

Office of Science
Notice 01-06

***Scientific Discovery through Advanced Computing:
National Collaboratories and High Performance Networks***

Department of Energy
Office of Science

**Office of Science Financial Assistance Program Notice 01-06: Scientific
Discovery through Advanced Computing: National Collaboratories and High
Performance Networks**

Agency: U.S. Department of Energy

Action: Notice Inviting Grant Applications

SUMMARY: The Office of Advanced Scientific Computing Research (ASCR) of the Office of Science (SC), U.S. Department of Energy (DOE), hereby announces its interest in receiving applications for grants in support of the National Collaboratories and High Performance Networks Programs, which include scope supportive of the Scientific Discovery through Advanced Computing Initiative. Collaboratories link geographically dispersed researchers, data, and tools via high performance networks to enable remote access to facilities, access to large datasets, shared environments, and ease of collaboration. This announcement is focused on research and development to support DOE-specific activities in three areas: 1) high performance middleware services that include, but are not limited to, software to allow applications to adapt to changing network conditions and software that provides ease of collaboration for distributed teams; 2) innovative, high performance network research that includes, but is not limited to, high performance transport protocols, network measurement and analysis, and traffic engineering tools and services which are focused on improving the end-to-end performance for data intensive scientific applications; and 3) collaboratories to test and validate the enabling technologies for discipline-specific applications. Collaborations across organizations that include networking researchers, middleware developers and discipline-specific scientists are encouraged. The full text of Program Notice 01-06 is available via the Internet using the following web site address: <http://www.science.doe.gov/production/grants/grants.html>.

DATES: Preapplications referencing Program Notice 01-06 should be received by January 31, 2001. Formal applications in response to this notice should be received by

4:30 p.m., E.S.T., March 15, 2001, to be accepted for merit review and funding in FY 2001.

ADDRESS: Preapplications referencing Program Notice 01-06 should be sent via e-mail using the following address: preapplications@er.doe.gov. Formal applications referencing Program Notice 01-06, should be forwarded to: U.S. Department of Energy, Office of Science, Grants and Contracts Division, SC-64, 19901 Germantown Road, Germantown, MD 20874-1290, ATTN: Program Notice 01-06. This address must be used when submitting applications by U.S. Postal Service Express Mail or any commercial mail delivery service, or when hand-carried by the applicant.

FOR FURTHER INFORMATION CONTACT: For further information on this notice contact:

National Collaboratories: Dr. Mary Anne Scott, Office of Advanced Scientific Computing Research, SC-31, Office of Science, U.S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874-1290, telephone: (301) 903-6368, e-mail: scott@er.doe.gov.

High Performance Networks: Dr. Thomas D. Ndousse, Office of Advanced Scientific Computing Research, SC-31, Office of Science, U.S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874-1290, telephone: (301) 903-9960, e-mail: tndousse@er.doe.gov.

SUPPLEMENTARY INFORMATION:

Background: Scientific Discovery Through Advanced Computing

Advanced scientific computing will be a key contributor to scientific research in the 21st Century. Within the Office of Science (SC), scientific computing programs and facilities are already essential to progress in many areas of research critical to the nation. Major scientific challenges exist in all SC research programs that can best be addressed through advances in scientific supercomputing-designing materials with selected properties, elucidating the structure and function of proteins, understanding and controlling plasma turbulence, and designing new particle accelerators. To help ensure its missions are met, SC is bringing together advanced scientific computing and scientific research in an integrated program entitled "Scientific Discovery Through Advanced Computing."

The Opportunity and the Challenge

Extraordinary advances in computing technology in the past decade have set the stage for a major advance in scientific computing. Within the next five to ten years, computers that are 1,000 times faster than today's computers will become available. These advances herald a new era in scientific computing. Using such computers, it will be possible to dramatically extend our exploration of the fundamental processes

of nature (e.g., the structure of matter from the most elementary particles to the building blocks of life) as well as advance our ability to predict the behavior of a broad range of complex natural and engineered systems (e.g., the earth's climate or an automobile engine).

To exploit this opportunity, these computing advances must be translated into corresponding increases in the performance of the scientific codes used to model physical, chemical, and biological systems. *This is a daunting problem.* Current advances in computing technology are being driven by market forces in the commercial sector, not by scientific computing. Harnessing commercial computing technology for scientific research poses problems unlike those encountered in previous supercomputers, in magnitude as well as in kind. As noted in the 1998 report (See Footnote Number 1) from the NSF/DOE "National Workshop on Advanced Scientific Computing" and the 1999 report (See Footnote Number 2) from the President's Information Technology Advisory Committee, this problem will only be solved by increasing investments in *computer software* - in research and development on scientific simulation codes as well as on the mathematical and computing systems software that underlie these codes.

Investment Plan of the Office of Science

To meet the challenge posed by the new generation of terascale computers, SC will fund a set of coordinated investments as outlined in the long plan for scientific computing, "Scientific Discovery through Advanced Computing," (See Footnote Number 3) submitted to Congress on March 30, 2000. First, it will create a *Scientific Computing Software Infrastructure* that bridges the gap between the advanced computing technologies being developed by the computer industry and the scientific research programs sponsored by the Office of Science. Specifically, the SC effort proposes to:

- Create a new generation of *Scientific Simulation Codes* that take full advantage of the extraordinary computing capabilities of terascale computers.
- Create the *Mathematical and Computing Systems Software* to enable the Scientific Simulation Codes to effectively and efficiently use terascale computers.
- Create a *Collaboratory Software Environment* to enable geographically separated scientists to effectively work together as a team and to facilitate remote access to both facilities and data.

These activities will be supported by a *Scientific Computing Hardware Infrastructure* that has been tailored to meet the needs of its research programs. The *Hardware Infrastructure* is *robust*, to provide the stable computing resources needed by the scientific applications; *agile*, to respond to innovative advances in computer

technology that impact scientific computing; and *flexible*, to allow the most appropriate and economical resources to be used to solve each class of problems. Specifically, the SC proposes to support:

- *A Flagship Computing Facility*, the National Energy Research Scientific Computing Center (NERSC), to provide the robust, high-end computing resources needed by a broad range of scientific research programs.
- *Topical Computing Facilities* to provide computing resources tailored for specific scientific applications and to serve as the focal point for an application community as it strives to optimize its use of terascale computers.
- *Experimental Computing Facilities* to assess the promise of new computing technologies being developed by the computer industry for scientific applications.

Both sets of investments will create exciting opportunities for teams of researchers from laboratories and universities to create new revolutionary computing capabilities for scientific discovery.

The Benefits

The *Scientific Computing Software Infrastructure*, along with the upgrades to the hardware infrastructure, will enable laboratory and university researchers to solve the most challenging scientific problems faced by the Office of Science at a level of accuracy and detail never before achieved. These developments will have significant benefit to all of the government agencies who rely on high-performance scientific computing to achieve their mission goals as well as to the U.S. high-performance computing industry.

Background: National Collaboratories and High Performance Networks

The current core programs in ASCR are intended to enhance the Department's ability to satisfy mission requirements through advanced technologies such as distributed computing, national collaboratories, high performance networks, remote access to facilities, and remote access to petabyte-scale datasets with complex internal structure. Within this context, the National Collaboratories and High Performance Networks Programs provide a coordinated program of technology research and development that leverages the strengths of computer and computational science research programs and partners with science application pilot projects. Likewise, these programs support the Scientific Discovery through Advanced Computing by enabling integration of multi-institutional, geographically-dispersed researcher into effective, efficient teams and by providing distributed computing environments and tools to support the use of remote computers and access to data and facilities.

Advances in high performance network capabilities and collaboration technologies are making it easier for large geographically dispersed teams to collaborate effectively. This is especially important for the teams using the major computational resources, data resources, and experimental facilities supported by DOE. With leadership from DOE, these geographically distributed laboratories or collaboratories have begun to play an important role in the Nation's scientific enterprise. The importance of collaboratories is expected to increase in the future. However, significant research questions must be addressed if collaboratories are to achieve their potential: namely, to enable remote access facilities that produce petabytes/year; to provide remote users an experience that approaches the same as "being there;" to provide remote visualization of terabyte to petabyte data sets from computational simulation; and to enable effective remote access to advanced scientific computers.

Solving the challenging network and distributed computing problems calls for new modalities of scientific research. Many scientific applications when deployed on existing networks fail to meet the end-to-end expectations for performance. This is especially true for distributed high-end applications such as remote visualization and high capacity data transfer. Recent advances in optical networks brought about by Dense Wave Division Multiplexing (DWDM) are resulting in unprecedented increases for bandwidth in the core networks. However, many challenging protocol engineering, traffic engineering, and high-performance middleware problems must be addressed before complex scientific high-end applications and collaboratories can benefit from this increase in bandwidth. Harnessing this bandwidth at the application level poses some important and challenging problems.

Research is needed to understand what services collaboratories require and how these services should be integrated with the large number of network devices and network-attached devices that must work together. Examples of the components and services that need to be integrated include: data archives on tape, high performance disk caches, visualization and data analysis servers, authentication and security services, directory services, network resources, and computational systems including the computer on a scientist's desk. All of these physical and software services must be tied together by common software framework building blocks or "middleware" to enable the collaboratories of the future to succeed.

Further, at the network level, research is needed for advanced services to develop advanced network services and tools to deliver high end-to-end performance to distributed scientific applications. There are several areas that can contribute to improving the end-to-end performance for secure multi-gigabits/sec transport that some of DOE's advanced scientific applications require. These include: enhancement of existing transport protocols, the development of accurate measurement and analysis techniques, and the network services that can provide online performance predictions.

These challenges will be addressed through an integrated program of fundamental research in high performance networking and collaboratory technologies in partnership with key scientific disciplines that provide the applications--the research may be focused for short-term results (within three years) or long term (five-years and greater). This announcement seeks applications in three areas:

1. **Middleware:** research and development projects that will address individual technology elements to enable universal, ubiquitous, easy access to remote resources or that will contribute to the ease with which distributed teams work together. Enabling high performance for scientific applications is an important consideration.
2. **Collaboratory Pilots:** research and development of enabling technologies that is integrated with and required by distributed scientific applications. An example of such a distributed application is the real-time data acquisition, reduction and visualization for macromolecular crystallography using a high intensity X-ray light source facility remotely. Another distributed application could be an extensive network measurement and analysis infrastructure employed to diagnose and predict end-to-end performance.
3. **High Performance Network Engineering:** research, development, and testing of advanced network protocols, traffic engineering, and network services that can significantly improve capabilities, end-to-end performance, and controllability of networks infrastructures designed to support distributed scientific applications.

To the extent that software and/or infrastructure development is involved, all applications to this notice should address the issues that characterize a successful research lifecycle. That is, technology transfer strategies should be provided for the transition of research code and/or infrastructure into robust production. Long term software evolution and maintenance and end user support should also be considered.

Integration of work efforts across all projects funded under this notice will occur following the awards, to preclude duplication of effort and to maximize leveraging and coordination. Projects are expected to work closely with other SciDAC teams, where identified during this integration. Coordination through a participatory management process will continue for the life of the projects.

(See <http://doecollaboratory.pnl.gov/> for a list of currently funded projects in National Collaboratories and background of the program that began as the DOE 2000 Initiative.)

(See http://www.er.doe.gov/production/octr/mics/network_research. For background on the High Performance Networks Program.)

Solicitation Emphasis Areas

1. **Middleware** technology research and development projects are to have certain characteristics. Products of this research and development are expected to provide services that interoperate and feature common interfaces. It should be easy to learn and use the tools. Applications in response to this notice should delineate an effective strategy for coupling with requirements from the scientific applications of the potential collaboratories. Applications in response to this notice should also provide a plan for software maintenance and support.

Middleware technology research and development projects that enable collaboration may focus on provide a broad set of tools or toolkits to support, but are not limited to, the following areas of interest:

- Collaborative Visualization
- Collaborative Problem Solving Environments
- Real-time Analysis
- Group Collaboration
- Data Management
- Science Portals
- On-line Instrumentation
- Data Grids

In addition, middleware technology research and development projects may address standard services and protocols that are needed to enable persistent, universal, and ubiquitous access to networked resources, such as, but are not limited to, the following:

- Directory Services
- Authentication/Authorization Services
- Co-scheduling Distributed Resources
- Multicast and efficient broadcast capabilities
- Automatic resource discovery protocols
- Remote data access services
- Network-attached memory and storage systems
- Communications services

For middleware technology research and development projects, it is estimated that between four and eight awards could be made in FY 2001, contingent upon the availability of appropriated funds. The scope of a single-focus project is expected to range from \$150K to \$500K.

Collaboratory pilots should have certain characteristics.

The project should:

- address a problem of national scientific or engineering significance clearly related to the mission of DOE and have high visibility.
- involve geographically separated groups of personnel and/or facilities that are inherently required to collaborate or be used remotely for success of the project.

The project may:

- focus on developing and providing a set of middleware services needed by a broad set of applications requiring distributed computing capabilities.
- focus solely on advanced network development and testing such as a measurement and analysis infrastructure to accurately measure, calibrate, diagnose performance related problems, and predict the end-to-end performance of operational high-speed networks.

All responses to this notice must provide a plan for transition to sustaining activities and services for end users on completion of the project. The scope of a collaboratory pilot is expected to be about \$0.5M to \$2.5M total per year. This is the total for all the institutions participating and it is expected that a single institution would be funded at a level of no more than \$600K. It is estimated that three to five awards will be made for this area during FY 2001.

It is also possible for middleware technology research and development projects and/or collaboratory pilots to address an element for evaluating systems and their impact on the process of science, namely identifying factors that facilitate or impede the adoption of technology.

2. High Performance Network Engineering is key to the DOE vision of collaborative scientific research environments in which geographically distributed research teams and computing resources are interconnected to form a virtual computing research environment. Emerging high-end scientific applications, when deployed on existing networks, fail to meet the expected end-to-end performance, latency, security, and guaranteed quality of service required for complex scientific investigations. The high-performance network program addresses these challenges in the current announcement by focusing in three major research areas of high performance network engineering:

- Network Measurement and Analysis: Focuses on the fundamental issues of end-to-end performance through measurement and analysis.
- High-performance Transport Protocols: Addresses the performance and security enhancement issues of traditional protocols operating in high-speed, high-performance networks.
- Advanced Traffic Engineering Tools And Services: Deals with advanced tools and service for managing, differentiating, and controlling network traffic in order to satisfy the end-to-end performance objectives.

(a) *Network Measurement and Analysis*: Applications may address innovative scalable network measurement and analysis infrastructures, tools, services, etc., that can be used to accurately measure, calibrate, diagnose performance related problems, and predict the end-to-end performance of operational high-speed network networks. This may involve passive and active measurement, SNMP derived data, or a combination and may include, but not be limited to, the following:

- Bandwidth estimation techniques for high-speed links (OC-12, OC-48)
- Measurement infrastructures to collect, store, and analyze traffic traces
- Distributed agent architecture for network measurement and analysis
- On-line analysis and data mining of measured data
- Dynamic end-to-end path selection based on online analysis
- Measurement and calibration of transport protocol performance

Applications focusing on measurement and analysis infrastructures are expected to work in close collaboration with DOE's Energy Science Network (ESnet) in the deployment measurement facilities. A network research testbed facility has been established, with the cooperation of ESnet, for experimental network research activities. Researchers requiring the use of this experimental facility are encouraged to work closely with the ESnet Research Support Subcommittee (ESRSC) chartered to coordinate the activities of the testbed. A complete description of this experimental facility can be found at <http://www.es.net>.

(b) *High-Performance Transport Protocols*: The performance expectation for the delivery of multi-gigabits/sec throughput to distributed scientific applications far exceeds the capability of current networks. This performance expectation raises some fundamental questions concerning the capability of conventional routing protocols optimized for low-speed, best-effort traffic. The current announcement addresses transport protocol performance issues by seeking innovative approaches that may include but are not be limited to the following:

- Transport protocol measurement, tuning, and calibration tools
- Adaptive extensions of transport protocols for high-speed networks

- High-performance network traffic characterization
- Transport protocol parallelization at high-speed

The objective is to reduce the contribution of transport protocol on end-to-end congestion. Potential applications must provide a sound mathematical analysis of the proposed enhancements when subjected to high-end scientific applications that potentially exercise its important features.

(c) *Advanced Traffic Engineering Tools and Services*: Addresses the resource and performance optimization of high-performance and high-speed networks, including advanced traffic management and control strategies, services, and tools that can be used for traffic differentiation and for steering traffic. Applications may focus on, but are not limited to, the following:

- QoS-based routing and source routing
- Dynamic routing and traffic control
- Congestion notification and avoidance
- Bandwidth brokering services
- Advanced traffic management tools and services
- Simulation of large traffic flows

Applications addressing these and other related issues should concentrate on those activities that lead to a significant improvement in end-to-end performance of applications running across high performance networks.

The high-performance network research program anticipates funding projects in these three areas in FY 2001. It should also be noted that a collaboratory pilot (as discussed under section 2.) may focus solely on advanced network development and testing such as a measurement and analysis infrastructure to accurately measure, calibrate, diagnose performance related problems, and predict the end-to-end performance of operational high-speed networks. The scope of a single project is expected to range from \$150K to \$500K.

Preapplications

Potential applicants are strongly encouraged to submit a brief preapplication that consists of two to three pages of narrative describing the research objectives and technical approach(s). Preapplications will be reviewed relative to the scope and research needs of the ASCR National Collaboratories and High Performance Networks Programs, as outlined in the summary paragraph and in the SUPPLEMENTARY INFORMATION. The preapplication should identify, on the cover sheet, the title of the project, the institution, principal investigator name,

telephone, fax, and e-mail address. The focus element (Middleware Technology, Collaboratory Pilots, or High Performance Network Engineering) for the preapplication should also be clearly identified. A response to each preapplication discussing the potential programmatic relevance of a formal application will be communicated to the Principal Investigator within 7 to 14 days of receipt.

Collaboration

Applicants are encouraged to collaborate with researchers in other institutions, such as: universities, industry, non-profit organizations, federal laboratories and Federally Funded Research and Development Centers (FFRDCs), including the DOE National Laboratories, where appropriate, and to include cost sharing wherever feasible. Additional information on collaboration is available in the Application Guide for the Office of Science Financial Assistance Program that is available via the Internet at: <http://www.sc.doe.gov/production/grants/Colab.html>.

Program Funding

It is anticipated that up to \$6 million will be available for all National Collaboratories and High Performance Networks Programs awards in Fiscal Year 2001; from ten to as many as fifteen awards are anticipated, contingent on availability of appropriated funds in FY 2001 and the size of the awards. Multiple year funding is expected, also contingent on availability of funds and progress of the research.

Awards are expected to be at most \$500,000 per year for individual middleware technology and network engineering R&D projects. Awards for collaboratory pilots are expected to be at most \$2.5 million per year. Since pilots are expected to be multi-institution projects, awards under this notice would range from \$200,000 to \$600,000 for participation in a pilot. The term for projects can be from one to three years.

Merit Review

Applications will be subjected to scientific merit review (peer review) and will be evaluated against the following evaluation criteria, which are listed in descending order of importance codified at 10 CFR 605.10(d):

- 1) Scientific and/or Technical Merit of the Project;
- 2) Appropriateness of the Proposed Method or Approach;
- 3) Competency of Applicant's Personnel and Adequacy of Proposed Resources;
- 4) Reasonableness and Appropriateness of the Proposed Budget.

The evaluation under item 1, Scientific and/or Technical Merit of the Project, will also consider the following elements:

- a) The potential of the proposed project to make a significant impact in the effectiveness of SciDAC applications researchers.
- b) The degree to which an application area can benefit from collaborative technology.
- c) The extent to which the project will test important collaborative technologies.
- d) The extent to which the results of the project are extensible to other program or discipline areas.

The evaluation under item 2, Appropriateness of the Proposed Method or Approach, will also consider the following elements:

- a) The degree to which the project adheres to the management philosophy of incorporating collaboration into the project execution.
- b) The quality of the plan for ensuring interoperability and integration with software produced by other SciDAC efforts.
- c) The extent to which the project incorporates broad community (industry/academia/other federal programs) interaction.
- d) Quality and clarity of proposed work schedule and deliverables.
- e) Knowledge of and coupling to previous efforts for collaborative technologies such as DOE 2000.

The evaluation will include program policy factors such as the relevance of the proposed research to the terms of the announcement and the agency's programmatic needs. Note, external peer reviewers are selected with regard to both their scientific expertise and the absence of conflict-of-interest issues. Non-federal reviewers will often be used, and submission of an application constitutes agreement that this is acceptable to the investigator(s) and the submitting institution.

Submission Information

The Project Description must be 20 pages or less, exclusive of attachments. It must contain an abstract or project summary on a separate page with the name of the applicant, mailing address, phone, FAX and E-mail listed. The application must include letters of intent from collaborators (briefly describing the intended contribution of each to the research), and short curriculum vitae for the applicant and any co-PIs.

To provide a consistent format for the submission, review and solicitation of grant applications submitted under this notice, the preparation and submission of grant applications must follow the guidelines given in the Application Guide for the Office of Science Financial Assistance Program, 10 CFR Part 605. Access to SC's Financial Assistance Application Guide is possible via the World Wide Web at:

<http://www.science.doe.gov/production/grants/grants.html>.

The Catalog of Federal Domestic Assistance number for this program is 81.049, and the solicitation control number is ERFAP 10 CFR Part 605.

John Rodney Clark
Associate Director of Science
for Resource Management

Footnotes:

1) This workshop was sponsored by the National Science Foundation and the Department of Energy and hosted by the National Academy of Sciences on July 30-31, 1998. Copies of the report may be obtained from:

<http://www.er.doe.gov/production/octr/mics/index.html>

2) Copies of the PITAC report may be obtained from: <http://www.ccic.gov/ac/report/>.

3) Copies of the SC computing plan, Scientific Discovery through Advanced Computing, can be downloaded from the SC web site at:

<http://www.sc.doe.gov/production/octr/index.html>.

Published in the Federal Register December 19, 2000, Volume 65, Number 244,
Pages 79350-79354.