Resource Framework，它通过一个隐含的资源模式为在Web服务之间创造有状态的资源定义了一个系统。该架构定义了使用Web服务来访问有状态资源的一系列规范，包括Web服务资源特性（WS-Resource Properties）、Web服务资源生命周期（WS-Resource Lifetime）、Web服务基本故障（WS-Base Faults）和Web服务服务组（WS-Service Group）规范。这些新的规范在于使数据的值通过Web服务交互得以持久化，并且

作为Web服务交互的结果而保存。因此，伴随网格技术体系的不断变迁，虚拟天文台也会处于不断的变化之中。可以说虚拟天文台的建立和实现需要网格技术的支持，同时，虚拟天文台也将为网格技术的发展提供最好的试验床。我们还可以看出，China-V0的体系结构是一个庞大的虚拟天文台范例，完全实现这样一个大型的系统也会存在很多客观的困难，比如它所依托的网格中间件技术尚不成熟，牵涉面广，特别是涉及IVOA的很多互操作

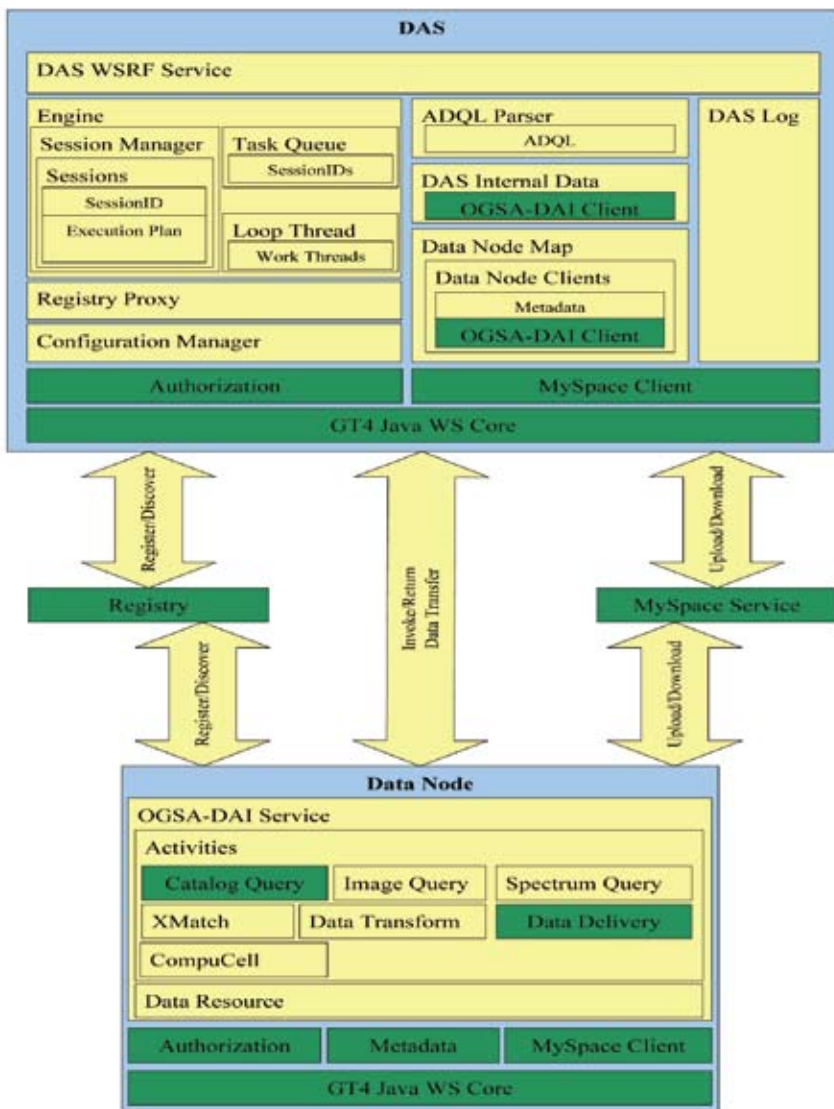


图2 虚拟天文台数据访问系统体系结构

► 协议并不成熟甚至尚未出现, 以及系统开发者多数欠缺网格技术以及天文背景知识等。中国虚拟天文台在不能一口气完成一个大系统的研究与开发的情况下, 选择了在这个框架指导下进行小项目的应用开发作为中国虚拟天文台发展的方向。

3. 虚拟天文台数据访问

世界各国天文数据中心的数据存储主要有三个特点: 首先, 数据的分布性使天文学家对它们的访问变得更加复杂, 对来自不同数据库的数据进行交叉认证花费的时间较长; 其次, 数据的异构性使得用户必须熟悉不同数据库的操作方法, 是非常耗时耗力的过程; 最后, 天文学产生的海量数据使得天文学家基于这些数据难以完成得更加精确或有新的发现, 且存储管理的成本高。那么如何提供给天文学家一个统一访问这些异地异构的数据资源的方案, 是虚拟天文台的一个重要课题。因此, 我们设计了数据访问服务系统 (VO-DAS) [5]。该系统要实现数据节点的自动资源发现, 能够统一访问异地异构的天文数据库, 支持天文数据的交叉认证或联合查询, 能够访问不同类型的数据资源, 包括星表数据、图像数据和光谱数据, 支持尽可能多的数据访问量以及数据结果的多种存储格式。

VO-DAS是一个建立在网格技术基础之上的异地异构天文数据访问平台。它以Globus Toolkit [6]

网格系统为基础环境, 利用OGSA-DAI [7] 提供的基本数据访问服务, 结合国际虚拟天文台联盟 (IVOA) 制定的数据访问服务规范和本项目的实际需求, 以Web Service和Grid Service的形式实现对异地异构数据库系统、文件系统数据的访问功能。VO-DAS体系结构 [8] 如图2所示。VO-DAS系统主要分成两个部分: DAS和Data Node。DAS是一个基于Web服务资源架构 (简称WSRF) 的任务管理服务。一方面它是一个用户作业调度和管理服务, 另一方面它管理Data Node, 保证用户的数据查询请求找到正确的Data Node, 找到正确的数据资源。Data Node是一个基于OGSA-DAI WSRF中间件的数据资源节点。数据 (星表, 图像或光谱) 都通过Data Node提供给DAS服务。数据的拥有者不需要了解DAS和Data Node之间的通讯方法, 只要按照要求将自己的数据配置到一个就近的Data Node之下即可发布到DAS服务上。

4. 虚拟天文台服务

这几年来, 中国虚拟天文台紧随计算机技术和网格技术日新月异的变化, 把天文应用产品的研究与开发作为自身的发展方向。团队经过多年来的努力, 目前已经开发完成并向国内外用户发布了一系列实用的工具和服务。

2005年, China-VO发布了OpenOffice用VOTable转换插件VOFilter [9]。这是一个非常小

的转换工具, 没有一行执行代码, 而是用XSLT实现了VOTable和OpenOffice表格文档格式的相互转换。这个转换工具可以让OpenOffice也能够浏览和编辑VOTable格式的天文星表数据。在VOTable浏览工具还比较缺乏的时候, 这样一个小小的插件有助于天文学家从天文数据网站下载了VOTable格式的数据以后迅速浏览和编辑它们, 还可以借助OpenOffice将VOTable格式的数据文件转换成自己需要的文本格式, 以便于自己编写的程序的读取。

同一年, China-VO发布了VOImpat [10], VOImpat适合进行快速的星空位置查找和主要星表的联合查询。它支持DSS的图片检索, 只要给出天空的坐标就可以快速导入DSS的剪裁图像。同时, 为了能够获得天体的更多信息, VOImpat还可以借助SkyPortal (JDL数据挖掘原型中一个数据访问模块) 查询USNO、2MASS、NVSS等不同观测波段的主要巡天星表, 将查询到的星表数据和图像数据结合, 提供每个天体的综合信息。此外, 在要求不高的情况下, VOImpat还可以浏览本地FITS格式的图像文件以及VOTable格式的星表文件, 并具有一定的图像处理能力。

2006年, China-VO又发布了SkyMouse, 一个桌面天文信息搜索工具, 它可以通过鼠标选取屏幕上的词组, 从多个天文信息数据库中查找对应名称的天体数据、图像和相关文献 [11]。该工



图3 SkyMouse的客户端



图4 FitHAS主界面

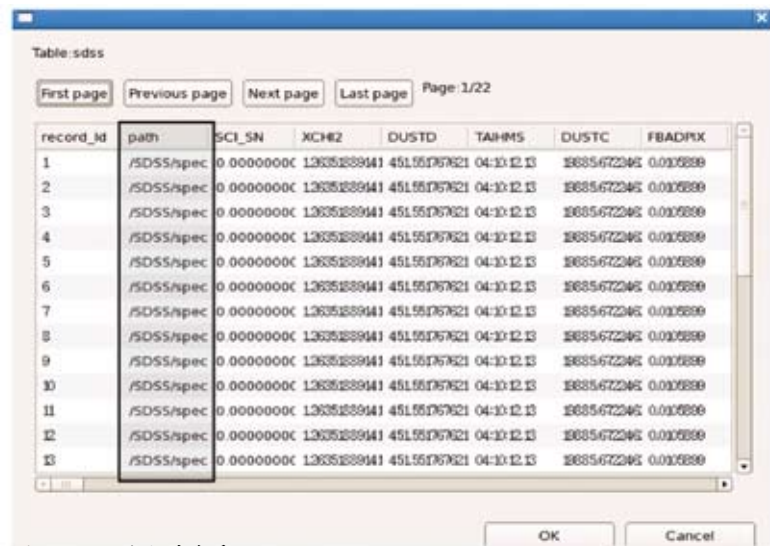


图5 FITHAS数据库查看

具可以让天文学家在阅读文献的时候快速浏览某个感兴趣的天体的数据信息，简单而实用。它同时还可以成为天文爱好者的工具，了解更多的天体知识。该软件客户端界面如图5所示，由图可知它集成了Simbad、NED、VizieR Catalogue等多个天文数据服务。

2007年又完成了FitHAS^[12]开发。FitHAS是一个批量读取FITS文件头并导入数据库的工具，主界面如图4所示。对于历史观测遗留下来的成千上万的FITS文件，由于以前没有用数据库进行管理，如果需从中找到有用的内容是非常困难的。FitHAS会自动读取一个文件夹中的所有FITS文件头信息，将它们装载到指定的数据库中，图5所示为装载后的数据库内容。这样，如果需要查找某个特定天体相关的FITS文件，可以直接利用数据库以特定查询条件进行查询，找到相关的FITS文件的相对存储位置。该工具对于整理和管理FITS文件有很大的帮助，从而从数据收集的角度支持了天文资源共享。

5. 结论及展望

虚拟天文台的建设与发展，为e-Science在天文学领域的应用提供了一个很好的试验田。中国虚拟天文台的产生正是由于国际e-Science的大浪潮以及国内LAMOST现实需求的驱动才得以在短短的几年内迅速成长，开发了一系列的工具服务，为天文

- 学家的科学研究工作方式提供了便捷高效的手段。但是中国虚拟天文台在很长时间内要保持可持续发展,正如赵永恒研究员所说的要坚持“四项基本原则”,即“与国际虚拟天文台联盟标准兼容原则”、“以网格为基础架构原则”、“开放的软件、廉价的硬件原则”、“在战争中学习战争原则”,还有很多工作需要做。因此,我们也将原有的工作基础之上,不断地利用本土的优势为其注入新生的活力,并着力加强与国际方面的合作与交流。



参考文献:



- [1] IVOA.国际虚拟天文台联盟. <http://www.ivoa.net>.
- [2] Y.H.Zhao.LAMOST project and its scientific goals. Publications of the Yunnan Observatory,1999,12:1-7.
- [3] 崔辰州.中国虚拟天文台系统设计. PhD thesis.中国科学院研究生院, 2003.
- [4] C.Z.Cui,Y.H.Zhao.Architecture of Chinese Virtual Observatory. Astronomical Research and Technololgy. Publications of National Astronomical Observatories of China (ISSN 1672-7673),2004,1(2):140-151.
- [5] 刘超,田海俊,高丹,等.异地异构天文数据资源的统一访问.天文技术与方法, 2007.
- [6] Globus Toolkit.<http://www.globus.org>.
- [7] OGSA-DAI.<http://www.ogsadai.org.uk>.
- [8] 刘超. 基于虚拟天文台的数据挖掘技术及其在银河系晕结构研究中的应用. PhD thesis.中国科学院研究生院, 2003.
- [9] C.Z.Cui,M.Dolensky,P.Quinn,Y.H.Zhao,F.Genova. VOFilter:Bridging Virtual Observatory and Industrial Office Applications.Chinese Journal of Astronomy and Astrophysics,2006,6:379-386.
- [10] D.Wang,Y.H.Zhao.VO-IMPAT:Image Processing and Analysis Toolkit for the Virtual Observatory of China.Publ. Nat. Astron.Obs.China,2006,3:295-303.
- [11] C.Cui,H.Sun,H.Zhao.SkyMouse,An Integrated On-line Astronomical Information Access System.The Virtual Observatory in Action:New Science,New Technology,and Next Generation Facilities. 26th meeting of the IAU. Special Session 3,17-18,21-22 August,2006 in Prague,Czech Republic,SPS3,#55,3,2006,8.
- [12] Cui C. Z.,Li W.,Yu C.,Xu Z.,Zhao Y.H..Search and Location of FITS Data Files. Astronomical Research and Technology. Publications of National Astronomical Observatories of China,2007.

收稿时间:2009年4月3日

作者信息



杨阳

国家天文台, 硕士, 助理研究员, 研究方向为虚拟天文台、天文信息技术。



崔辰州

国家天文台, 博士, 副研究员, 研究方向为虚拟天文台、天文信息技术。