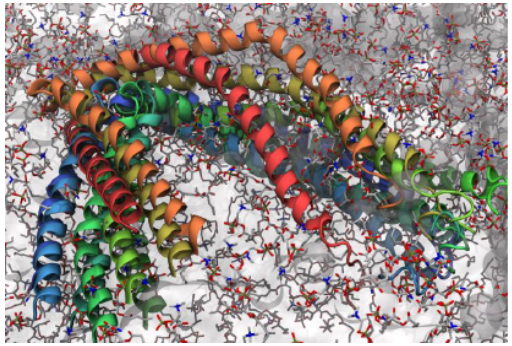
The Value of HPC for Scientific Research and the Roles of NSF and NIST

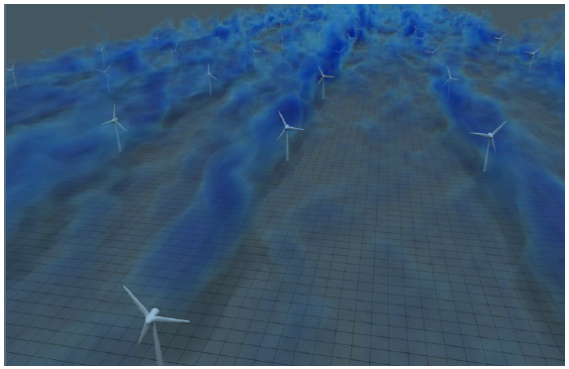
Rudi Eigenmann, Program Director
Division of Advanced Cyberinfrastructure (ACI)
NSF CISE
June 30, 2015

Computational Science Success Stories

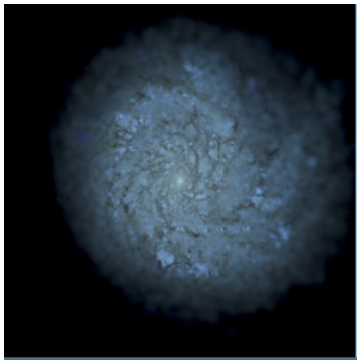
Thanks to computational methods and resources we can now better understand …



how dementia forms



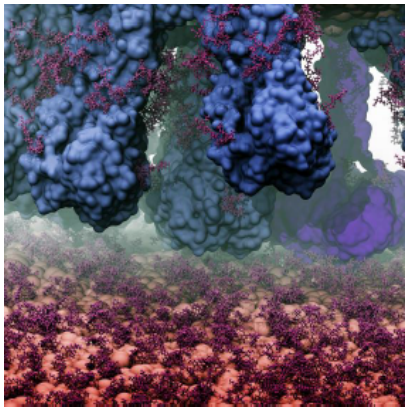
in what way windfarms impact the environment



why there are stellar gas clouds

If gravitational waves exist

how humans developed

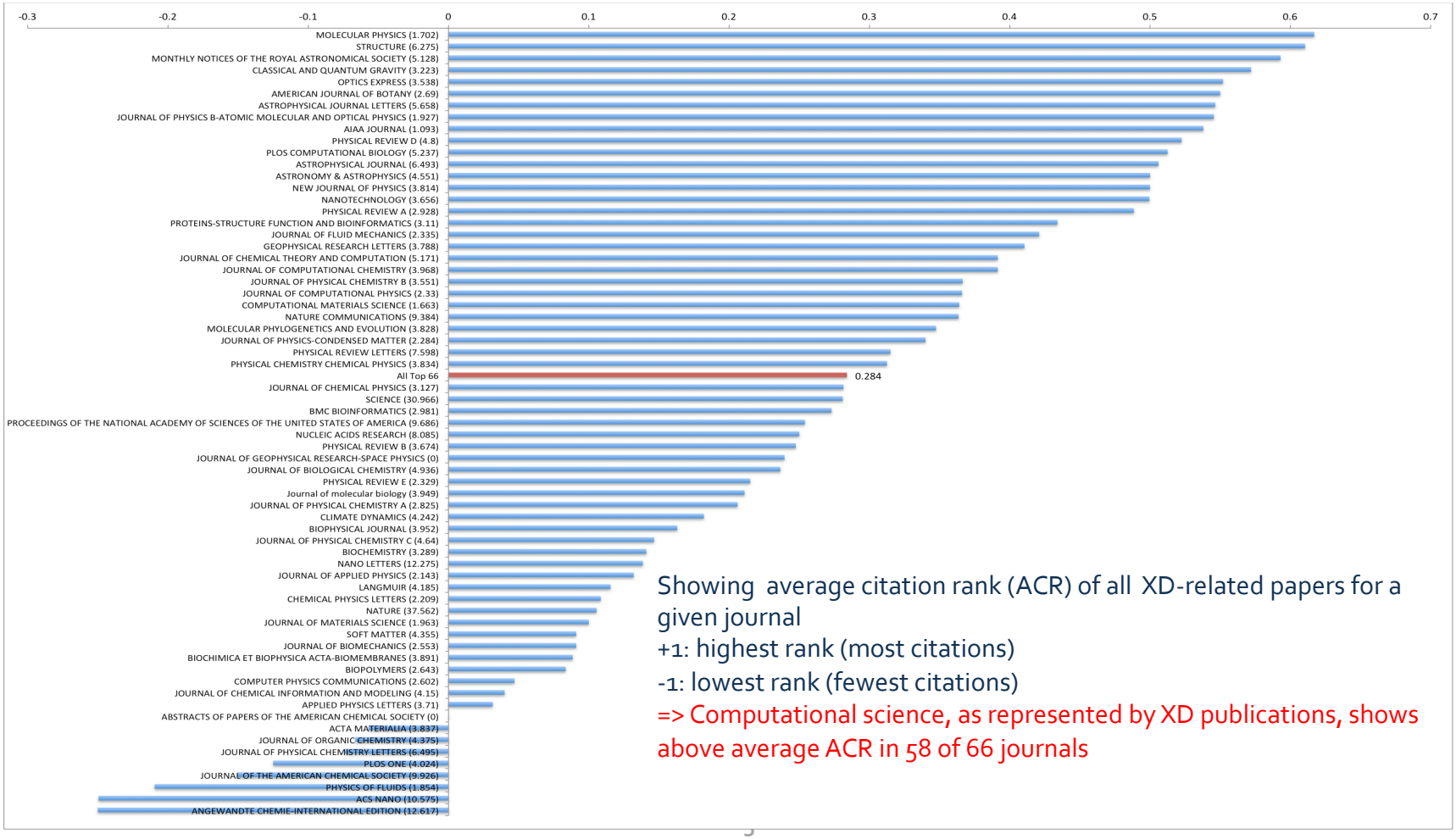


how the influenza virus behaves

Computational Research has higher impact across sciences

Analyzed all journals with >10 XD publications since 2005

(Courtesy of Gregor von Laszewski, publication in preparation)



Showing average citation rank (ACR) of all XD-related papers for a given journal

+1: highest rank (most citations)

-1: lowest rank (fewest citations)

=> Computational science, as represented by XD publications, shows above average ACR in 58 of 66 journals

NSF HPC Infrastructure

XD Network of Computational Resources and Services

Yellowstone
Geosciences

Blue Waters
Leadership Class

Open Science Grid
High throughput

Blacklight
Shared Memory
4k Xeon cores

Bridges (2016)
large, coherent
shared-memory

Comet (new)
"Long Tail Science"
47k cores/2 PF
High throughput

Gordon
Data intensive
64 TB memory
300 TB Flash Mem

Darter
24k cores

Nautilus
Visualization
Data Analytics

Jetstream (2016)
Cloud-based

Stampede
460K cores
w. Xeon Phi
>1000 users

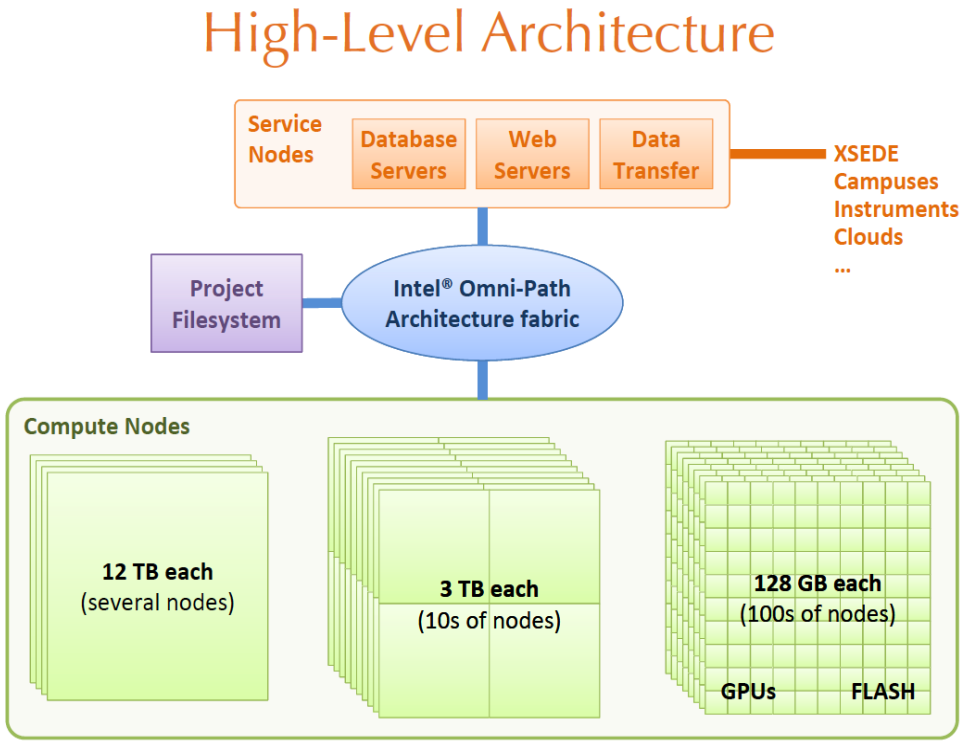
Wrangler (new)
Date Analytics

Maverick
Visualization
Data Analytics

SuperMIC
380 nodes – 1PF
(Ivy bridge, Xeon Phi, GPU)

Coordination though XSEDE
- Resource Allocation
- Advanced User Support
- Education and Outreach
- Digital Services Architecture

Bridges – Large-Memory Compute Resource
Pittsburgh Supercomputing Center - 2016

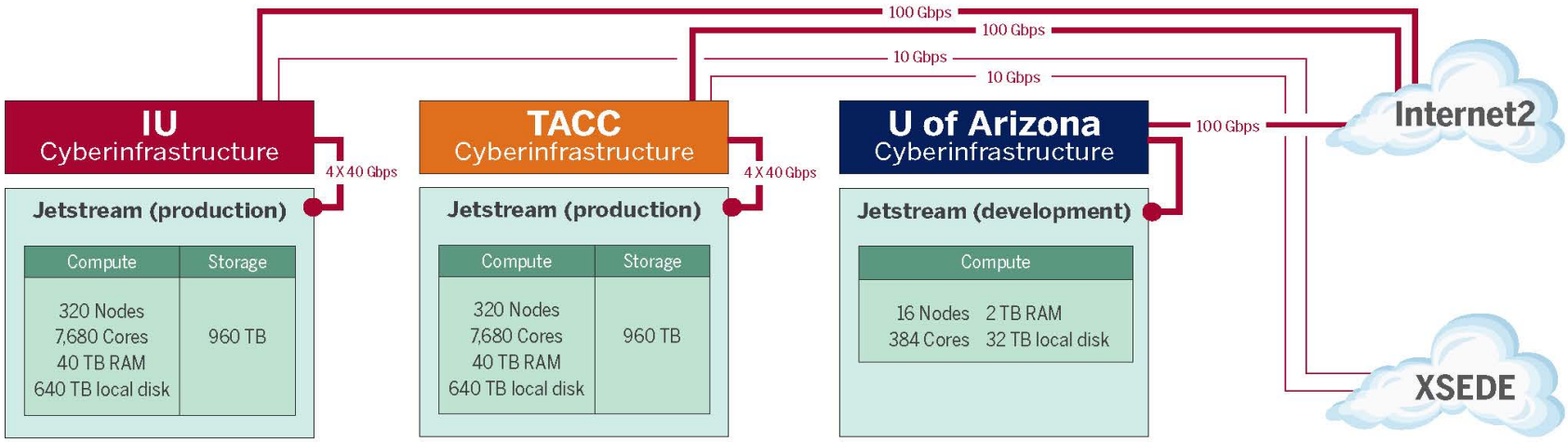


BRIDGES
A PITTSBURGH SUPERCOMPUTING CENTER RESOURCE

Data-intensive computation
- 3 tiers of node types from HP
- Latest Intel Xeon CPUs
- NVIDIA Tesla K80 dual-GPUs and next-generation Tesla

Data manag. & movement
- Database nodes with fast local storage
- Web and data transfer nodes with fast networking
- Shared, parallel, high-performance Project File System
- Intel Omni-Path Architecture fabric

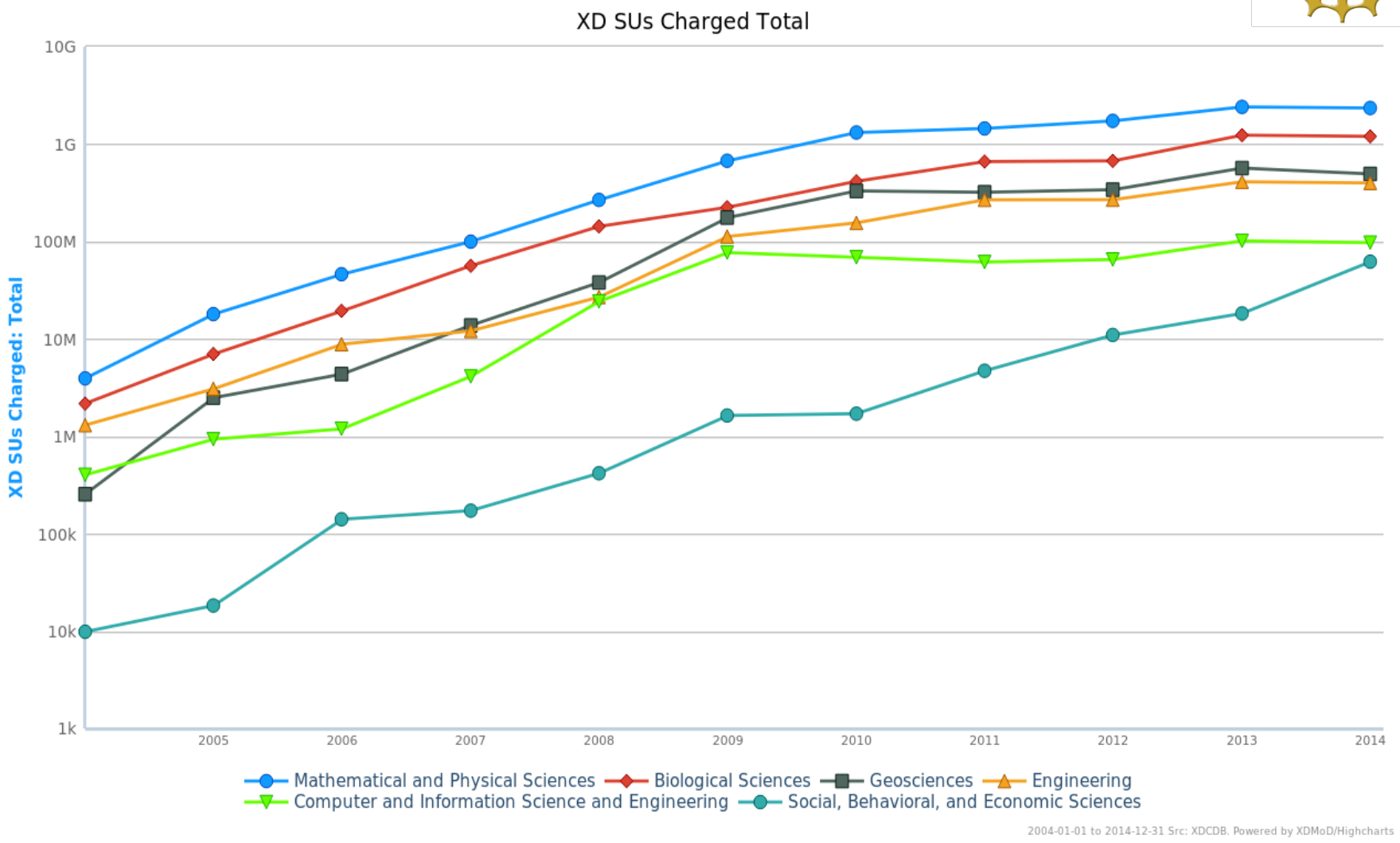
Jetstream – cloud computing resource Indiana University - 2016



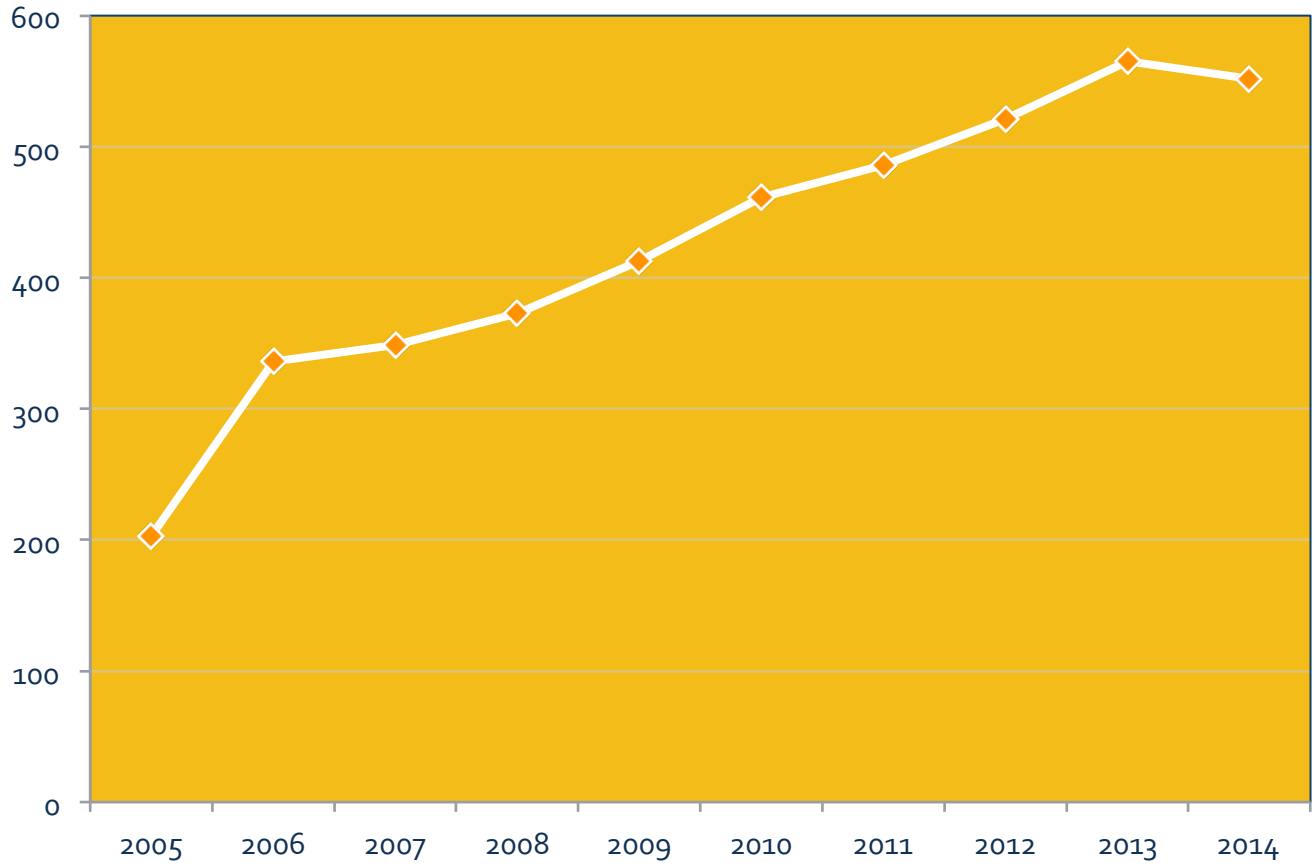
- NSF's first cloud for science and engineering research
- Interactive computing and data analysis "on demand."
- User-selectable library of virtual machines
- Customizable virtual machines – build your "private computing system" within Jetstream
- Broad support across disciplines including biology, atmospheric science, economics, network science, observational astronomy, and social sciences.
- Special support for the iPlant and Galaxy platforms

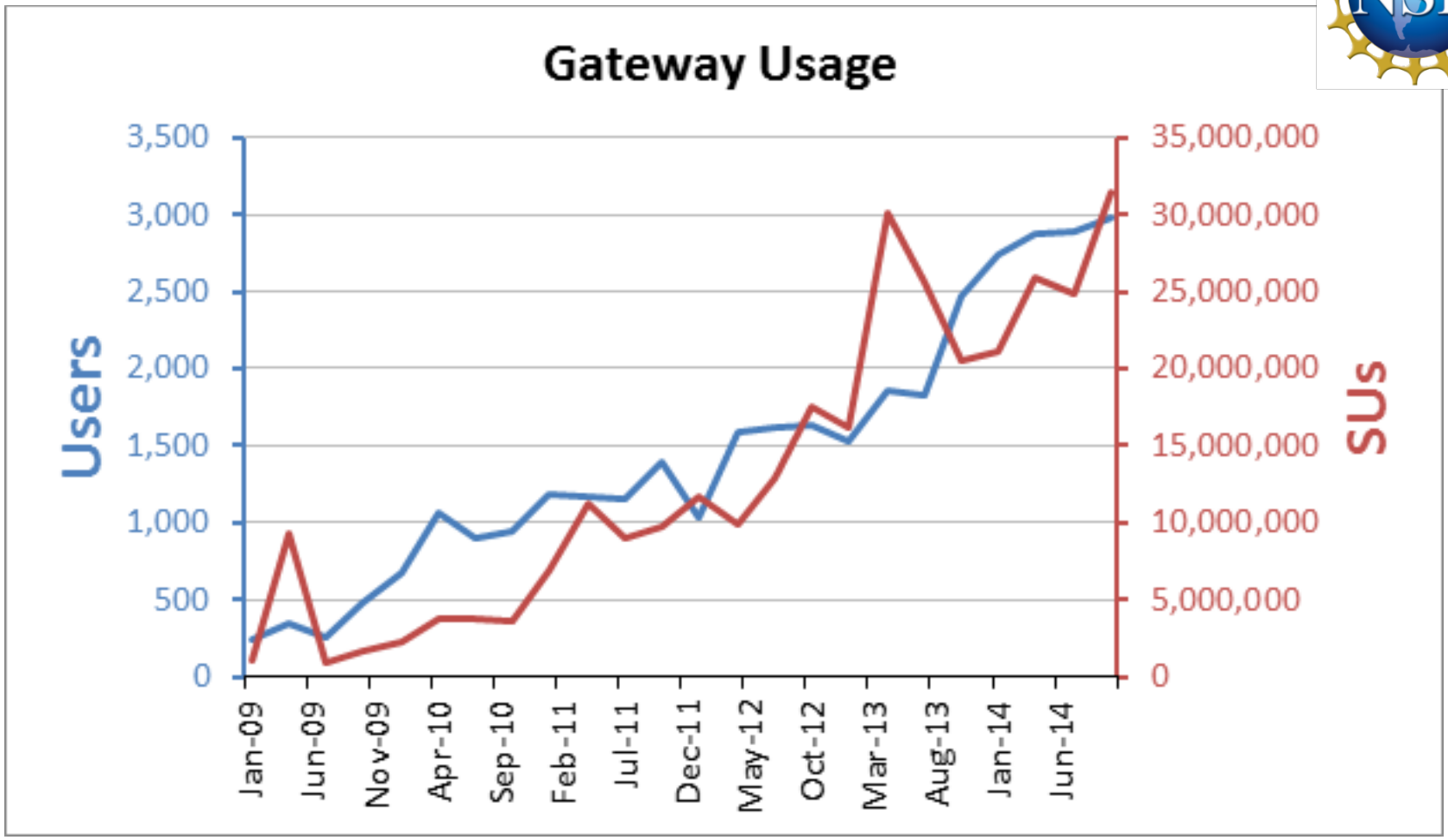
Network	Billion CPU-Hours provided in 2014
XDNet	1.4
BlueWaters	2.3
NCAR	0.57
OSG	0.8

Compute Power Usage Trends by Discipline



10 Year trend in # of Institutions Represented in XDNet





By comparison (approximate):

- Total open XSEDE user accounts: 8,000 => 45%
- Total SUs used in XDNet: 1B => 3%

XD Integrative Services

Projects: XSEDE, TAS(XMS)

Productivity enhancement

- Coordinating resource providers
- Creating uniform user experience, one-stop shop
- Providing expertise

Community building

- Developing the workforce
- Engaging new science communities
- Broadening participation
- Connecting to other CI organizations/projects

Performance Monitoring, Metrics (XDMoD)

- Operational Metrics on XDNet
- Job performance viewer
- Scientific impact metrics

Value of Integrative Services

Investment in integrative services is large:

about 1/3 of ACI's HPC budget -> ~ $30M/year

Cost savings to resource providers:

- Approx. 40% of XD investment, according to a recent analysis by Stewart et. al.

Savings and enhanced productivity for user community:

- Uniformity across sites and one-stop shop for users
- Coordinated education and training
- Better balancing of human and digital resources
- Better decision making based on information from across the entire network
- Reduced duplication of effort
- Enabled multi-site science

Evolving XD Services

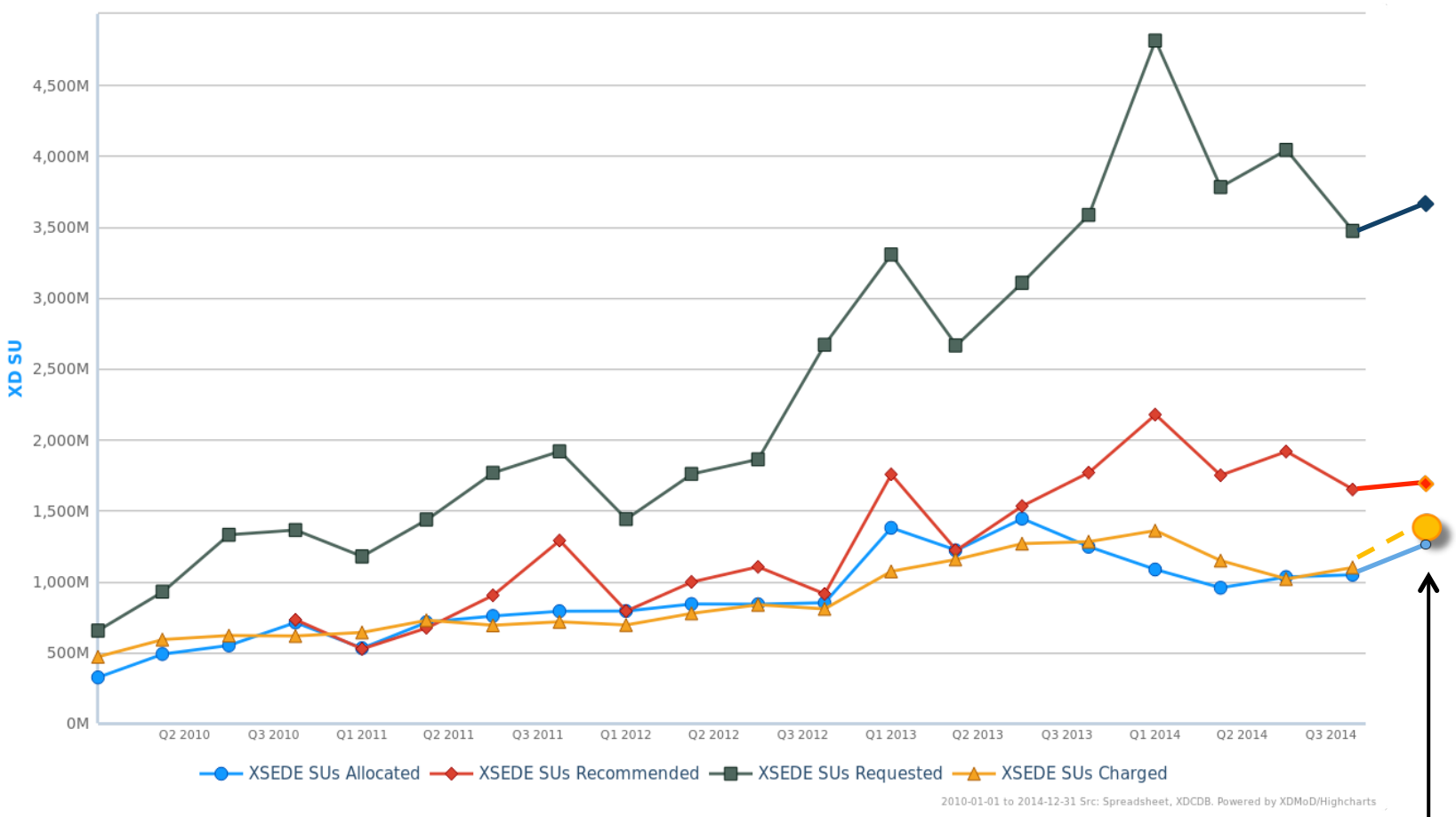
+ Current Human Services
+ Current Digital services
+ What is needed in the future?

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

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

- New machine access methods (non-batch, interactive, real-time, Data streaming ?)
- New workflows ?
- New architectures in support of non-traditional science?

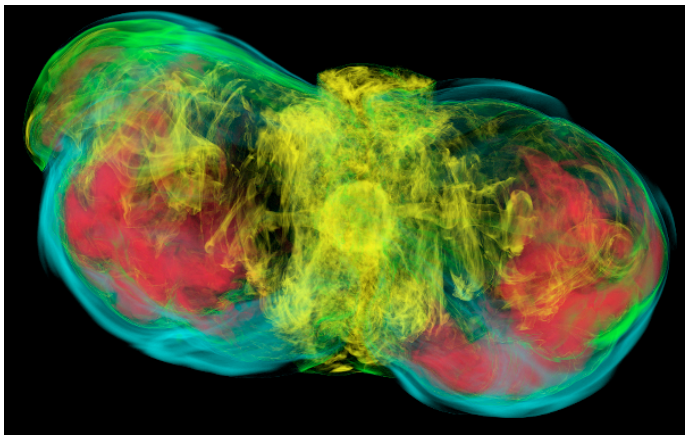
Oversubscription of XDNet ?



June 2015, SDSC COMET online

Blue Waters, Leadership-Class Computational Resource

- Continues to award allocations for breakthrough science projects that need petascale resources.
- PRAC November 2014 proposal close to being awarded
- Next PRAC deadline: November 2015



Christian Ott, Caltech:
First set of full 3D, dynamical-spacetime GR-magnetohydrodynamic simulations of magnetorotational core-collapse supernovae.

Complementing Programs in NSF/ACI

Benchmarking Program
PD 15-7685

- BENCHMARKS OF REALISTIC SCIENTIFIC APPLICATION PERFORMANCE OF LARGE-SCALE COMPUTING SYSTEMS (BRAP)
- How to truly measure an HPC system's capability in executing realistic applications?
- Deadline was Feb 2, 2015 – proposals currently under review.

Scientific Visualization

CADENS – Donna Cox, U. Illinois:

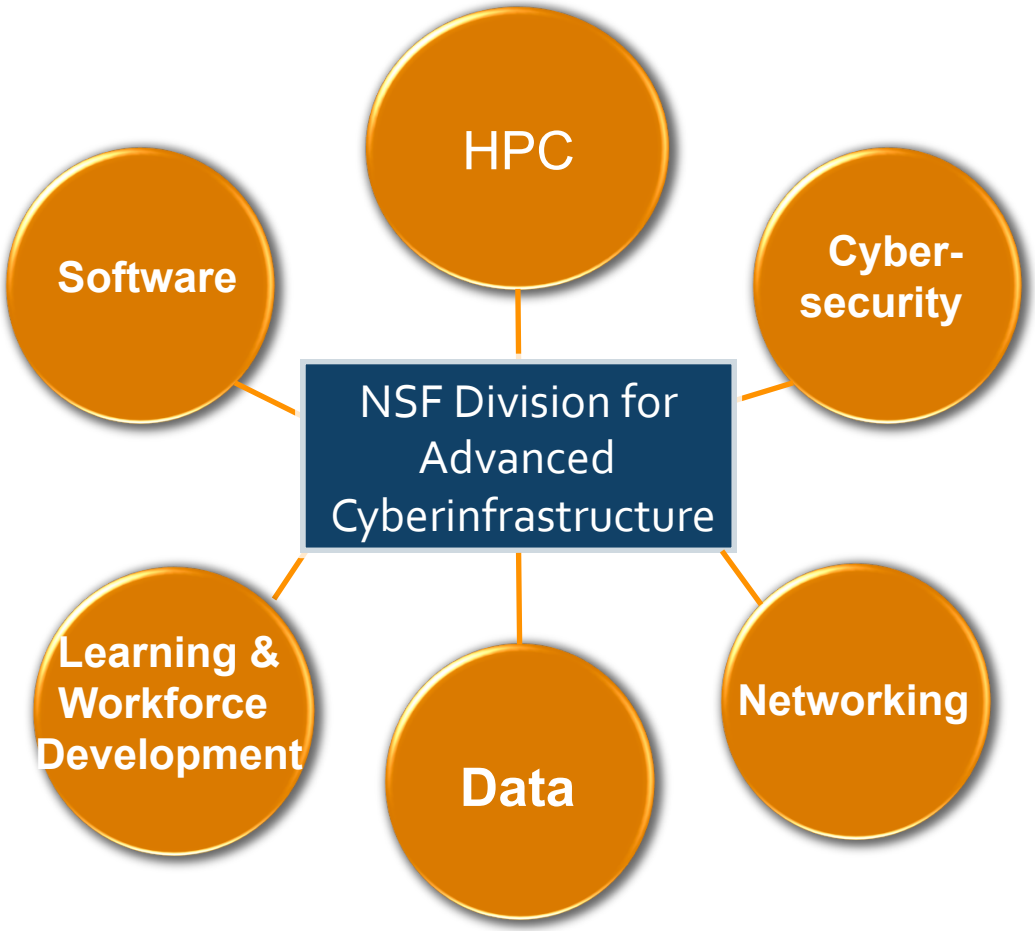
- "increase digital literacy and inform the general public about computational and data-enabled scientific discovery."
- Three full-dome (IMAX) shows + 12 4k videos
- First full-dome show in production: Solar Superstorms
  - Trailer at

    https://vimeo.com/127872610

Data

- 4-th Paradigm and hot topic
- Observation vs. modeling – combination?
- Observation is old; what's new?
  - scale
  - diverse, interconnected data sources
- Merger of data and compute ?

Other "Clusters" in NSF/ACI



NSF's Role: Future Directions for NSF Advanced Computing Infrastructure

- NAS study:
  - received over 60 comments (some from groups) on interim report
  - currently preparing final report, to be released this summer
- Dear Colleague Letter NSF DCL 15-053 solicits proposal for community workshops to
  - identify needs for advanced computing infrastructure in support of future computational research
  - NSF's role
  - Broad range of topics possible

NIST's Role ?

- Metrics
  - XDMoD – XD Metrics on Demand, xdmod.ccr.buffalo.edu
- Benchmarks
  - Measuring HPC system performance
- Standards for Data and Computation
  - Metadata
  - Increasing interconnection

Thank you!

Contact Information:
reigenma@nsf.gov
+1 703 292 2598