$7.3 billion FY 2015 appropriation

94% funds research, education and related activities

48,100 proposals

11,000 awards funded

1,826 NSF-funded institutions

320,900 NSF-supported researchers

214 Nobel Prize winners

All S&E disciplines funded

Funds research into STEM education

Other than the FY 2015 appropriation, numbers shown are based on FY 2014 activities
NSF Support of Academic Basic Research in Selected Fields
(as a percentage of total federal support)

<table>
<thead>
<tr>
<th>Field</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Science and Engineering Fields</td>
<td>24%</td>
</tr>
<tr>
<td>Engineering</td>
<td>40%</td>
</tr>
<tr>
<td>Physical Sciences</td>
<td>44%</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>55%</td>
</tr>
<tr>
<td>Mathematics</td>
<td>59%</td>
</tr>
<tr>
<td>Environmental Sciences</td>
<td>60%</td>
</tr>
<tr>
<td>Biology</td>
<td>66%</td>
</tr>
<tr>
<td>Computer Science</td>
<td>87%</td>
</tr>
</tbody>
</table>

Note: Biology includes Biological Sciences and Environmental Biology; excludes National Institutes of Health.
The Conduct and Practice of Science are Changing

• Increased complexity inherent in research questions
  – Multidisciplinary approaches and expertise
  – Collaborations – geographically distributed

• Rapid advances in technology and deployment
  – Instrumentation at all scales (large to small) producing data
  – Research Cyberinfrastructure

• Unprecedented growth in data (observation & simulated)
  – Enrich and advance research and society
  – Open new areas and approaches of investigation

• Greater societal expectations
  – Access and credibility of research results

• Global Investments in Research
  – Flat US research budgets
  – More opportunities to collaboration

• Demands for greater technical skills of workforce
NSF embraces an expansive view of cyberinfrastructure motivated by research priorities and the scientific process.

**Cyberinfrastructure Ecosystem**

- Scientific Instruments
- Computational Resources
- Data
- People, organizations, & communities
- Networking & Cybersecurity
- Software

**Scientific Discovery & Innovation**

1. **Observe**
2. **Theorize**
3. **Hypothesize**
4. **Experiment**
5. **Analyze**
Research Trends influence CI strategy

Research Context

- Globalization of Research Communities
- Digital technologies deeply permeating both research and education
- Research expectations high and funding flat
- Research frontiers increasingly multidisciplinary
- National demographics increasingly diverse

Cyberinfrastructure Response

- Cooperating, dynamic, multilayer, national and global CI at all scales
- Sustainability focus (especially software, data) in addition to Innovation
- Ubiquitous, capable, secure and facile CI access for more researchers, educators, institutions, communities
- CI Collaborations
- Learning and Workforce development for both CI creators and users
FY 2016 Budget Request

- **NSF**
  - FY 2016 Budget Request: $7723.55 Million
  - Increase over FY 2015 Est: $379.34 Million, +5.2%

- **CISE Directorate**
  - Computer Information Science & Engineering
  - FY 2016 Budget Request: $954.41 Million
  - Increase over FY 2015 Est: $32.68 Million, +3.5%
  - ACI Division at $227.29 (+3.9%)

- CISE FY 2016 request is shaped by investments in **core research, education, and infrastructure programs** as well as critical investments in **NSF cross-directorate priorities and programs**.
ACI Mission: To support advanced cyberinfrastructure to accelerate discovery and innovation across all disciplines

- CI Coordination role across NSF
- Supports Use-inspired Cyberinfrastructure
  - Research and Education
  - Science and Engineering
- Inherently multidisciplinary with strong ties to all disciplines/directorates
Laying a National Foundation for Research Innovation

• ACI supports
  • National-scale capabilities
  • CI innovation & collaboration
  • Inherently Multidisciplinary
    + International
    + Multi-institutional
    + Domain-specific

• Research CI: Networking, security, software, data, advanced computing, learning and workforce development
NSF-supported national resources provide a forward looking, cohesive computational foundation for researchers/educators.
“Blue Waters”/UIUC/PI Kramer
Has Enabled Science Not Possible Any Other Way

- Computational Microscope – K. Schulten (UIUC)
  - HIV capsid required 65 million atoms
  - Chromatophore requires 100 million atoms
- Largest universe evolution simulation – T. Di Matteo (CMU)
  - Required 1.2 PB of DRAM
  - Formation of the first quasars and galaxies.
  - Used by WFIRST and JWST for observation predictions.
- Transistor roadmap projections – G. Klimeck (Purdue)
  - Support for CPU/GPU codes.
  - Rapid turn-around times.
- Plasma simulations – W. Mori (UCLA)
  - High sustained floating point performance needed
  - 150 million grid points and 300 million particles
  - (2 cm)$^3$ of plasma
- Earthquake response modeling – T. Jordan (USC)
  - CyberShake workloads using CPU and GPU nodes, sustained, for weeks.
  - Seismic hazard maps (NSHMP) and building codes.
25% of Blue Waters projects involve researchers and/or students from EPSCOR states

Map of Principal Investigator Institutions. Orange states indicate EPSCoR states with at least one Blue Waters Project principal investigator. Puerto Rico is not shown. The dots show the institutions for principal investigators leading one or more projects.
2015 Deployments Recognize Increasing Scientific Breadth and Computational Diversity

Comet: UCSD/SDSCC/PI Norman
- Increase in access, usability and capacity
- Problem reach beyond campus scale
- Bridging to science gateways/portals

Wrangler: UT Austin/TACC/PI Stanzione
- Groundbreaking data analytics at scale
- Novel configuration (I/O, file system)

Sample Project: PaleoCore
- Numerous projects in East Africa
- Custom databases of paleoanthropological and archeological data
- How to share and analyze data across multiple projects?

Sample Project: Colloids and self-assembling systems

Glotzer/UM uses Comet to simulate colloids of hard particles, including spheres, spheres cut by planes, ellipsoids, convex polyhedra, convex spheropolyhedra, and general polyhedra.
FY2015 MRI CI Awards Stimulate Novel Technical and Operating Model Approaches

Awards made to universities in six states (CO, IL, MI, NY, OK, TX)

- Collaborators and participants included many other institutions
- 14 states
- 28 universities
- 10 laboratories / facilities
2015 ACI MRI Awards are Diverse in Scope and Size

<table>
<thead>
<tr>
<th>ID</th>
<th>PI</th>
<th>Institution</th>
<th>Title</th>
<th>$Award</th>
</tr>
</thead>
<tbody>
<tr>
<td>1531128</td>
<td>Brunson</td>
<td>Oklahoma State University</td>
<td>MRI Acquisition of Shared High Performance Compute Cluster for Multidisciplinary Computational and DataIntensive Research</td>
<td>$951,570</td>
</tr>
<tr>
<td>1531492</td>
<td>Harrison</td>
<td>SUNY Stony Brook</td>
<td>MRI Acquisition of SeaWulf A reconfigurable computer system for Research and Education</td>
<td>$1,400,000</td>
</tr>
<tr>
<td>1531594</td>
<td>Hibbs</td>
<td>Trinity University</td>
<td>Acquisition of High Performance Scientific Computing Cluster at Trinity University</td>
<td>$623,730</td>
</tr>
<tr>
<td>1531752</td>
<td>Duraisamy</td>
<td>University of Michigan</td>
<td>MRI Acquisition of Conflux A Novel Platform for DataDriven Computational Physics</td>
<td>$2,422,972</td>
</tr>
<tr>
<td>1531814</td>
<td>Cheung</td>
<td>University of Houston</td>
<td>MRI Acquisition of a High Performance Computing System for Science and Engineering Research and Education at the University of Houston</td>
<td>$950,000</td>
</tr>
<tr>
<td>1532133</td>
<td>Catlett</td>
<td>University of Chicago</td>
<td>MRI Development of an Urban-Scale Instrument for Interdisciplinary Research</td>
<td>$3,110,488</td>
</tr>
<tr>
<td>1532235</td>
<td>Seigel</td>
<td>Colorado State University</td>
<td>COLLABORATIVE PROPOSAL: MRI Collaborative Consortium Acquisition of a Shared Supercomputer by the Rocky Mountain Advanced Computing Consortium</td>
<td>$700,000</td>
</tr>
<tr>
<td>1532236</td>
<td>Hauser</td>
<td>University of Colorado Boulder</td>
<td>COLLABORATIVE PROPOSAL: MRI Collaborative Consortium Acquisition of a Shared Supercomputer by the Rocky Mountain Advanced Computing Consortium</td>
<td>$2,030,000</td>
</tr>
</tbody>
</table>

Total $12,188,760
High Performance Computing (HPC) to support a regionally distributed and diverse set of computing- and data-intensive research and research training.

• Broad set of research disciplines: bioinformatics, environment and ecosystems, proteins, physics and medical physics, chemistry, computer science (data and cybersecurity) and sociological modeling.

• Collaborating research teams include faculty across many departments, significant participation of postdocs, graduate students and undergraduates.

Broader Impacts:
• Researchers across the state of Oklahoma
• Collaborating institutions: Indiana University, Kansas State University, Oklahoma Innovation Institute, OneNet, University of Texas at Austin, and the USDA/Agricultural Research Service.
• Diverse team and student population. Link to REU and RET activities.
• Analysis of social and biological networks.
MRI: Development of an Urban-Scale Instrument for Interdisciplinary Research

Charles Catlett, University of Chicago [Award #ACI-1532133]
Co-funded by CISE/OAD, ENG/CBET, ENG/CMMI

The ‘Array of Things’ instrument allows researchers to rapidly deploy sensors, embedded systems, computing, and communications systems at scale in an urban environment.

• Funds the development and installation of AoT ‘nodes’ -- enclosures containing instruments for measuring temperature, barometric pressure, light, vibration, carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, ambient sound intensity, pedestrian and vehicle traffic, and surface temperature.
• All data collected by the nodes will be free and publicly available through the City of Chicago Data Portal and other open data platforms.

Broader Impacts:
• In partnership with the City of Chicago, 500 nodes will be mounted around the city by 2017.
• Public health researchers will be able to study the relationship between diseases, which occur at higher rates in urban areas, and environmental conditions.
• Climate researchers will have higher resolution data than currently provided by existing weather stations to study urban micro-climates, with benefits for hyper-local weather forecasting and energy efficiency.
• Social scientists can study the dynamics of urban activity in public spaces and the effects on economics and livability.
COLLABORATIVE PROPOSAL: MRI: Collaborative Consortium by the Rocky Mountain Advanced Computing Consortium

Siegel, Colorado State University [Award #1532235]
Hauser, University of Colorado Boulder [Award #1532236]
Co-funded by GEO

Regional high performance computing (HPC) as part of the Rocky Mountain Advanced Computing Consortium (RMACC) of more than 14 universities, laboratories and research institutions.

- Supports computing for 2,000 researchers in a broad range of domains: physics, astrophysics, engineering, materials science, renewable energy, earth system science, computational fluid dynamics, bioinformatics, and social science.

- A major emphasis as a test, porting, and scaling platform -- a valuable complement to large national centers, where research applications must have completed porting and scaling tests to compete for computing allocations.

Broader Impacts:
- Collaborating organizations include major research institutions and laboratories in six states (Colorado, Idaho, Montana, New Mexico, Utah, and Wyoming).
- The consortium includes universities in multiple states, multiple federal agency participation.
ConFlux is designed to enable High Performance Computing (HPC) clusters to communicate seamlessly and at interactive speeds with data-intensive operations.

- Addresses a major research challenge: developing data-driven multiscale computational techniques that provide quantitative predictions of the behavior of physical systems with quantifiable uncertainties on the predictions.
- Enables predictive modeling innovations in a number of fields, including computational materials physics, turbulence, blood flow modeling, climate modeling and cosmology.

Broader Impacts:
- Educational outreach includes a new undergraduate data science degree program at the University of Michigan, classes, training, massive open online courses (MOOCs), and workshops.
Research and Education Network – Advancing the Foundation: 2012-15

2012-2015 CC-NIE/CC*IIE/CC*DNI Programs
- 46 states/jurisdictions
- >$80M
- >170 awards
Updating International Network Collaborations:
IRNC AmLight project: PI Ibarra/Florida International University

- 300GHz of configurable spectrum: Santiago-São Paulo, and São Paulo-Miami
- Regional resilience for U.S.-Latin America, and U.S.-Europe connectivity, supporting global science research
- Science drivers include telescopes in Chile and High Energy Physics requiring 100’s Gbps by end of decade:
  - Large Synoptic Survey Telescope (LSST)
  - Atacama Large Millimeter Array (ALMA)
  - LHC Open Network Environment (LHCONE)
<table>
<thead>
<tr>
<th>PI</th>
<th>Organization</th>
<th>Proposal Title</th>
<th>Science Areas / CI Issues</th>
</tr>
</thead>
</table>
| Lifka, David | Cornell University            | Federated Cloud (#1541215)                                         | • Cornell and its partners (SUNY Buffalo and UC Santa Barbara) create a federated cloud in New York and California. Includes seven science use cases, emphasizing a ‘time to science’ metric.  
• Serves as a model for campus cyberinfrastructure at other institutions; includes an allocation model that provides a fair exchange mechanism for resource access between and across multiple institutions. |
| McKee, Shawn | University of Michigan         | Multi-Institutional Open Storage Research InfraStructure (MI-OSiRIS) (#1541335) | • Deploys a storage platform across the three largest research universities in the state of Michigan, based on the CEPH object based file system.  
• Advances the shared knowledge and understanding of data in genomics, healthcare, geology and physics.                                                                                     |
| Peterson, Larry | University of Arizona       | Give Your Data the Edge: A Scalable Data Delivery Platform (#1541318) | • A nine-university collaboration (Arizona, Hawaii, Indiana, North Carolina, Northern Arizona, Princeton, Texas, UC Davis, Wisconsin) .  
• Leverages "Syndicate" platform to facilitate global, scalable and secure access to data. Demonstrates how scientific applications can effectively and easily access data, independent of where it is stored.  
• Explores pay-to-play model of sustainability. Strategic partnerships with Google and Akamai.                                                                                                           |
| Smarr, Larry | University of California – San Diego | The Pacific Research Platform (PRP) (#1541466) | • A collaborative, domain-science-driven environment across 20 campuses along the Pacific region of the US.  
• Tests a new networking appliance to enhance access speeds for sharing/accessing data and models.                                                                                                         |
Pacific Research Platform: Enabling a science-driven, high-capacity data sharing network

- 5 year, $5M CC*DNI grant
- Digital fabric to support Big Data across 10 UC campuses and 10 other research institutions
- Builds on ESNET Science DMZ

---

Research Scientists to Use Network Much Faster Than Internet

By John Markoff  July 31, 2015

SAN FRANCISCO — A series of ultra-high-speed fiber-optic cables will weave a cluster of West Coast university laboratories and supercomputer centers into a network called the Pacific Research Platform as part of a five-year $5 million dollar grant from the National Science Foundation.
NSF Advanced Computing Vision & Strategies

• **Vision:** To position and support the entire spectrum of NSF-funded communities at the cutting-edge of advanced computing technologies

• **Strategies:**
  – Foundational research to fully exploit parallelism and concurrency
  – Application of research and development of high-end computing resources
  – Build, test, and deploy resources into a collaborative ecosystem
  – Support comprehensive education and workforce development
  – Develop transformational, grand-challenge communities
NSF’s vision for cyberinfrastructure has been informed by years of community input, development, and experience.

**Initial Vision Developed (2007-2010)**

**NSF-Wide Task Force Reports (2009-2011)**
Looking to the Future: Continuing Community Engagement Nationally

- **NSF Advanced Computing Infrastructure for 21st Century Science and Engineering: Vision and Strategic Plan (Feb 2012)**
  - Position, support spectrum of NSF-funded communities at cutting edge of advanced computing technologies, hardware, software, services

- **Future Directions of NSF Advanced Computational Infrastructure to Support US Science in 2017 – 2022**
  - National Academy of Sciences (NAS)
  - Interim Report (Oct 2014), Final Report (Fall 2015)

- **National Strategic Computing Initiative (Executive Order)**
  - DOD, DOE, NSF designated as lead agencies

- **Advisory Committee for Cyberinfrastructure (ACCI)**
  - November 5-6, 2015

Interim report Co-chairs:
W. Gropp/UIUC
R. Harrison/Stony Brook
National Strategic Computing Initiative (NSCI) Executive Order signed by President Obama on July 29, 2015

*Maximize HPC benefits for economic competitiveness and scientific discovery*

**NSF Called on to Provide Leadership**
- Scientific Discovery Advances
- Broader HPC Ecosystem for Scientific Discovery
- Workforce Development

**Goals**
- 100x performance increase in HPC simulations
- **Technical synergy in platform for modeling/data analytics**
- Research to take scaling beyond current device limits
- Increase capacity and capability of national HPC ecosystem
- Public/private partnership

**Executive Council**
- Co-chairs: OSTP and OMB Directors
- Initial Implementation Plan Submitted (90 days)

**NSF Plan Builds on the current foundation while setting an ambitious Future**
- Co-led by MPS/CISE with participation by all directorates
- Investment in Software, Technology Platforms, People
NSCI Executive Order calls on NSF to play a leadership role

Scientific discovery advances

The broader HPC ecosystem for scientific discovery

Workforce development

Co-lead with DOD and DOE
NSCI Objectives

1. Accelerate delivery of a capable exascale computing system (hardware, software) to deliver approximately 100X the performance of current 10PF systems across a range of applications reflecting government needs

2. Increase coherence between technology base used for modeling and simulation and that used for data analytic computing.

3. Establish, over the next 15 years, a viable path forward for future HPC systems in the post Moore’s Law ...

4. Increase the capacity and capability of an enduring national HPC ecosystem, employing a holistic approach ... networking, workflow, downward scaling, foundational algorithms and software, and workforce development.

5. Develop an enduring public-private partnership to assure that the benefits .. are transferred to the U.S. commercial, government, and academic sectors
NSCI Objective 2

Increase coherence between technology base used for modeling and simulation and that used for data analytic computing

Modeling and Simulation
- Multi-scale
- Multi-physics
- Multi-resolution
- Multidisciplinary
- Coupled models

Data Science
- Data Assimilation
- Visualization
- Image Analysis
- Data Compression
- Data Analytics
- Machine Learning

NSF Role: Support foundational research and research infrastructure within and across all disciplines (across all NSF directorates)
Data Science: Emerging ... inherently multidisciplinary.

Engineering Practice

Theoretical Foundation

Technical Skills

Math & Statistics Knowledge

Domain Knowledge

Substantive Expertise

Danger Zone!

Traditional Research

Machine Learning

Adapted from Drew Conway’s 2010 Venn Diagram
NSCI Objective 3

Establish, over the next 15 years, a viable path forward for future HPC systems in the post Moore’s Law era

Happening now
- Multi-core and many-core processors
- Domain-specific integrated circuits
- Energy-aware computing
- Hierarchical memories
- High-speed Interconnects

 Longer term
- Usable parallelism, concurrency, and scalability
- Resiliency at scale
- Decreased power consumption
- Architectures that reduce data movement
- New materials (e.g., carbon nano-tubes, graphene-based devices)
- Non-charge transfer devices (e.g., electron spin)
- Bio, nano, and quantum devices

NSF Role: Support foundational research
Leadership by CISE, ENG, MPS, and BIO
Build on existing interagency and industry partnerships
NSCI Objectives 2 and 4

Increase synergy between technology base used for modeling and simulation and that used for data analytic computing

Increase the capacity and capability of an enduring national HPC ecosystem, employing a holistic approach ... networking, workflow, downward scaling, foundational algorithms and software, and workforce development.

Emphasis on re-use, agility, interoperability, sustainability

NSF Role: Accelerate scientific discovery advances
Participation by all NSF directorates
Expand international, interagency, public sector, and industry collaborations

NSF Investments: (Scientific) Software, Platform Deployments, Workforce Development
Evolve Service Architectures (XD)

- Current Human Services
- Current Digital services
- What is needed in the future?

- Resource Allocation (XRAC)
- Computational Experts (ECSS)
- Educational Services (TEOS)

- Compute cycles and storage
- XDMoD metrics tool
- File transfer
- Common web portal
- Authentication services

- New economic models for access/allocation?
- Support for dynamic workflows?
- New models and types of collaboration (ACI-REF)?
- APIs to access services (e.g. iPlant, HubZero, Gateways)?
Optimizing the Research Infrastructure Ecosystem

Science Frontiers
- data-intensive science;
- software sustainability;
- diverse priorities

Technology Advances
- end of Moore’s law;
- commoditization opportunities;
- cohesive platform for simulation and data analytics

Operating Models
- efficient shared services;
- stimulating innovation and effectiveness;
- collaborations

Human Considerations
- workforce diversity;
- career paths;
- education;
- community development
THANK YOU