

努力探索中国特色科研信息化之路

——中国科学院科研信息化实践与案例

努力探索中国特色科研信息化之路

——中国科学院科研信息化实践与案例

中国科学院信息化工作领导小组办公室

2009年12月3日，北京

CONTENTS

目录

2	前言
4	面向需求，加强科研信息化基础设施建设
5	1. 快速发展的中国科技网
9	2. 不断完善的超级计算环境
17	3. 稳步提升的科学数据应用能力
24	服务科研，深化科研信息化创新实践
25	1. 世界水平的数据密集型计算网格
26	2. 网格技术支撑的空间天气数字实验室
28	3. 碳收支集成研究的科研信息化应用示范
30	4. 网络环境支持下的科学考察和合作研究
32	5. 虚拟天文台探索浩瀚太空
33	6. 野生鸟类资源及疫病监测与风险评估
35	7. 新一代工业生物技术知识环境
37	8. 生物大分子研究e-Science服务中心
38	9. 科研信息化综合性协同工作平台
40	展望

前言

20世纪中叶以来，信息技术的快速发展与广泛应用加速了社会信息化的进程，使人类社会迈入了信息社会的新阶段。信息化是现代化的助推剂，一个国家的信息化程度已经成为其现代化程度的重要标志之一。信息化源于科学技术的进步，同时又推进了科学技术自身的发展。科学研究活动的信息化在世界范围内方兴未艾，各发达国家都制定了明确的国家科技活动信息化战略，投入巨资支持建设科研信息化基础设施，为提升本国的科技创新和国家竞争能力发挥了重要作用。

中国科学院院长路甬祥指出，信息化基础设施是当代最重要的科研基础设施。

中国科学院高度重视信息化建设与发展，逐步形成了全面推进科研和管理信息化，打造“有创意、有特色、功效一流”

的中国科学院信息化发展愿景。

中国科学院江绵恒副院长指出：“e-Science的实质就是‘科学研究的信息化’，是信息时代中科学研究环境和科学研究活动的典型体现。它不仅包括采用最新的信息技术，如Grid等，建设起来的新一代的信息基础设施，更有在这种基础设施和相关支撑技术构成的平台上开发的科学研究的应用，以及科学家们在这样一个前所未有的环境中进行的科学研究活动。e-Science的实现将为科学家们提供一个信息化的科学研究环境，改变他们从事科学研究活动的方法和手段，甚至直接影响到一些学科的发展。”

近年来，中国科学院大力加强科研信息化基础设施建设，组织实施了互联网络、超级计算和科学数据应用等三大环境的建设，全面推进科研和管理信息化应用，部署了网络化科学

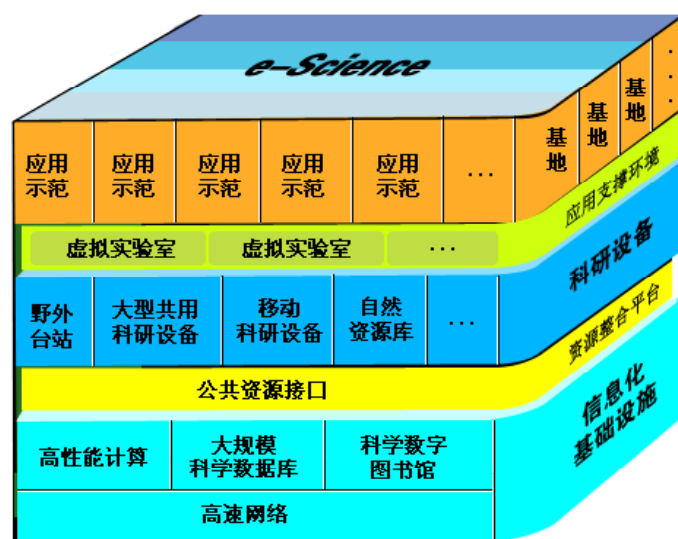


研究 (e-Science)、网络化运行管理 (ARP)、网络化信息发布、网络化教育培训和网络化科学传播等五大平台的建设, 组织实施了一批e-Science示范项目。

几年来的实践证明, 信息化对促进中国科学院科研与管理水平的提高, 推进各学科的创新跨越, 提升科技创新能力, 形成高效的国际国内合作, 逐步展现出不可替代、至关重要的作用。

科研信息化设施与应用的基本要素与内涵包括:

- 宽带网络, 支撑一切基于网络的科研信息流动。
- 主要的科研信息化基础设施, 包括超级计算中心、科学数据中心和数字图书馆, 它们是支持各类科研活动的共用技术设施, 并向用户提供应用公共资源的接口。
- 各类信息化的科研活动设施和工具、方法, 包括大型科学装置、实验室仪器设备、野外观测台站、数字标本馆、移动数字终端等等, 直接服务于各自领域的科研活动。
- 虚拟实验室, 提供基于网络和所有这些资源的应用环境, 为虚拟研究组织或团队的科研活动提供支持与服务, 是科研工作者与科研设施间的信息化界面。



e-Science 结构模型图

简言之, 科研信息化就是在科研活动的各个环节采用先进信息技术提供支持与服务的环境和由此而产生的新的科研模式。科研信息化将随着各个学科和技术领域信息化应用的深化, 随着信息技术不断进步, 其内涵将不断丰富, 并将对科技活动本身产生革命性的影响。

• 面向需求，加强科研信息化基础设施建设

中国科学院在国家有关部委的支持下，经过多年建设，逐步形成了由高速科技网络、超级计算环境和科学数据中心构成的先进科研信息化基础设施及其服务环境，不断改革运行体制和机制，得到进一步的完善和快速发展，不仅服务于科技创新，而且服务于国家重大工程建设、经济社会发展与企业技术创新。

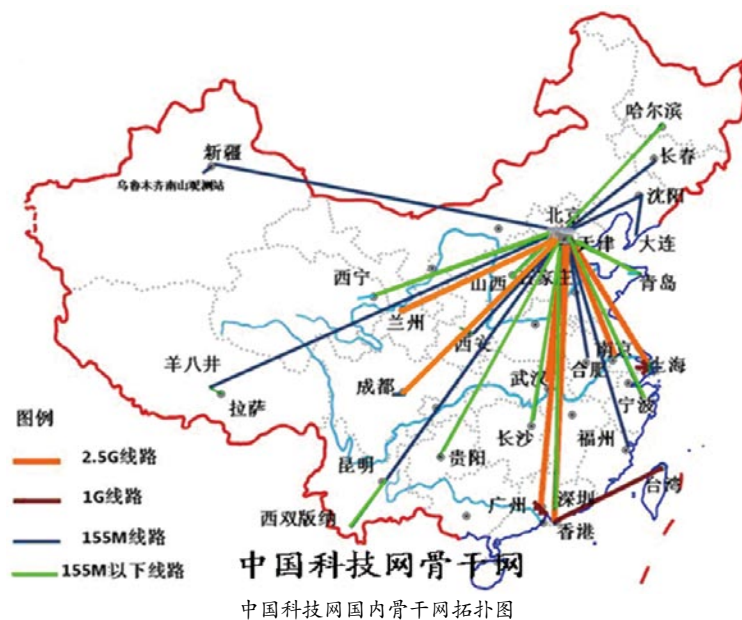
1. 快速发展的中国科技网

中国科技网（CSTNET）是我国最早的国际互连网络接入方，是我国互联网创建时期的四大网络之一。目前，CSTNET由北京、广州、上海、昆明、新疆等十三家地区分中心组成国内骨干网，拥有多条通往美国、俄罗斯、韩国、日本等国际出口，并与中国香港、中国台湾等地区以及中国电信、中国教育和科研计算机网、中国网通、BJNAP（北京国家互联网交换中心）等国内主要互联网运行商分别通过光纤高速互联，中国科技网已成为中国互联网行业快速发展的一支主要力量。

【网络能力】

★ 中国科技网网络

北京至6个分院为2.5Gbps带宽，北美方向622Mbps带宽。



★ 中美俄环球科教网络(GLORIAD)

实现环绕北半球光网络基础设施，骨干带宽2.5Gbps以上。

★ 香港开放交换节点 (HKOEP)

连接日本、韩国和中国香港、中国台湾等国家或地区的科研网络，是亚太地区互联网的汇聚中心和国际互联网在亚太地区的交换中心。

★ 下一代互联网 (CNGI)

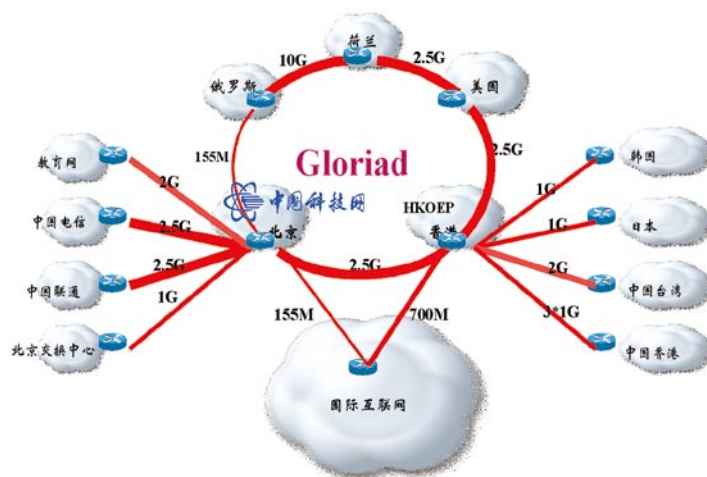
参加了国家下一代互联网 (CNGI) 示范工程项目，建设了北京、上海、广州、沈阳、长春、成都、兰州等7个CNGI核心网节点，完成了103个CNGI科研机构驻地网建设任务。

【 提供服务 】

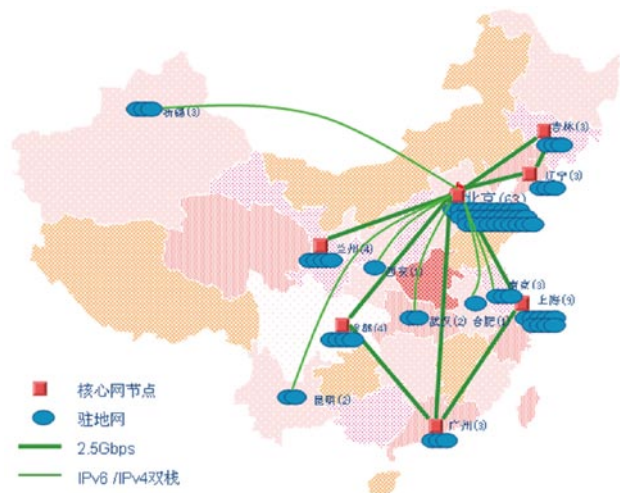
★ 为科研活动的各个环节提供网络支撑，提供海量科学数据、计算设施、野外台站等科学资源数据传输服务，接入中国科学院100多家单位，及院外科研单位100多家。

★ 实现院内开放互联。

★ 提供视频、邮件、IPv6服务等增值服务。



中国科技网国内国际互连拓扑

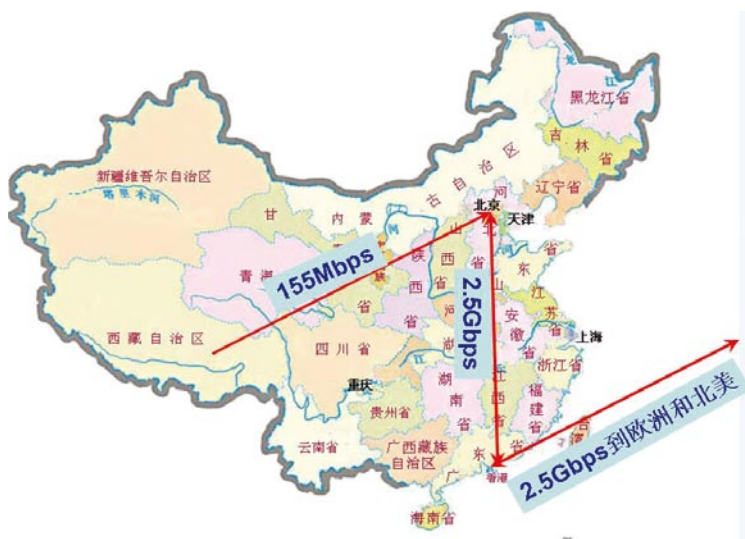


中国科学院CNGI核心与科研机构驻地网

【应用案例】

支持宇宙线数据预处理中心的应用

实现了羊八井宇宙线数据实时采集、传输、预处理和国际共享服务。基于中国科技网这一基础设施，中国科学院高能物理研究所加入了欧洲核子研究中心（CERN）的高能物理研究网络，成为该网络的重要节点。



羊八井的网络链接

支持天文学科射电天文望远镜实时甚长基线干涉测量（e-VLBI）系统应用

提供国际观测专用光通道和长距离、高带宽实时海量数据传输技术服务。服务中国e-VLBI观测网。e-VLBI是在甚长基线干涉测量（Very Long Base Line Interferometry, VLBI）领域出现的新技术。它将互联网络环境、高性能计算和VLBI信息处理技术相结合，采用高速通信网络，将观测数据从距离遥远的观测站直接传送至VLBI数据处理中心进行处理，能极大缩短VLBI数据信息处理周期。

2009年1月15日，支持在法国巴黎举行2009年国际天文年开幕式上，包括中国佘山和南山站在内的世界各国的17个射电望远镜同步进行全球最大规模e-VLBI观测演示。

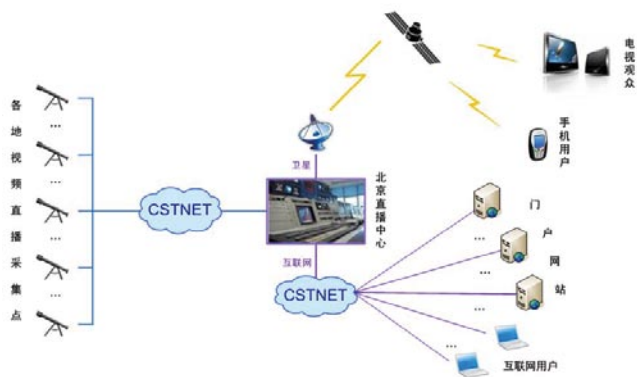


2009国际天文年e-VLBI观测演示

全程助力“2009日全食异地多路联合直播”活动

2009年7月22日，本世纪最壮观的日全食天象——长江“大日食”隆重上演，多家单位联合进行了“异地多路网络直播”。

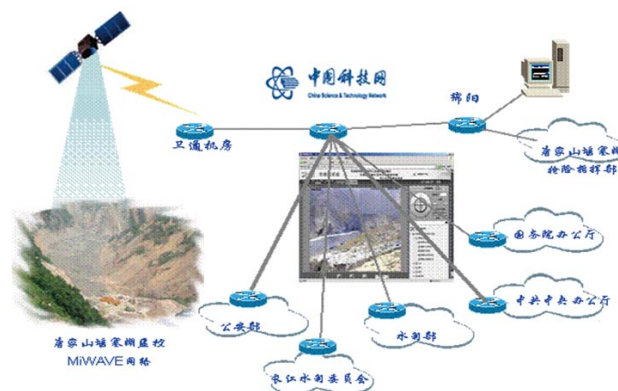
中国科技网利用与各大网络运营商之间G级的高速互联链路，为直播活动提供最优带宽链路解决方案，使得位于多个地区、不同网络的直播点实现与直播中心的高速互联，将公共信号通过互联网、手机WAP网进行发布，利用P2P等技术实现异地多路信号清晰流畅的网络日全食直播。



2009日全食异地多路联合直播

为汶川抗震救灾提供有利保障

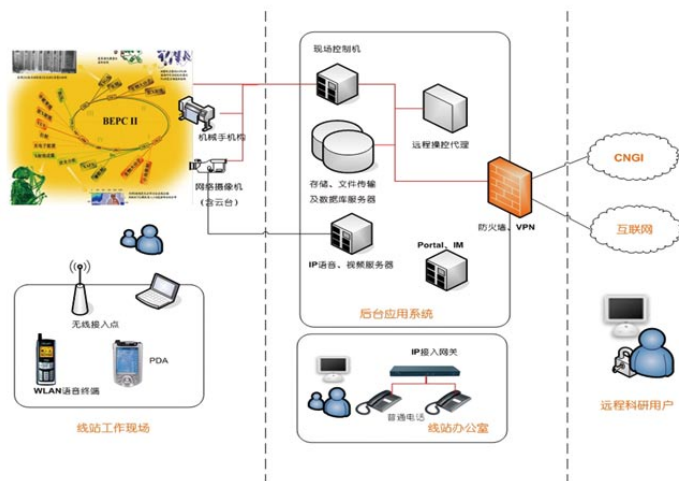
四川汶川大地震后，中国科技网紧急搭建了唐家山堰塞湖监控专网，结合中国科学院上海微系统所提供的现场宽带无线网络，确保了抗震一线信息及时传递给当地政府领导人以及有关部门，为抗震救灾提供了技术保障。



唐家山堰塞湖大坝及泄洪流域实时监控系统

建设面向科研应用的多媒体通信系统

充分利用现有网络资源、通信资源和应用系统，依托中国科学院邮件系统现有用户资源和认证信息，融合邮件、语音（VoIP+程控交换固定电话）、视频会议系统，即时通讯、短信功能，建设面向科研应用的通信系统。该系统已开始在中国科学院院属科研单位逐步部署。



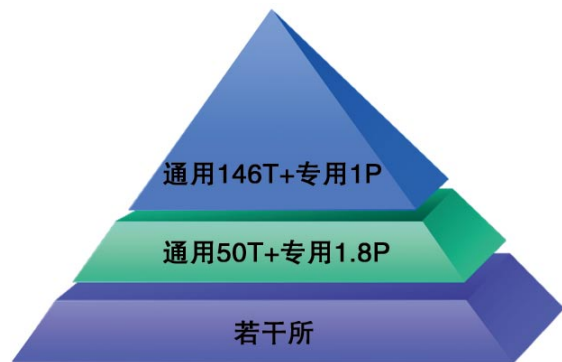
应用于北京同步辐射装置大分子实验线站

2. 不断完善的超级计算环境

中国科学院“十一五”统筹规划建设的“院总中心-分中心-所级中心”三层结构的超级计算环境，整合院内外有关超算资源，通用计算与专用计算相结合，CPU计算与GPU计算相结合，具备国内一流、国际先进的超级计算环境。顶层目前拥有联想深腾7000超级计算机，理论峰值达每秒146万亿次，是中国国家网格（CNGrid）的北方主节点和运行中心。至2009年10月底，超级计算中心已累计为全社会近千科研用户提供了逾6000万CUP小时的计算服务。

【计算能力】

初步建成的三层结构的中国科学院超级计算环境，顶层计算能力是通用146万亿次、专用1000万亿次，中间层的分中心和第三层的所级中心约20个左右，总聚合计算能力是通用100万亿次以上、专用1800万亿次，通过中国科技网实现计算能力和计算数据的共享。



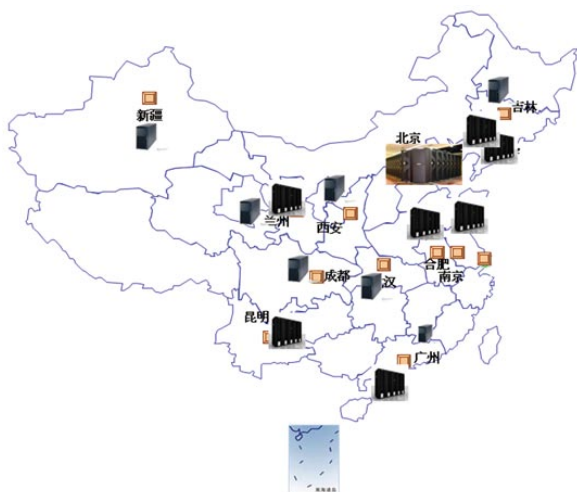
中国科学院超级计算环境三层网格架构

深腾7000超级计算机主要性能指标

- ★ 集群结构，理论峰值146Tflops
- ★ 采用64位CPU，节点机采用刀片服务器
- ★ 2台NUMA结构胖节点用于共享内存的并行应用软件
- ★ 内存总量66TB，磁盘阵列裸容量354TB
- ★ MPI通讯带宽 $\geq 2\text{GBps}$ ，延迟 $\leq 3\mu\text{s}$
- ★ 持续IO聚合带宽50GBps
- ★ 磁带库不压缩容量1PB，聚合带宽 $\geq 800\text{MBps}$



深腾7000效果图



中国科学院初步建成的超级计算运行环境三层结构分布示意图

【提供服务】

针对不同学科领域的需求提供丰富的超级计算软件，包括常用商用软件、开源软件及自主研发软件。为高性能计算环境长期高效、稳定的运行提供保障服务，同时，为院内外用户提供高性能计算技术与应用培训。

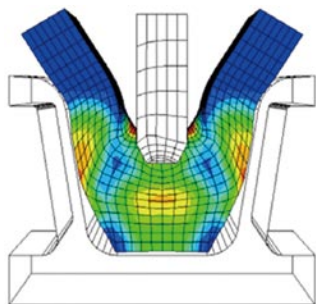
深腾7000超级计算机主要软件列表

移植开发的软件	领域	商业应用软件	领域	开源应用软件	领域
Lammps	分子动力学	MOLPRO	计算化学	GROMACS	分子动力学
DL_Poly	分子动力学	GAUSSIAN	计算化学	GrADS	可视化
TAO	数值计算	Q-CHEM	计算化学	GMT	气象地理数据处理
MM5	气象	MATLAB	分析和方 针工具	BLAST	基因比对
Nwchem	计算化学	VASP	材料科学	FFTW	数值计算
Abinit	计算物理	TURBOMOLE	计算化学	Hypre	数值计算
NAMD	分子动力学	WIEN2K	计算化学	OpenMX	材料计算
GAMESS	从头量化计 算	SPARTAN	计算化学	VENUS	量子力学
DOCK	制药	Materials Studio	材料科学	AutoDock	分子对接程序
CPMD	分子动力学	LS_DYNA	计算力学	GRAPES	气象数据处理
Elmer	有限元计算	Fluent	流体计算	MM5	气象数据处理
LIS	生态学	CFX	流体计算	NAMD	分子动力学
PHG	有限元计算	Crystal	材料科学	VMD	分子可视化
NetCDF	数据处理	ADF	计算化学	TINKER	分子动力学

【应用案例】

可视化铸锻技术

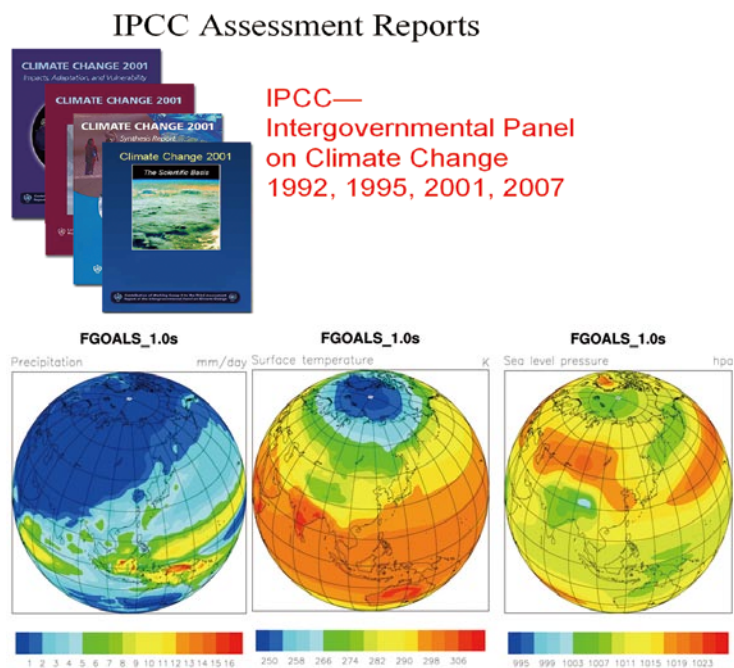
利用高性能计算模拟，将材料成形过程“可视化”，指导大型铸锻件的加工工艺过程，成为提升我国铸造、锻造等传统材料加工工艺水平的有效方法，已经在我国重点工程的船用曲轴毛坯、大型铸钢支撑辊、大型空心钢锭和核电用关键铸件等的制造技术上成功应用，解决了我国不能加工特大型铸锻件的难题。



大型船用曲轴曲柄的弯曲成形过程

气候变化数值模拟

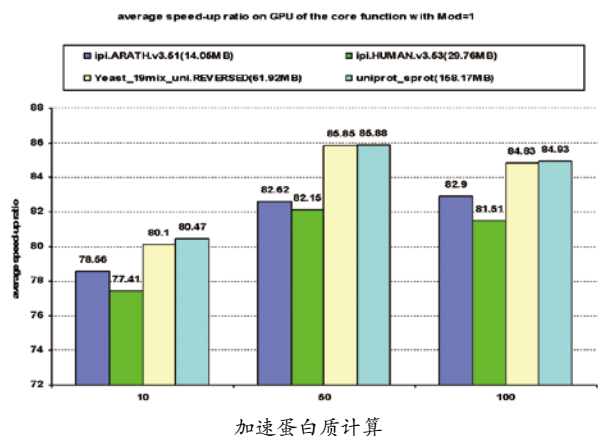
发展了独具中国特色的四代气候系统模式，在高性能计算机上模拟了人类活动对全球变化的可能影响以及未来气候的可能演变趋势，其结果被政府间气候变化委员会（IPCC）已有的四次气候评估报告所采纳。



我国第四代气候系统模式FGOALS模拟的降水、温度和海平面气压

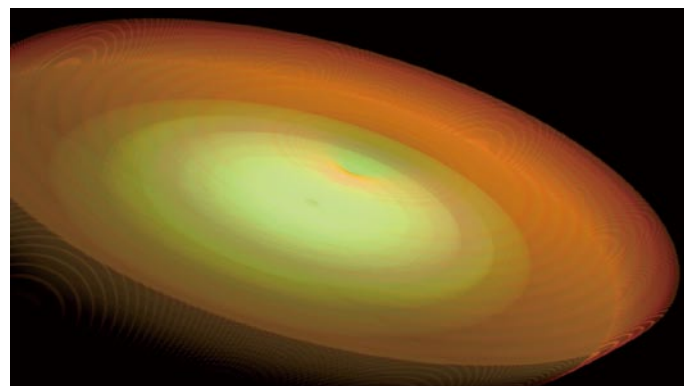
加速蛋白质计算

基于开源软件，开发了并行蛋白质翻译后修饰与鉴定软件，实现了1024核计算，使原来一个月才能完成的计算任务缩短到10分钟之内完成。



天文气体动力学模拟

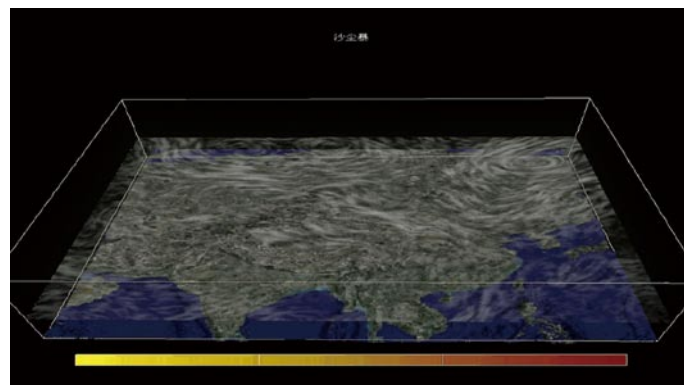
基于中国科学院紫金山天文台的星系演化串行模型，开发大规模并行程序Parawidgeon，实现了8192核天文计算，极大地提高了国内天文学家的气体动力学模拟能力，使得他们可以驾驭更高的分辨率，获得更精细的天文模拟现象，计算规模达到国际先进水平。



星系演化大规模并行程序模拟

沙尘暴业务数值预报

通过中国气象局业务化评估的沙尘暴业务数值预报系统，将原耗时10多小时的3天沙尘暴预报的时间降至20多分钟。



沙尘暴业务数值预报系统

● GPU超算在产业和学科中的研究与应用

GPU (Graphic Processing Unit) 在高效能低成本的大规模并行计算方面显示了强劲发展势头。中国科学院过程工程研究所2009年研制并使用了我国第一套单精度1000万亿次超级计算系统，该系统在目前国际上基于GPU的超级计算系统中处于领先行列，并在软硬件的总体设计思路上有鲜明的创新特色。针对众多行业和学科的应用需求，承担了多项国家重大专项、国家科技支撑计划、国家自然科学基金重大基金和多家大型企业的计算任务，涉及化工、冶金、石油、矿产、动力等多个产业和物理、化学、材料、生物等多个学科，展示了广阔的应用前景。

【计算能力】

- ★ 中国科学院过程工程研究所研制的GPU超算系统：单精度理论峰值超过1Pflops



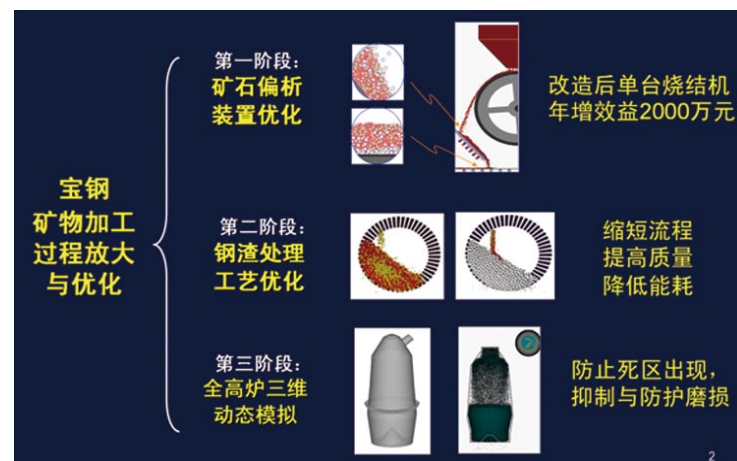
- ★ 中国科学院开展GPU超算应用单位：

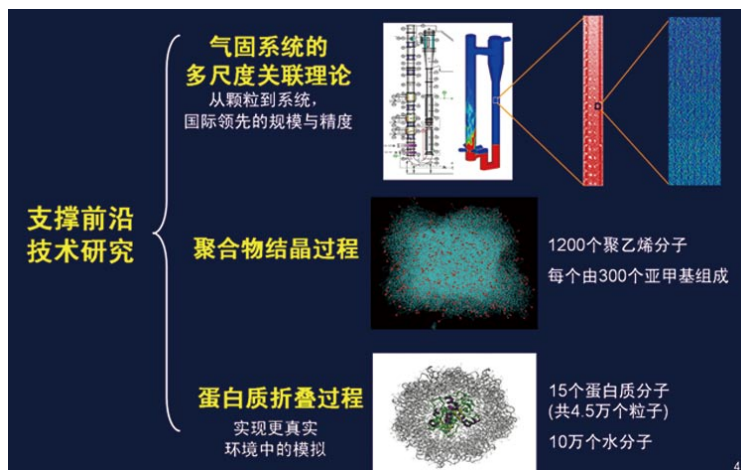
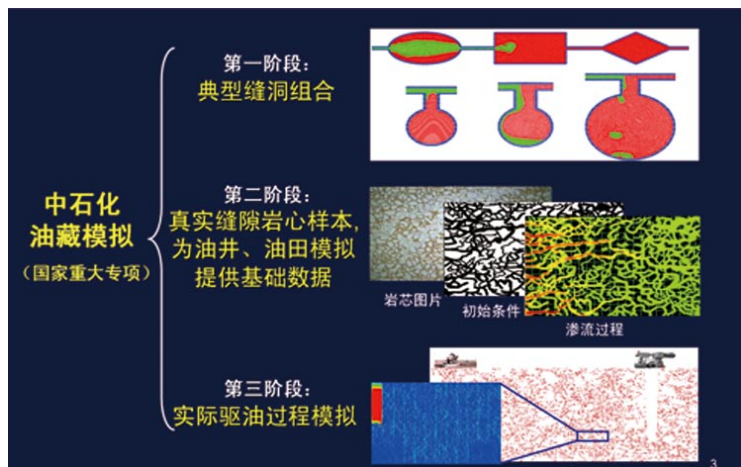
研究所	GPU计算能力 (Tflops)
电工研究所	112
深圳先进技术研究院	200
中国科技大学	183
计算机网络信息中心	200
国家天文台	158
地质与地球物理所	200
近代物理研究所	202.5
高能物理研究所	195.5
金属研究所	183
紫金山天文台	180
合计	1814

【应用案例】

1. 宝钢集团有限公司矿物加工过程模拟及放大优化。宝钢集团有限公司与中国科学院过程工程研究所合作先后利用该所的『高效低成本超级计算系统』开展矿石分级和钢渣处理新工艺开发。
2. 中国石油化工股份有限公司与中国科学院过程工程研究所利用『高效低成本超级计算系统』模拟了从简单缝洞、岩芯样本到油井尺度的驱油过程，已能部分代替昂贵费时的物理实验，并将在国家重大专项支持下服务于尚无成熟理论和开发技术的缝洞型碳酸盐岩油藏的开发，为油井、油田模拟提供基础数据。
3. 中国科学院地质与地球物理研究所联合有关单位在石油勘探的GPU应用方面取得一定突破。在叠前时间偏移计算上，对GPU的峰值速度利用率已经突破25%，使得运算速度提高100—400倍；成本低廉，不足常规并行机的10%；节电95%以上、占地空间省90%以上。该技术已经经过20个油田试验，并开始在大庆油田、胜利油田、吐哈油田、江汉油田、恒泰爱普公司进行商业化应用。

4. 利用『高效低成本超级计算系统』取得了世界前沿水平的基础研究成果，如气固系统的多尺度关联理论、聚合物结晶过程和蛋白质折叠过程。





3. 稳步提升的科学数据应用能力

科学数据是支撑国家科技创新的战略资源。“中国科学院科学数据库”是我国国内最早开展的基于现代信息技术建设的综合性科学数据库系统。从上世纪八十年代开始建设，迄今已有62家中国科学院科研单位参加科学数据库建设，并建成了分布式的网络数据服务体系。数据资源得到了持续积累和不断整合，覆盖了物理、化学、空间、天文、材料、生物、地理、资源环境、能源、海洋等基础科学领域，形成了支撑院内外科研机构的开放服务平台，服务正常率达到97.78%，访问累计人数超过800万人次。为国家天文台大天区面积多目标光纤光谱天文望远镜（LAMOST）、合肥超导托卡马克核聚变实验装置（EAST）、北京正负电子对撞机（BEPC）和北京谱仪（BES）、兰州重离子加速器冷却存储环（HIRFL-CRS）、西藏羊八井宇宙射线观测站（ARGO）以及空间环境地基综合监测系统“子午工程”等重大科学工程提供基于海量数据研究的支撑。



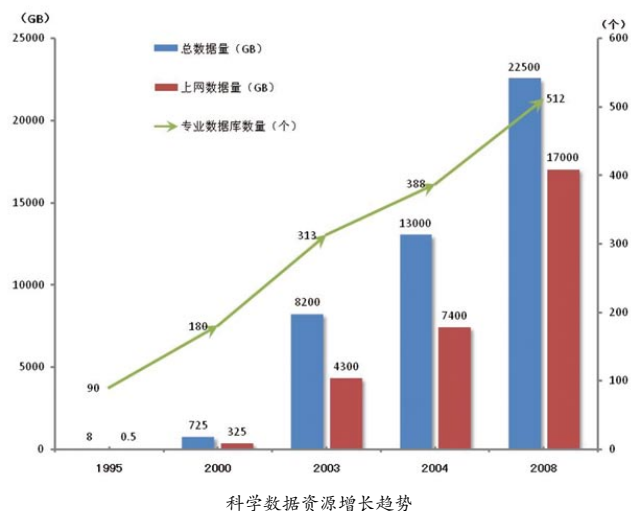
【建成能力】

★ 数据资源

长期累积、整合建设了丰富的数据资源，目前重点整合形成了：2个参考型数据库、4个专题数据库、8个主题数据库、37个专业数据库，提供40TB可共享的数据。

★ 存储资源

拥有海量数据存储能力，已建成2.3PB存储容量（2010年将形成6PB的存储能力），提供数据保存、数据灾备和在线存储服务。提供中国科学院院内1Gbps、院外200Mbps的带宽链接。



【提供服务】

★ 数据服务

通过科学数据库门户系统 (www.csdb.cn) 和国际科学服务平台提供多种数据服务，包括：用户自助式数据服务、下载和主动式数据服务。提供数据预定和委托查询服务功能，为数据传递提供通道，并提供数据预处理，及数据规范和制定服务。

★ 存储服务

提供海量数据网络存储环境和技术支持。

	学科领域	数据源	数据量
Landsat	遥感影像	USGS 美国地质调查局	6TB
SRTM	数字高程 (90米分辨率)	CIAT 国际热带农业中心	168GB
Genome	生物信息学	UCSC 加州大学 圣克鲁兹分校	3.5TB
生物医学	生物医学	NCBI、EBI、 DDBJ、SIB等	2.5TB
NCAR	大气海洋等	NCAR 美国大气研究中心	1TB

数据资源中心开通优质数据资源的镜像服务

★ 专员服务

在地学数据、生物信息学数据、天文与空间数据、理化与材料数据等学科领域匹配专业数据服务专员提供主动服务。

★ 技术服务

对数据存储、数据库建设、科学数据应用等方面提供专业服务和专业技术支持。

自主开发了数据库定制及存储管理系统，包括：可视化关系型数据库管理系统 (VisualDB)，文件型数据、远程存储管理和共享工具 (Visualfolder)。

★ 培训服务

提供数据库技术及数据标准规范培训。

【应用案例】

中国科学院科学数据库作为科研信息化基础设施的重要组成部分，在科学研究、国家经济建设和社会宏观决策服务中发挥了重要作用。

中国科学院科学数据库支持服务的重要项目列表（部分）

科学数据库名称	支撑服务项目
中国自然资源数据库	支持国家发展与改革委员会与中国科学院重大创新课题“中国西部生态——经济区划及典型区可持续发展模式与对策研究”。 国家基金委重大研究计划“中国西北近50年城市与环境互动作用机理研究”。
大气科学与环境数据库	支持大气物理研究的一系列国家和科学院重要项目，如“我国重大气候和天气灾害形成机理和预测理论的研究”、“亚洲季风区海-陆-气相互作用对我国气候变化的影响”等。
空间环境数据库	密切结合“双星计划”、“子午工程”等重大空间内环境探测计划，提供实时空间环境数据，支持了“空间环境预报模式研究”等重要研究项目，特别是为载人航天工程提供了及时可靠的空间环境预报服务，为保障“神州”系列飞船发射飞行的空间环境安全做出了贡献。

科学数据库名称	支撑服务项目
病毒资源数据库	支持多个研究所在禽流感方面的联合重大研究计划
遥感卫星图像检索数据库	支持全国国土资源大调查，3年来共计提供了1122万平方公里的卫星数据，所提供的数据服务为我国遥感应用各相关领域，特别是在农业估产、林业调查、土壤、水文、地质分析、海洋环境监测、城市土地利用、国土资源调查、多种自然灾害监测与评估等方面发挥了显著作用。

● 应对全球变化和自然灾害研究的数据共享平台建设

中国科学院对地观测和数字地球科学中心是中国科学院对地观测数据接收、处理、分发和研究的重要基地，也是我国主要卫星和航空对地观测的数据中心，对于我国和我国应对全球变化和自然灾害等重大问题的研究有重要的支撑作用。灾害数据共享和全球变化数据共享平台建设是对地观测中心信息化建设的重点。

【全球变化数据共享平台】

构建了一个可以动态进行全球变化资源汇聚和共享的数据云服务环境，初步具备了本地100TB的存储能力，并通过IPv6和中国科学院计算机网络信息中心的PB级存储互联，使用其中大规模存储，直接服务于多个973、863和国际合作等全球变化研究项目。



大规模数据存储环境

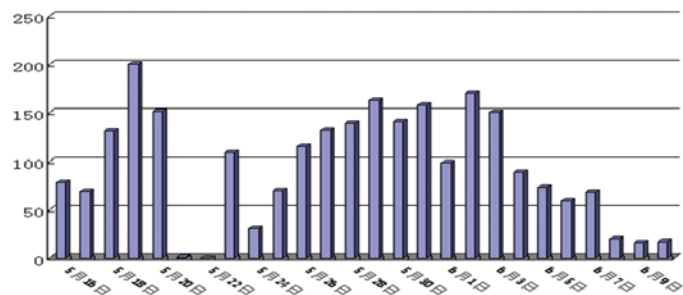


全球变化对地观测数据共享平台

【灾害数据共享平台】

在2008年汶川地震期间，该平台仅使用两天时间就形成了众多国际救援数据基于网格的标准化服务能力，累计为用户提供了数万次访问，下载数据1TB左右，在地震减灾的快速响应中发挥了极其重要的作用。

灾害数据共享平台还将进一步规范和扩大，逐步为全国用户提供标准化的服务能力。



汶川地震国际对地观测数据援助网格平台及下载量统计

中国生态系统研究网络（CERN）数据信息系统建设

中国生态系统研究网络（CERN）是国家野外台站体系的重要组成部分，是我国生态与环境、资源与地球科学领域的野外观测和试验研究的公共平台，在相关学科的基础研究和应用研究中发挥着不可替代的作用。该项目建设的数据库信息管理系统，取得了明显成效，促进了国家野外观测科技平台的建设。



目前CERN网络体系包括1个综合中心、5个专业分中心、40个不同生态系统野外站（农田、森林、草地和水体、城市）。经过20多年建设，形成了台站-分中心-综合中心不同尺度相对完善的数据资源体系，开发了不同尺度的专题数据库，建立了综合中心-分中心-台站三级数据共享平台，逐渐由现在集中的、孤立的数据信息系统向分布的、联网的数据信息系统转变。

建成的数据库数据量3TB，年访问量5.7万人次，开发的空
间栅格化数据为中国气象局、水利部、国家环保局（原）、
中国农业科学院、中国测绘科学研究院、黄河水利委员会、
中国科学院多个研究所承担的多个国家项目和业务运行提供
基础数据支撑。开发的多个数据信息系统支持了CERN的数
据管理和共享服务工作，并在中国科学院、国家林业局、农
业部和教育部的53个国家野外台站中推广使用，促进了国家
野外观测科技平台的e-Science基础设施的建设。



国家标准、数据专著、软件著作权与奖励



国家生态系统研究网络分布式信息系统和专题数据信息系统



● 服务科研，深化科研信息化创新实践

科研信息化的基础设施和核心技术为科研活动提供了信息化的平台和环境，而不同的科学研究领域及具体的科学研究活动，必需开展针对相应需求的应用，才能够使科研信息化真正成为现实。为此，中国科学院坚持以应用需求为出发点，在“十一五”信息化建设中部署实施了一批“科研信息化（e-Science）应用示范”项目，目的是充分发挥中国科学院信息技术和基础设施的优势，探索科研信息化的体系结构、标准规范和关键技术，总结内在规律，形成运维机制和应用服务模式，为全面实施科研信息化提供经验，在推进我国科研信息化方面发挥引领示范作用。

1. 世界水平的数据密集型计算网格

科学研究规模的扩大，使人们需要采集海量的数据用于研究和解决复杂的科学问题。这些科学问题包括认识自然界物质的微观结构、天体的演化、全球气候变化、基因的奥秘等。但是科学数据呈指数级增长又对计算技术提出了更大的挑战。数据密集型网格平台建立在网格技术基础上，是对海量数据进行高效分析处理的有效工具，在物理、化学、生物学、医学、系统科学、经济学、气象、地震、核能技术、石油探勘、航天工程等众多领域具有广泛的应用前景，并正得到快速推广。

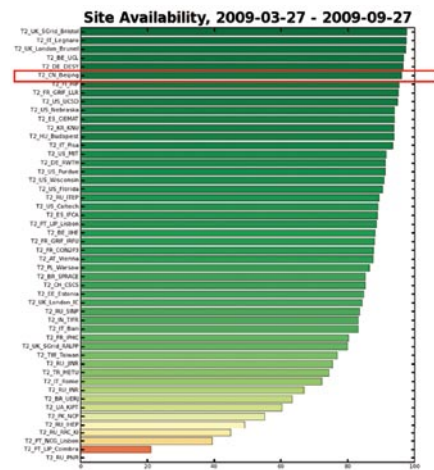
【项目名称】

数据密集型网格平台示范站点与应用

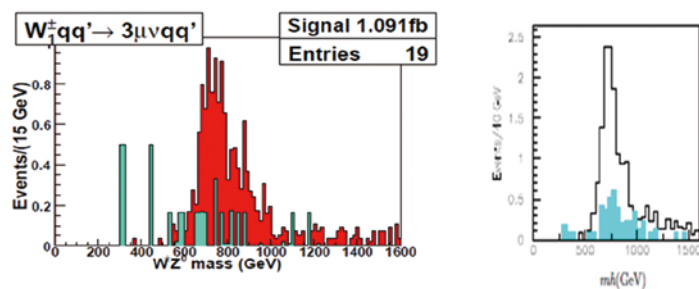
【主要成效】

目前，中国科学院高能物理研究所的数据密集型网格站点加入了世界上最大的网格系统WLCG，成为全球200多个网格站点中运行水平最好的站点之一。

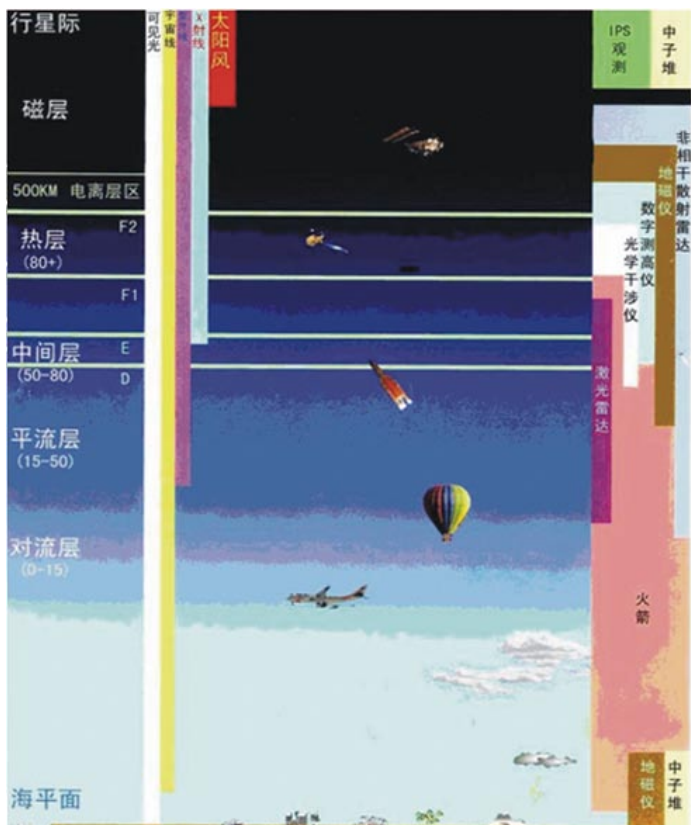
为大型强子对撞机（LHC）国际合作承担了大量的计算任务，帮助中国用户使用国际上更大规模的网格计算资源为科研服务。中国物理学家利用网格平台进行了数十个物理前沿课题的Monte Carlo模拟研究，为未来实验数据分析打下良好基础。



网格站点（T2_CN_Beijing）可靠性统计



LHC实验寻找新玻色子的物理模拟



子午工程检测范围



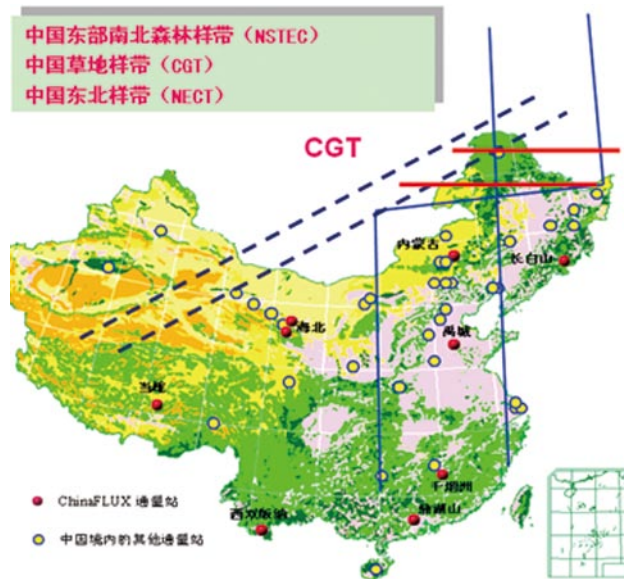
中国科学院空间科学与应用研究中心卫星数据地面接收站天线



中国科学院空间科学与应用研究中心卫星监控大厅

3. 碳收支集成研究的科研信息化应用示范

陆地生态系统碳循环已成为地球系统科学和全球变化科学的前沿问题。碳循环研究一方面可以为我国参与应对全球气候变化国际行动和制定行动方案的谈判提供坚实的科学依据；另一方面也可通过对生态系统结构和功能的有效管理，增加陆地碳汇能力，减缓我国的减排和限排压力，保障生态安全，维持国家经济与环境可持续发展。



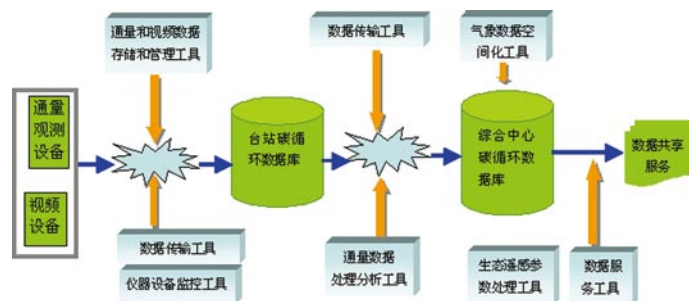
【项目名称】

服务于生态系统碳收支集成研究的e-Science环境建设及应用示范

【主要成效】

以现有的8个野外观测台站为基础，通过由“三个工作环境”（数据集成环境、模型模拟环境、可视化协同环境）和“四个层次的应用系统”（典型生态系统、典型区域、中国区域、东亚区域）构成的中国陆地生态系统碳循环科学研究的e-Science应用示范体系，以满足我国生态系统碳循环科学和生态信息学科发展的需求，实现全国/区域/典型生态系统（站点）不同尺度碳收支状况的快速评估，构建生态系统碳循环联网观测和联网实验的数据采集与在线远程传输、数据分析与多源数据管理、模拟分析与可视化表达一体化的信息化环境。

为实现全国/区域/典型生态系统（站点）不同尺度碳收支状况的快速评估提供了支撑，也为建立基于视频技术的生态要素观测系统奠定了基础，推动我国生态信息学科的发展。



陆地生态系统碳循环数据集成与服务环境框架



通量数据处理系统

4. 网络环境支持下的科学考察和合作研究

将东北亚各区联系起来开展全球变化的综合考察是事关“全球变化区域响应问题”的重要课题，将为我国典型区域在全球变化背景下的合理发展提供对策和决策依据。各国科学家对在联合科学考察中的地学资源共享、主要科考活动的网络化协同有着强烈而现实的需求。



e-GeoScience集成界面

【项目名称】

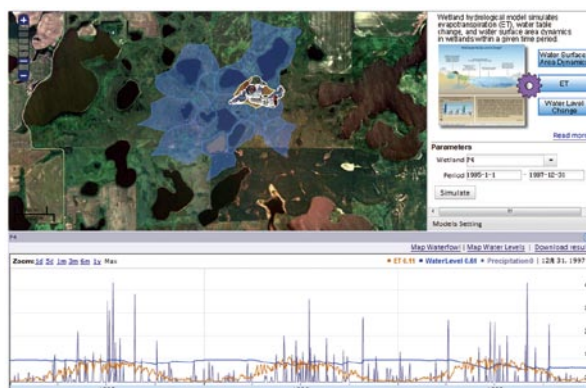
东北亚联合科学考察和合作研究平台

【主要成效】

- ★ 集成了一批东北亚数据资源、地学模型和地学文献，推动了跨组织、跨区域（中国、俄罗斯、蒙古）的资源共享和合作研究。已经为科技部基础性工作专项“中国北方及毗邻地区综合科学考察项目”各课题组提供了基础数据共享。
- ★ 目前中国科学院地理科学与资源研究所已经分别与俄罗斯科学院远东分院太平洋地理研究所、蒙古科学院地理研究所签署了节点部署的协议，并于2009年9月完成了太平洋地理研究所节点的部署，同时将在国际山地中心推广此应用。



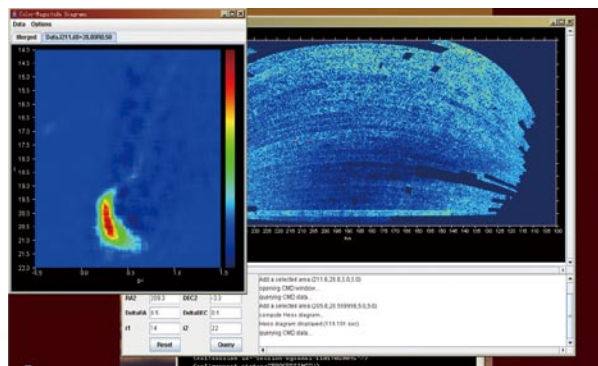
考察路线预选择与考察点预设置



地学模型动态调用

5. 虚拟天文台探索浩瀚太空

虚拟天文台是天文信息技术的代表，是通过先进的信息技术将全球范围内的天文研究资源（包括天文数据、天文文献、计算资源、各种软件工具，甚至天文望远镜等观测设备）无缝透明连接在一起形成的数据密集型网络化天文研究和科普教育环境。它将打破现有科研工作的地域与时空界限，通过对资源的统一使用和天文学家的协同工作为天文学带来新的突破。天文学家只需登录到虚拟天文台系统就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摆脱出来，而把精力集中在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，成为开创“天文学发现新时代”的关键性因素。



利用虚拟天文台开展银河系结构研究

【项目名称】

虚拟天文台（Virtual Observatory, VO）

【主要成效】

- ★ 整理和集成了超过300GB的天文数据，实现BATC多色巡天数据产品发布平台。
- ★ 实现天文数据的统一访问，提供一系列分析处理软件工具和可视化工具，可用于科学普及和公众教育。
- ★ 开展在科学数据网格环境下真实的天文学应用研究。
- ★ LAMOST是世界上光谱观测效率最高的望远镜，其全球最大天文光谱库将通过China-VO进行存储、管理和全球共享。

中国虚拟天文台将以中国科学院科研信息化基础设施为依托，在国家重大项目助力的同时，引领天文学研究进入e-Science时代。



中国虚拟天文台重点服务对象：LAMOST国家大科学工程

6. 野生鸟类资源及疫病监测与风险评估

青海湖是中国最大的鸟类保护区之一，2006年以来，中国科学院的科研人员利用信息化的利器改变了传统的原始观测、跟踪研究手段，在青藏高原高海拔的极端环境下，建立起第一个国家级自然保护区野生鸟类视频监控系统，产生了对青海湖生物和生态的保护实践具有指导意义的科研成果。

【项目名称】

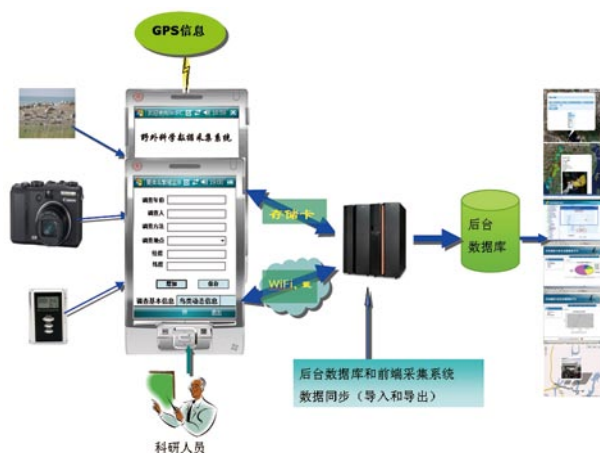
青海湖区域重要野生鸟类资源及疫病监测与风险评估研究
e-Science应用

【主要成效】

- ★ 建成网络化的科研信息化环境，包括三个平台，分别是青海湖基础数据平台、青海湖科研协同平台和青海湖科研应用平台。
- ★ 物种鉴别与资源量调查，对监测站点的常见候鸟种群数量变化进行了监测、统计和数据库建立。



中国科学院青海湖联合科研基地协同平台



基于移动设备的野外科学数据采集系统

- ★ 流感病毒跨种传播建模与预测，共获得239个毒株基因组数据，共计1434条蛋白相应核酸序列。
- ★ 建立完善的野外数据获取系统。
- ★ 建立覆盖核心区域的网络视频监控系统：共在青海湖部署了19套网络视频监控设备，支持科研人员定制观测任务，支持观测任务的自动执行及系统故障的自动诊断和修复。
- ★ 鸟类栖息地数据挖掘与结果分析。



7. 新一代工业生物技术知识环境

工业生物技术作为生物技术发展的第三个浪潮，以微生物或酶为催化剂进行物质转化，与现代工程技术有机结合，大规模生产人类所需的化学品、医药、能源、材料等，是解决人类目前面临的资源、能源及环境危机的有效手段。目前，中国科学院“先进工业生物技术基地”面向国际工业生物技术科学发展前沿和国家能源、资源、环境重大需求，在生物能源、生物炼制、绿色工业过程等领域布局了一批重要的项目，迫切需要利用e-Science的手段，建设工业生物技术的知识环境，促进资源共享和合作研究。



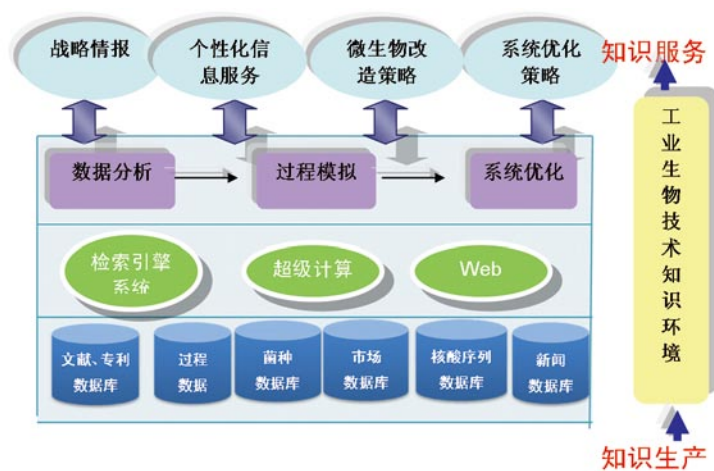
工业生物技术信息网首页

【项目名称】

工业生物技术知识环境建设及其e-Science应用

【主要成效】

- ★ 建设网络协同化工作环境，使协同环境的工作用户覆盖先进工业生物技术基地的18个研究所主要研究队伍。
- ★ 整合数据资源、形成服务于工业生物技术领域基础研究和应用的知识仓库，形成针对重点项目“纤维素乙醇高温发酵与生物炼制”、“能源植物选育”的综合信息平台。
- ★ 开发生物信息学并行计算程序，形成超级计算资源与高通量筛选平台、大规模发酵平台等关键技术平台的无缝链接。
- ★ 开发智能型的“工业生物技术检索引擎”，跨库跨平台的综合检索功能，实现独特的引文检索、主题检索、化学结构检索、基于内容与引文的跨库交叉浏览、检索结果的信息分析、检索结果的信息管理。
- ★ 为中国科学院院内18个研究所，1800人的研发队伍和全国从事工业生物技术研究的70余家研究机构、60余家企业提供信息定制、战略研究、信息推送等个性化信息服务。



工业生物技术知识环境架构

8. 生物大分子研究e-Science服务中心

多蛋白分子组成的复合体和膜蛋白在行使生物学功能的重要性和与人类健康的密切相关性方面已成为目前结构生物学的最新热点。目前，我国启动了多个相应的重要蛋白质复合体及膜蛋白的结构和功能研究课题，越来越多的研究机构和科学家投身于系统生物学的研究。将中国科学院科研人员发展的研究方法和国际上最新的研究方法整合在一起，构建一个有自己特色的、基于网络的、服务于国内外同行的系统生物学分析网络平台对相关领域的研究具有重要意义。

【项目名称】

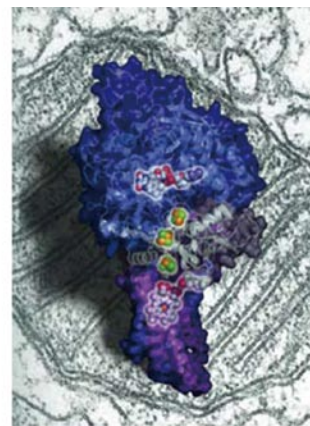
生物大分子研究e-Science服务中心

【主要成效】

- ★ 建成包含最完整的ncRNA和蛋白质结构以及其他重要公共数据库的生物大分子信息资料库
- ★ 编制能对相关信息资料进行快速查询以及有效分析的软件包
- ★ 相关大规模计算的e-Science网络服务平台
- ★ 自动化晶体结构解析网络平台
- ★ 系统生物学分析网络平台



神经营养因子3与其受体
p75NTR胞外区复合物的晶体
结构



线粒体膜蛋白复合物II三维
结构

9. 科研信息化综合性协同工作平台

根据科研活动的需求，建立一个基于互联网的虚拟科研平台，包括一些通用及学科领域的工具集，可提供安全、稳定的服务，对支撑信息化环境下跨地域、跨学科、跨组织科研活动的开展，以及虚拟科研组织的科研活动，具有重要意义。



协同工作环境套件结构

【项目名称】

科研信息化虚拟科研平台

【主要成效】

★ 虚拟科研平台软件工具包

2009年11月，发布虚拟科研平台在线服务版本（Duckling 2.0）。至2009年9月底，协同工作环境套件用户案例已达35个，为中国科学院院内外共26家用户单位提供服务。

★ 数据集成与共享插件

提供多种数据来源的集成和数据发布方式，以及数据处理流程的定制和可控数据清洗功能。

★ 计算插件

提供网格接口和应用用于地学研究的计算插件界面。

★ 数据分析与可视化插件

可便捷绘制常规数据图表；可根据数据（大气环境、蛋白质等）含义和类型不同，扩展实现高级分析和可视化。

★ 文献插件

可支持多源自动收集，协助文献管理和团队共享。

★ 设备管理插件

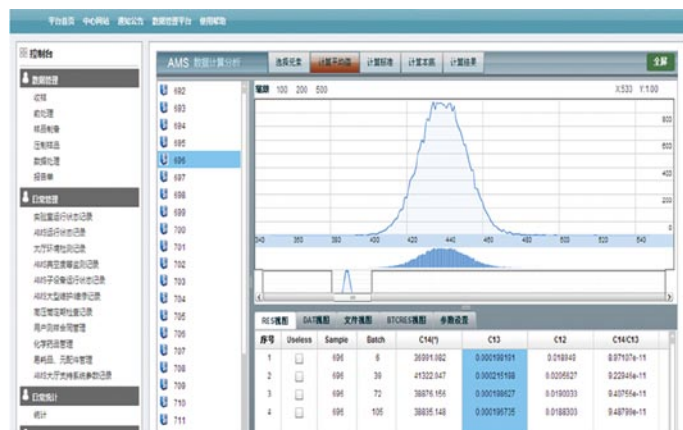
可围绕大型设备的数据处理和管理设定流程。目前已开发应用于加速器质谱仪管理的设备管理插件界面。

★ 会议管理插件

具备学术会议组织功能。2009年成功支持上海生化国际会议，3000余人注册参加了此次会议，提交论文摘要950余篇，来自于114个国家/地区的用户访问了网站，访问量超过12万次，页面综合浏览量41万余次。



应用于大气环境监测的数据分析与可视化插件界面



应用于加速器质谱仪管理的设备管理插件界面

展望

展望未来，人类社会信息化进程将进一步加快。互联网络、超级计算、虚拟现实、网络制造与网络增值服务，极大丰富了信息的形式与内容，拓展了信息的传播范围，提升了信息的应用价值，使人们将突破语言文字的屏障，更便捷地进行跨文化、跨地域的信息交流。下一代互联网、传感网、物理网和后IP时代网络正在蓬勃发展，将使人类有可能突破时空的界限，不断创造个人和社会生活的新模式，加快全球化的进程。

到2015年前后，中国科学院将实现全院上下的科研信息化体系，包括：

- 基本建成信息化的科研环境，实现中国科学院科研环境的数字化和网络化，先进的宽带网络和基于宽带网络的多媒体通信设施覆盖全院科研活动的每一个环节与每一个场所，支持无处不在的科研活动；

- 形成一个具有万万亿次（通用+专用）计算能力的国家计算中心和若干重要网格节点组成、支撑全院的网格化超级计算环境，整体满足中国科学院高性能科学计算的需求，为

国家经济社会发展提供支撑服务，成为国际一流的高性能计算中心之一；

- 全面集成整合中国科学院科学数据资源，实现科学数据的自动获取、存储、传输和处理，建成由一个总中心和若干分中心组成，统筹规划、科学规范的国家基础科学数据共享服务网格；

- 以虚拟实验室为平台的信息化协同工作环境普遍应用，形成跨地域、跨行业、跨学科的信息化科研活动环境，有效支持科研活动的全过程。

中国科学院将始终坚持面向国家战略需求，面向世界科学前沿，在大幅度提升中国科学院信息化应用及服务能力的基础上，形成功效一流的科研和管理信息化环境及服务模式，在“感知中国”的信息化大潮中，发挥科研信息化的示范和引领作用，有力支撑国家和中国科学院的科技创新，迎接新一轮科技革命的到来！

