

ccidnet????

出版日期：2004-09-13 总期号：1347 本年期号：68

## 中国计算机报

本期导读 0

要闻综合 0

中国信息化 0

网络与通信 0

软件与服务 0

产品与应用 0

渠道与市场 0

华南专刊 0

西北专刊 0

东北专刊 0

存储世界 0

网吧新天地专刊 0

## 虚拟天文台——网格技术最好的试验场

中国科学院国家天文台 崔辰州

## 走近虚拟天文台

天文学是一门古老而又永远充满新发现的科学。然而，“靠天吃饭”的状况一直困扰着天文学家。变幻莫测的天气经常打乱天文学家们的研究计划。同样，我们普通人要想观测有趣的天象也首先要查询天气预报。这两年的日月食、流星雨、火星大冲、水星凌日、双彗奇观，就有许多朋友由于天气的原因而无法欣赏。

20世纪90年代，一场新的革命性的变化开始在天文学领域进行。这一变化是由前所未有的科技进步所推动的，特别是望远镜的设计和制造技术、高性能计算和互联网技术。

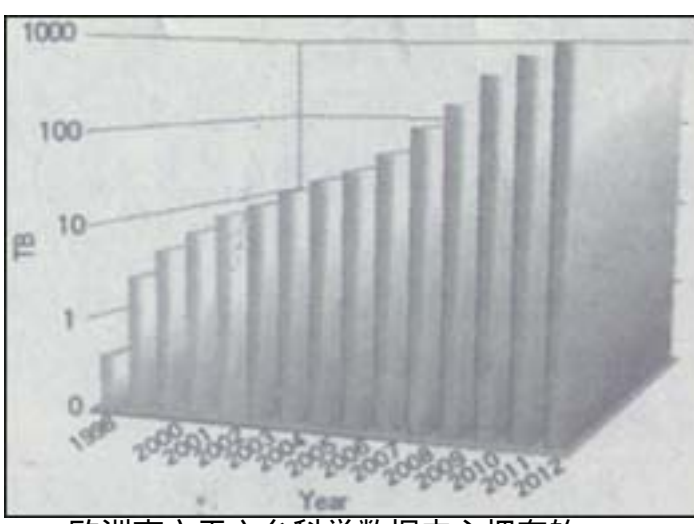
望远镜技术的进步使得人类可以建造更大、更具威力的望远镜，比如8米的下一代空间望远镜、30米甚至100米的地面光学望远镜、500米的射电望远镜等等。

不但计算机的发展遵循摩尔定律，在过去十多年中技术进步使得天文学的数据收集能力也在遵循摩尔规律。图中显示的是欧洲南方天文台科学数据中心所拥有的数据的增长情况，这在很大程度上代表了整个天文界数据量的增长趋势。从图中可以看出，天文数据增加一倍的时间还不到一年，比摩尔定律还快。巨大的数据产出，在天文学历史上第一次使天文学家得到的数据多得用不了。如何高效、充分的挖掘这些数据的科学价值？天文学家正面临数据爆炸所带来的巨大挑战。

同时，天文研究方式的变化也离不开高性能计算和互联网技术的快速发展，它使得在不同地点间进行海量天文数据的交换与传输成为可能，使得世界各地的天文学家都能够访问和使用这些数据。

虚拟天文台（Virtual Observatory，简称VO）是古老的天文学与最新的信息技术相结合的产物。受制于天的天文观测方式将随着虚拟天文台的诞生而发生彻底的改变。

虚拟天文台将利用最先进的计算机和网络技术将各种天文研究资源，包括各种观测数据、天文文献、计算资源、存储资源、各种软件工具，甚至天文观测设备，以某种统一的服务模式无缝地汇集在系统中。天文学家只需登录到VO系统便可以享受其提供的丰富资源和强大的服务，使自己从数据收集、数据处理这些繁琐的事务中彻底摆脱出来，从而把精力集中在自己感兴趣的科学问题上。虚拟天文台将使天文学研究取得前所未有的进展，它将成为开创“天文学发现新时代”的关键性因素。虚拟天文台可被全球各种各样的群体访问，特别是那些没有经济能力建造和运行大型观测设备的国家，能大大促进发展中国家和不发达国家的天文研究。虚拟天文台是非常好的公众教育设施。它让普通公众和天文学家一样能接触到真实的天文资源和服务，让公众也有可能做出新的天文发现。



欧洲南方天文台科学数据中心拥有的

数据的增长情况

## 虚拟天文台研究现状

1999年，美国科学院天文与天体物理发展规划委员会在名为“新千年的天文学和天体物理学”的未来十年发展规划中把建立（美国）国家虚拟天文台(NVO)作为最优先推荐的项目之一，首次提出了虚拟天文台的概念。此后，世界各国的天文学界迅速响应，纷纷提出了各自的虚拟天文台研究计划，在全球掀起了一场虚拟天文台的研究浪潮。

虽然这些计划来自不同的国家，有着不同的天文和信息技术背景，但是他们之间有许多非常重要的共同点：每个项目都在寻求数据密集型天文研究的出路，都在力图挖掘现有以及未来海量天文数据的潜力。从整个天文界的眼光来看，这些项目的最终目标就是建立国际虚拟天文台（IVO）。为了将各国在虚拟天文台方面的努力联合在一起，2002年6月成立了国际虚拟天文台联盟（IVOA）。中国、英国、欧洲、法国、加拿大、印度、德国、澳大利亚、日本、俄罗斯、意大利、韩国、匈牙利虚拟天文台，这14个项目目前都是国际虚拟天文台联盟的成员。

“中国虚拟天文台（China-VO）”计划是以中国科学院国家天文台为首的中国天文学界在2002年提出的。中国虚拟天文台将作为未来国际虚拟天文台的有机组成部分，为中国天文学界提供一个全新的网络化研究平台，引领中国天文学进入数据密集型在线科学研究新时代。

## 虚拟天文台与网格技术

通过上面的介绍我们可以很明显的看出网格技术在很大程度上与虚拟天文台的发展目标一致，利用网格技术作为基础设施发展虚拟天文台是必然趋势。

一方面，虚拟天文台的建立和实现需要网格技术的支持。虚拟天文台的最终发展目标就是实现全球天文数据的高级共享，同时提供一整套的智能化工具。TB量级甚至PB量级大型天文数据产出项目的不断涌现，对数据存储、数据管理、数据传输、数据检索等技术提出了更高的要求。在如此海量分布式数据的基础上进行科学研究，就必须有全新的数据共享、数据互操作、作业调度、数据可视化、数据统计分析、数据挖掘、数据安全等工具的支持。

虚拟天文台的这些需求正是网格技术要实现的目标。网格技术将实现把整个互联网整合成一台巨大的超级计算机，实现计算资源、存储资源、数据资源、信息资源、知识资源、专家资源的全面共享，为用户提供一步到位的服务。因此，虚拟天文台把网格技术作为自己的技术基础将是可行而明智的选择。

另一方面，虚拟天文台将为网格技术的发展提供最好的试验场。为什么这样说呢？这是因为e-Science时代的天文数据有着其他学科数据所无法比拟的特点：

\* 天文数据绝大部分是开放数据。国际上许多大型天文观测项目的观测数据都会及时在互联网上公布，这为数据共享提供了良好基础。

\* 天文数据很少有商业价值。这意味着相互之间不存在保密性的问题，可以将数据和结果进行自由共享，非常适合国际性的联合研究与试验。

\* 天文数据是海量的数据。例如美国正计划建造的“暗物质望远镜”，每天的观测数据高达18TB。这非常适合网格环境下海量数据处理和数据挖掘算法的研究。

\* 天文数据是真实而归档完好的数据，并提供在线访问服务。

\* 天文数据是高度复杂的数据。海量高度复杂的数据对数据处理、数据挖掘、可视化等研究提出了挑战。

## 虚拟天文台对网格技术的需求

为了比较明确地定义虚拟天文台对网格的功能需求，我们以“网格环境下的银河系化学演化研究”为例探讨虚拟天文台需要网格提供哪些功能。

它需要用户提供的是：恒星名称或者代码。网格系统输出的结果包括：轨道参数、轨道图、化学成分数据、化学成分与轨道参数的相关分析结果。在网格环境下，研究工作可以按照如下几步进行。

1. 用户通过浏览器上按照一定的格式编辑好的恒星名称列表。这需要网格系统支持：安全认证、临时用户空间的申请与分配。

2. 网格系统根据用户的要求和当前资源、服务的可用情况生成工作流程。这需要网格系统支持：工作流生成、作业监控、资源和服务的动态注册与发现、访问控制。

3. 网格系统通过天体名称解析服务将用户输入的恒星名称解析为统一的命名格式。这需要网格系统支持：语义网功能。

4. 网格系统以恒星名为关键字在全球范围内可用的数据库中检索所需数据，将检索结果合并为统一的文件保存到临时空间中。这需要网格系统支持：数据互操作、分布式并行检索和传输、检索结果的融合、元数据服务。

5. 从天文专家库中查找最佳的银河系引力场模型，利用上步检索到的数据计算恒星在银河系中的运行轨道，得到轨道参数。这需要网格系统支持：专家系统、应用软件开发环境、现有计算工具的网格集成环境。

6. 利用三维绘图服务将计算得到的恒星轨道在三维空间中进行绘制。这需要网格系统支持：现有可视化工具的网格集成环境。

7. 对恒星的化学组成和轨道参数进行相关分析。这需要网格系统支持：现有数据挖掘工具的网格集成环境。

8. 将上面各步的中间结果以及最终结果保存在用户临时空间，通知用户工作完成。这需要网格系统支持：用户通告、临时空间的生命期管理、衍生数据的复用。

虚拟天文台最终要实现对各种天文科研资源的融合。这样的发展目标为网格技术提供了独一无二的试验场。从网格基础设施的构建，到网格操作系统的开发，最后到网格应用服务的实现，虚拟天文台为网格技术提供了一整套的测试空间。