NATIONAL STRATEGIC COMPUTING INITIATIVE UPDATE: PIONEERING THE FUTURE OF COMPUTING

A report by the

FAST TRACK ACTION COMMITTEE ON STRATEGIC COMPUTING

NETWORKING & INFORMATION TECHNOLOGY
RESEARCH & DEVELOPMENT SUBCOMMITTEE

COMMITTEE ON SCIENCE & TECHNOLOGY ENTERPRISE

of the

NATIONAL SCIENCE & TECHNOLOGY COUNCIL

NOVEMBER 2019

About the National Science and Technology Council

The National Science and Technology Council (NSTC) is the principal means by which the Executive Branch coordinates science and technology policy across the diverse entities that make up the Federal research and development enterprise. A primary objective of the NSTC is to ensure that science and technology policy decisions and programs are consistent with the President's stated goals. The NSTC prepares research and development strategies that are coordinated across Federal agencies aimed at accomplishing multiple national goals. The work of the NSTC is organized under committees that oversee subcommittees and working groups focused on different aspects of science and technology. More information is available at https://www.whitehouse.gov/ostp/nstc.

About the Office of Science and Technology Policy

The Office of Science and Technology Policy (OSTP) was established by the National Science and Technology Policy, Organization, and Priorities Act of 1976 to provide the President and others within the Executive Office of the President with advice on the scientific, engineering, and technological aspects of the economy, national security, homeland security, health, foreign relations, the environment, and the technological recovery and use of resources, among other topics. OSTP leads interagency science and technology policy coordination efforts, assists the Office of Management and Budget with an annual review and analysis of Federal research and development in budgets, and serves as a source of scientific and technological analysis and judgment for the President with respect to major policies, plans, and programs of the Federal Government. More information is available at https://www.whitehouse.gov/ostp.

About the Networking and Information Technology Research and Development Program

The Networking and Information Technology Research and Development (NITRD) Program is the Nation's primary source of federally funded coordination of pioneering information technology (IT) research and development (R&D) in computing, networking, and software. The multiagency NITRD Program, guided by the NITRD Subcommittee of the NSTC Committee on Science and Technology Enterprise, seeks to provide the R&D foundations for ensuring continued U.S. technological leadership and meeting the needs of the Nation for advanced IT. The National Coordination Office (NCO) supports the NITRD Subcommittee and the Interagency Working Groups (IWGs) that report to it. More information is available at https://www.nitrd.gov/about.

About the Fast Track Action Committee on Strategic Computing

The NITRD Subcommittee's Fast Track Action Committee (FTAC) on Strategic Computing was created to update and disseminate the objectives and approaches necessary for sustaining and enhancing U.S. scientific, technological, and economic leadership in strategic computing research, development, and deployment. More information is available at https://www.nitrd.gov/news/Updating-US-Strategic-Computing-Objectives.aspx.

Acknowledgments

This document is the result of contributions from many Federal agency representatives, members of the FTAC on Strategic Computing, the staff of the NITRD NCO, public responses to a Request for Information on Update to Strategic Computing Objectives, and observations from the Future Computing Community of Interest meeting that took place August 5-6, 2019. The FTAC on Strategic Computing gratefully acknowledges all individuals who contributed to the content of this National Strategic Computing Initiative Update.

Copyright Information

NATIONAL SCIENCE & TECHNOLOGY COUNCIL

Chair

Kelvin Droegemeier, Director, OSTP

Executive Director

Chloé Kontos, NSTC

FAST TRACK ACTION COMMITTEE ON STRATEGIC COMPUTING

Co-Chairs

Jake Taylor, OSTP
Barry Schneider, NIST
Bill Vanderlinde, DOE
Manish Parashar, NSF

Executive Secretary

Ji Lee, NITRD

Members

Erin Cheese, OMB
Stephen Dennis, DHS
Jeff Falgout, USGS
Susan Gregurick, NIH
Bill Harrod, ODNI
Thuc Hoang, DOE
Frank Indiviglio, NOAA
Tsengdar Lee, NASA
Richard Linderman, DOD
David Mountain, NSA
Andreas Olofsson, DARPA
Jesse Taylor, FBI

Table of Contents

Abbreviations

DARPA	Defense Advanced Research Projects Agency
DOD	Department of Defense
DOE	Department of Energy
ED	Department of Education
DHS	Department of Homeland Security
FBI	Federal Bureau of Investigation
FTAC	Fast Track Action Committee on Strategic Computing
HEC	high-end computing
NASA	National Aeronautics and Space Administration
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology
NITRD	Networking and Information Technology Research and Development (NSTC Subcommittee and Program)
NOAA	National Oceanic and Atmospheric Administration
NSA	National Security Agency
NSF	National Science Foundation
NSTC	National Science and Technology Council
ODNI	Office of the Director of National Intelligence
OSTP	Office of Science and Technology Policy
R&D	research and development
RFI	Request for Information
USGS	United States Geological Survey

Executive Summary

The national computing landscape is undergoing rapid evolution along multiple dimensions due to the introduction of new and potentially disruptive technologies and the demands of new classes of data-intensive applications. Computer architectures and systems are more heterogeneous and complex, and the challenges associated with the complexity and sustainability of software are significant. Application developers are being called to meet demands for improved usability and productivity. Therefore, to address this new environment, it is essential to update the 2016 U.S. Strategic Computing Plan[1] in a holistic and synergistic fashion that involves government, academia, and industry. The Office of Science and Technology Policy and the National Science and Technology Council Subcommittee on Networking and Information Technology Research and Development created a Fast Track Action Committee on Strategic Computing in June 2019 to address this imperative.

The National Strategic Computing Initiative Update builds on the 2016 Strategic Computing Plan to enable focused and collaborative research, development, and deployment of future computing technologies to benefit the Nation. Realizing the successes of the National Strategic Computing Initiative to date, this update examines steps beyond those already implemented for realizing a computing ecosystem that combines heterogeneous computing systems (from extreme-scale to edge-centered systems and beyond) with the networking, hardware, software, data, and expertise required to support national security and defense as well as U.S. scientific, engineering, and economic leadership.

The refocused objectives identified are:

1. Pioneer new frontiers of digital and non-digital computation to address the scientific and technological challenges and opportunities of the 21st century.
2. Develop, broaden, and advance the Nation's computational infrastructure and ecosystem.
3. Forge and expand partnerships for the future of computing to ensure American leadership in science, technology, and innovation.

What follows includes key goals for a strategic approach to the future of computing, and a set of recommendations to achieve those goals, which will expand our scientific knowledge, enable economic growth, and improve our national security. These recommendations and goals were determined via engagement with government, academia, and industry through various means, including a Request for Information (RFI)[2] and a Future Computing Community of Interest Meeting.[3]

[1] https://www.whitehouse.gov/sites/whitehouse.gov/files/images/NSCI%20Strategic%20Plan_20160721.pdf.pdf

[2] https://www.federalregister.gov/documents/2019/06/18/2019-12866/request-for-information-on-update-to-strategic-computing-objectives

[3] https://www.nitrd.gov/nitrdgroups/index.php?title=FC-COI-2019

1. Enabling the Future of Computing

Objective: Pioneer new frontiers of digital and nondigital computation to address the scientific and technological challenges and opportunities of the 21st century.

The computational landscape is rapidly evolving as a consequence of the slowing progression of technology advances within Moore's law, the end of Dennard scaling, and the unsustainable power requirements of current technology systems. Significant disruptions from emerging technologies and paradigms can be found at all levels of the system, from hardware devices to system architectures and software stacks. The growing complexity of algorithms and programming models (designed to overcome these disruptions and integrate them into existing applications to enhance performance) exacerbates the challenge. At the same time, application workflows are evolving with new requirements that necessitate the integration of heterogeneous platforms, including those within a given architecture as well as network-centric and edge computing.

> *As the Association for Computing Machinery A.M. Turing Award laureates John Hennessy and David Patterson wrote, "The next decade will see a Cambrian explosion of novel computer architectures, meaning exciting times for computer architects in academia and industry."* [4]

Although exciting, these technological advances have far-ranging consequences that must be considered when reexamining the U.S. Strategic Computing Plan.

Future Computing Hardware Considerations and Expectations

Recent cost performance dynamics are resulting in leading-edge computing hardware being "commoditized" and ubiquitous, both in its access and use, across much of the science and engineering enterprise. This democratization facilitates tackling many scientifically challenging problems, for example, through the emergence of cloud service providers. Future computing systems will be highly multicore and heterogeneous within a single node or cabinet, and will increasingly involve explorations into heterogeneous processors, heterogeneous memories and models, new interconnect technologies, special-purpose and energy-efficient architectures, and a number of non-von Neumann computing elements such as those based on neuromorphic and quantum technologies.

> **NVIDIA response to the RFI:** *"Looking forward, the practical realization of slowing technology scaling will likely require a range of approaches including: (1) architectures that incorporate increasingly specialized accelerator hardware; (2) packaging, signaling, and interconnect technologies that enable greater scaling at both the "node" and "system" level; (3) novel devices (e.g., carbon nanotube FETs) that can provide smaller digital devices at lower power; and (4) novel computing technologies such as analog, quantum, and neuromorphic that may require fundamental changes to algorithms."*

[4] https://cacm.acm.org/magazines/2019/2/234352-a-new-golden-age-for-computer-architecture/fulltext

Software Considerations and Challenges

Software supporting the future computing ecosystem must balance the following: efficiency of development, debugging, verification, and validation; usability, reproducibility, manageability, extensibility, and sustainability; and performance and scalability. Software must be capable of operation in multiple modes and high degrees of parallelism, with efficient management of memory and input/output (I/O), while also supporting workflow composability and execution. Emerging future computing technologies present new opportunities (e.g., real-time response, embedded processing of data from sensors or actuators, and steering of models and experiments), but will also require new algorithms, models of computation, data, programming environments, and software stacks. This transformation will require considerable support and a radical rethinking of the software infrastructure used in standard programming models. Software development, including migrating (often legacy) applications to new systems, continues to be a challenge and prevents the community from taking advantage of existing and emerging processing capabilities. It is likely, due to the growing complexity of applications and hardware environments, that some future software development and testing will lie well beyond human timescale development and thus be performed by software itself.

AMD response to the RFI: "*It is possible to abstract and simplify much of the complexity so that programmers can utilize simpler models of a system and increase their productivity. Useful software abstractions improve developer productivity and reduce execution risk by muting the cognitive noise produced by complexity.*"

Recommendations

- To enable effective use of the Nation's computing ecosystem from the edge to exascale and beyond, the U.S. Government can:
    - Embrace a diversity of hardware and software approaches for the future of computing and leverage the innovation ecosystem to: stay at the forefront of computing, increase focus on usability and productivity, lower barriers to use for research and applications, and support the integration of edge resources and data with traditional computing platforms, including emerging data-driven applications;
    - Provide early access to novel hardware, software, and systems platforms that are needed to identify and support promising research approaches and to reduce time to deployment;
    - Identify and prioritize the software research needed for the future of computing;
    - Encourage development, deployment, and sustainment of software tools, frameworks, and systems; and
    - Engage industry, academia, and U.S. Government laboratories to coordinate software development and sustainability efforts via consortia and other forms of partnerships.
- To enable the development of end-to-end application workflows and integrated systems to tackle pressing science, engineering, and national security challenges, the U.S. Government should:
    - Foster a fully engaged future computing community by working with stakeholders from across the Nation, poised to ensure that upcoming hardware and software developments are leveraged in a timely manner to support applications;
    - Encourage novel solutions that leverage in-network and edge-processing capabilities to process data close to the source as part of end-to-end application workflows;

  - Encourage multidisciplinary teams of application specialists, end-users, developers, and researchers to enable new, integrated solutions with appropriate security considerations, to address pressing computational challenges and to widen the user base; and
  - Promote timely access for developers of technologies, architectures, and systems to carry out the research needed to create the future computing software ecosystem. Developers should also have a mechanism to provide early input into discussions about upcoming hardware developments to ensure design decisions help users take advantage of new hardware more rapidly.
- To explore critical fundamental scientific and technical limits to computing in order to maximize the benefits of novel computational hardware, software, architectures, and new computing paradigms on applications, and to enable translation of this research to deployed technologies, the U.S. Government can:
  - Support continuous, long-term research and development (R&D) in the basic science and technology of computation to ensure American leadership in computing for the coming decades;
  - Support the rapid translation to practice of basic R&D and technology to address scientific challenges requiring the effective integration of advanced software and hardware; and
  - Support the integration and interoperability of application software via the development and refinement of scientific gateways, portals, and associated workflow tools to enable more efficient approaches to solving challenging scientific and technical problems.
- Leverage the diversity of research opportunities and systems emerging from R&D on computation to enable prosperity, ensure national security, and provide a stronger technological and scientific foundation for the Nation.

2. Providing a Strategic Foundation for Computing

Objective: Develop, broaden, and advance the computational infrastructure and ecosystem.

The confluence of disruptive architectural and technological trends across the evolving future computing landscape, and emerging trends in application workflows that increasingly integrate simulation, traditional data analytics, and machine learning models, demand an agile, robust, secure, useable, capable, and sustainable computational ecosystem. Such an ecosystem must integrate emerging and future hardware platforms with the necessary software, data, and networking expertise.

Next-Generation Preparations: Infrastructure for Hardware and Software

The emerging highly multicore, heterogeneous, and energy-efficient architectures promise a revolution in computing capabilities to solve challenging scientific and technological problems. It is also critical to explore new software approaches that will enable us to effectively exploit this hardware to advance science. Continuing this progress requires a sustained investment into the exploration of new materials, devices, paradigms, technologies, and infrastructure in the form of foundries, testbeds, experimental systems prototypes, and the supply chain. Services required to support the use of these new computing paradigms also require consideration.

> **Computing Community Consortium response to the RFI:** "*Computational reproducibility is another area that deserves expanded efforts. Computational reproducibility is a requirement for science to progress, and currently, we lack the necessary tools and infrastructure to make computations reproducible.*"

Data Considerations and Challenges

Data production from next-generation instrumentation as well as "long tail" data (smaller in size but large in number and heterogeneity) continues to increase exponentially, creating a demand for a distributed but federated network of data preservation systems. Solutions to many global challenges require interdisciplinary data integration based on agreed-upon standard interfaces for data discovery, access, compatibility, and reusability. International organizations are advancing these capabilities, and U.S. and European policies exist to make publicly funded research data openly available for technologies such as artificial intelligence and machine learning applications that need access to large bodies of carefully curated, interoperable, and validated training data.

> **IBM response to the RFI:** "*Focus on development of capabilities that enable important applications rather than on building even larger peak performance systems.*"

Cybersecurity: Issues, Challenges, and Opportunities

Effective security ensures that computer systems are available for use, resources are not misused, data are valid and trustworthy, and sensitive information is protected. Key challenges in this area include: the impact of cybersecurity on system performance, a diverse user base that often spans multiple networks and organizations, and the lack of clear security policies at the management level. However, the opportunity to protect strategic data assets (e.g., intellectual property, economic competitiveness, national security, and personal privacy) and ensure data integrity is critical for the adoption and success of future computing.

Conversely, cybersecurity functions can be accelerated and improved using future computing technologies. The volume, velocity, variety, and veracity of data within information pipelines combined with the need to rapidly ingest, index, assess, and query that data demands future computing systems that have real-time cybersecurity functionality. Opportunities exist for combining innovative analytics and artificial intelligence techniques to address problems such as network event detection, user behavior analytics, and network mapping. Near-real-time approaches, enabled by advanced computing capabilities and high-bandwidth, low-latency networks, allow defenders to understand and minimize the threat of an ongoing security breach. Advanced computing resources can be used to model the cyber landscape and simulate attack events, allowing defenders to minimize attack vectors.

Cyberinfrastructure Services: Opportunities and Challenges

Cyberinfrastructure coordination services will be integral to the future computing ecosystem and applications. Users depend on these services to ensure productivity, reduce time-to-solution, and:

- Discover, provision, and access available resources and services;
- Receive relevant and timely expert technical guidance and focused instruction in efficient and effective use of available resources and services; and
- Engage in developmental efforts to ensure the effective use of the evolving computing landscape.

Equally important are services to ensure effective management, operation, and use of computing resources, and integrating these services into a coherent, coordinated national cyberinfrastructure ecosystem. A modern computing ecosystem must securely increase user accessibility (including web-based and mobile platforms); enable collaboration; simplify the use of high-end computing (HEC) in dynamic, system-of-systems scenarios; allow persistent access to relevant data; and offer advanced, high-performance service delivery mechanisms such as secure containers and virtual machines. By lowering the barriers to high-productivity computing, the ecosystem can deliver future computing capabilities to a broader workforce, allowing workforce development to focus on innovation and increasing diversity of thought.

Workforce Considerations and Challenges

Leveraging new computing capabilities and turning them into practical and usable forms for applications is an important and challenging issue. It requires a skilled workforce, trained in the current state of the art, that is able to anticipate and exploit future technologies and solutions. This involves a wide range of training and skills, from cable laying to power and heat management to research and development to marketing and communication. Developing and sustaining a strong, diverse workforce involves not only training within educational institutions or on the job, but refreshing skills throughout a career as technologies, platforms, and applications evolve. Developing the necessary tools for workforce training, re-skilling, productivity, and collaboration is equally important. Finally, establishing synergies across government, academic, and industry stakeholders and developing creative incentives and rewards mechanisms are essential to sustain this workforce.

Micron response to the RFI: "*The number of application programmers has remained relatively constant. This is a challenge. Promotion and funding for STEM- related careers need to go into overdrive.*"

Recommendations

- To provide a strong hardware and software foundation, the U.S. Government should:
    - Ensure that investments in infrastructure such as foundries, testbeds, experimental systems, and prototypes, as well as in relevant domains such as materials science, microwave engineering, and supply chains, are made to enable the future of computation;
    - Support critical cyberinfrastructure services including discovery, allocation, provisioning, user support, and monitoring and management of the computational ecosystem;
    - Prioritize the development of a robust software ecosystem composed of shared, sustainable software stacks, libraries, frameworks, and services that are essential for accelerating access to and use of early-stage systems;
    - Support improved usability and productivity by lowering barriers to broaden HEC use and application areas; and
    - Encourage stakeholders to develop common interfaces, align solutions, and establish shared best practices and standards for the future of computing.
- To prioritize improved cybersecurity, the U.S. Government can:
    - Recognize and emphasize the importance of cybersecurity for the computing ecosystem and of using advanced computing for enhancing and accelerating cybersecurity;
    - Encourage community members to work together to raise cybersecurity awareness, provide tools to assess and evaluate cybersecurity, establish effective best practices, and develop cybersecurity control baselines;
    - Develop near-real-time approaches (enabled by HEC and high-bandwidth, low-latency networks) to understand and minimize threats and to model and simulate the cyber landscape; and
    - Prioritize achieving cybersecurity situational awareness to provide early identification of security challenges, provide the technical foundation for a more resilient cyber defense for computing, and bolster computing-enhanced cybersecurity to make systems significantly more secure and easier to use.
- To support data usage and management for computing, the U.S. Government should:
    - Develop strategies for the management, access, and curation of data sets for research and applications;
    - Support end-to-end data stewardship to increase the efficiency of scientific workflows and ensure the dissemination of research results, including data; and
    - Develop common interfaces, knowledge networks, tools, and services for data discovery, access, transport, and processing, and timely or real-time data flows.
- To recognize that the entire strategic approach to the future of computing is dependent on a capable and flexible workforce, the U.S. Government should work both directly and with stakeholders to:
    - Create a diverse workforce necessary to achieve the goals of the U.S. Strategic Computing Plan and to support the broader U.S. innovation ecosystem at the leading edge of computation.

3. Ensuring Collaborative and Coordinated Approaches

Objective: Forge and expand partnerships for the future of computing to ensure American leadership in science, technology, and innovation.

Partnerships across Industry, Academia, and U.S. Federal Agencies

The computing ecosystem is extremely broad and diverse, and Federal agencies, academia, non-profit organizations, and industry address and optimize different, often complementary aspects of this ecosystem that align with their individual missions and priorities. As a result, cross-sector partnerships can lead to efficiencies and synergies that can significantly benefit all stakeholders; such partnerships should be encouraged and actively facilitated. Long-term Federal agency engagement with industry and academia is necessary to explore, develop, and produce technologies for mission-specific support; these partnerships are also an important vehicle for recruitment and hiring. Finally, there exist several unexplored opportunities for closer collaboration among Federal agencies to increase the productivity and efficiency of the user base.

The Defense Advanced Research Projects Agency (DARPA) Joint University Microelectronics Program (JUMP) is an example of a partnership between the nonprofit Semiconductor Research Corporation and DARPA at a half-dozen university-based research centers. Project 38, a vendor-agnostic collaboration between the Department of Energy and the National Security Agency, is another example of Federal agencies exploring new architectures based on a limited set of applications of interest.

Federal Coordination

Recognizing that the future of computing exists in a competitive global landscape, coordination of R&D in the future of computing should ensure resources are available and efficiently used for both individual agency efforts and to enable cross-agency opportunities and collaborations. At the same time, coordination activities should provide an appropriate venue for unclassified and classified discussions of both R&D and strategic considerations for the Nation's approach to the future of computing. Adopting a tiered structure that leverages existing working groups and expertise at NITRD and other interagency bodies—along with appropriate decision makers in a new NSTC subcommittee that focuses on the strategic implications of computing—can provide the necessary multifaceted coordination.

Recommendations

- To encourage partnerships and sustained engagement, the U.S. government should:
  - Foster Federal agency, academic, and industry collaborations that are both deep and broad to leverage investments, realize synergies, and enable next- and next-next-generation technology. Engagements with like-minded international partners should be included where appropriate;
  - Provide multiyear mechanisms to explore, develop, and potentially deploy new technologies; and
  - Encourage private sector research, development, and coordination, and work to integrate new approaches and technologies generated through such activities into agency efforts.

- To ensure effective coordination, the U.S. Government should:
  - Implement an interagency governance structure focused on major future computing initiatives across agencies, including:
    - An Executive Council comprising senior management of the member agencies. Consistent with individual agency missions, the Council can identify priorities and support future computing objectives;
    - A new NSTC subcommittee to ensure cross-agency coordination that includes defining strategic implications for the future of computing at the classified level, in alignment with the NSTC's NITRD Subcommittee and other subcommittees as appropriate, to support the Council by producing a yearly report on progress towards objectives and performing other duties as assigned by the Council; and
    - Existing and ad hoc working groups or other entities (as identified by the subcommittee) to be housed within NITRD and other NSTC bodies and reporting up to the cognizant bodies, as needed, to support the subcommittee's efforts.
  - Track global efforts in future computing.
  - Align U.S. future computing initiatives with other major national initiatives.

THE WHITE HOUSE