U.S. Department of Energy's
Office of Science

Program Area Presentation

Network Environment Research

Advanced Scientific Computing Research
Strategic Planning workshop

Mary Anne Scott

July 22, 2003

In coordination with
Thomas Ndousse
George Seweryniak

U.S. Department of Energy

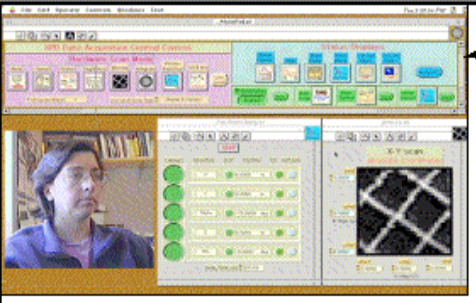
Office of Science

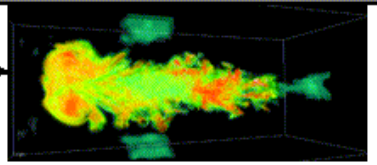
Network Environment Research

Contribution of Program Element to Overall ASCR Strategic Goal

- A scalable, secure, integrated network environment for ultra-scale distributed science is being developed to make it possible to combine resources and expertise to address complex questions that no single institution could manage alone. It is creating the means for research teams to integrate unique and expensive DOE research facilities and resources for remote collaboration, experimentation, simulation and analysis.

Building an Integrated Network Environment for Distributed Science



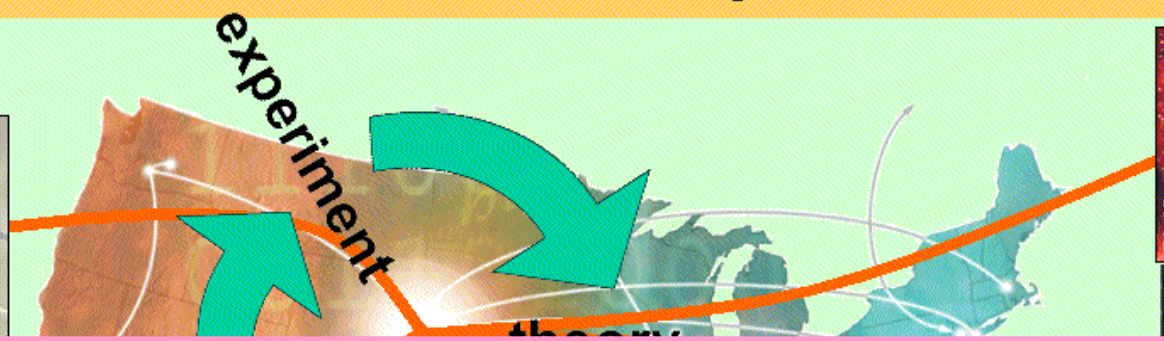


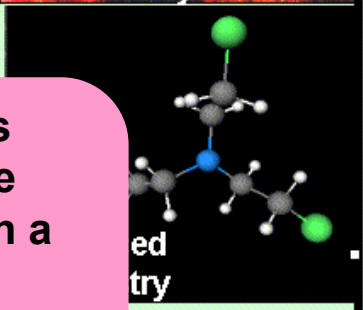
Science Portals: collaboration and problem solving

Application building services

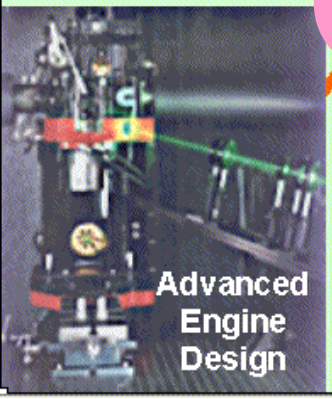
Grid Services: secure and uniform access and management for distributed resources

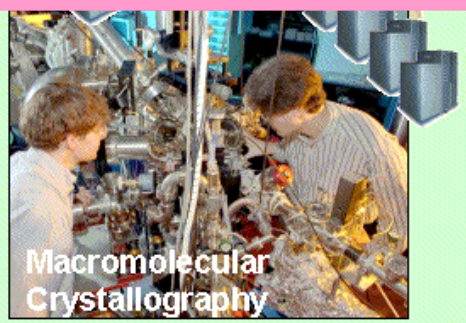
Supercomputing and Large-Scale Storage

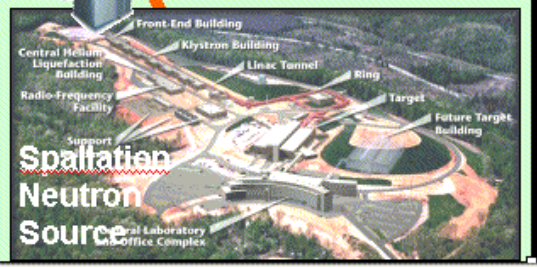


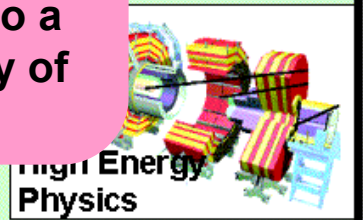


Science applications and specialized experimental facilities are n-way interconnected to terascale computing, petascale storage, high-end visualization, and remote collaborators in a seamless environment that provides the performance level required to move science, especially large-scale science, to a new regime—rapid scientific progress through the interplay of theory simulation, and experiment.









U.S. Department of Energy

Office of Science

Network Environment Research

Planning horizon for Program Element

- Short term 1-3 yrs
  - integration, prototyping, testing, and accelerated deployment of advanced computing, communications and middleware technologies
- Long term 1-5 yrs
  - Fundamental research issues for advanced collaborative and network capabilities addressing key issues for projected future applications requirements
- Keep the pipeline moving—research through development and prototyping

U.S. Department of Energy

Office of Science

Network Environment Research

Areas of research Program Element currently invests in

- **Middleware - $7.4M**
  - Facilitate the discovery and utilization of scientific data, computers, software, and instruments over the network in a controlled fashion
  - Integrate remote resources and collaboration capabilities into local experimental and computational environments
  - Services to support collaborative scientific work
- **Network research - $5.9M**
  - Network measurement and analysis
  - High performance transport protocols
  - Multicast and secure group communication
- **Collaboratory pilots - $9.4M**
  - Early implementations of virtual laboratories to test and validate the enabling technologies that unite distributed expertise, instruments, and computers for discipline-specific applications.

Base – $11.8M SciDAC - $10.9M

U.S. Department of Energy

Office of Science

Network Environment Research

Major Accomplishments

- Access Grid – group-to-group collaboration tool
  - large-scale distributed meetings, collaborative work sessions, seminars, lectures, tutorials, and training
  - over 150 institutions worldwide
- Globus – world-wide de facto standard for base software for grids (partially funded)
  - GridFTP – effective movement of very large files
  - CAS – community authorization service for group collaborations
  - Production quality Monte Carlo simulation (200K events) for CMS/LHC
- Electronic Notebook - record of ideas, data, and events for joint experiments and research programs
  - Used by large number of disciplines
  - Over 1000 software downloads.
- Net100 - modifies operating systems to respond dynamically to network conditions and make adjustments in network transfers, sending data as fast as the network will allow

Network Environment Research

How does Program Element transfer knowledge or provide services to application scientists?

- Collaboratory pilots are partnerships between CS/network researchers and discipline scientists and drive the initial technology transfer
- Goal of supporting collaborative/cooperative work is incorporated into the program and leads to high level of interaction, integration, and cooperation
- Incorporate new services as they mature into ESnet
  - DOE Science Grid
- Outreach
  - Process involvement
  - Demonstrations
  - Workshops, conferences
  - Newsletters, reports
- Support standards development

U.S. Department of Energy

Office of Science

Network Environment Research

Program Element Strengths

- Fully integrated, real partnerships—application scientists, middleware developers, network researchers
- Lab leadership in distributed computing technology (grids)
- Effective pipeline from research through development
- Core groups developed over last eight years
  - dedicated to program philosophy of collaboration/cooperation
- Early adopters of technology in a production setting accelerates tool development and deployment

Network Environment Research

Program Area Weaknesses

- **Cultural inertia**
  - Community building is hard
  - No mechanism for institutionalizing new services
  - Success metrics for production conflicts with success metrics for research and development
- **Technology barriers**
  - Perceived cost of adopting new technology vs benefit
  - Lack of investment in fault-tolerance, error detection/recovery, etc
  - Limited ability to test at scale in a production network
  - Less than mature code with no support guarantee
- **Organizational barriers**
  - Ineffective integration of program responsibilities and accountabilities moving from research to production
  - Minimal migration of new technologies into production network

U.S. Department of Energy

Office of Science

Network Environment Research

Program Element Opportunities

- Adopt abundant cost-effective capacity in the optical core networks to develop cost-effective agile network infrastructures to support high-impact science applications
- Develop terabit networks and services for interconnecting data analysis and management centers associated with Petascale computers
- Partnerships between scientists and network/middleware researchers can be exploited to increase scale and productivity of science in areas like bioinformatics and nanotechnology
- Apply research network testbeds to accelerate the development of advanced networking technologies
- Develop experimental network testbeds to test and validate new network technologies using real world applications

Network Environment Research

Program Element Threats

- Heterogeneity of devices, mechanisms and polices
- Chicken and egg – advanced infrastructure and services must be persistent before scientists are willing to invest time in using, but infrastructure is justified through demand
- Long-term software evolution and support
- Long-term planning for facilities and programs
- Onerous policies promulgated without a clear understanding of the open science environment
- Maintain critical mass and momentum
- Highly specialized network technologies beyond scope of industry
- Graduates of traditional CS curriculums no exposure to unique DOE challenges

U.S. Department of Energy

Office of Science

Network Environment Research

Program Element Gap Analysis

- End-to-end performance
  - Multi-domain
  - Ultra high-speed transport protocol
  - Network measurement and prediction
- Cyber security
  - scalable distributed authentication and authorization systems
  - Ultra high-speed network components
- High-Performance Middleware
  - Network caching and computing
  - Real-time collaborative control and data streams
  - Fault-tolerance, error detection/correction
- Integrated testbeds and networks
  - Network research to accelerate advanced technologies
  - Experimental deployment of high-impact applications