Office of Science
Notice 01-07

Scientific Discovery through Advanced Computing: Integrated Software Infrastructure Centers

Department of Energy
Office of Science

Office of Science Financial Assistance Program Notice 01-07: SciDAC - Integrated Software Infrastructure Centers

Agency: U.S. Department of Energy

Action: Notice Inviting Research Grant Applications

SUMMARY: The Office of Advanced Scientific Computing Research (OASCR) of the Office of Science (SC), U.S. Department of Energy (DOE), hereby announces its interest in receiving applications for projects in the Integrated Software Infrastructure Centers (ISIC) component of the Scientific Discovery through Advanced Computing (SciDAC) research program. The software infrastructure vision of SciDAC is for a comprehensive, portable, and fully integrated suite of systems software and tools for the effective management and utilization of terascale computational resources by SciDAC applications. This infrastructure will provide maximum performance, robustness, portability and ease of use to application developers, end users, and system administrators. Successful ISIC activities must establish and maintain close interactions with other ISIC activities and SciDAC efforts, and it is essential that they address the complete software lifecycle including transition of successful research software to robust production software and appropriate mechanisms for long term software support and evolution. Partnerships among universities, national laboratories, and industry are encouraged. The full text of Program Notice 01-07 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 01-07 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.

ADDRESSES: Preapplications referencing Program Notice 01-07 should be sent via e-mail using the following address: preapplications@er.doe.gov.

Formal applications referencing Program Notice 01-07, 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 01-07. 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: Dr. Frederick C. Johnson, Office of Science, U.S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874-1290, telephone: (301) 903-5800, E-mail: fjohnson@er.doe.gov, fax: (301) 903-7774.

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, e.g., 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 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 increased 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 its long-range 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 are supported by a *Scientific Computing Hardware Infrastructure* that will be 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: Integrated Software Infrastructure Centers

This solicitation addresses the Mathematical and Computing Systems Software Environment element of the SciDAC Scientific Computing Software Infrastructure. ISIC envisions a comprehensive, integrated, scalable, and robust high performance software infrastructure, which overcomes difficult technical challenges to enable the effective use of terascale systems by SciDAC applications. ISIC addresses needs for: new algorithms which scale to parallel systems having thousands of processors; methodology for achieving portability and interoperability of complex high performance scientific software packages; operating systems tools and support for the effective management of terascale and beyond systems; and effective tools for feature identification, data management and visualization of petabyte-scale scientific data sets. ISIC provides the essential computing and communications infrastructure for support of SciDAC applications. The ISIC effort encompasses a multi-discipline approach with activities in:

- Algorithms, methods, and libraries -- Algorithms, methods and libraries that are fully scalable to many thousands of processors with full performance portability.
- Program development environments and tools -- Component-based, fully integrated, terascale program development and runtime tools, which scale effectively and provide maximum utility and ease-of-use to developers and scientific end users.

- Operating system software and tools -- Systems software that scales to tens of thousands of processors, supports high performance application-level communication and provides the highest levels of fault tolerance, reliability, manageability, and ease of use for system administrators, tool developers and end users.
- Visualization and data management systems -- Scalable, intuitive systems fully supportive of SciDAC application requirements for moving, storing, analyzing, querying, manipulating and visualizing multi-petabytes of scientific data and objects.

The complexity of these challenges and the strong emphasis on scalability, interoperability and portability requires novel approaches in the proposed technical research and the research management structure. ISIC emphasizes the formation of Enabling Technologies Centers (ETC) as an organizational basis for successful applications. An ETC is a virtual multi-institution, multi-disciplinary team which will:

- Create mathematical and/or computing systems software to enable scientific simulation codes to take full advantage of the extraordinary capabilities of terascale computers;
- Work closely with application teams and other SciDAC teams to ensure that the most critical computer science and applied mathematics issues are addressed in a timely and comprehensive fashion; and
- Address all aspects of the successful research software lifecycle including transition of a research code into a robust production code and long term software evolution and maintenance and end user support.

Solicitation Emphasis

This notice is one of several that addresses the initial requirements of the SciDAC program. The focus is on four topics: 1) algorithms, methods and libraries; 2) program development environments and tools; 3) operating systems software and tools; and 4) visualization and data management. Responses to this notice may propose work in one or more of these areas and may be single institution efforts or partnerships that involve many organizations. It is expected that most, if not all, of the proposed activities will be organized as ETCs. Specific areas of interest include, but are not limited to, the following examples listed for each subtopic:

1) Algorithms, Methods and Mathematical Libraries

a) Mesh generation and discretization technology. Tools to facilitate the generation and partitioning of all types of meshes (structured, unstructured, and chimera (overlapping)) designed for many thousands of processors.

b) Mathematical analysis and scalable numerical algorithms. Mathematical methods to help SciDAC applications achieve high performance on hierarchical memory terascale computers such as multiscale analysis, multilevel methods, and fast transforms capable of spanning multiple spatial and temporal scales. Resultant algorithms must be deployed in component-based mathematical software and made available to a broad range of DOE mission areas.

2) Program Development Environment and Tools

a) High Performance Component Architectures. Component technology that builds upon and extends commercial component architectures to support high performance parallel components, low-latency, high bandwidth communication among components, and efficient data and work redistribution.

b) Code Design and Development Tools. Scaling methodology to deploy existing parallel code development environments on multi-teraflops SciDAC systems. Support for multi-language applications including C, C++, UPC, Fortran, Co-Array Fortran, Python and Java; parallel programming libraries, such as MPI, OpenMP, thread libraries, the Global Array library; and multi-level hierarchical memory programming models.

c) Code Correctness and Validation. Debugging tools that implement emerging community standards in parallel debuggers and automated data dependency analysis. Relative debugging methodology for comparing at run time the execution of two versions of a code.

d) Performance Tools. Evaluation of existing research and commercial performance analysis tools, both tracefile-based and dynamic, for scalability and suitability for SciDAC applications. Performance metrics and benchmarks which enable reliable and credible performance predictions of application codes on terascale and larger systems. Tools which link hardware counters to meaningful terascale system performance characteristics and application performance.

3) Operating System Software and Tools

a) Terascale System Resource Management. Modular infrastructure for resource management on terascale clusters including resource scheduling, meta-scheduling, node daemon support, comprehensive usage accounting and user interfaces that also emphasizes portability to terascale vendor operating systems.

b) Terascale System Support. Scalable checkpointing and improved runtime steering for early deployment. Methodology for analyzing tradeoffs between fault tolerance and peak performance. Support for robust runtime job management and I/O systems

that are tolerant of component failure. Scalable tools for system administration including initial system boot, system updates, job launch and system utilities.

c) High Performance Communication. Operating system support for application level communication which scales to thousands of processors, provides minimum latency and maximum bandwidth between parallel application processes. Innovative approaches to terascale operating system architectures including non-uniform kernel support for computational, service, interactive and i/o nodes.

4) Visualization and Data Management

a) Data Management Systems. Data exchange methods and standardizations that facilitate collaborative applications. Innovative Database Management Systems (DBMS) approaches for high throughput parallel I/O and complex queries of large scientific databases. Hierarchical data storage systems involving tertiary storage media that are sequential. Agent methodology for feature extraction and complex query operations. Tools for user-driven and automatic clustering, reclustering or replication of objects to maximize retrieval efficiency. Collaborations with the DBMS and tertiary storage vendor industry are encouraged.

b) Visualization. Vector/tensor field visualization in 3-D. Modes of visualization for interpretation and understanding of large datasets. Remote and collaborative visualization methods. Characterization of simulation, experimental and visualization errors/uncertainties. Adaptive, multiresolution, parallel and scalable visualization algorithms. Innovative techniques for exploring multi-dimensional, multi-discipline data sets.

Collaborations with the high performance hardware and software vendor industry are encouraged wherever appropriate.

Integration of Software Components and Tools

Responses to this notice should cover the full range of activities from basic research to development of software that can be deployed to the SciDAC applications communities. It is critical that these submissions demonstrate effective strategies for coupling with requirements from applications researchers and ensuring that software developed will interoperate with software developed by other ISIC activities and be effectively deployed to SciDAC computing facilities and applications groups.

ISIC envisions a fully integrated software environment that provides both robustness and ease of use to the end user application scientist. Implementation of this vision will be coordinated through a participatory management process with input from ISIC teams and other key participants of SciDAC. As component and tool implementations

mature, each team will be expected to develop the necessary technology to fully and smoothly incorporate their software tools into the ISIC environment.

ISIC activities play a critical cross-cutting role in the SciDAC. ISIC goals require significant interactions, ranging from the joint development and deployment of tools and technologies into the applications community, to the incorporation of needed capabilities into new products and systems. ISIC researchers will need to interact closely with diverse groups including: applications scientists, vendor providers, the DOE ASCI program, and other federal agency programs addressing complementary goals. To support and facilitate the maximum impact of the SciDAC Scientific Computing Software Infrastructure, high emphasis will be placed on ensuring that source code is fully and freely available for use and modification throughout the scientific computing community.

This solicitation is focused on larger ETC efforts in support of the SciDAC program. Applications to the OASCR base program through the Continuing Solicitation for all Office of Science Programs Notice 01-01, found at http://www.science.doe.gov/production/grants/grants.html, which may have the potential for contributing to the ISIC software infrastructure, should so indicate.

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 $7 million annually will be available for multiple awards for these components of the ISIC program. Initial awards will be made in FY 2001 in the categories described above, and applications may request project support for up to five years. All awards are contingent on the availability of funds, research progress, and programmatic needs. Annual budgets for successful ISIC projects are expected to range from $2,000,000 to $4,000,000 per project. Annual budgets may increase in the out-years but should remain within the overall annual maximum guidance. Any proposed effort that exceeds the annual maximum in the out-years should be separately identified for potential award increases if additional funds become available.

Preapplications

Preapplications are strongly encouraged but not required prior to submission of a full application. However, notification of a successful preapplication is not an indication that an award will be made in response to the formal application. The preapplication should identify on the cover sheet the institution, Principal Investigator name(s), address(s), telephone, and fax number(s) and E-mail address(es), title of the project, and the field of scientific research. A brief (one-page) vitae should be provided for each Principal Investigator. The preapplication should consist of a two to three page narrative describing the research project objectives, the approach to be taken, and a description of any research partnerships. Preapplications will be reviewed by DOE relative to the scope and research needs of the ISIC program.

Merit Review

Applications will be subjected to scientific merit review (peer review) and will be evaluated against the following evaluation criteria listed in descending order of importance as 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 of applications under item 1, Scientific and Technical Merit, will pay particular attention to:
a) The potential of the proposed project to make a significant impact in the effectiveness of SciDAC applications researchers;
b) The demonstrated capabilities of the applicants to perform basic research related to ISIC and transform these research results into software that can be widely deployed;
c) The likelihood that the algorithms, methods, mathematical libraries, and software components that result from this effort will have impact on science disciplines outside of the SciDAC applications projects;
d) Identification and approach to software integration and long term support issues, including component technology, documentation, test cases, tutorials, end user training, and quality maintenance and evolution.
The evaluation under item 2, Appropriateness of the Proposed Method or Approach, will also consider the following elements related to Quality of Planning:
a) Quality of the plan for effective coupling to applications researchers;
b) Quality of plan for ensuring interoperability and integration with software produced by other ISIC and SciDAC efforts;
c) Viability of plan for deployment of software to SciDAC facilities and applications

groups;
d) Knowledge of and coupling to other efforts in high performance scientific computing software such as the DOE ACTS program, the DOE ASCI program and the NSF ITR program;
e) Quality and clarity of proposed work schedule and deliverables.
Note that external peer reviewers are selected with regard to both their scientific expertise and the absence of conflict-of-interest issues. Non-federal reviewers may be used, and submission of an application constitutes agreement that this is acceptable to the investigator(s) and the submitting institution. Reviewers will be selected to represent expertise in the technology areas proposed, applications groups that are potential users of the technology, and related programs in other Federal Agencies or parts of DOE, such as the Advanced Strategic Computing Initiative (ASCI) within DOE's National Nuclear Security Administration.

Information about the development and submission of applications, eligibility, limitations, evaluation, selection process, and other policies and procedures including detailed procedures for submitting proposals from multi-institution partnerships may be found in 10 CFR Part 605, and in the Application Guide for the Office of Science Financial Assistance Program. Electronic access to the Guide and required forms is made available via the World Wide Web at: http://www.science.doe.gov/production/grants/grants.html. The Project Description must be 20 pages or less, including tables and figures, but exclusive of attachments. The application must contain an abstract or project summary, letters of intent from collaborators, and short vitae.

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 website at: http://www.sc.doe.gov/production/octr/index.html.

Printed in the Federal Register December 19, 2000, Volume 65, Number 244, Pages 79346-79350.