



Fostering a Cyberinfrastructure Ecosystem that Transforms Scientific Discovery

Manish Parashar
Office Director

Office of Advanced Cyberinfrastructure,
Directorate for Computer & Information Science &
Engineering
National Science Foundation

CC* & CICI PI Workshop
September 25, 2018

Outline



Outline

```
graph LR; A[NSF/OAC Update] --> B[Emerging Science & Role of CI]; B --> C[Rethinking the CI Ecosystem]; C --> D[Conclusion]
```

NSF/OAC
Update

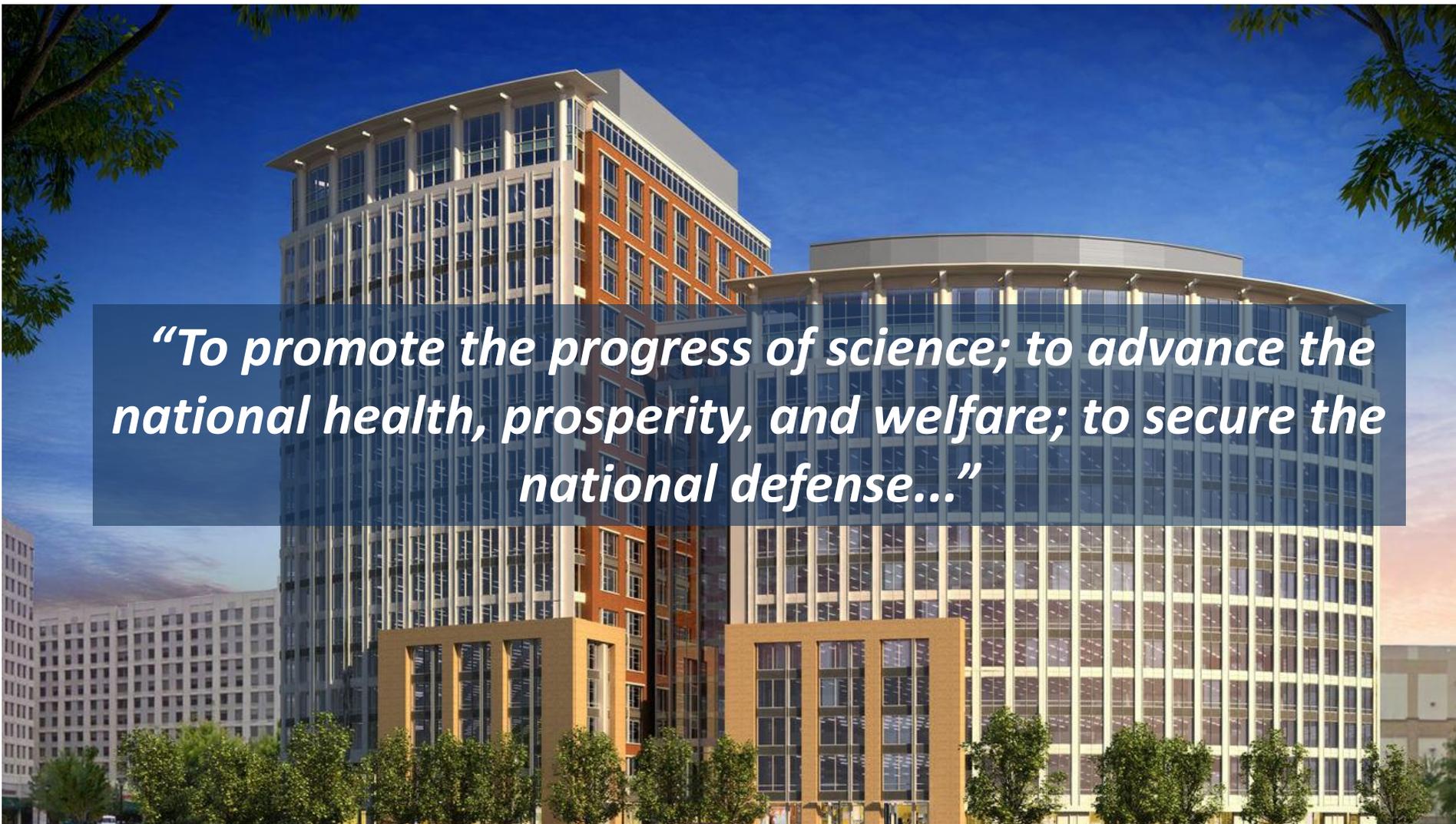
Emerging
Science &
Role of CI

Rethinking the
CI Ecosystem

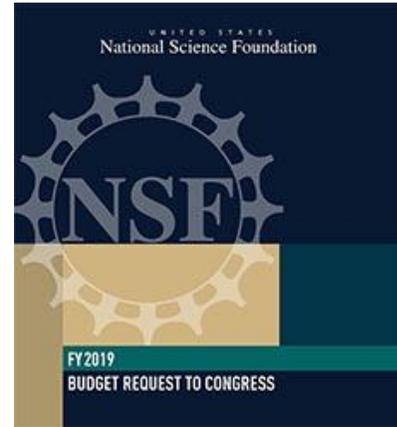
Conclusion



National Science Foundation's Mission



NSF Budget Update



FY 2018

- FY18 Congressionally Appropriated Budget
 - +4%
 - +\$300M over FY17 budget!
- 2nd largest increase in NSF research budget increase in 15 years! (not counting ARRA)

FY 2019 (Request)

- NSF: \$7.47 billion
 - Flat with respect to FY 17 Enacted
- CISE: \$925.4 million
 - -1.1% from FY 17 Enacted
- Big Ideas
 - Research Ideas: \$30 million each
 - Process Ideas: Midscale infrastructure: \$60 million
- Convergence Accelerators
 - \$60 million
 - HDR, FW-HTF: \$30 million each



The NSF Big Ideas

RESEARCH IDEAS

 <p>Harnessing Data for 21st Century Science and Engineering</p>	<p>Work at the Human-Technology Frontier: Shaping the Future</p>   <p>Navigating the New Arctic</p>	<p>Windows on the Universe: Multi-messenger Astrophysics</p>  	<p>Quantum Leap: Leading the Next Quantum Revolution</p>  <p>Understanding the Rules of Life: Predicting Phenotype</p> 
---	---	---	---

PROCESS IDEAS

<p>Mid-scale Research Infrastructure</p> 	<p>NSF 2026</p> 
 <p>Growing Convergence Research at NSF</p>	 <p>NSF INCLUDES: Enhancing STEM through Diversity and Inclusion</p>

“ ... bold questions that will drive NSF's long-term research agenda -- questions that will ensure future generations continue to reap the benefits of fundamental S&E research. ”



Big Ideas => Big Cyberinfrastructure Challenges & Opportunities

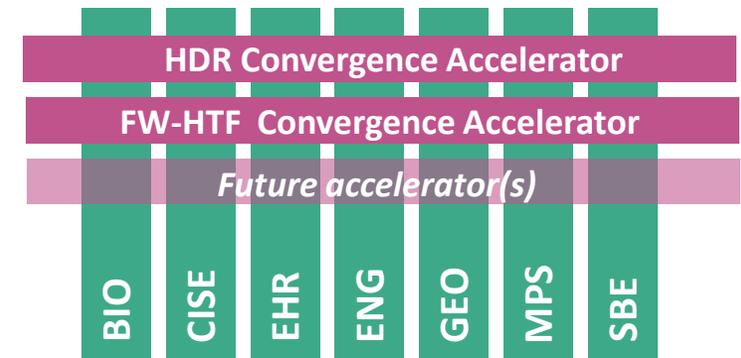


Convergence Accelerators

Accelerating Discovery through Convergence Research

Motivation: Changing nature of science research - research frontiers at intersection of existing disciplines

- Time-limited entities: accelerating impactful *convergence* research in areas of national importance
- Innovating in organizational structure to better enable frontier research
- Separate (from directorates) in leadership, budget, and programmatic; but aligned with, relying on, and building on foundational disciplinary research
- Emphasis on translational research, partnerships



National Strategic Computing Initiative (NSCI)

Maximizing benefits of HPC for scientific discovery and economic competitiveness

- *Multi-agency effort to maximize the **benefits of High Performance Computing (HPC)** for scientific discovery and economic competitiveness*
- NSF/CISE/OAC leads interagency effort to:
 - Increase coherence between technology base used for modeling/simulation and for data analytics;
 - Establish viable path forward for HPC systems in post-Moore's Law era; and
 - Increase capacity, capability, and sustainability of an enduring national HPC ecosystem



Includes CISE investments in the following programs: High Performance Computing (HPC), Scalable Parallelism in the Extreme (SPX), Quantum Leap Big Idea, and Cyberinfrastructure for Sustained Scientific Innovation – Software (CSSI)





EXECUTIVE OFFICE OF THE PRESIDENT
WASHINGTON, D.C.



American Leadership in Artificial Intelligence, Quantum Information Sciences, and Strategic Computing

Continued leadership in AI, quantum information science (QIS), and **strategic computing** is critically important to our national security and economic competitiveness. Advances in these areas promise opportunities for major scientific breakthroughs and are quickly transforming American life and industry. Agencies should invest in **fundamental and applied AI research**, including machine learning, autonomous systems, and applications at the human-technology frontier. Agencies should prioritize QIS R&D, which will build the technical and scientific base necessary to explore the next generation of **QIS theory, devices, and applications**. Agencies should **prioritize investment in research and infrastructure to maintain U.S. leadership in strategic computing, from edge devices to high-performance computing, that accelerates delivery of low-power, high performance devices; supports a national high-performance computing ecosystem; and explores novel pathways to advance computing in a post-Moore's Law era.**



NSF Office of Advanced Cyberinfrastructure

Program Staff



Manish Parashar*
Office Director



Amy Friedlander
Deputy Office Director



Bill Miller
Science Advisor

Computing

Data

Software

Networking &
Cybersecurity

Learning & Workforce
Development



Beth Plale*
Science Advisor
Public Access



Bob
Chaddock



Amy Walton



Vipin
Chaudhary*



Micah Beck*



Sushil Prasad*



Alejandro
Suarez
Cooperative
Agreements



Ed Walker



Stefan
Robila*



Rajiv
Ramnath*
(Part-Time)



Kevin
Thompson



Julie Stalhut
AAAS S&T
Policy Fellow

Join NSF/OAC: Multiple Program Officer openings

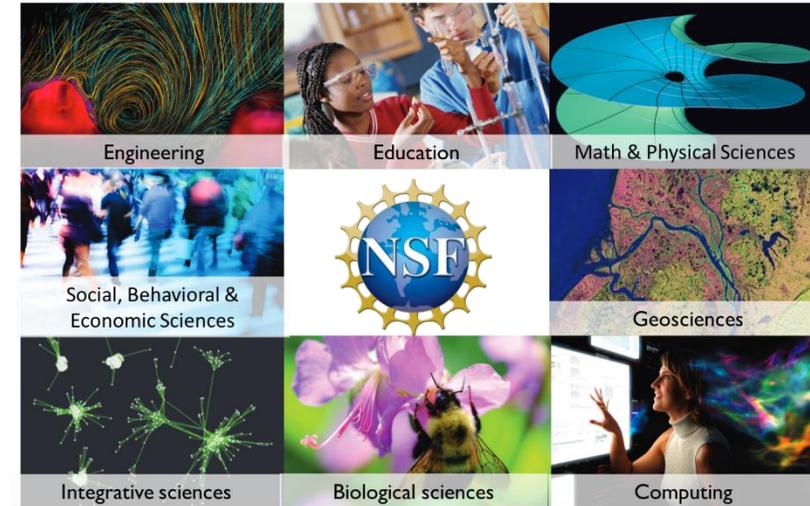
* IPA Appointment



CISE/OAC – Transforming the Frontiers of Science & Society

Foster a cyberinfrastructure ecosystem to transform computational- and data-intensive research across all of science and engineering

- Cyberinfrastructure Research & Research Cyberinfrastructure



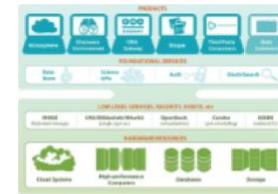
CI-Enabled Instrumentation



Computing Resources



Data Infrastructure



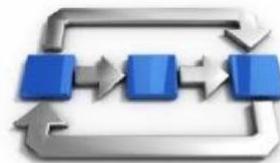
Gateways, Hubs, and Services



R&E Networks, Security Layers



Coordination & User support



Software and Workflow Systems



Pilots, Testbeds



People, organizations, and communities



CISE/OAC – Transforming the Frontiers of Science & Society

Computing

Advanced resources and services at all scales – MRI (clusters); Innovative HPC; Leadership Class; XSEDE coordination and user services; Research

Data

Data Building Blocks (DIBBS) Program

Software

Software Infrastructure for Sustained Innovation (SI2)

Networking & Cybersecurity

Campus Cyberinfrastructure (CC*), International Research Network Connections (IRNC), Cybersecurity Innovation for CI (CICI)

Learning & Workforce Dev.

Training-based Workforce Development for Advanced Cyberinfrastructure (CyberTraining), CAREER, CRII

Cyberinfrastructure for Sustained Scientific Innovation (CSSI)

OAC Core Research Program
NSF 18-567; Due 11/15/18

Emerging Opportunities

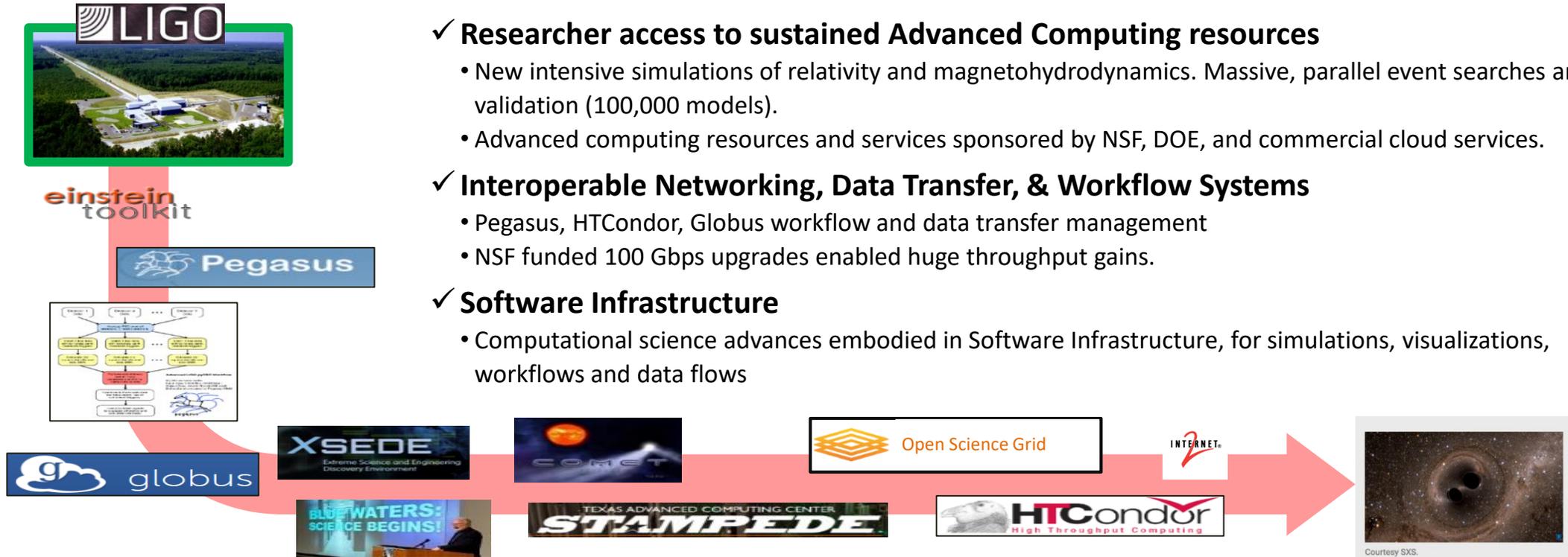
Cyberinfrastructure for Emerging Science and Engineering Research (CESER), Public Access



Transforming Science through a Cyberinfrastructure Ecosystem: Dynamic discovery pathways at scale

CI enables Big Science

Gravitational wave detection enabled by NSF investments across the cyberinfrastructure ecosystem



✓ **Researcher access to sustained Advanced Computing resources**

- New intensive simulations of relativity and magnetohydrodynamics. Massive, parallel event searches and validation (100,000 models).
- Advanced computing resources and services sponsored by NSF, DOE, and commercial cloud services.

✓ **Interoperable Networking, Data Transfer, & Workflow Systems**

- Pegasus, HTCondor, Globus workflow and data transfer management
- NSF funded 100 Gbps upgrades enabled huge throughput gains.

✓ **Software Infrastructure**

- Computational science advances embodied in Software Infrastructure, for simulations, visualizations, workflows and data flows

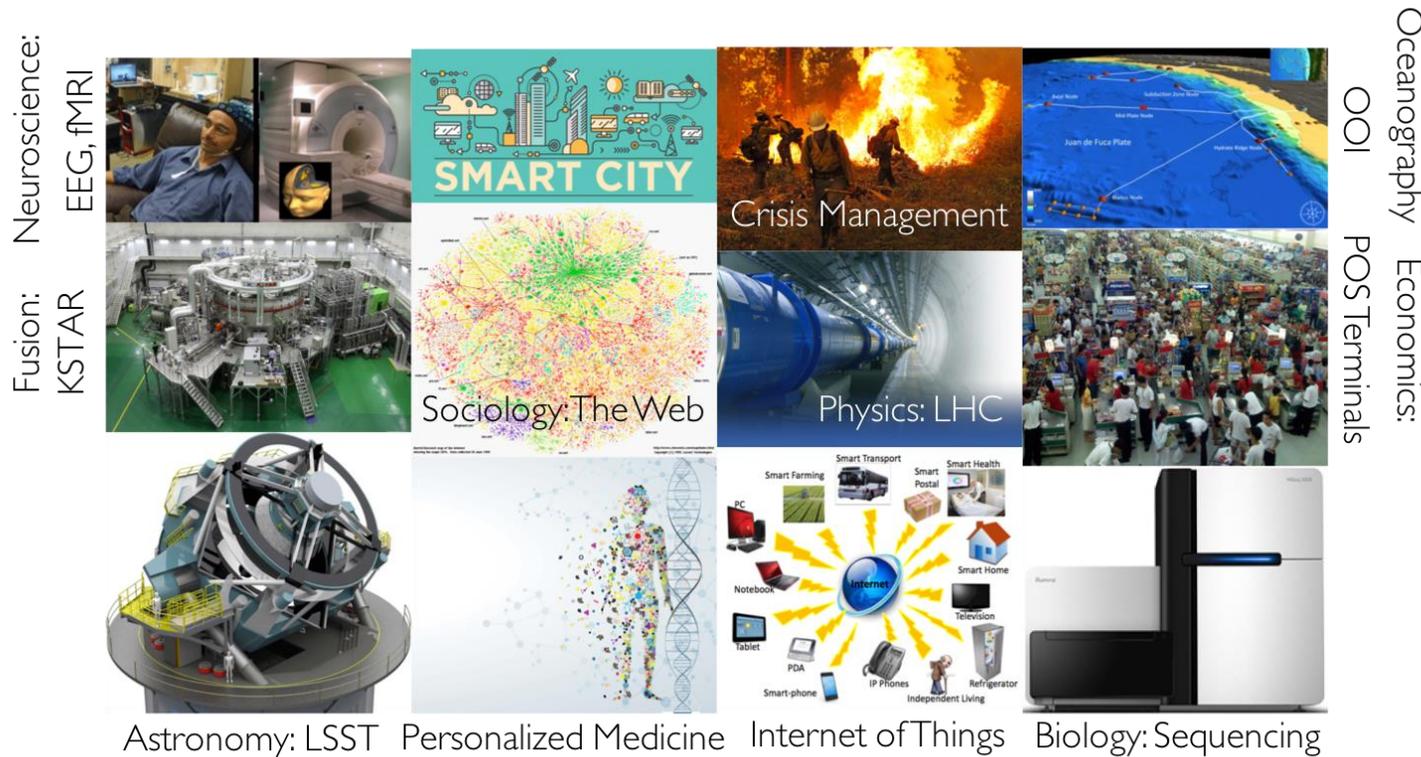
NSF programs: Data Building Blocks (DIBBs), Software Infrastructure (SI²), Campus Cyberinfrastructure Network Infrastructure and Engineering (CC*NIE, DNI), and others. OSG and Pegasus are also supported by the Department of Energy.



Outline



Science / Society Transformed by Compute & Data



Science and Engineering in 21st Century

- New paradigms and practices in science and engineering
- Inherently multi-disciplinary
- Extreme scales, data-driven, data and compute-intensive
- Collaborative (university, national, global)

- *Nearly every field discovery is transitioning from “data poor” to “data rich”*
- *The scientific process has evolved to include computation & data*



Evolving Science, CI Landscapes



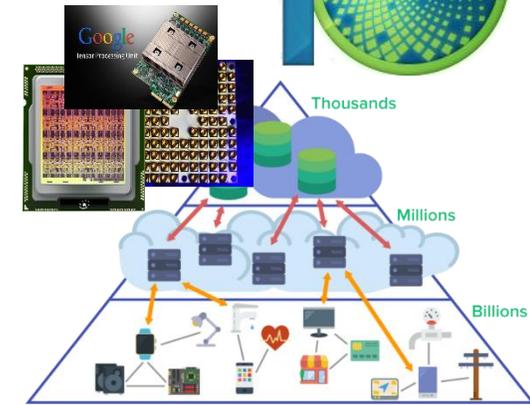
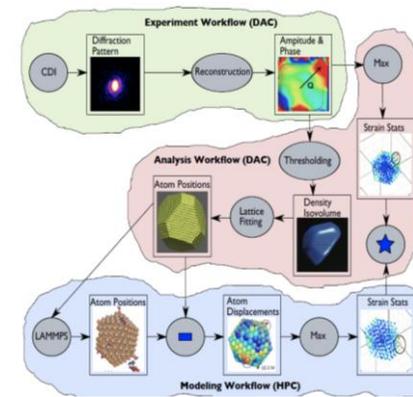
Evolving Science/Engineering Landscape

- Large scales, high-resolution, multi-scale, multi-physics simulations / Complex, dynamic workflows
- Emerging data-driven (ML-based) models
- Streaming data from observatories, instruments
- Rapidly growing “long-tail”, “gateway” jobs; “small” jobs dominating; increasing use of clouds
- Heightened emphasis on robust results (transparency, credibility, correctness, security, ...)

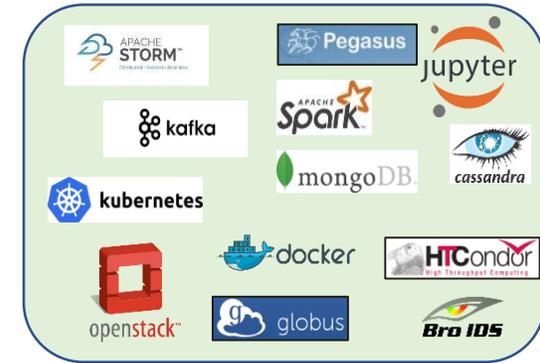
Evolving Technology Landscape

- Extreme scales / pervasive computing and data
- Diverse / disruptive technologies increasing
- Role of (non-traditional) software in taming complexity
- High throughput/low-latency networks
- Novel paradigms / Increasing role of clouds, multiclouds, hybrid environments / Growing capabilities & capacities at the edges

End-to-end Workflows



Instrument, Observatories, Experimental Facilities

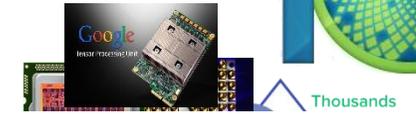


Evolving Science, CI Landscapes



Evolving Science/Engineering Landscape

End-to-end Workflows

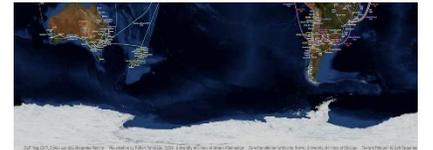
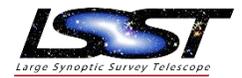


The cyberinfrastructure ecosystem must evolve to provide new levels of agility, interoperability, collaborations & reuse

*How do we catalyze a **Cyberinfrastructure Continuum**, from the sensors to science...?*

multiclouds, hybrid environments / Growing capabilities & capacities at the edges

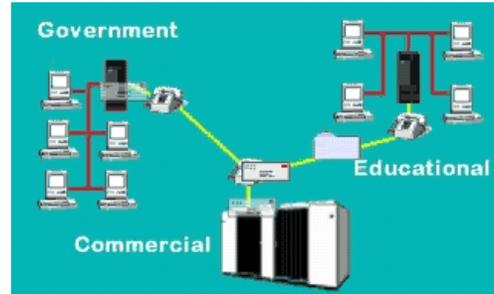
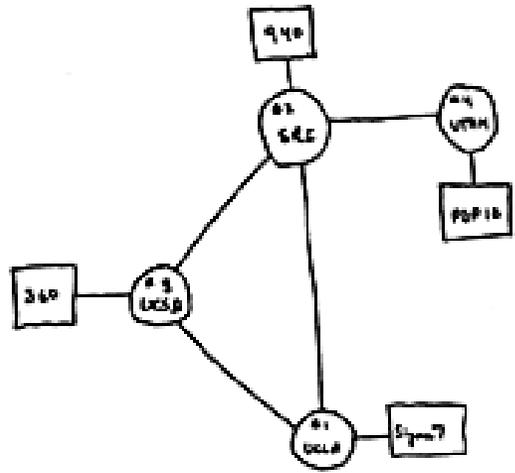
ions



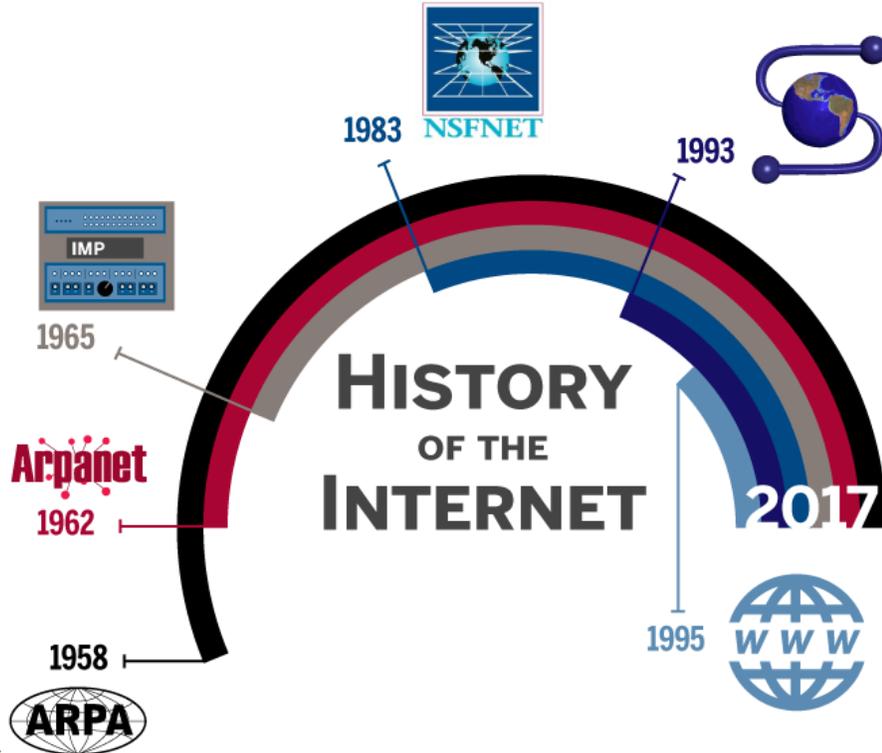
Outline



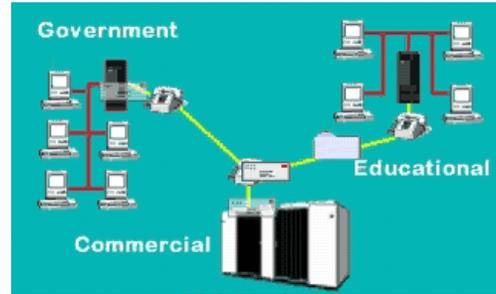
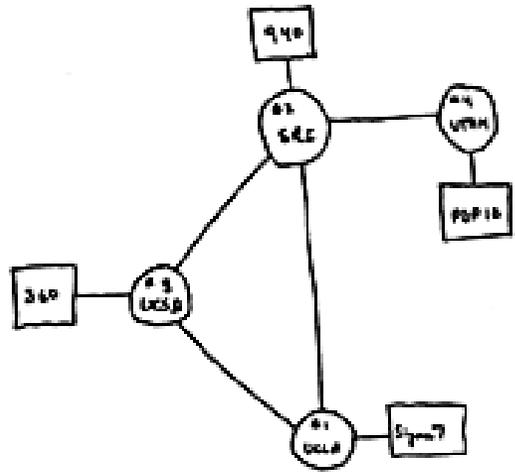
Looking Back: NSF's Role the Internet



- **1981:** NSF establishes the Computer Science Network (CSNET) to provide networking services to all university computer scientists.
- **1984:** NSF creates NSFNET
- **1985:** NSF links research universities to provide greater access to computing resources at its supercomputer centers.
- **1987:** NSF merges BITNET & CSNET to form the Corporation for Research and Educational Networking (CREN).
- **1990:** NSFNET replaces the obsolete ARPANET using the same “backbone” infrastructure.
- **1991:** NSF creates the National Research and Education Network (NREN) for high speed networking research.
- **1993:** NSF creates InterNIC to provide specific Internet services: directory and database services (by AT&T), registration services (by Network Solutions Inc.), and information services (by General Atomics/CERFnet).
- **1995:** NSFNET backbone decommissioned
- **1998:** End of NSF's direct role in the Internet.



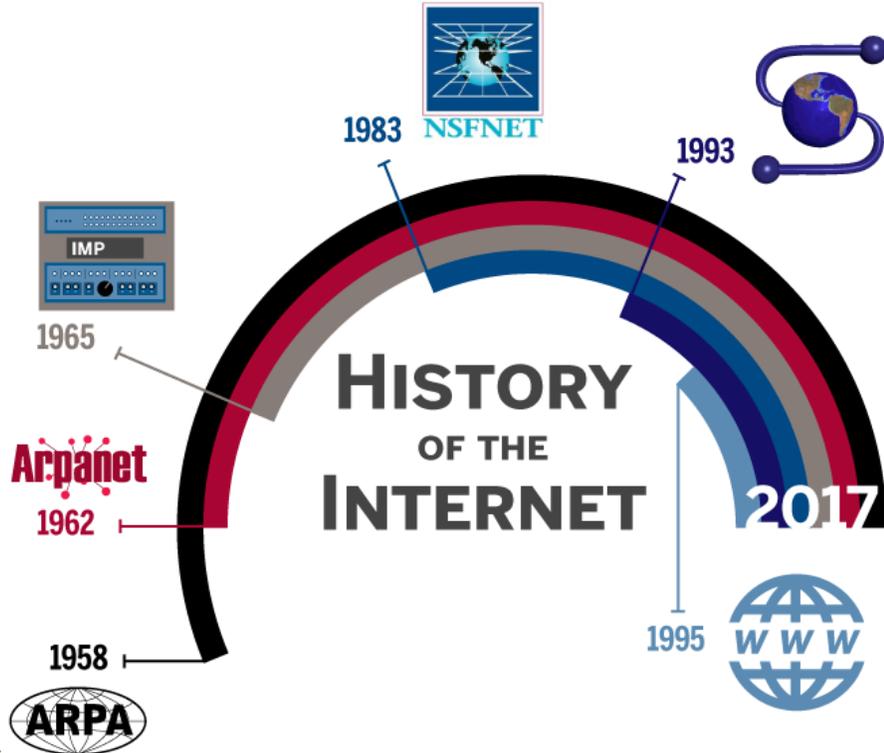
Looking Back: NSF's Role the Internet



- Creation of NSFNET was an intellectual leap
 - First *large-scale implementation of Internet technologies* in a complex environment of many independently operated networks.
 - Forced the Internet community to iron out technical issues arising from the rapidly increasing number of computers and address.

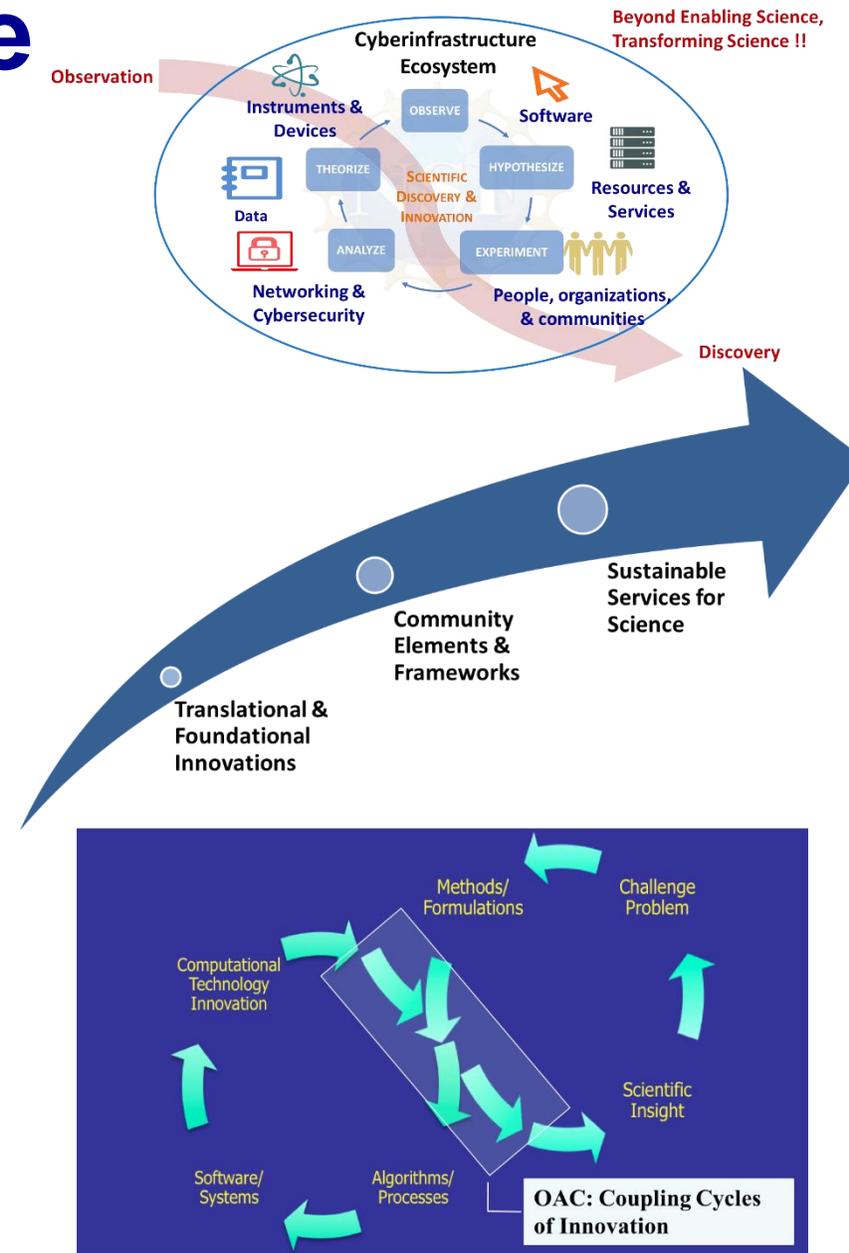
- NSF helped catalyze a self-governing and commercially viable Internet
 - Commercial firms built their own networks.
 - Outlined a new Internet architecture that largely remains in place.

- NSF continues to support evolutions
 - Helps research and education institutions make and enhance their connections to the Internet.
 - Instrumental in providing international connections services.

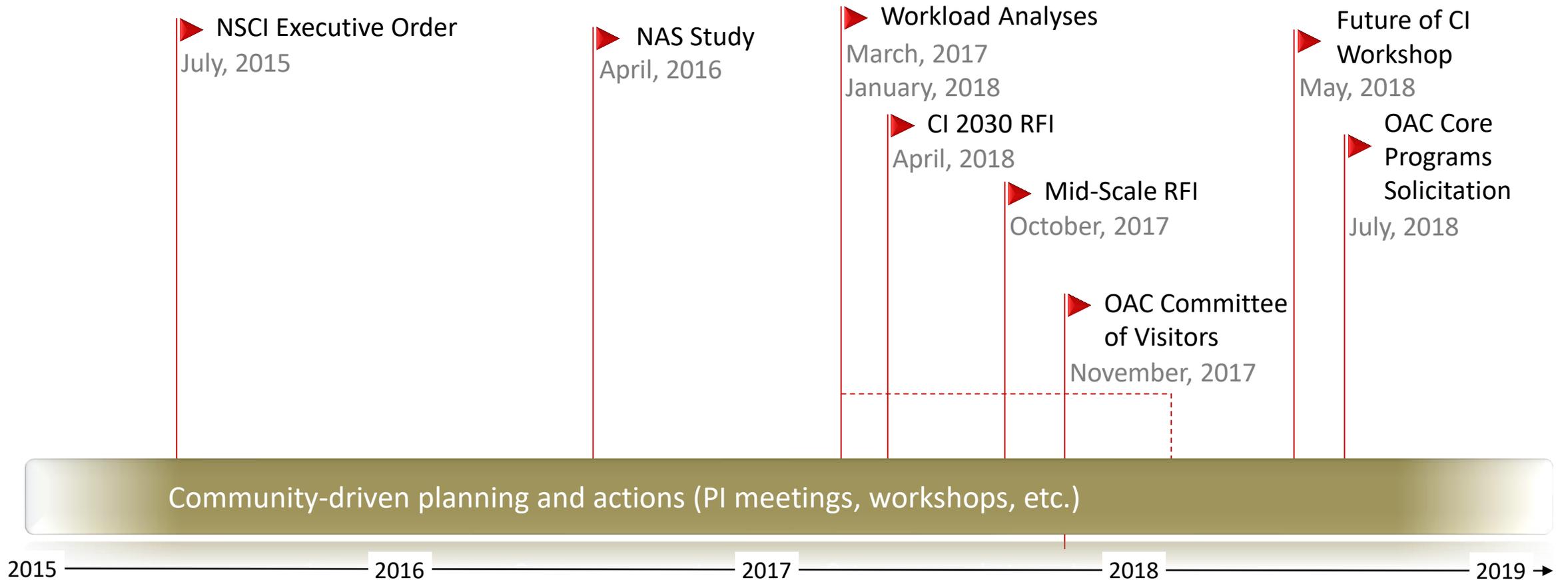


Rethinking CI: Realizing a CI Ecosystem to Transform Science

- Realize a holistic and integrated and secure cyberinfrastructure ecosystem aimed at transforming science
- Support the translational research continuum, from catalyzing core innovations, through fostering the community tools and frameworks, and enabling sustainable cyberinfrastructure services
- Work closely with science and engineering communities, and other stakeholders to tightly couple the cycles of discovery and innovation



Rethinking CI: Building on Community Input



Key Drivers

- Changing application landscape & workload profile

- Changing technology, services landscape
- Increasing availability of (exp., obs.) data
- Growing role of ML, data-driven approaches



Rethinking CI – Computational Infrastructure*

* Not mutually exclusive.

- Highly specialized instruments
- Extreme-scale capabilities
- NSF 17-558

Leadership
Class
Systems





NSF 17-558: Towards a Leadership-Class Computing Facility – Phase 1

Resolution

RESOLVED, that the National Science Board authorizes the Director at her discretion to make an award, OAC-1818253, to the Texas Advanced Computing Center (TACC) at the University of Texas at Austin for the acquisition of the system described in proposal “Computation for the Endless Frontier,” in an amount not to exceed \$60,000,000 for a period of 60 months. Pending appropriate approval associated with NSF MREFC policies, an additional amount not to exceed \$8 million may be made available to TACC in the form of supplemental funding to this award to advance the design of the Phase 2 leadership-class system.

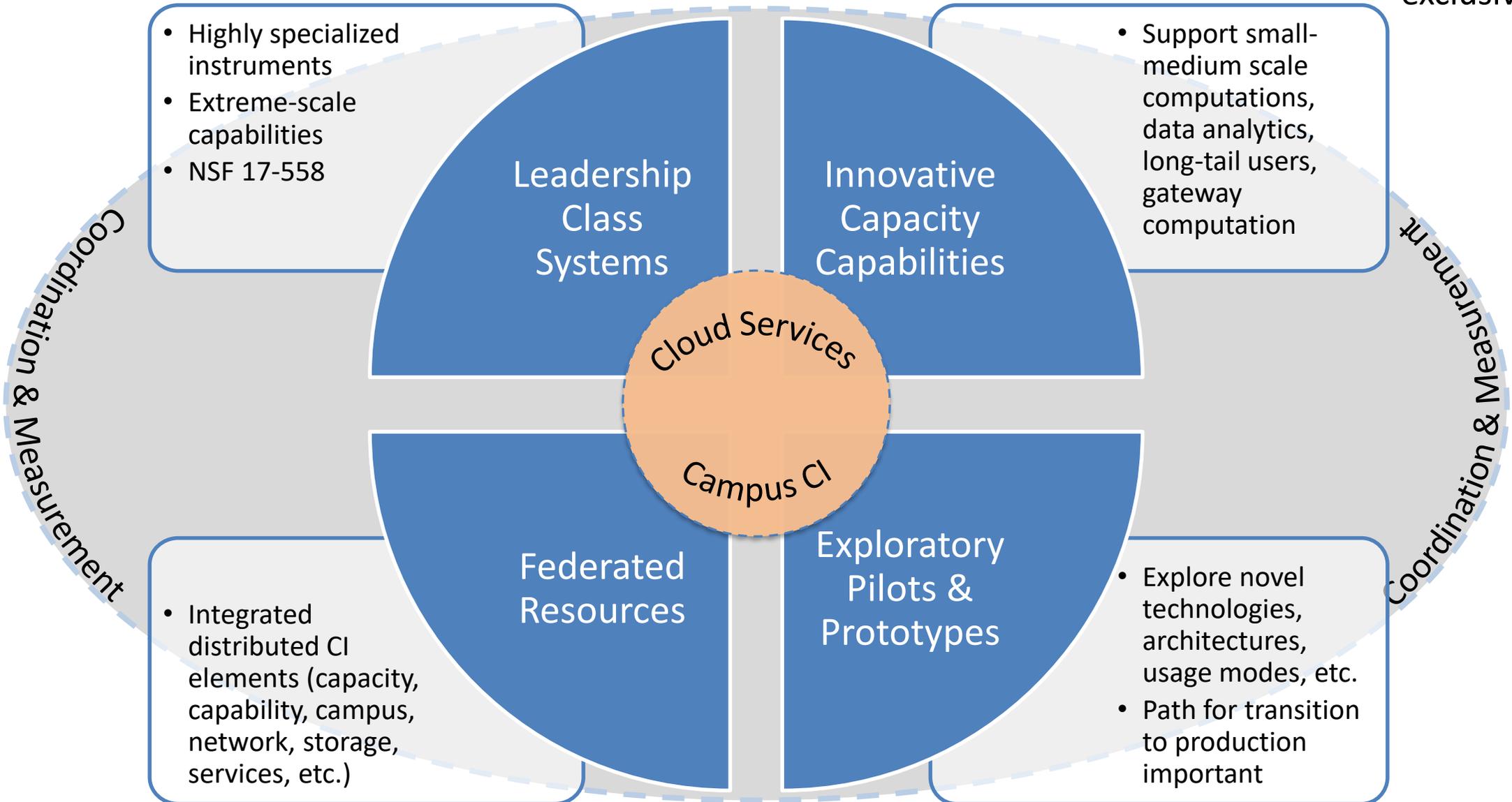


- To be deployed 2019
 - Primary compute: DellEMC and Intel
 - 35-40 PetaFlops Peak Performance
 - Interconnect: Mellanox HDR & HDR-100
 - Fat Tree topology, 200Gb/s bet. switches.
 - Storage: DataDirect Networks
 - 50+ PB disk, 3PB Flash, 1.5TB/sec peak.
 - Single Precision Subsystem: Nvidia
 - Front end for data movers, workflows
- Potential 10x Phase 2 system.



Rethinking CI – Computational Infrastructure*

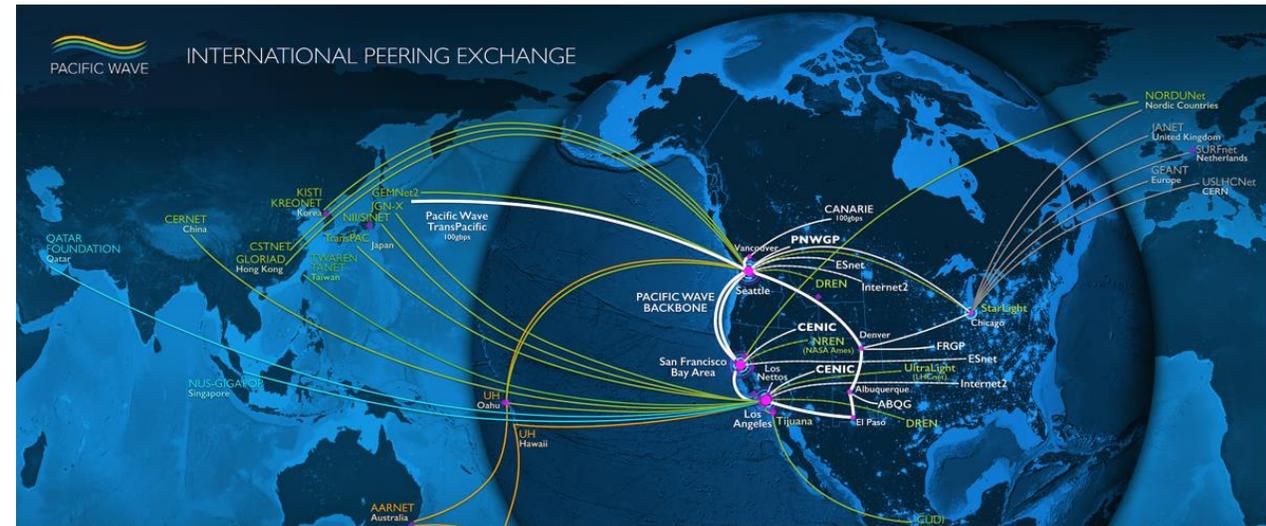
* Not mutually exclusive.



Laying a Robust, Secure Foundation: Networking & Cybersecurity

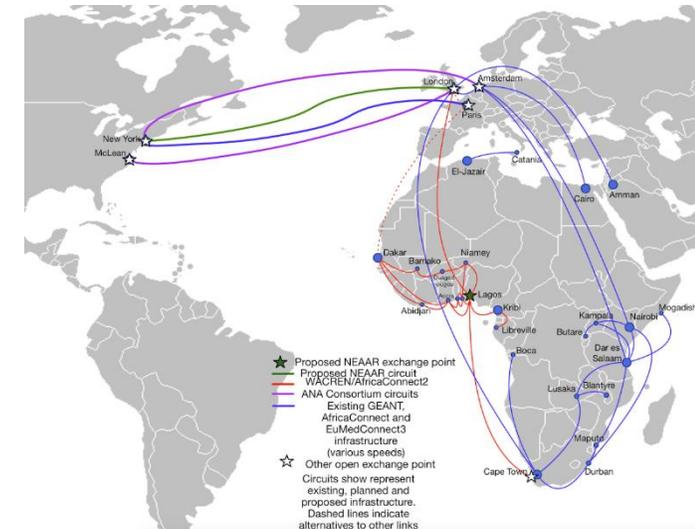
- Networking as a fundamental layer and underpinning of the CI Continuum
 - CC* - Campus Cyberinfrastructure
 - IRNC – International R&E Network Connections
 - Includes performance flow measurement, monitoring, training

- Distinct cybersecurity requirements of science and engineering
 - Secure and Trustworthy Cyberspace (SaTC)
 - Focus on Transition to Practice (TTP) projects, and use in operational CI environments
 - Cybersecurity Innovation for Cyberinfrastructure (CICI)

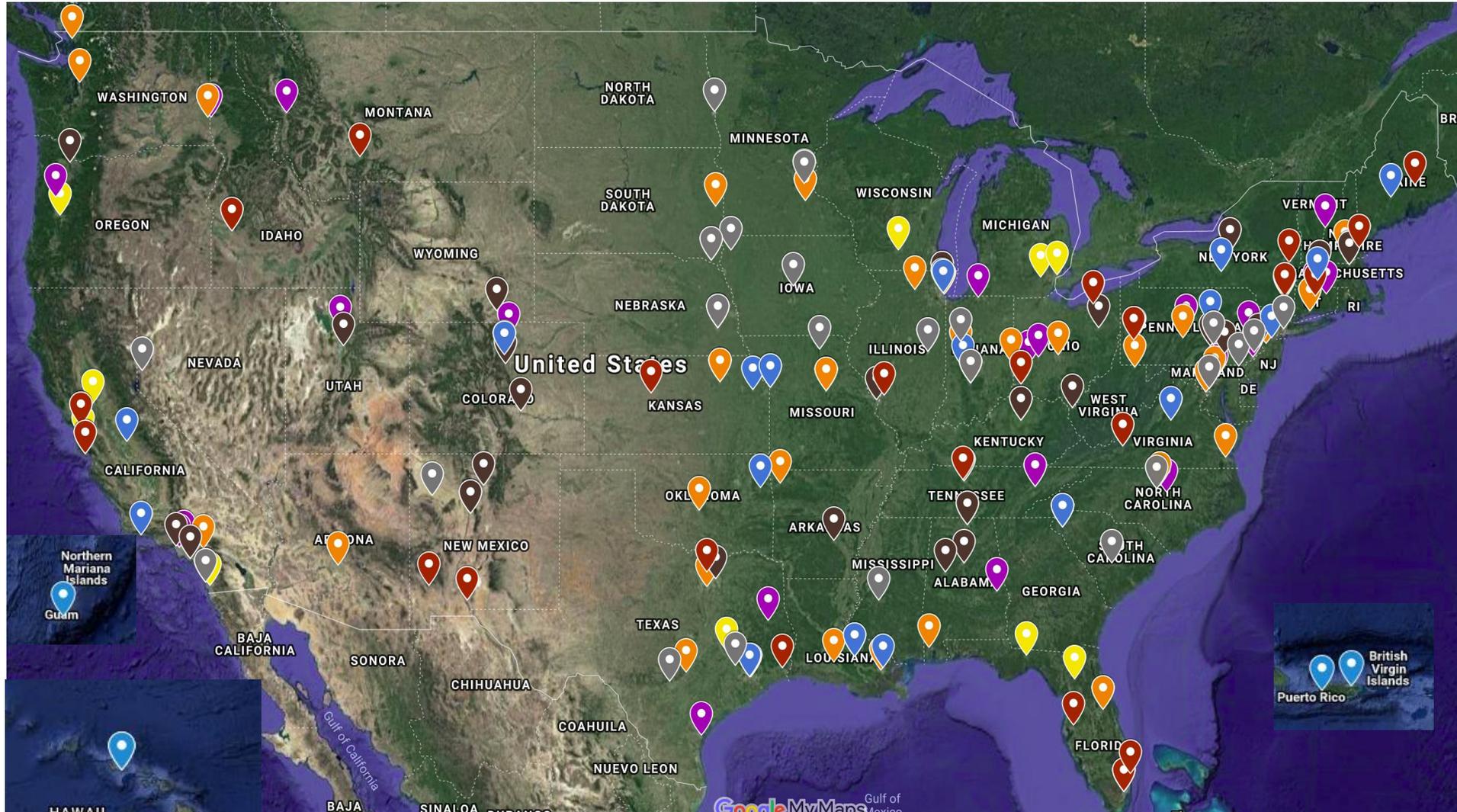


CTSC

CENTER FOR TRUSTWORTHY
SCIENTIFIC CYBERINFRASTRUCTURE
The NSF Cybersecurity Center of Excellence



Building the Research Network Substrate at the Campus Level CC* (2012-2018): ~250 awards across 44 states



Outline

NSF/OAC
Update

Emerging
Science &
Role of CI

Rethinking the
CI Ecosystem

Conclusion



Some Personal Observations

- Unprecedented opportunities for insights, discoveries: **CI poised for transformative impact**
- Compute and data are essential to all areas of science, engineering: **easy/pervasive/cost-effective access to CI is critical**
- Dramatic changes in application, method, technology, CI solution landscapes: **an inflection point in cyberinfrastructure**
- Growing demands across the spectrum; flat budgets: **how to maximize the impact of NSF investments? How to leverage other investments (including industry)?**
- CI ecosystem is a national asset: **where can NSF investments add unique value as part of a larger CI ecosystem?**



Conclusion

- Science and society are being transformed by compute and data
 - a connected, robust and secure cyberinfrastructure ecosystem is essential
- Rapidly changing application requirements; resource and technology landscapes
 - Our cyberinfrastructure ecosystem must evolve in response
- Lets build a cyberinfrastructure ecosystem aimed at transforming science



Join the conversation

- OAC Webinar Series
 - 3rd Thursday @ 2PM ET
- OAC Newsletter
- Follow us on Twitter @NSF_CISE

Stay informed

- Join the OAC, CISE Mailing Lists
 - Learn about NSF events, programs, webinars, etc.
- Send email to:
 - oac-announce@listserv.nsf.gov
 - cise-announce-subscribe-request@listserv.nsf.gov

Get involved

- Reviews proposals, serve on panels
- Visit NSF, get to know your programs and Program Officers
- Participate in NSF workshops and visioning activities
- Join NSF: serve as Program Officer, Division Director, or Science Advisor

NSF Office of Advanced Cyberinfrastructure (OAC) Newsletter

Table of Contents

- [About the Office](#)
- [Project Highlights](#)
- [OAC Program and Updates](#)
- [Related Events/Programs](#)
- [Subscribe to OAC Mailing List](#)



The screenshot shows the Twitter profile for NSF Comp & Info (@NSF_CISE). The profile picture is the NSF logo. The bio reads: "Exploring the frontiers of computing" and "Arlington, Virginia · nsf.gov/dir/index.jsp?..." The statistics show 1,075 tweets, 27 following, and 2,050 followers. A "Following" button is visible in the top right corner of the profile card.



“Make no little plans; They have no magic to stir men's blood ...”

Daniel H. Burnham, Architect and City Planner Extraordinaire, 1907.

*“If you want to travel fast, travel alone;
if you want to travel far, travel together”*

African Proverb.

THANKS!

Manish Parashar

Office Director, Office of Advanced Cyberinfrastructure

Email: mparasha@nsf.gov

To subscribe to the OAC Announce Mailing List

Send an email to: OAC-ANNOUNCE-subscribe-request@listserv.nsf.gov

