



NASA SBIR 2009 Phase I Solicitation

S6 Information Technologies

NASA Missions and Programs create a wealth of science data and information that are essential to understanding our earth, our solar system and the universe. Advancements in information technology will allow many people within and beyond the Agency to more effectively analyze and apply this data to create knowledge. In particular, modeling and simulation are being used more pervasively throughout NASA, for both engineering and science pursuits, than ever before. These are tools that allow high fidelity simulations of systems in environments that are difficult or impossible to create on Earth, allow removal of humans from experiments in dangerous situations, and provide visualizations of datasets that are extremely large and complicated. In many of these situations, assimilation of real data into a highly sophisticated physics model is needed. Information technology is also being used to allow better access to science data, more effective and robust tools for analyzing and manipulating data, and better methods for collaboration between scientists or other interested parties. The desired end result is to see that NASA science information be used to generate the maximum possible impact to the nation: to advance scientific knowledge and technological capabilities, to inspire and motivate the nation's students and teachers, and to engage and educate the public.

Subtopics

S6.01 Technologies for Large-Scale Numerical Simulation

Lead Center: ARC

Participating Center(s): GSFC

NASA scientists and engineers are increasingly turning to large-scale numerical simulation on supercomputers to advance understanding of complex Earth and astrophysical systems, and to conduct high-fidelity aerospace engineering analyses. The goal of this subtopic is to increase the mission impact of NASA's investments in supercomputing systems and associated operations and services. Specific objectives are to:

- Decrease the barriers to entry for prospective supercomputing users;
- Minimize the supercomputer user's total time-to-solution (e.g., time to discover, understand, predict, or design);

-
- Increase the achievable scale and complexity of computational analysis, data ingest, and data communications;
 - Reduce the cost of providing a given level of supercomputing performance on NASA applications; and
 - Enhance the efficiency and effectiveness of NASA's supercomputing operations and services.

Expected outcomes are to improve the productivity of NASA's supercomputing users, broaden NASA's supercomputing user base, accelerate advancement of NASA science and engineering, and benefit the supercomputing community through dissemination of operational best practices.

The approach of this subtopic is to seek novel software and hardware technologies that provide notable benefits to NASA's supercomputing users and facilities, and to infuse these technologies into NASA supercomputing operations. Successful technology development efforts under this subtopic would be considered for follow-on funding by, and infusion into, NASA's high-end computing (HEC) projects (<http://www.hec.nasa.gov/>): the High End Computing Capability project at Ames and the Scientific Computing project at Goddard. To assure maximum relevance to NASA, funded SBIR contracts under this subtopic should engage in direct interactions with one or both HEC projects, and with key HEC users where appropriate. Research should be conducted to demonstrate technical feasibility and NASA relevance during Phase 1 and show a path toward a Phase 2 prototype demonstration.

Offerors should demonstrate awareness of the state-of-the-art of their proposed technology, and should leverage existing commercial capabilities and research efforts where appropriate. Open source software and open standards are strongly preferred. Note that the NASA supercomputing environment is characterized by: HEC systems operating behind a firewall to meet strict IT security requirements, many applications requiring tight coupling and high concurrency, complex computational workflows and immense datasets, and the need to support hundreds of complex application codes - many of which are frequently updated by the user/developer. As a result, solutions that involve the following must clearly explain how they would work in the NASA environment: Grid computing, web services, client-server models, embarrassingly parallel computations, and technologies that require significant application re-engineering. Projects need not benefit all NASA HEC users or application codes, but demonstrating applicability to an important NASA discipline, or even a key NASA application code, could provide significant value.

Specific technology areas of interest include:

- **Integrated Environments:** The user interface to a supercomputer is typically a command line in a text window. This subtopic element seeks more intuitive, intelligent, user-customized, and integrated interfaces to supercomputing resources, enabling users to more completely leverage the power of HEC to increase their productivity. Such an interface could enhance many essential supercomputing tasks: accessing and managing resources, training, getting services, developing codes, running computations, managing files and data, analyzing and visualizing results, transmitting data, collaborating, etc.
- **Efficient Computing:** In spite of the rapidly increasing capability and efficiency of supercomputers, NASA's HEC facilities cannot purchase, power, and cool sufficient HEC resources to satisfy all user demands. This subtopic element seeks dramatically more efficient and effective supercomputing approaches in terms of their ability to supply increased HEC capability or capacity per dollar and/or per Watt for real NASA applications. Examples include novel computational accelerators and architectures, more capable storage/interconnect/visualization technologies, improved algorithms for key codes, and power-aware "Green" computing technologies and techniques.

-
- HEC Ecosystem Modeling: NASA endeavors to maximize the productivity of its world-class HEC activities. To identify and prioritize improvement initiatives, this subtopic element seeks tools and techniques to routinely monitor and model the productivity of NASA's HEC ecosystem, including modeling change scenarios. The technology should model the workflows of HEC users, facility staff, and resources (supercomputers, storage, networks, etc.), and it should reflect constraints such as budget, power, and space. Offerors should minimize the effort of HEC staff to provide process information.
 - Archive Data Use: NASA has a vast and rapidly growing wealth of Earth and space observational data, stored in various archives around the U.S. NASA's supercomputers could extract more value from this data and advance NASA's science missions through large-scale data analysis and visualization, and ingest into high-fidelity models. This subtopic element seeks technologies that facilitate efficient, automated use of data in NASA's observational data archives by its HEC centers and users.

Proposals should show an understanding of one or more relevant science needs, and present a feasible plan to fully develop a technology and infuse it into a NASA program.

S6.02 Earth Science Applied Research and Decision Support

Lead Center: SSC

Participating Center(s): ARC, JPL

The NASA Applied Sciences Program (<http://nasascience.nasa.gov/earth-science/applied-sciences>) seeks innovative and unique approaches to increase the utilization and extend the benefit of Earth Science research data to better meet societal needs. One area of interest is new decision support tools and systems for a variety of ecological applications such as managing coastal environments, natural resources or natural disasters.

This subtopic seeks new, advanced information systems and decision environments that take full advantage of multiple data sources and platforms. Tailored distribution networks and timely products delivered to a broad range of users are needed to support applications in disaster management, resource management, energy and urban sustainability.

- Development of new integrated multiple user requirements knowledge data bases and archival library tools to support researchers and promote infusion of successful technologies into existing processes.
- Development of new decision support strategies and presentation methodologies for applied earth science applications to reduce risk, cost, and time.

This subtopic is also soliciting proposals for utilities, plug-ins or enhancements to open source geobrowsers that improve their utility for earth science research and decision support. Examples of geobrowsers include NASA World Wind, World Wind Java (http://worldwindcentral.com/wiki/Main_page) and COAST (<http://www.coastal.ssc.nasa.gov/coast/COAST.aspx>). Special consideration will be given to tools for COAST. Examples of specific interest are:

-
- Tools and utilities to support creation or simplify the import and integration of new datasets;
 - Tools and utilities to discover and integrate existing web-enabled sensor data (e.g., webcams, meteorology stations, beach monitors);
 - Innovative output mechanisms for data layer sharing and collaboration;
 - Enhancements to visualization of custom 3rd dimensional data;
 - Enhancements to real time animation capabilities, or incorporation of existing animations into a geobrowser;
 - Plug-ins that enable visualization of high resolution imagery in a COAST accessible data viewer;
 - Utilities that enable regional estuarine or bay data compilations that are of interest to the major coastal ecosystem managers in those areas;
 - Applications that subset, filter, merge, and reformat existing spatial data; provide links to attribute data; or visualize spatial or temporal analytic results in innovative value added fashion within the application.

Proposals should present a feasible plan to fully develop and apply the subject technology.

S6.03 Algorithms for Science Data Processing and Analysis

Lead Center: GSFC

Participating Center(s): ARC, JPL, LaRC, MSFC, SSC

This subtopic seeks technical innovation and unique approaches for the processing and the analysis of data from NASA science missions. Analysis of NASA science data enables insights into dynamic systems such as the sun, oceans, and earth's climate in addition to looking back in time to explore the origins of the universe. Complex algorithms and intensive data processing are needed to understand and utilize this data. Advances in such algorithms will support science data analysis and decision support systems related to current and future missions and mission concepts such as:

Current operational missions listed at <http://www.nasa.gov/missions/current/index.html>

- Landsat Data Continuity Mission (LDCM) (<http://ldcm.nasa.gov/>),
- NPOES Preparatory Project (NPP) (<http://jointmission.gsfc.nasa.gov/>),
- Lunar Reconnaissance Orbiter (LRO) (<http://lunar.gsfc.nasa.gov/>),

-
- Orbiting Carbon Observatory (OCO) (<http://oco.jpl.nasa.gov/>),
 - Lunar Atmosphere and Dust Environment Explorer (LADEE) (<http://nasascience.nasa.gov/missions/ladee>),
 - Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) (<http://crism.jhuapl.edu>),
 - HypsIRI Earth orbiting hyperspectral instrument (<http://hyspiri.jpl.nasa.gov/>),
 - Visual Infrared Mapping Spectrometer (VIMS) on Cassini (<http://wwwvims.lpl.arizona.edu/>),
 - Moon Mineralogy Mapper (M3) on Chandrayaan (<http://moonmineralogymapper.jpl.nasa.gov/>),
 - James Webb Space Telescope (JWST) (<http://www.jwst.nasa.gov/>).

Research proposed to this subtopic should demonstrate technical feasibility during Phase 1, in partnership with scientists, and subsequently show a path toward a Phase 2 prototype demonstration, with significant communication with missions and programs to ensure a successful Phase 3 infusion. Innovations are sought in data processing and analysis algorithms in the following areas:

- Optimization of Algorithms and Computational Methods that increase the utility of scientific research data, models, simulations, and visualizations. Of particular interest are innovative computational methods that will dramatically increase algorithm efficiency as well as the performance of scientific applications. Success will be measured by both speed improvements and output validation.
- Improvement of Data Collection, by identifying data gaps in real-time, and/or derive information through synthesis of data from multiple sources. The ultimate goal is to increase the value of data collected in terms of scientific discovery and application.
- Frameworks and Related Tools for Processing, Analyzing and Fusing image and vector data for the purpose of analyzing NASA's astrophysics, heliophysics, planetary and earth science mission data and therefore enable the advancement of NASA's scientific objectives. Of particular interest are open source frameworks that would enable sharing and validation of tools and algorithms.

Tools and products developed under this subtopic may be used for broad public dissemination or for use within a narrow scientific community. These tools can be plug-ins or enhancements to existing software or on-line data/computing services. They also can be new stand-alone applications or web services, provided that they are compatible with most widely used computer platforms and exchange information effectively (via standard protocols and file formats) with existing, standard or prevalent applications. To promote interoperability, tools shall use industry standard protocols, formats, and Application Programming Interfaces (APIs), including compliance with the Federal Geographic Data Committee (FDGC) and Open Geospatial Consortium (OGC) standards as appropriate.

It is highly desirable that the proposed projects lead to software that is infused into NASA programs and projects.

S6.04 Data Management - Storage, Mining and Visualization

Lead Center: GSFC

Participating Center(s): JPL, LaRC

This subtopic focuses on supporting science analysis through innovative approaches for managing and visualizing collections of science data which are extremely large, complicated, and highly distributed in a networked environment that encompasses large geographic areas. There are specific areas for which proposals are being sought:

- Collaborative visualization tools that enable data exploration, data sharing, and data manipulation among scientists worldwide that make use of innovative hardware and software technologies for data manipulation and display, including the use of large multi-touch input devices or 3 dimensional display devices.
- Social networking tools that enable secure high bandwidth scientific collaboration among scientists worldwide that promote the development of online communities for sharing thoughts and ideas and for arriving at consensus opinions and understanding.
- Tools for science data discovery, data mining, data search, and data subsetting in extremely large data sets in clustered processing and storage environments, cloud computing environments, or shared data and computation environments.
- Storage systems, file systems, and data management systems that promote the secure long term preservation of data in a distributed online storage environment, provide for recovery from system and user errors, and provide dynamically configurable high speed access to data shared over wide area high speed networks.

Research should be conducted to demonstrate technical feasibility during Phase 1 and show a path toward a Phase 2 hardware/software demonstration, and when possible, deliver a demonstration unit for functional and environmental testing at the completion of the Phase 2 contract.

S6.05 Software Engineering Tools for Scientific Models

Lead Center: GSFC

This subtopic seeks to improve the productivity and quality of NASA's scientific modeling endeavors through customized tools, which enable and encourage improved software engineering practices. Because many of NASA's principal scientific models have evolved over decades to be hundreds of thousands of lines long with contributions from a wide variety of scientists, much of the software has become "brittle" in the sense that it has become difficult to extend, couple, and optimize. In other software communities (and other programming languages), access to modern software tools has enabled large gains in productivity by providing high-level tools for isolating software defects (bugs) as well as by automating common, albeit tedious, software processes. The goal is to extend these capabilities to support the Fortran programming language so that NASA's scientific models can extract similar benefits.

Target Programs, Missions and Mission Classes

Advances in developer productivity would be of significant benefit to several research and analysis programs within the Science Mission Directorate including:

- High-End Computing Program (<http://hec.nasa.gov>)
- Modeling, Analysis, and Prediction Program (<http://map.nasa.gov>)

Technology Areas

The objective is to create a suite of software tools, which directly ameliorate the most significant bottlenecks to productivity in the development of scientific models:

- Tools that assist in the construction of fine-grained unit-level software tests based upon existing functionality in a legacy Fortran application. Although tests written by developers are desirable, such tests are exceedingly difficult to create for legacy numerical software. Suites of these tests could provide a significant element of risk-reduction for maintenance and extension of these models, and would be incorporated into some sort of unit-testing framework.
- Tools that enable high-level source code transformations ("refactorings"). Although refactoring support for other programming languages, most notably Java, has shown significant gains in productivity, similar support for Fortran is rather limited. (<http://www.eclipse.org/photran/>).
- Integration of a Fortran unit-testing frameworks within an Integrated Development Environment (IDE). Although multiple Fortran unit-testing frameworks have been developed (<http://sourceforge.net/projects/pfunit>), adoption by the community has been slow in part due to lack of integration within IDE's. Integration of other Fortran capabilities is also encouraged.

Tools and products developed under this subtopic may be used for broad public dissemination or for use within a narrow scientific community. These tools can be plug-ins or enhancements to existing software or on-line data/computing services. They also can be new stand-alone applications or web services, provided that they are compatible with most widely used computer platforms and exchange information effectively (via standard protocols and file formats) with existing, standard or prevalent applications. To promote interoperability, tools shall use industry standard protocols, formats, and APIs (Application Programming Interfaces).

It is highly desirable that the proposed projects lead to software that is infused into NASA programs and projects.

