Supporting Research Communities with XSEDE

John Towns
PI and Project Director, XSEDE
Executive Director, Science & Technology, NCSA
Deputy CIO, Research IT, University of Illinois
jtowns@ncsa.illinois.edu

XSEDE
Extreme Science and Engineering Discovery Environment

License terms

- Please cite as: Towns, John. *Supporting Research Communities with XSEDE,* October 2015, [http://www.slideshare.net/jtownsil/supporting-research-communities-with-xsede-54245889]
- ORCID ID: http://orcid.org/0000-0001-7961-2277
- Except where otherwise noted, by inclusion of a source URL or some other note, the contents of this presentation are © by the Board of Trustees of University of Illinois. This content is released under the Creative Commons Attribution 3.0 Unported license (http://creativecommons.org/licenses/by/3.0/). This license includes the following terms: You are free to share – copy and redistribute the material in any medium or format; and to adapt – remix, transform, and build upon the material for any purpose, even commercially.
- This can be done under the following conditions: attribution – You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

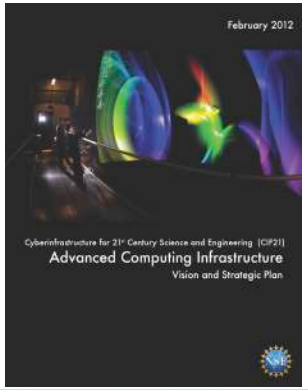
XSEDE in Context

- XSEDE is an award made under the eXtreme Digital solicitation
  - TeraGrid Phase III: eXtreme Digital Resources for Science and Engineering (XD), NSF 08-571
- Consistent with NSF's vision and strategy statements

- *NOTE: internationalization*
  - *Cyberinfrastructure == eScience Infrastructure == eInfrastructure == Digital Research Infrastructure*

NSF's Strategic Planning Documents

- *Investing in Science, Engineering, and Education for the Nation's Future - National Science Foundation Strategic Plan for 2014-2018*
  - www.nsf.gov/pubs/2014/nsf14043/nsf14043.pdf
  - Vision: *A Nation that creates and exploits new concepts in science and engineering and provides global leadership in research and education.*
- *Cyberinfrastructure Framework for 21st Century Science and Engineering*
  - www.nsf.gov/cif21
- NSF's *Advanced Computing Infrastructure: Vision and Strategic Plan*
  - www.nsf.gov/pubs/2012/nsf12051/nsf12051.pdf



Original Motivation for XSEDE

- Scientific advancement requires a variety of resources and services
  - and thus availability of comprehensive cyberinfrastructure composed of heterogeneous digital resources
- Computational science better served if we leverage aggregate expertise of a small number of leading institutions
  - not fully centralized at a single institution; not fully decentralized
  - full centralization less agile, single point of failure
  - different sites each offer a unique perspective and talent to address a particular suite of community needs
  - best to have several leadership perspectives for addressing the broad range of disciplinary needs

XSEDE – accelerating scientific discovery

XSEDE's Vision:

a world of digitally-enabled scholars, researchers, and engineers participating in multidisciplinary collaborations while seamlessly accessing computing resources and sharing data to tackle society's grand challenges.

XSEDE's Mission:

to substantially enhance the productivity of a growing community of scholars, researchers, and engineers through access to advanced digital services that support open research;

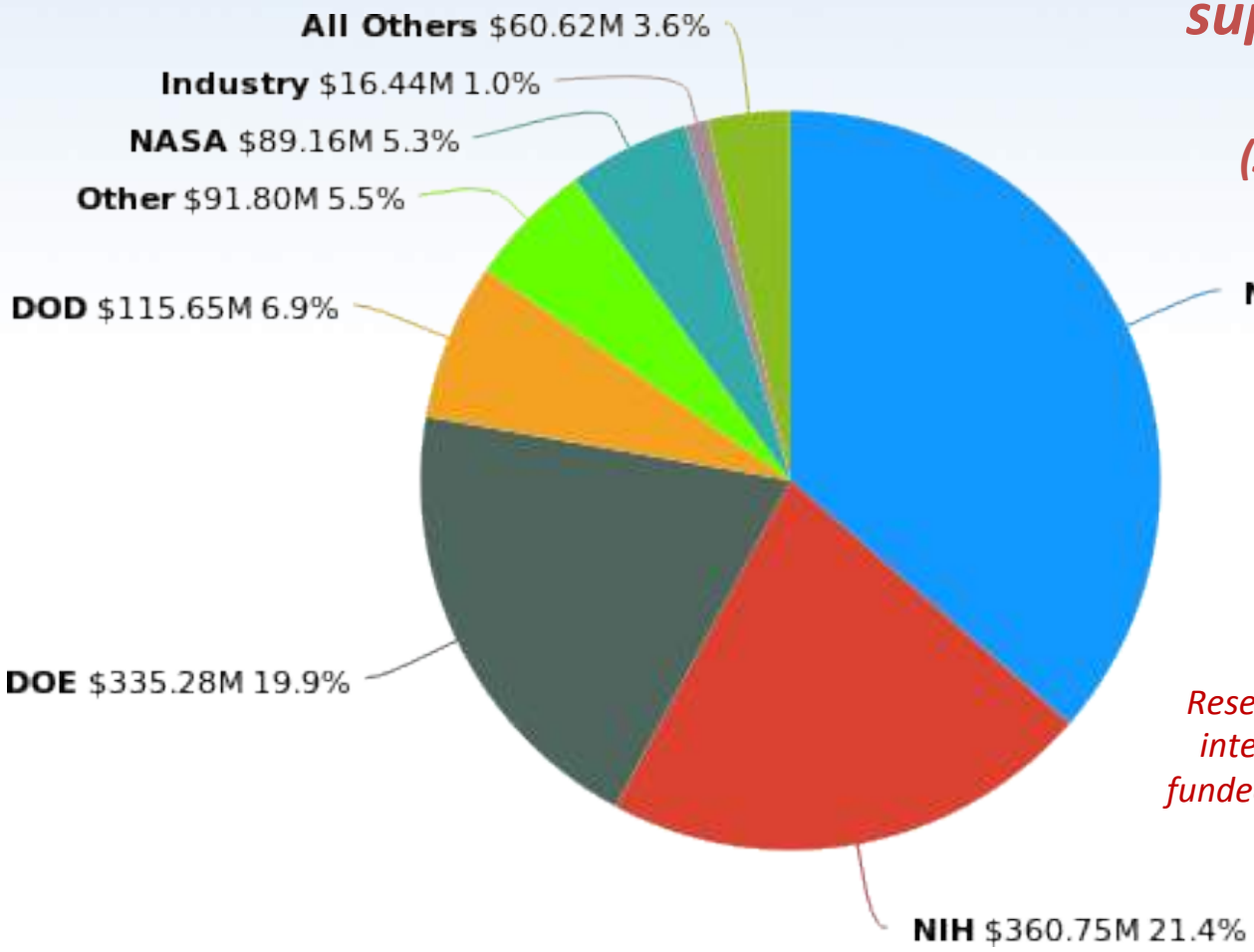
and to coordinate and add significant value to the leading cyberinfrastructure resources funded by the NSF and other agencies.

XSEDE Factoids: high-order bits

- 5 year, US$121M project
  - plus US$9M, 5 year Technology Investigation Service
    - separate award from NSF
  - option for additional 5 years of funding upon major review after PY3
- No funding for major hardware
  - coordinate, support and create a national/international cyberinfrastructure
  - coordinate allocations, support, training and documentation for >US$100M of concurrent project awards from NSF
- ~112 FTE /~240 individuals funded across 20 partner institutions
  - this requires solid partnering!

Total Research Funding Supported by XSEDE in Program Years 1-4

$1.68 billion in research supported by XSEDE in PY1-PY4
(July 2011-June 2015)



Research funding only. XSEDE leverages and integrates additional infrastructure, some funded by NSF (e.g. Track 2 systems) and some not (e.g. Internet2).

What is XSEDE?

- An ecosystem of advanced digital services accelerating scientific discovery
  - support a growing portfolio of resources and services
    - advanced computing, high-end visualization, data analysis, and other resources and services
    - interoperability with other infrastructures
- A virtual organization (partnership!) providing
  - dynamic distributed infrastructure
  - support services and technical expertise to enable researchers engineers and scholars
    - addressing the most important and challenging problems facing the nation and world
- More than just a project funded by the National Science Foundation
  - XSEDE is a path-finding experiment in how to develop, deploy and support e-science infrastructure

XSEDE Offers Efficient and Effective Integrated Access to a Variety of Resources

- Leading-edge distributed memory systems
- Very large shared memory systems
- High throughput systems, including Open Science Grid (OSG)
- Visualization engines
- Accelerators like GPUs and Xeon PHIs
- Virtualization
- Cloud-based resources

Many scientific problems have components that call for use of more than one architecture.

Direct interactions with the Community

- Facilitate broad range of ground-breaking research
  - provided in-depth support contributing to improved user productivity
  - supported over 15,000 publications to date
- Seamlessly integrate and retire resources
  - transition community smoothly
- Pursue new disciplinary areas
  - increasing the diversity of disciplines utilizing advanced digital services
- Campus Champions continue to reach new heights
  - over 250 Champions at more then 200 institutions
  - expanding program: Regional, Student, and Domain Champions

Centralized/Coordinated Services Provide Value Add

- User productivity enhancements
  - XSEDE User Portal, single sign-on, allocation processes
- Centralized/coordinated support services
  - coordination of problem resolution, extended support disciplinary breadth and depth
- National leadership function
- Training, Education, Outreach
  - national scope

Centralized/Coordinated Support Services

- Coordinated problem resolution
  - field and route over 10,000 tickets annually
  - work with Service Providers to resolve all reported issues
- Extended Collaborative Support Services (ECSS)
  - single, coordinated effort to bring the right expertise to bear on issues raised by any user on any resource(s)
  - no unnecessary replication across Service Providers
    - disciplinary breadth of expertise, allows coverage of domains composed of diverse sub-domains
  - Novel and Innovative Projects
    - support of emerging & innovative research
  - optimization of widely used community codes
    - prioritizing and coordinating effort
    - often optimized for multiple architectures
    - improving code substantially is as better than buying more hardware

Diverse ECSS Expertise Possible Because of Scale

- ***Fields of expertise***: astrophysics, bioinformatics, CFD, chemistry, computer science, climate modeling, engineering, **genomics**, hydrology, **humanities , machine learning,** molecular dynamics, **phylogenetics**, physics, seismology, **statistics**.
- ***Technologies:*** clusters, large shared memory systems, **MICs, GPUs**
- ***Languages:*** C, C++,Fortran, MPI, OpenMP, **Java, JavaScript,** shell programming, CUDA, OpenACC, **Python, R, MATLAB**

- ***Techniques***: benchmarking, **cloud computing**, Condor, **data mining, databases,** FFTs, finite element methods, grid generation, grid middleware**,** Lattice Boltzmann methods, libraries, linear algebra, Monte Carlo methods, parallel debugging, parallel I/O, petascale computing, scheduling, science gateways, visualization, **workflows**

Training, Education, Outreach

- Single set of programs of national scope
  - Training & Education
  - Underrepresented Community Engagement
  - Campus Champions
- Programs serve a more diverse community
  - single coordinated set of programs without competition
  - one consistent message and set of technical information makes it easier for technology adoption to spread organically
- Better ability to cover the entire nation in outreach:
  - XSEDE Conference
  - users in all 50 states, D.C., and US territories
  - Campus Champions – all 50 States
  - XSEDE staff physically located in 18 states + DC
- Over 32,000 training registrations over PY1-PY4!
  - HPCU and CI-Tutor, as well as center trainings, have been used in universities around the country to prepare students to use the nation's pre-eminent computational resources

Campus Collaborations and Communications

- Outreach to a diverse range of higher ed institutions
  - 2- and 4-year, MSI, EPSCoR, PhD granting universities
  - meet with CIOs, researchers, faculty, staff, and students
  - campus visits and discussions at conferences and workshops
- Establish institutional commitments
  - Champions identified for sustained local support
    - over 200 member institutions, over 250 Champions
  - Campus Bridging
    - 4 pilot sites; over 100 campuses using tools
    - working with Champions to enhance campus infrastructure
  - sustained curricular change
    - hosting workshops on campuses – 25 to date; 55 new courses
    - developing CS&E certificate programs; 3 new programs in place
  - increased under-represented community engagement
    - 54 minority faculty with allocations; 927 with portal accounts

Campus Champions:
Support as a local effort on the campus

- Champions are selected by administrators to assist local researchers, faculty and students
  - commit time of the local staff member to support this effort
- Source of local information on local and national HPC resources and services, including XSEDE
  - able to get users started with XSEDE access
- Attend XSEDE conference to expand skills
- Ability to apply to Champions Fellows Program
- Peer support community
- Champions are volunteers

MSU Champions are
Andrew Keen <keenandr@msu.edu>
and Yongjun Choi <choiyj@msu.edu>

Campus Engagement

- Historically the Campus Champions program has been focused on expanding breadth – 200th member institution announced recently
- Adds focus on expanding depth by engaging campuses through researchers, CIOs, VPRs, IT staff, trainers and educators
- Continue to foster the creation of subgroups to achieve “small discussion size” groups
  - regional champions have started
  - other subgroups based on common interests and goals

Campus Engagement

- Continue engaging CASC as a means to reach key campus CI decision makers
  - CASC provides mechanism to identify workforce development opportunities and best practices being adopted at the campus level
- Engage with organizations such as Internet2 to engage CIOs and VPRs
- Anecdotally, we have observed exemplar campuses having membership in multiple groups
  - Campus Champions
  - CASC
  - SP Forum
  - organizations such as ACI-REF, Great Plains Network (GPN), …

Campus Engagement

- Recruit additional Level 3 Service Providers to facilitate engagement by individual institutions in the national CI ecosystem
- Engage and collaborate with other complementary organizations
  - upcoming workshop: Improving Access to the National Computational Infrastructure
  - collaboration involving XSEDE (Campus Champions), ACI-REF, Great Plains Network, Internet2, OSG,…

Convenience Requirements will Always Increase

No, his mind is not for rent
To any god or government.
Always hopeful, yet discontent,
He knows changes aren't permanent,
But change is.

– Rush - Tom Sawyer

- Each generation of users requires more convenience than the former
- We must always be adding new capabilities while maintaining and extending existing reliability
- XSEDE has learned from the past
  - adds value in how we address going forward

Change is the only Constant

– Heraclitis 535BC-475BC

A Relevant Event Coming Soon: *ARCC16, March 22-24, 2016 @ NCSA*

- Advancing Research Computing on Campuses
  - for professionals involved in operating and supporting campus shared research computing infrastructure
- Planned topics:
  - best practices for running advanced computing resources in a higher education environment
  - the condo model
  - business models
  - collaboration with researchers
  - return on investment
  - interactions with national data centers and infrastructures
- Third event in series
  - 2014: first ARCC held @ NCSA; 85 attendees
  - 2015: second ARCC held at Clemson; 175 attendees
  - *2016: third ARCC returns to NCSA*

Questions?


Our reach will forever
exceed our grasp, but,
in stretching our horizon,
we forever improve our world.
XSEDE
Extreme Science and Engineering
Discovery Environment