Transforming Science in the 21st Century: A Vision for a National Cyberinfrastructure Ecosystem

Manish Parashar
Office Director

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

PRAGMA 37, San Diego, CA
September 13, 2019

Outline

NSF Mission & Priorities

OAC Overview & Update

Towards a National CI Ecosystem

Conclusion

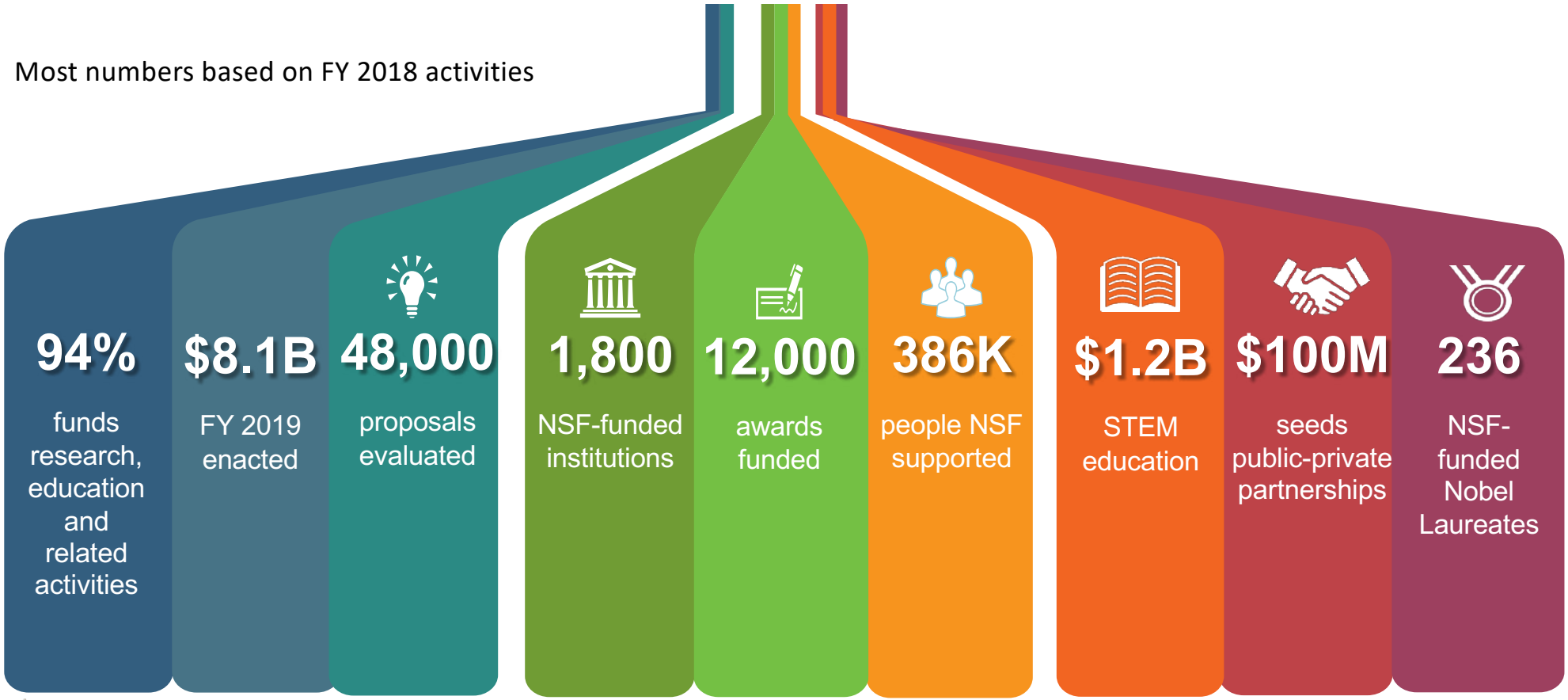
Outline


NSF Mission & Priorities
OAC Overview & Update
Towards a National CI Ecosystem
Conclusion

The National Science Foundation

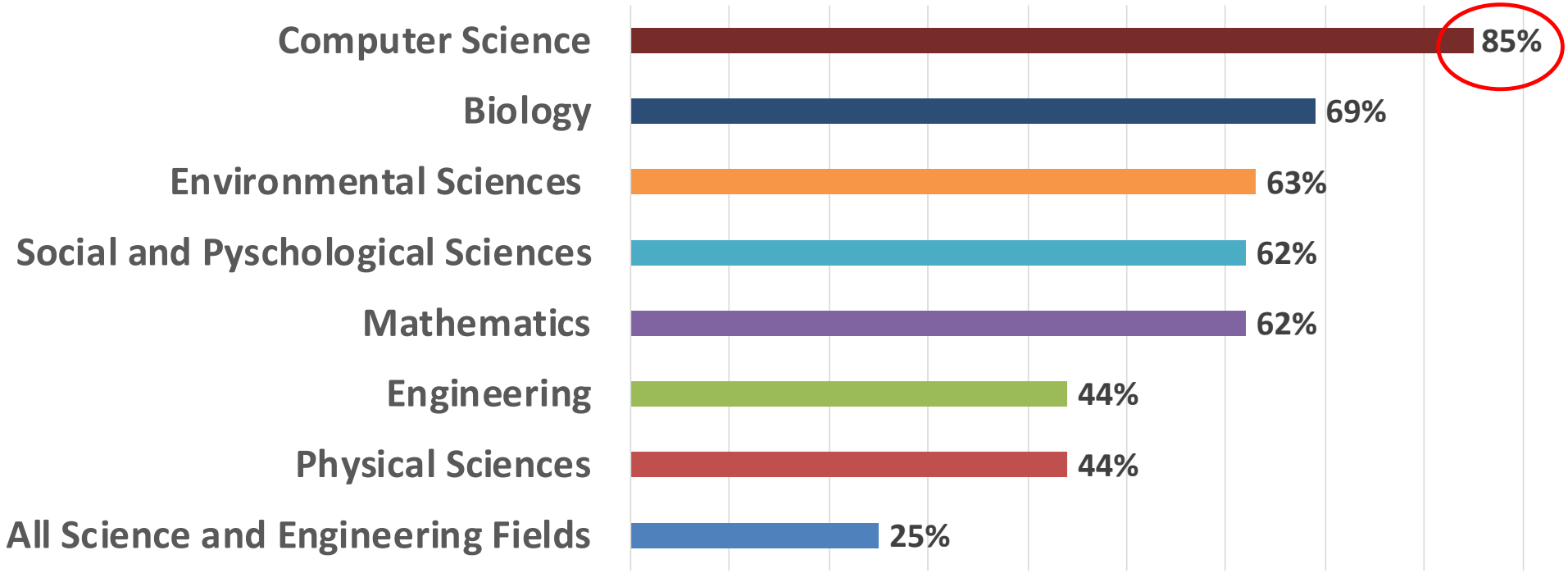
NSF by the Numbers

Most numbers based on FY 2018 activities



NSF Supports All Areas of Fundamental Research

NSF support as a percentage of total federal support for basic academic research



Source: NSF/NCSES, "Survey of Federal Funds for Research and Development." In FY20 NSF Budget Request to Congress

NSF Big Ideas

RESEARCH IDEAS

Harnessing Data for 21st Century Science and Engineering

Work at the Human-Technology Frontier: Shaping the Future

Navigating the New Arctic

Windows on the Universe: Multi-messenger Astrophysics

Quantum Leap: Leading the Next Quantum Revolution

Understanding the Rules of Life: Predicting Phenotype

PROCESS IDEAS

Mid-scale Research Infrastructure

NSF 2026

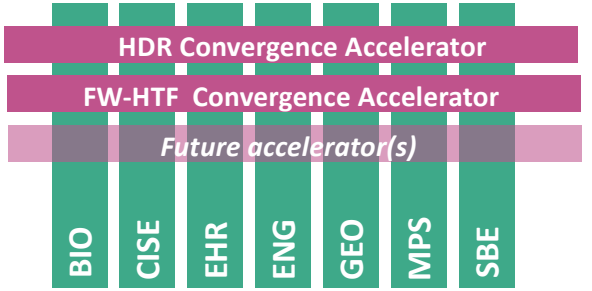
Growing Convergence Research at NSF

NSF INCLUDES: Enhancing STEM through Diversity and Inclusion

" … 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. "

Convergence Accelerators

Accelerating Discovery through Convergence Research



Cyberinfrastructure is a key enabler for NSF Big Ideas

NSF Big Ideas: full steam ahead in FY 19

- Convergence research: many disciplines required
- Budget model: 5-year funding, $30M/idea/yr, *outside* directorates

Harnessing the Data Revolution (HDR)

- HDR: TRIPODS Phase I (2/19)
- HDR: Institutes for Data-Intensive Research in Science and Engineering - Frameworks (2/19); Ideas Labs (12/18)
- HDR: Data Science Corps (DSC) (10/18)

Future of Work at the Human-Technology Frontier (FW-HTF)

- FW-HTF: Core Research (2/19)
- "advancing fundamental understanding of future work, and potential improvements to work, workplaces, workforce preparation, or work outcomes for workers and society"

Quantum Leap (QL)

- QL: Challenge Institutes (2/19)
- QL: Idea Incubator for Transformational Advances in Quantum Systems (10/18)
- QL: Quantum Materials Science, Engineering, and Information (8/18)

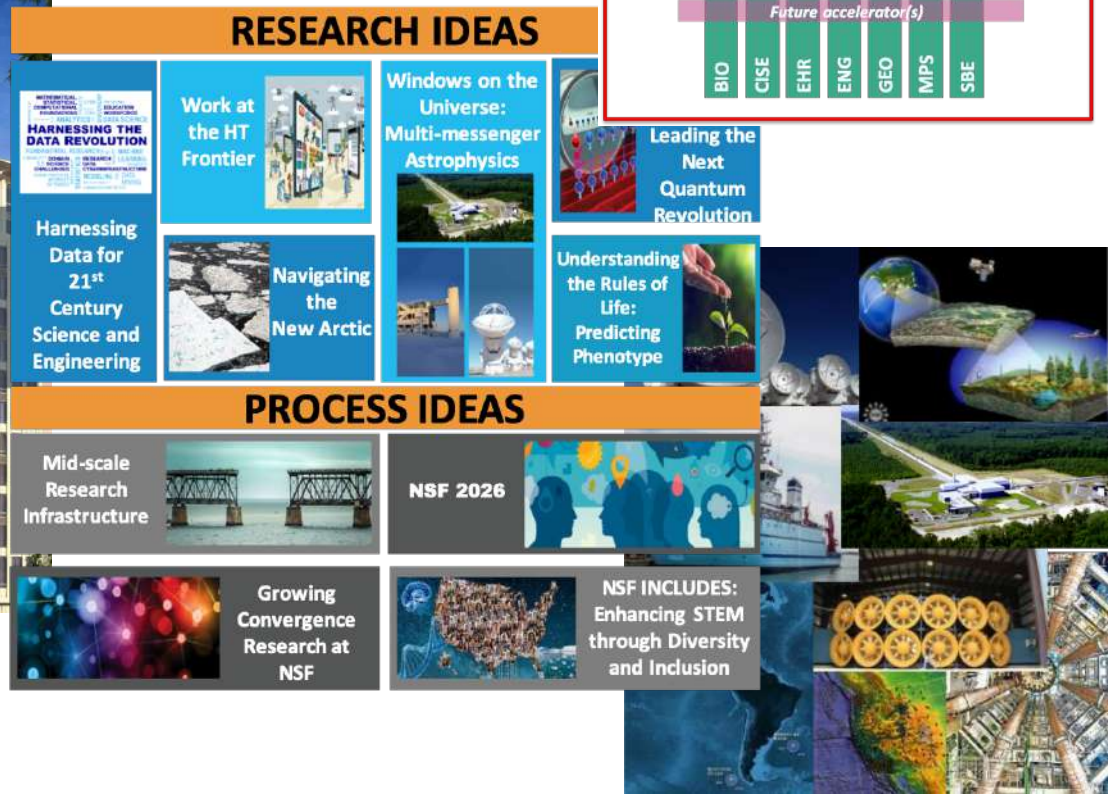
Mid-scale Research Infrastructure

- Mid-scale Research Infrastructure-2 (12/18)
- Mid-scale Research Infrastructure-1 (11/18)

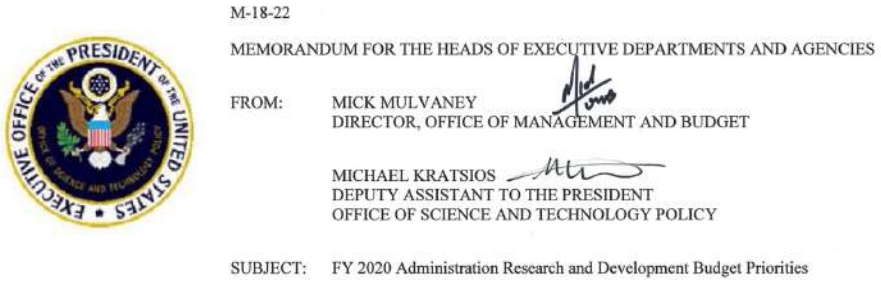
Outline


NSF Mission & Priorities
OAC Overview & Update
Towards a National CI Ecosystem
Conclusion

Cyberinfrastructure is Central NSF's Mission & Priorities



… and aligned with US Administration and Congressional Priorities



M-18-22

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

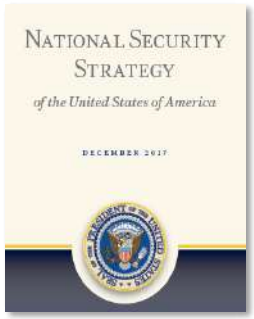
FROM: MICK MULVANEY
DIRECTOR, OFFICE OF MANAGEMENT AND BUDGET

MICHAEL KRATSIOS
DEPUTY ASSISTANT TO THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY

SUBJECT: FY 2020 Administration Research and Development Budget Priorities

FY 2020 R&D Budget Priorities Memo

"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, … Agencies should prioritize investment in research and infrastructure to maintain U.S. leadership in strategic computing, from edge devices to high-performance computing, … use of embedded sensors, data analytics, and machine learning"



National Security Strategy

National Defense Strategy

National Quantum Initiative Act

AI Executive Order

... and aligned with US Administration and Congressional Priorities

M-18-22

MEMORANDUM FOR THE HE

FROM: MICK MULVANE
DIRECTOR, OFFI

MICHAEL KRAT
DEPUTY ASSIST
OFFICE OF SCIE

SUBJECT: FY 2020 Administ

FY 2020 R&D Budget Prior

"Agencies should invest in fun
research, including machine le
systems, and applications at t
frontier. Agencies should prio
… Agencies should prioritize i
and infrastructure to maintair
strategic computing, from edg
performance computing, … u
data analytics, and machine le

EXECUTIVE OFFICE OF THE PRESIDENT
WASHINGTON, D.C.

August 30, 2019

M-19-25

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: RUSSELL T. VOUGHT
ACTING DIRECTOR, OFFICE OF MANAGEMENT AND BUDGET

DR. KELVIN K. DROEGEMEIER
DIRECTOR, OFFICE OF SCIENCE AND TECHNOLOGY POLICY

SUBJECT: Fiscal Year 2021 Administration Research and Development Budget Priorities

Artificial Intelligence, Quantum Information Science, and Computing: Departments and agencies should prioritize basic and applied research investments that are consistent with the 2019 *Executive Order on Maintaining American Leadership in Artificial Intelligence*[5] and the eight strategies detailed in the 2019 update of the *National Artificial Intelligence Research and Development Strategic Plan*.[6] Consistent with the 2018 *National Quantum Initiative Act*[7] and the 2018 *National Defense Authorization Act*,[8] departments and agencies should prioritize R&D advancing fundamental QIS, building and strengthening the workforce, engaging industry, and providing infrastructure supporting QIS while coordinating relevant activities to ensure intelligence, defense, and civilian efforts grow synergistically. In terms of computing, departments and agencies should work together to explore new applications in and support R&D for high performance future computing paradigms, fabrication, devices, and architectures

efense Strategy

AI Executive Order

NSF Office of Advanced Cyberinfrastructure (OAC)

Directorate for Computer & Information Science & Engineering (CISE)

Mission: Foster a cyberinfrastructure ecosystem to transform science and engineering research … through Research CI and CI research

* Manish Parashar
Office Director

Amy Friedlander
Deputy Office Director

* IPA Appointment

Bob Chadduck

Amy Walton

* Vipin Chaudhary

Kevin Thompson

Bill Miller
Science Advisor

Beth Plale*
Science Advisor
Public Access

Ed Walker

* Stefan Robila

* Micah Beck

* Alan Sussman

Alejandro Suarez
Cooperative Agreements

We're hiring!

NSF Office of Advanced Cyberinfrastructure (OAC)

Foster a cyberinfrastructure ecosystem to transform science and engineering research…

… through Research CI and CI research

$224M FY 2018 research budget

950 proposals

305 awards

32% Success Rate

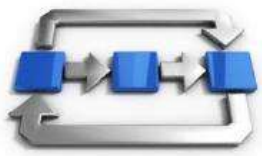
Source: https://dellweb.bfa.nsf.gov/starth.asp

People, organizations, and communities

Coordination & User support

Gateways, Hubs, and Services

Data Infrastructure



Software and Workflow Systems

CI-Enabled Instrumentation

Pilots, Testbeds

Computing Resources

R&E Networks, Security Layers

Cloud Resources & Services

NSF Office of Advanced Cyberinfrastructure (OAC)

Foster a cyberinfrastructure ecosystem to transform science and engineering research…

… through Research CI and CI research

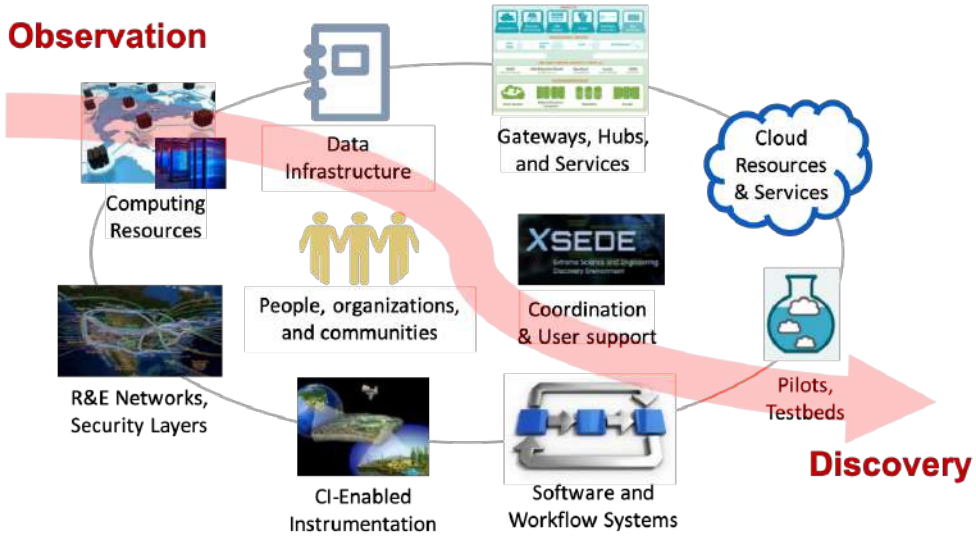
$224M FY 2018 research budget

950 proposals

305 awards

32% Success Rate

Source: https://dellweb.bfa.nsf.gov/starth.asp



Transforming OAC's Investment Strategy

Towards an Integrative, Holistic and Robust CI Ecosystem

Computing	HPC, Services (XSEDE) MRI (clusters)	NSF 19-587 Advanced Computing Systems & Services: *Adapting to Rapid Evolution of S&E Research*	**Expanding and Integrating**
Data	Data Building Blocks (DIBBS)	Cyberinfrastructure for Sustained Scientific Innovation (CSSI)	
Software	Software Infrastructure (SI^2)		
Networking & Cybersecurity	Campus CI (CC*) International Networks (IRNC) Cybersecurity (CICI, SATC)	OAC Core Research Program (NSF 18-589)	**Translational Research**
		CI Centers of Excellence (TrustedCI, Facility CI CoE Pilot), NFS Public Access	**Community Practices**
Learning & Workforce Dev.	CyberTraining CAREER, CRII	CI for Emerging S&E Research (CESER)	**Emerging Opportunities**

Outline


NSF Mission & Priorities
OAC Overview & Update
Towards a National CI Ecosystem
Conclusion

Evolving Science, CI Landscapes

Evolving Science/Engineering Landscape

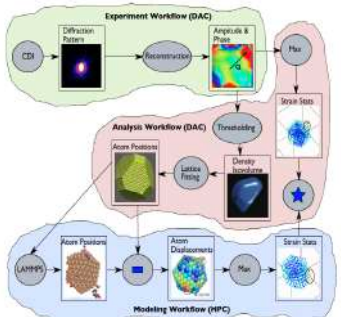
- Large scales / Complex, dynamic workflows
- Data-driven and data intensive
  - Streaming data from observatories, instruments
  - Increasing use of ML
- Heightened emphasis on robust results

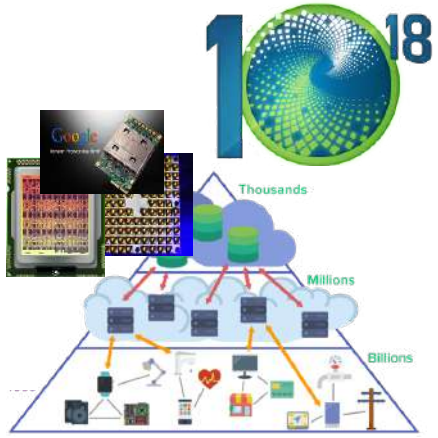
Evolving Technology Landscape

- Diverse / disruptive technologies
- Role of (non-traditional) software in taming complexity
- Novel paradigms / Increasing role of clouds / Growing capabilities & capacities at the edges

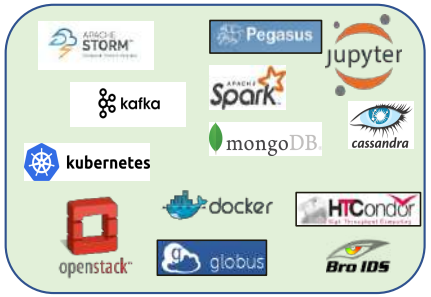
Our cyberinfrastructure ecosystem must evolve!

End-to-end Workflows





Instruments, Observatories, Experimental Facilities



Rethinking CI: Responding to a Changing Landscape

Rapid (disruptive) changes in S&E and CI landscapes → *Our cyberinfrastructure ecosystem must evolve!*

- NSCI EO, NASEM Report, CI2030 RFI, Workload studies, COV, etc. (2015-2018)
- Future of CI Workshop (May '18)
- Facilities CI Pilot (September '18)
- Exploring Clouds for S&E Pilot (November '18)
- Large Facilities Workshop – Special Event on Future of Facility CI (April '19)
- National CI Coordination Services Workshop (June '19)
- NSF/DOE Software Interoperability "Hackathon" (July '19)
- 2019 Cyberinfrastructure for Facilities Workshop (September'19)
- CSSI Future Directions Workshop (October 2019)
- *Ongoing strategic planning activities across all aspects of the CI ecosystem*

Draft
04/19

Transforming Science Through Cyberinfrastructure

NSF's Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century

"…. an agile, integrated, robust, trustworthy and sustainable CI ecosystem that drives new thinking and transformative discoveries in all areas of S&E research and education"

Community-informed blueprints focused on different elements of the CI ecosystem

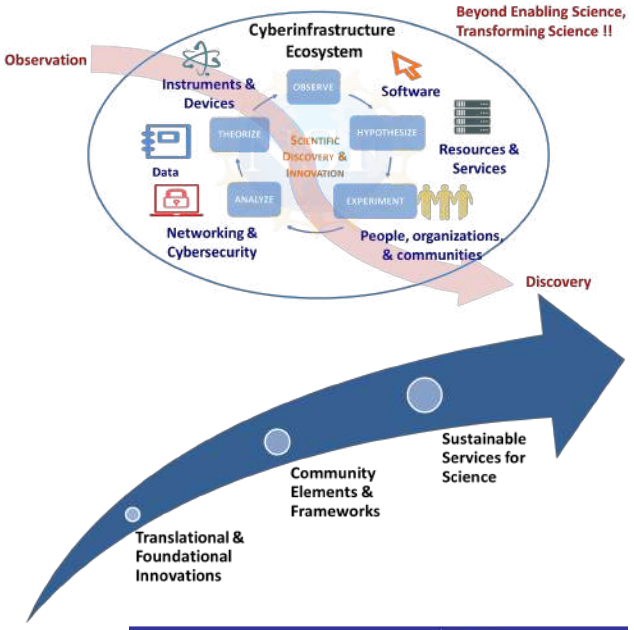
http://go.usa.gov/xm8bU

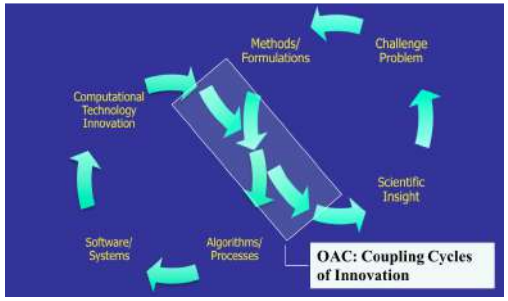
A new vision…

An agile, integrated, robust, trustworthy and sustainable CI ecosystem that drives new thinking and transformative discoveries in all areas of S&E research and education.

Overarching principles:

- View CI more holistically
- Support translational research
- Balance innovation with stability
- Couple discovery and CI innovation cycles
- Improve usability





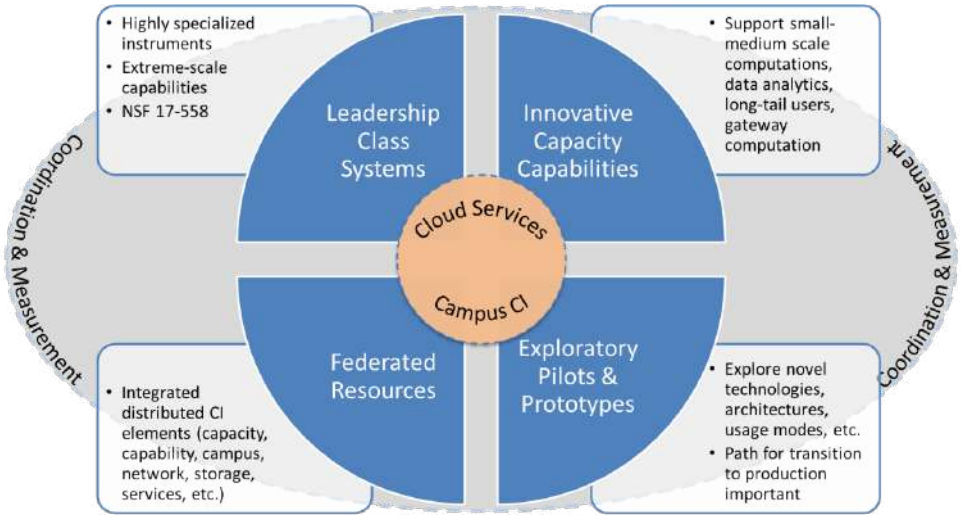
Computational blueprint.

Implement extensions and enhancements to current investments and new programs and opportunities in 2019 and beyond.

Two strategies

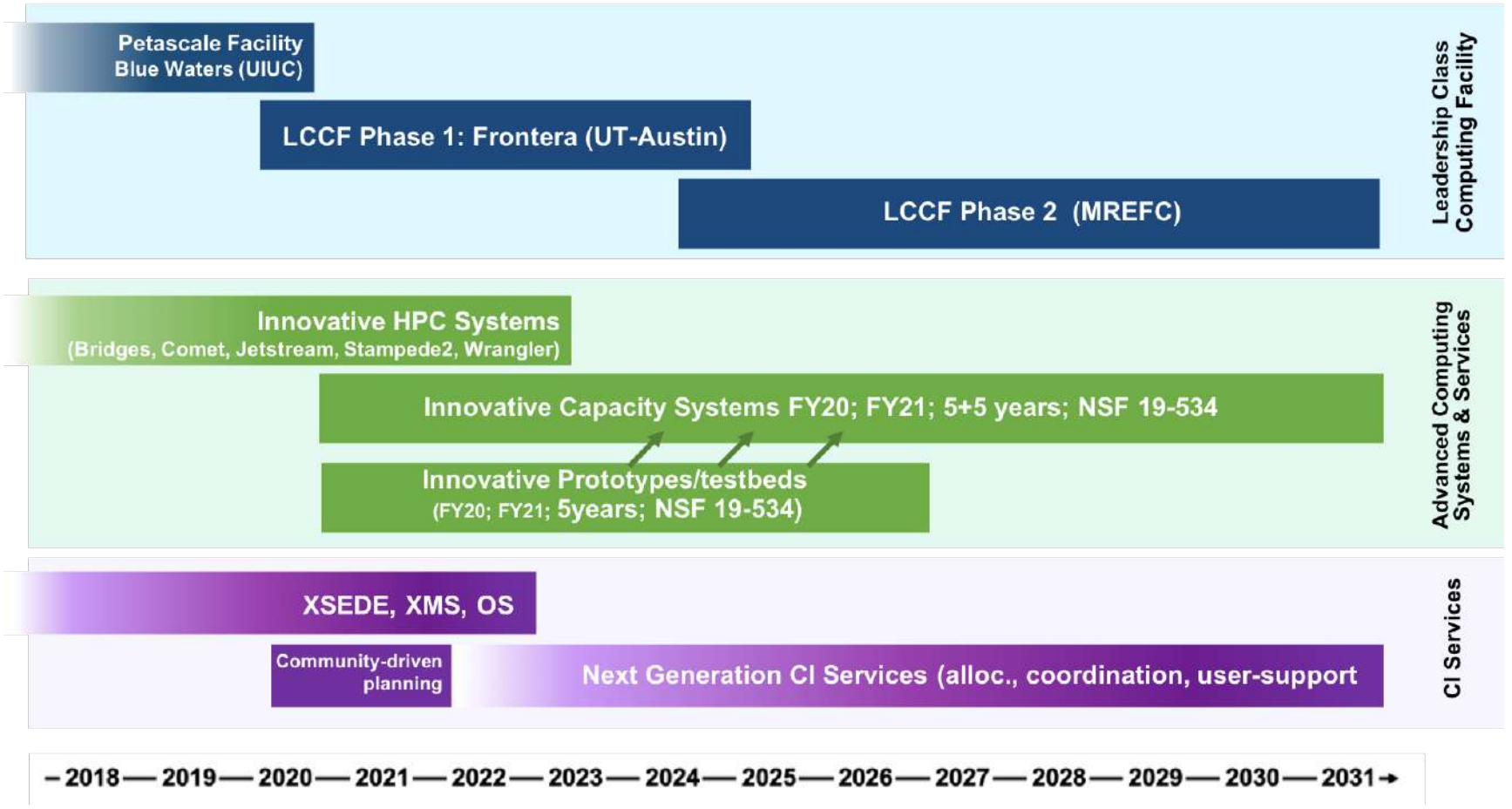
Deploy a balanced computational ecosystem that supports broad and diverse requirements, users and usage modes

Achieve maximal impact from the array of computational capabilities and expertise



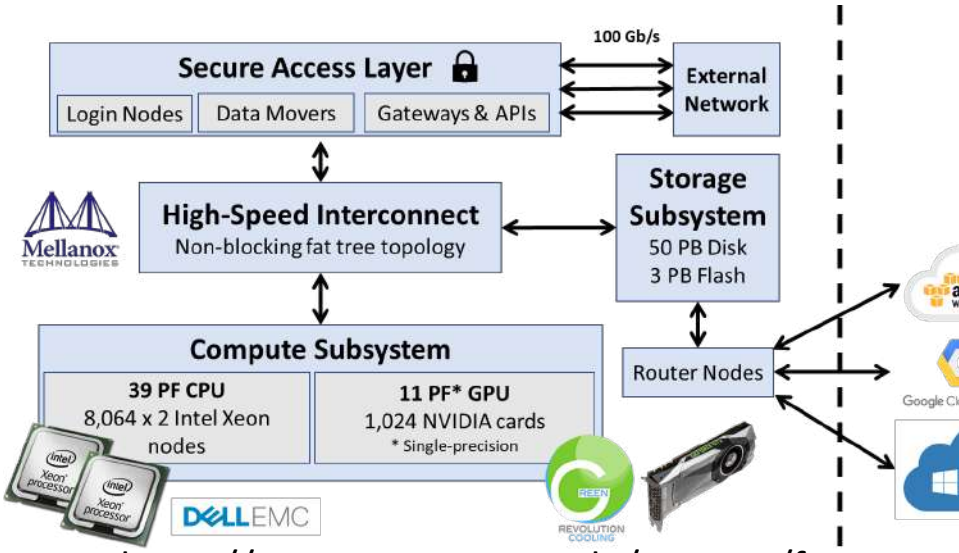
First of several blueprints focused on different elements of the CI ecosystem

Computational Blueprint: Elements of a balanced computational ecosystem



Computation for the Endless Frontier

FRONTERA



THE OHIO STATE UNIVERSITY · PRINCETON UNIVERSITY · Caltech · Georgia Tech · THE UNIVERSITY OF UTAH · THE UNIVERSITY OF CHICAGO · UC DAVIS UNIVERSITY OF CALIFORNIA

https://www.tacc.utexas.edu/systems/frontera

Early user access started in May 2019

#5 on Top500 (06/19)

Frontera will be:

- A leadership-class computational instrument with the **broadest utility for all of S&E applications**
- The **largest CPU system** on a **US academic campus**
- A national asset that **complements** other leadership-class computing investments in the US research ecosystem

Advanced Computing Systems & Services (ACSS): FY19 Awards

Category I

Computing without Boundaries: Cyberinfrastructure for the Long Tail of Science

- *Increased capacity and performance for users of batch-oriented and science gateway computing;* integration with the public cloud and the Open Science Grid
- *PI: Michael Norman, University of California-San Diego*

Bridges-2: Scalable Converged Computing, Data, and Analytics for Rapidly Evolving Science and Engineering Research

- *High capacity, large memory system targeting high-performance data driven analytics with machine learning / deep learning / artificial intelligence applications*
- *PI: Nicholas Nystrom, Carnegie-Mellon University*

Category II

Ookami: A high-productivity path to frontiers of scientific discovery enabled by exascale system technologies

- *Explores the Fujitsu A64fx processor with ultra-high memory bandwidth to better support memory-intensive applications*
- *PI: Robert Harrison, SUNY at Stony Brook*

Clouds and the NSF CI Ecosystem

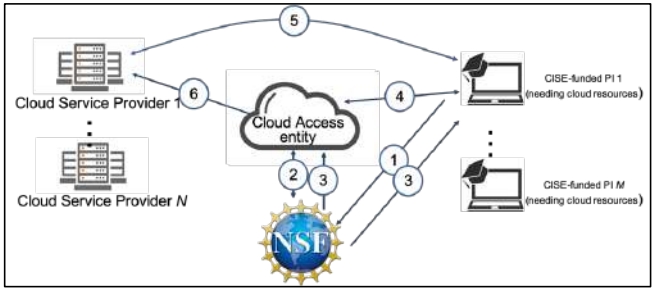
- CISE CloudAccess
  - Explore models for providing (CISE) researchers access to Cloud services

  CloudBank, UCSD, UC B, UW, PI, M. Norman

- Exploring Clouds for Acceleration of Science (ECAS)
  - Explore clouds as platforms for leading edge science

  Internet2, PI, H. Pfeffer

- CC*: Clouds and Campus Computing
  - Integrated Cloud services/expertise into campus CI



NSF INTERNET2 aws Google Cloud Platform ORACLE

www.internet2.edu/ecas

Data-Intensive Discovery Pathways – The "missing middle"

Facility / Discipline Specific CI Solutions

Data Sources

Instrument/Facility Portals & Data streams

Sources & Repositories

Data-Driven S&E

Opportunities

New science drivers, users and usage modes
End-to-end Workflows

Challenges

Data Access: Realtime, streaming, on-demand

Data Discovery: Knowledge networks, Intelligent data delivery

Data Fusion: Data integration & interoperability

Science Outcomes & Results Dissemination

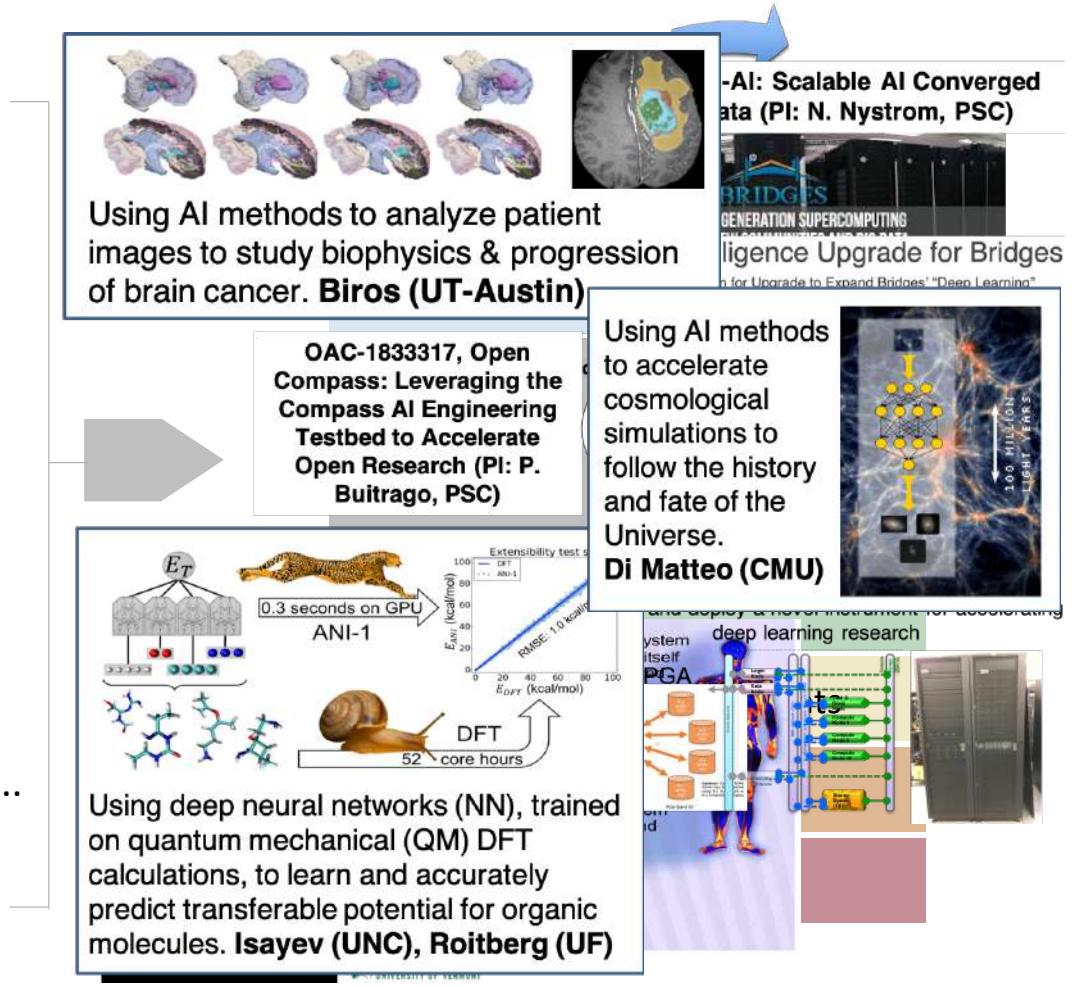
Data lifecycle

Transdisciplinary Enabling CI

CI Ecosystem: Computing, Data, Networking, Software, People

Artificial Intelligence & Cyberinfrastructure

- Providing infrastructure, services for AI
- Supporting/accelerating large-scale AI, S&E+AI
- Enabling new models and paradigms for S&E discovery
- Fostering intelligent (self-managing) CI systems and services
  - Robust, secure, performant, agile, resilient, …..
  - Platform for explainability, fairness, trust, privacy, …



National Strategic Computing Update

- ***Envisioning a national CI ecosystem:*** *"Ecosystem"* view of computing, data investments
  - Cross-agency sharing: resources; software/services, practices; LWD investments
- ***Update NSCI Objectives:***
  - Engage community
    - RFI on Strategic Computing
    - Community of Interest workshop (August 5 & 6)
    - BoF at SC'19
  - Produce and disseminate report: *Strategic Computing Update: Enabling the Future of Computing*

Outline


NSF Mission & Priorities
OAC Overview & Update
Towards a National CI Ecosystem
Conclusion

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
- NSF/OAC strives to build a cyberinfrastructure ecosystem aimed at transforming science

Join the conversation

- OAC Webinar Series
  - 3rd Thursday @ 2PM ET
- OAC Newsletter
- OAC Townhalls (CASC, LFW, PEARC, SC)
- 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

"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