NSF Assessment of Responses to the Joint NIH-DOE-NSF Request for Information (RFI) on Science Drivers Requiring Capable Exascale High Performance Computing

National Strategic Computing Initiative (NSCI)

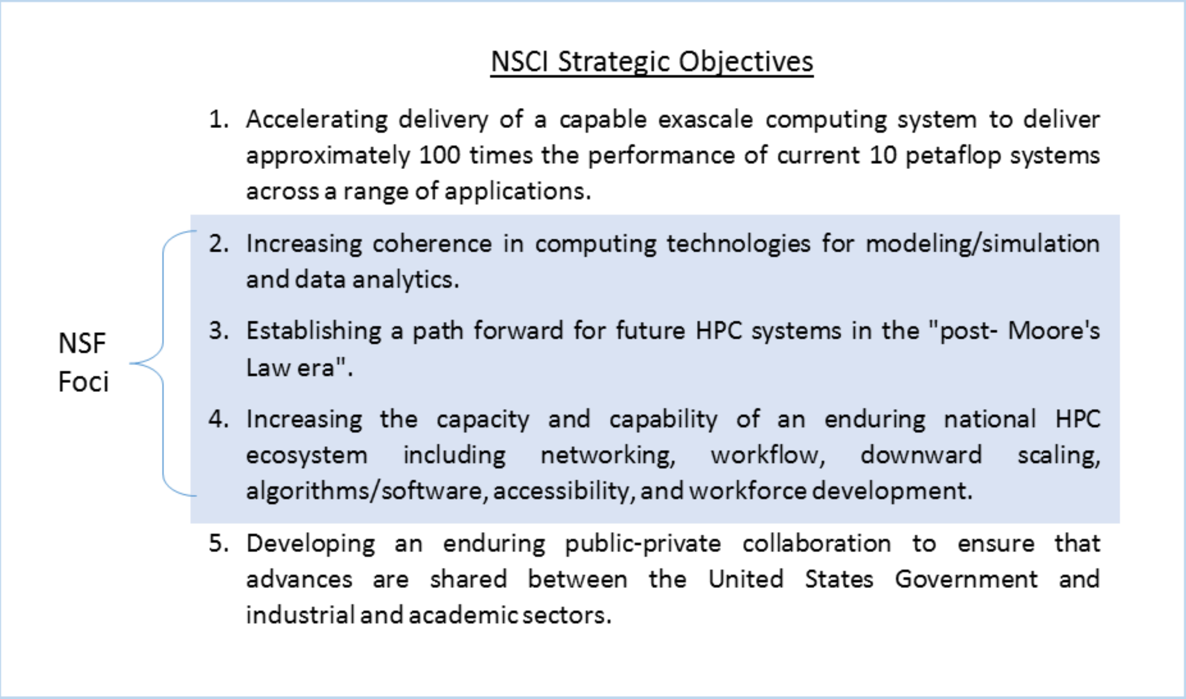
July 29, 2016

Contents

Background

The National Science Foundation (NSF) supports all fields of fundamental scientific and engineering research, including investing in advanced instrumentation and research cyberinfrastructure (CI) that enable discoveries in all those fields. Given the increasing scientific dependence on advanced computational infrastructure, and the breadth and complexity of evolving scientific need, NSF pursues a comprehensive and dynamic investment strategy in advanced computing and related ecosystem capabilities and services (i.e., *cyberinfrastructure*, comprising computational, data, and software infrastructure as well as networking, cybersecurity, and workforce development). This strategy is defined and coordinated at the Foundation level through crosscutting initiatives such as *Cyberinfrastructure for the 21st Century Science and Engineering (CIF21)*.

At the national level, NSF is a designated lead agency, together with the U.S. Department of Energy (DOE) and U.S. Department of Defense (DOD), in the National Strategic Computing Initiative (NSCI), which was announced by the White House on July 29, 2015 as a *whole-of-nation effort designed to create a cohesive, multi-agency strategic vision and Federal investment strategy, executed in collaboration with industry and academia, to maximize the benefits of High Performance Computing (HPC) for the United States*. NSF's responsibilities under NSCI are to invest in fundamental research that will drive and enable transformative advancements in methods and technologies across the spectrum of computational science needs, and to advance the capabilities of the HPC ecosystem for the science and engineering community, including education, learning and workforce development in NSCI areas. This responsibility is operationalized through designated focus on three of the five defined NSCI Strategic Objectives summarized below.



NSCI Strategic Objectives

1. Accelerating delivery of a capable exascale computing system to deliver approximately 100 times the performance of current 10 petaflop systems across a range of applications.
2. Increasing coherence in computing technologies for modeling/simulation and data analytics.
3. Establishing a path forward for future HPC systems in the "post- Moore's Law era".
4. Increasing the capacity and capability of an enduring national HPC ecosystem including networking, workflow, downward scaling, algorithms/software, accessibility, and workforce development.
5. Developing an enduring public-private collaboration to ensure that advances are shared between the United States Government and industrial and academic sectors.

NSF Foci

As disciplinary science and engineering requirements and cyberinfrastructure technologies continue to rapidly evolve, NSF regularly seeks advice and guidance on its cyberinfrastructure investments and future plans from the scientific community at many levels, including through the NSF Advisory Committee on Cyberinfrastructure (ACCI), which serves the whole Foundation; commissioned studies by high-level scientific bodies such as the National Academies of Science (NAS); and community-level workshops and requests for information. For example, the NAS recently completed and released a commissioned report entitled, "Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020"[1], which will help to inform the Foundation's strategy for future HPC investments.

Concurrent with the NAS study, and as a first effort to directly assess scientific requirements related to NSCI strategic objectives, NSF joined DOE and the National Institutes of Health (NIH) in releasing a joint "Request for Information (RFI) on Science Drivers Requiring Capable Exascale High Performance Computing"[2] (hereafter, "RFI") on September 15, 2015. The RFI requested input from the broad science and engineering research communities as well as HPC communities on research areas and applications that would require HPC capabilities extending 100 times beyond today's application performance in 10 years. The RFI also requested information on computational and technical parameters, current barriers, and other needed resources and capabilities across the HPC ecosystem such as applications development, data infrastructure, analysis and visualization tools, workflows and collaboration environments, networking, and training and workforce development.

In a Dear Colleague Letter (DCL)[3] issued on October 13, 2015, NSF encouraged the broad science and engineering community to respond to the joint RFI. The DCL emphasized NSF's particular interest in identifying research areas that would require capable computing resources offering 100-fold increase over today's application performance for large-scale numerically-intense analysis and also for deriving fundamental understanding from large-scale data such as image analysis, data integration, visualization and data analytics.

Over 100 responses to the joint RFI were received[4]. Per prior agreement, and as noted in the RFI, each of the three participating agencies is assessing and utilizing the RFI responses according to its respective missions and objectives. The three agencies are also currently working together to identify opportunities for interagency collaboration represented by the RFI responses and NSCI goals.

---

[1] http://www.nap.edu/catalog/21886/future-directions-for-nsf-advanced-computing-infrastructure-to-support-us-science-and-engineering-in-2017-2020.

[2] Original RFI: https://grants.nih.gov/grants/guide/notice-files/NOT-GM-15-122.html. An update with a revised response deadline: https://grants.nih.gov/grants/guide/notice-files/NOT-GM-15-123.html.

[3] NSF 16-008, Dear Colleague Letter: http://www.nsf.gov/pubs/2016/nsf16008/nsf16008.jsp.

[4] Separately, the DOE commissioned an internal survey of National Laboratories on science drivers for capable exascale computing; those responses are not addressed in the present report but are under ongoing joint analysis.

Context for NSF assessment of RFI responses

NSF undertook its analysis of the RFI responses with interest in uncovering patterns and emerging trends in the projected needs of the science and engineering community for advanced HPC resources and the broader HPC ecosystem (i.e., NSCI Objective 4), including needs for coherence in computational technologies for numerical modeling/simulation and data-intensive analysis and analytics (i.e., NSCI Objective 2). NSF was also interested in assessing the breadth of responding disciplines – i.e., including fields in which computational and data science at scale is well developed as well as those in which such activities are nascent and emerging. Given the great range of scientific and engineering research domains that NSF supports, identifying any gaps in disciplinary coverage in the responses was of interest. Finally, NSF sought to identify commonalities and differences in both scientific and technical "pressure points" across disciplines for scaling up computational and data science and engineering as systems capable of orders of magnitude more performance for applications become available.

Disciplinary coverage and gaps

The ensemble RFI responses represent a rich diversity of research fields. Table 1 shows the breakdown of RFI responses by primary research domain and associated disciplines. RFI responses spanned most NSF-supported scientific and engineering domains and many disciplines within those domains. The disciplines included those with established computational efforts (e.g., molecular and cell biology, physics, and earth and climate studies) as well as those in which computational efforts are just now emerging (e.g., neuroscience and materials science).

Table 1. Breakdown of RFI responses by domain and disciplines
(Coverage gaps are discussed below)

Primary Domain	Disciplines responding	Proportion of Total*
Life Sciences	Molecular and Cell Biology, Biophysics, Neuroscience, Omics, Systems Biology, Social Sciences	42%
Physical Sciences	Particle, Plasma, Space, and Planetary Physics; Materials Sciences	16%
Health Sciences	Imaging, Population Health, Precision Medicine	16%
Engineering	Aerospace, Biomedical, Chemical, Civil, Mechanical, and Energy Engineering	9%
Geosciences	Climate, Solid Earth, Subsurface	8%
Computer and Information Science	Data Science, Hardware, Software	7%

*Entries add to less than 100% due to round-off.

The largest portion of the received responses focused on topics in the life sciences, including an array of disciplines supported by both NSF and NIH; and in health sciences primarily funded by NIH but also of important relevance to NSF-supported fundamental research in a number of fields including computer and information science, engineering, and social, behavioral and economic sciences. The predominance of responses in the life and health sciences may be due in part to the fact that the RFI was hosted on NIH's website, but also likely reflects the rapidly expanding computational and data science efforts and needs in those fields.

However, *RFI response coverage was incomplete relative to NSF-sponsored disciplines with known or emerging computational efforts and needs*; for instance, responses in the established fields of computational chemistry and combustion science were absent, as were responses in the ecological, ocean, and polar sciences where computational needs are emerging. It is important that a full and thorough representation of disciplinary science drivers for advanced computing be available to inform future investment planning. Consequently, *NSF intends to provide opportunities and take other actions to ensure that its understanding of scientific and engineering needs for transformative advanced computational infrastructure is complete.*

Science and engineering frontiers represented in the RFI responses

The RFI responses represent a large spectrum of exciting science and engineering frontiers that would be enabled and advanced by deployment of significantly more capable computational resources. Highlights of the most ambitious science and engineering research opportunities identified in the RFI responses in NSF-relevant domains and disciplines are presented below.

Biological Sciences

Computational cell biology at scale: *Transform understanding of biological processes and whole systems by performing ambitious subcellular and whole-cell modeling, and coupling these with experimental data.* Ambitious applications include macromolecular simulations of systems, e.g., subcellular compartments, large organelles and entire living cells and viruses, and ensembles of living cells and tissues. Many responses highlighted the transformational value of extending biological simulations to realistic macroscopic time scales and cellular spatial scales for such simulations. Related goals are achieving detailed multi-physics, multi-scale simulations of biochemical processes; and molecular dynamics simulations linking genetics to biological processes in time and space. Simulation applications will require simultaneous solution of a billion or more differential equations and evaluation of complex interactions between neighboring elements.

Data-intensive biological imaging: *Achieve transformative discoveries concerning structure-function relationships across scales in biology and brain science through massive-scale image reconstruction and analysis.* Ambitious applications include high-throughput image analysis and reconstruction from electron microscopy images for structural biology, and from optical microscopy and magnetic resonance imaging for neuroscience; and large-scale data analysis of biological samples from light source instruments.

Computational genomics and related data science: *Achieve grand-scale, rapid health objectives through an advanced, tightly coupled computational infrastructure ecosystem*. Ambitious application include temporal dynamics of gene responses, sustained high-throughput whole genome sequencing (WGS) for precision medicine, simulation of the effects of drug candidates on virtual patients and populations, and high-throughput processing of sequencing data for human genome analysis and microbiome studies.

Engineering

Biomedical engineering: *Achieve dynamic engineering-relevant simulations of biophysically detailed models at the tissue-to-organ scale*. Ambitious applications include realistic hemodynamic modeling; mechanistic modeling of drug interactions at the cellular level; and optimization of the design of implantable biomedical devices through computational simulation.

Realistic virtual simulation environments. *Create a virtual wind tunnel environment for complex aerospace design via advanced computing, analysis and visualization capabilities*. Ambitious applications include high-fidelity simulations of large-scale systems with many interacting components, for instance, modeling the engine inlet flow of blended wing-body (BWB) aircraft.

Geosciences

Climate science: *Significantly enhance accuracy and predictive capabilities of climate models at local and global scales for scientific and societal benefit.* Ambitious applications include projecting changes in climate *from* and *on* human activity (e.g. ecosystem impacts, land use change, CO_2 and other atmospheric gases and particles, health, resource utilization, food production, security, transportation, infrastructure, communities, and ecosystems); developing routine forecasts of space and high-atmospheric weather and solar events; making predictions of climate change on decadal scales, and reconstruct paleoclimate; and understanding terrestrial and aquatic hydrodynamic and biochemical processes.

Earth system modeling: *Transform the study of Earth dynamics and increase predictive capabilities through realistic spatial and temporal scale simulations for scientific and societal benefit.* Ambitious applications include: Understanding the multi-scale convection within the Earth; developing a high-resolution multi-scale model of Earth's mantle over the full 4.6 billion year history; achieving accurate quantitative characterization of earthquakes/seismic hazards to inform policy and civic decision-making; and developing high-resolution hydrologic models to understand and predict climate effects on water resources, evaluate contamination mitigation strategies, and understand land surface dynamics in critical ecosystems at continental scales toward a predictive National Water Model.

Physics

High Energy and Nuclear physics: *Deepening understanding of the standard model and discover new physics beyond the standard model through advances in computational and data-intensive science.* Ambitious applications include expanding the capability to search for and model new particles and forces of nature and performing measurements of the fundamental laws of physics, with data from the Large Hadron Collider experiments ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) and from other accelerators; validating new theories against observations, such as new hadronic states, the emergence of light nuclei and the nuclear forces from Quantum Chromodynamics (QCD), and the transition from hadronic matter to a new form of matter – the quark-gluon plasma; and understanding spectroscopy and structure of Hadrons, nuclear forces and reactions, fundamental symmetries, and phases of strongly interacting matter.

Plasma and magnetic field physics: *Transform understanding of the cosmic origins of magnetic fields, their dynamics in astrophysical dynamos, and magnetic fields in engineered plasmas through computation.* Ambitious applications include understanding the nonlinear turbulent dynamo processes in astrophysical systems including the Earth-Sun system through holistic modeling of experiments on turbulent amplification of magnetic fields and physical mechanisms occurring deep within the interiors of stars and planets. Ultimately, predicting short- and long-term impacts of solar events on Earth's magnetic field.

Astrophysics and Planetary Physics: *Revolutionize understanding of cosmic evolution, lifecycle dynamics of astrophysical objects and ensembles, and resulting synthesis and evolution of the elements of life.* Advances in fundamental physics theory drive the ability to advance computational modeling in areas where direct observation of the physical phenomena are lacking or scant. Ambitious applications: Revolutionize understanding of cosmic synthesis and evolution of the elements of life, via simulating weakly interacting dark matter, core-collapse supernovae, binary neutron star mergers, and cosmological galaxy formation and evolution. Understand the creation of the elements beyond hydrogen and helium forged in stars, and expelled into space by their explosions. Establish how supermassive black holes grew and influence gas cooling, galaxy evolution and star formation. Follow the complete development of galaxies of different final masses, from dwarfs to giants, in different environments, from isolated galaxies to galaxy groups. Resolve uncertainties in the mechanisms of core collapse supernovae and thermonuclear supernovae.

Materials Science

Computational materials evaluation and design: *Establish the capability to predict and explore fundamental properties of materials via advanced computation to revolutionize materials design and control.* A plethora of ambitious modeling and simulation applications were identified, spanning many types of materials and scales: Achieve a fundamental understanding of nanoscale and molecular "building blocks" via simulations at length scales typical in experimental nanoscience with inclusion of quantal effects. Understanding excited state properties of advanced materials for design and control. Exploring materials phenomena ranging from dielectric behavior, electromagnetic field response across spectra; irradiation effects and behavior; and chemical reactions and catalysis. Predictive modeling of quantum-mechanical electron-ion systems including via quantum Monte Carlo approaches. Irradiation surface science and modification of interfaces.

A vast range of materials and engineering applications of interest were cited including candidate materials for quantum computing, cryptography, communications, and machines; interfacial materials within transistors, batteries, solar cells, and membranes for water desalination; next-generation nano-sensors and magnetic systems; processes for photo-catalysis and charge separation in solar cell materials; ultrafast electronics; electronic friction; surface chemical reactions; effects of high sub-atomic flux bombardment on astrophysical and man-made surfaces; radiation damage and radiation therapy; pump-probe spectroscopy; and high-temperature superconductors.

A diversity of emerging computational requirements

Responses across all research fields called for significantly more computational capacity to perform novel, longer or more complex tasks, but the nature of the stated needs varied significantly, reflecting a great diversity of requirements for HPC capabilities and resources. Responses from established numerics-oriented computational fields specifically cited the need for sustained application performance levels 100 times greater than today’s capabilities (i.e., NSCI Strategic Objective 1), with a number of such responses including detailed visions and architectures requirements based on current work with highly-developed, theory-based numerical modeling and simulation applications. Interestingly, as such simulations were anticipated to grow in complexity and scale, and be compared against large experimental datasets, many of these numerics-oriented responses also pointed to increasing data analysis needs and associated architectures – as well as burgeoning data management challenges.

Many other responses that called for more computational capabilities were not as specific as to amount or architecture. Such responses were often associated with disciplines whose computational components are just now emerging, and that are being challenged by data inundation rather than numerics-intensive computational needs. Indeed, it was clear from such responses that the emergence of a plethora of experimental measurement and imaging techniques that extend to the nano, atomic, and quantal scales have made these disciplines vastly more data-intensive in current and anticipated science, with computing needs firmly oriented towards analytics capabilities integrated tightly with data management. A good portion of the data-intensive responses were focused on societal benefits, e.g., materials design for biomedical and energy applications, real-time patient-specific diagnostics and omics, drug discovery, and biomedical imaging.

Several repeating themes emerged in the computational strategies and needs that respondents from different domains and disciplines cited to accomplish their ambitious science and engineering objectives:

1. *Numerical-intensive science: Enabling vast improvements in model/simulation spatial and temporal realism and associated predictive accuracy*. Approaches and associated disciplines included:
    - Multi-scale, multi-physics models of dynamic interacting natural phenomena incorporating more realistic physics and complex multi-modal coupling – e.g., geodynamics and climate modeling; planetary and cosmological phenomena particularly including magnetic field phenomena; and biochemical, molecular and genetic processes extending to new capabilities to study linkages among them.
    - Micro-structural models of physical systems incorporating realistic numbers of atoms, introducing quantum mechanical dynamics, and computing over realistic time-periods – e.g., instantiating realistic materials science and design, and extending biophysical and molecular modeling to macromolecular simulations of systems (whole cells, cell ensembles and tissues).
    - Virtual engineering simulation environments, for instance, virtual wind tunnels enabling high-fidelity simulations of large-scale aerospace systems with interacting components.

2. *Data-intensive science: Enabling analysis and analytics of massive multi-source, multi-scale data and returning results in near real-time.* Approaches and associated disciplines included:
    - High-throughput image analysis, three-dimensional reconstruction and associated analysis of electron and optical microscopy, and high-energy light source instruments for structural biology, brain connectivity mapping, and omics, at unprecedented scales.
    - Massive concurrent comparison of an individual test element (e.g., test molecule, individual behavioral or genetic profile) against large-scale reference data (molecules, population data) for ligand binding studies, smart and connected health, and precision medicine applications.
    - Big data analysis and analytics to enable high-luminosity particle physics, statistical large-scale population and behavioral studies, and materials science, as well as fundamental data science research.

3. *Data-model inter-comparison at scale, and with rapid response: Enabling the comparison and validation of large-scale realistic models with concomitantly scaled data.* Nearly all the respondent domains and disciplines included desire to compare large-scale experimental data and computational models in transformative ways, and in certain cases with the desire of short turn-around times – for instance, for event detection/identification and alerting in astronomical sciences and earth sciences/hazard prediction and response. The emergence of data-model inter-comparison goals reflected in the RFI responses implicates *the need to develop ways of connecting and harmonizing modeling/simulation and data analytics across this spectrum of resource types* (i.e., NSCI Strategic Objective 2).

An important conclusion from the above analysis is that *a spectrum of HPC capabilities, ecosystem resources and services will be needed* to support myriad and highly diverse anticipated computational efforts across all of fundamental science and engineering research. Indeed, NSF's investment strategy in computational infrastructure has been necessarily responsive to increasing scientific pressure to broaden the spectrum of available computational resources. NSF's portfolio of advanced computational resources includes general-purpose and leadership systems focused on simulation/model-oriented applications (e.g., Stampede and Blue Waters, respectively), distributed high-throughput computing (Open Science Grid), and new types of capabilities comprising recent deployments of HPC resources that focus on data-intensive science (Wrangler), cloud computing (Jetstream), long-tail science (Comet) and high-memory applications (Bridges).

To better define the anticipated needs across this HPC spectrum, further community exploration of the three above computational strategy themes might be useful towards development of *transdisciplinary computational science benefit scenarios* that could inform future advanced computational infrastructure programs. Research into device technologies and computer architectures will be paramount.

Priority needs across the HPC Ecosystem

In addition to computational resource needs, the following areas stood out prominently across the RFI responses – and indeed transcended research domains – as foundational needs for ecosystem resources, services and developments to achieve the benefits of capable computing delivering a one-hundred-fold increase in application performance.

Applications and Software Infrastructure. The overriding common concern voiced by respondents to the RFI was that substantial effort in both the science and engineering domains and software infrastructure fields will be required to capitalize on anticipated increases in performance/capabilities. Two areas were particularly emphasized:

- Finding ways of scaling existing applications and gaining the benefits of accelerators. This was uniformly cited as an issue for multi-physics, multi-scale simulations such as in earth system, climate, and cosmology/astrophysics modeling – where the currently programmed strategies, for instance in coupling the different physics model components, may need complete rethinking and re-envisioning to allow the overall simulations to function accurately when scaled by two orders of magnitude in performance. While there are ongoing improvements, the major scale-ups and addition of features (e.g., non-linear rather than linear coupling, moving from two- to three dimensions) will require new modeling strategies and theory, along with end-to-end validation.
- Model Validation and Uncertainty Quantification. This was mentioned as a major concern in areas as diverse as health science (precision medicine), biology (neuroscience, imaging), physics, engineering, and materials science. As more complexity is added to models and analyses, both the validation of models and results as well as the quantification of uncertainty grow in difficulty.

Learning and Workforce Development (LWD). Along with applications development, LWD was the most strongly voiced area, and in some cases considered as even more of a foundational bottleneck issue that, if not adequately addressed, would prevent realization of scientific benefits of the computational advancements embodied in the NSCI objectives. Across science and engineering domains, the responses represented a clear call for *a well-trained expert workforce with computational knowledge and expertise*, capable of developing HPC applications, using and maintaining the advanced computing framework, and appreciating the importance of these skills. An oft-cited concept was that of a new scientific workforce with hybrid skills, i.e., science and engineering domain expertise, computer science knowledge, and skills in highly parallel systems. Stable career paths for these experts was also highlighted as critical, as was the need to develop multidisciplinary collaborations that spanned both computational expertise and domain science expertise. Also cited repeatedly as critical was increased training for students and researchers in using HPC; computational and data science; error analysis and reproducibility; techniques and principles of application development; and data science and workflow best practices.

Information Infrastructure. Dealing with data – whether generated from simulations, performing analysis of big datasets, or comparing models to data – was of great concern across respondents. As mentioned earlier, anticipated extreme data computational and management issues were particularly emphasized in biological "omics", health, and social sciences**.** Even within a strictly modeling/simulation paradigm, respondents cited need for fast access to recently computed results to further advance a given simulation. These challenges portend requirements on sufficient memory and input/output (I/O) footprints in future computational architectures, but also for advanced data analysis and visualization capabilities, data storage, transfer, security/protection, and data access from multiple sources. New comprehensive workflow systems will be required for data management, curation and sharing, and for automated pre-processing and post-computational analysis and visualization tools. These requirements together additionally implicate expanded requirements for a ubiquitous, highly performant and trustworthy research data network, with adequate speed and capacity to support low-latency, large data-driven applications. Fundamental advances in data science will also be critical.

Facilitating broader utilization at all scales. Many respondents across scientific disciplines mentioned the need for improved systems and processes that ease access to, and preparation for, utilization of advanced computational resources. Suggested solutions included creation of seamless end-to-end workflow systems, simplified means of obtaining allocations and access to systems and data, easier programming environments for HPC, and dissemination of best practices. From these many such requests, it became clear that workforce development, particularly training, as discussed above, will be paramount to reach and support new as well as current users as the whole computational ecosystem continues to develop.

Moreover, a number of respondents pointed out that the NSCI goal of significantly increasing HPC application performance must also take place across the whole ecosystem and at all levels of the computational science chain down to even desktop computer performance, so that benefits of advanced computation become fully embedded into the practice of scientific and engineering research.

Conclusions

Responses to the joint agency RFI comprise an exciting range of ambitious, anticipated transformative advancements in scientific and societal impacts offered by increasing HPC application performance by several orders of magnitude. The RFI responses demonstrate that computational science is both deepening within established computational fields and rapidly expanding into new domains; and that data-intensive research is rapidly emerging to complement modeling and simulation as combined pressures on the design of future HPC capabilities and resources.

NSF's overriding assessment from the responses is that *a comprehensive, inclusive and holistic investment approach will be required to support the expanding scientific and engineering horizons in all research domains*. Broad investments will continue to be necessary to explore and deploy a rich spectrum of HPC capabilities, architectures and technologies, and to advance the full HPC ecosystem of software, data, and information infrastructure, along with human resources, to achieve NSCI goals:

- Ambitious, highly diverse plans in numerical- and data-intensive science, and comparing large-scale data with large-scale models will require the investments foreseen by NSCI in convergence of architectures and technologies used for modeling/simulation and data integration, analysis, and analytics;
- Significant concomitant investments will be required in software, application development and portability, validation, and quantifying uncertainty in increasingly complex multimodal and multi-spatial analyses; and
- Also crucial will be development of a national cadre of computational technologists and HPC experts with stable career pathways; training the next generation of computationally-capable, multidisciplinary researchers; and facilitating access to HPC by an increasingly broad base of savvy users across large-scale and long-tail science.

Finally, NSF notes that additional community input will be needed to address identified gaps in the disciplinary coverage of the RFI responses, and to refine both within-discipline and transdisciplinary computational and data science needs for HPC ecosystem capabilities, resources and associated services. Consequently, *NSF intends to provide additional opportunities to ensure its understanding of scientific and engineering needs for transformative advanced computational infrastructure is complete.* NSF will use the information gained from this assessment together with the recommendations of the recent NAS report described above as well as future contributions from the community in planning the Foundation's HPC investments to enable future scientific and engineering research and discovery.