Coupling Experimental User Facilities with Large Scale HPC Cyberinfrastructure: Experience, Lessons, and Prospects

John W. Cobb, Ph.D.

Computer Science and Mathematics
Cyberinfrastructure
Oak Ridge National Laboratory

In collaboration with many teams:

NSTG, SNS, McStas group, Open Science Grid, Earth Systems Grid, TechX Corp, OGCE, UK eScience effort, and the TeraGrid Partners teams.

NSF

TeraGrid

OAK RIDGE
National Laboratory

Outline

1. **The “Other” facility report: HPC Cyberinfrastructure**
2. **Cyberinfrastructure**
3. **TeraGrid and other grids**
4. **Science Gateways and the Neutron Science TeraGrid Gateway**
5. **Issues in engagement: Building the Bridge**
6. **What’s Next?**

Acknowledgements and thanks

Staff and Effort

- NSTG: **M. Chen[1], J. Cobb[1], D. Giles[1], V. Hazlewood[7], S. Hicks[1], G. Hinkel[1], D. Hudson[1], V. Lynch[1], P. Newman[1], J. Nichols[1], D. Pack[1], G. Pike[1], J Rome[1], J. Travierso[9], W Wing[1]**
- SNS: **J. Bilheux[1], M. Green[6], G. Granroth[1], M. Hagen[1], J. Kohl[1], D. Mikkelson[8], R. Mikkelson[8], S. Miller[1], P. Peterson[1], S. Ren[1], M. Reuter[1], J. Schwidder[1], B. Smith[7], T. Swain[7], J. Trater[1], S. Vazhkudai[1]**
- McStas group: **E. Farhi[3], K. Lefmann[5], P. Willendrup[5]**
- OSG: **R. Pordes[2] , M. Livny[2]**
- ESG: **D. Bernholdt[1], D. Middleton[4]**
- TeraGrid partners and others

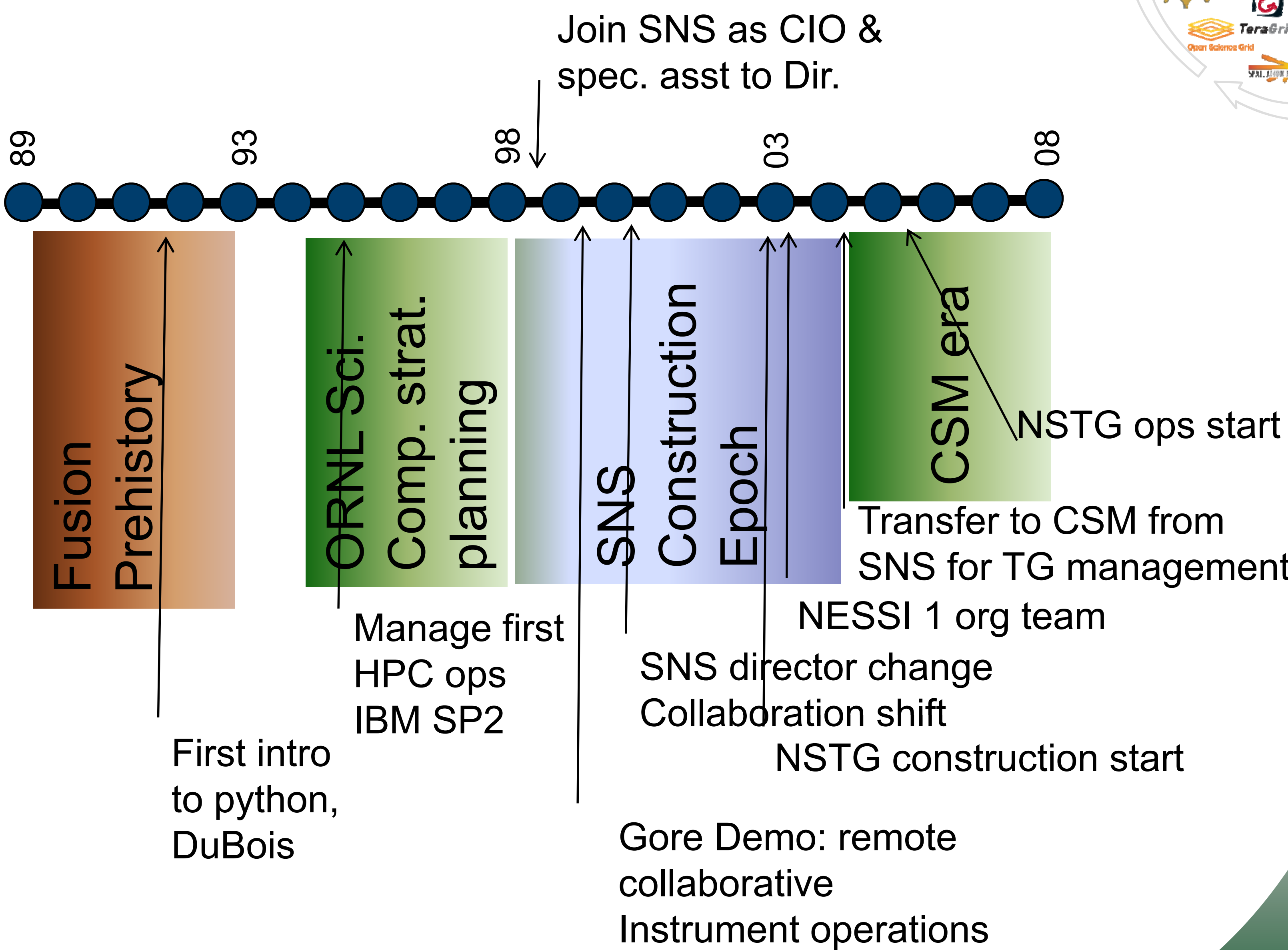
Support

- **National Science Foundation**
- **Department of Energy**
- **ORNL Laboratory Director's research and development fund**

Organization

[1]ORNL, [2]Fermilab, [3]Institut Laue-Langevin, [4]NCAR, [5]RISO, [6]TechX Corp., [7]University of Tennessee, [8]U Wisc., Stout, [9] Austin Peay State University, [10] U Wisc Madision

My Personal Timeline



Outline

1. The “Other” facility report: HPC Cyberinfrastructure
2. Cyberinfrastructure
3. TeraGrid and other grids
4. Science Gateways and the Neutron Science TeraGrid Gateway
5. Issues in engagement: Building the Bridge
6. What’s Next?

HPC is also a facility:
Ex Wigner LCF at Oak Ridge

- **Physical Plant**
  - **40,000 sq-ft of raised floor (slightly over an acre)**
  - **25 Megawatts power**
  - **6,000 tons chilled water**
- **Computing Centers**
  - **DOE/NCCS/Jaguar – 263 TF Cray (1.6 PF in 2009) 360 TB memory**
  - **NSF/NICS/Kraken – 166TF Cray (615 TF in 2009 and 962 TF in 2010. 100,224 cores)**
  - **Other machines: Glue-Gene/P, Linux cluster, viz system, NSTG, OIC**
  - **Storage: 10 PB lustre with 250 GB/sec I/O bandwidth**
  - **Tape: 3 PB stored. 20 PB current capacity. 30 PB capacity in 2009**
- **88 M$US ops/yr**

- **Compare with SNS**
  - **80 acre site**
  - **70 Megawatts**
  - **Accelerator**
  - **Target**
  - **Instruments**
  - **200M$US ops/yr**

But Wait: There's More!

- **ORNL is pursuing a US multi-agency strategy**
  - **US Dept. of Energy (NCCS/Jaguar)**
  - **US National Science Foundation (NICS/Kraken)**
  - **US Dept of Defense**
  - **Other customers (NOAA, EPA, …)**
- **DoD customer has "another one" similar in scope to Wigner LCF (another acre)**
- **CD-0 for third party construction of a 240,000 sq-ft new facility (scaling appropriately. ~ 100 MW power, etc)**
- **ORNL has key advantages:**
  - **Low cost of land**
  - **Excellence in operations**
  - **TVA**
- **8 Acres may only be the start!**

Implications

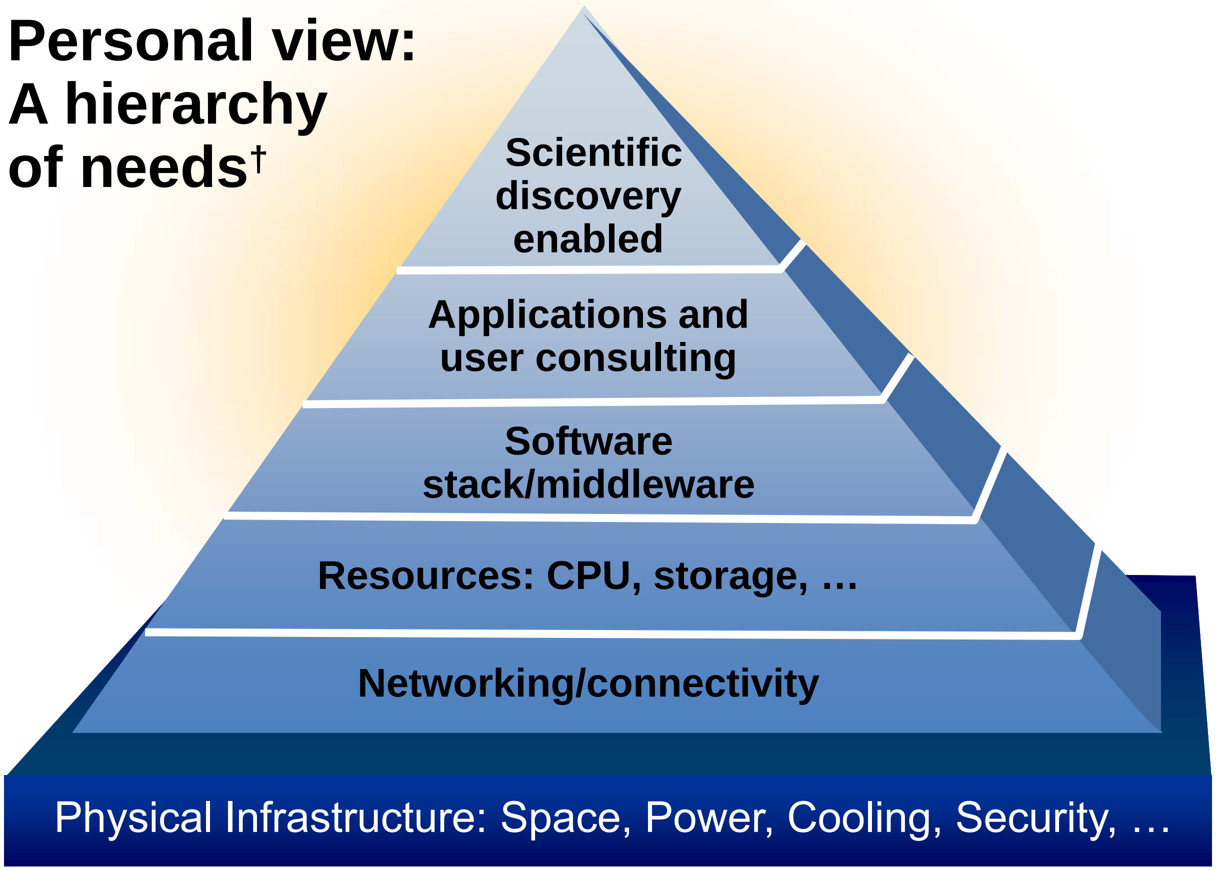
- **The infrastructure challenges of computing are at the same scale as experimental facilities and operations share many of the same issues.**
- **Experimental facilities are supercomputer centers "users". HPC centers must figure out how to support facilities and not just PI's (see NSTG section below)**
- **Experimental Facilities must engage cyberinfrastructure on an equal basis. It is not "just IT support". The converse is also a requirement (Choir preaching)**

Outline

1. The “Other” facility report: HPC Cyberinfrastructure
2. Cyberinfrastructure
3. TeraGrid and other grids
4. Science Gateways and the Neutron Science TeraGrid Gateway
5. Issues in engagement: Building the Bridge
6. What’s Next?

Cyberinfrastructure overview

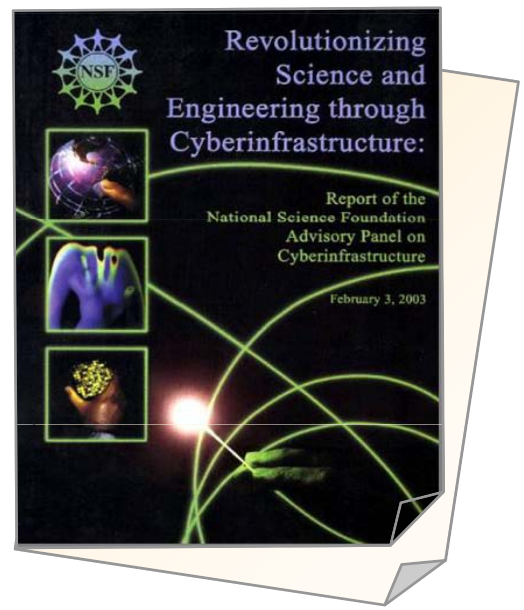
Personal view: A hierarchy of needs†



†Maslow, A. H. (1943). *A Theory of Human Motivation.* Psychological Review, 50, 370-396.

‡ David Hart in NCSA News release “National Science Foundation Releases New Report from Blue-Ribbon Advisory Panel on Cyberinfrastructure” February 3, 2003 http://access.ncsa.uiuc.edu/Releases/03Releases/02.03.03_National_S.html as quoted on Cyberinfrastructure Wikipedia entry.

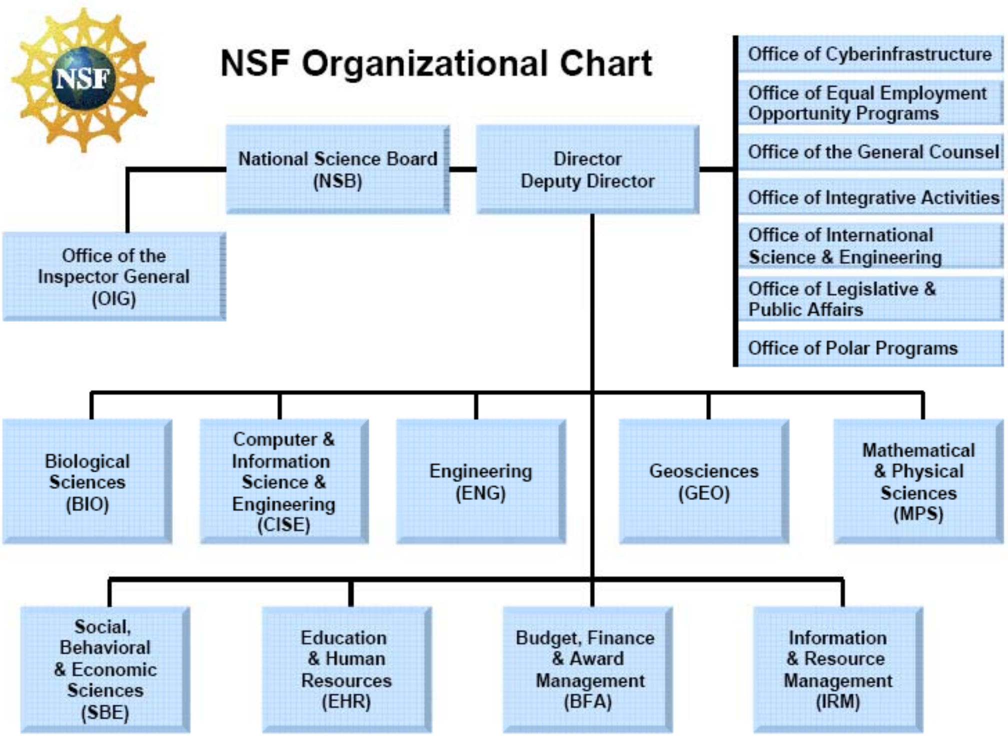
2003 Blue Ribbon Panel: “Atkins Report”



“Like the physical infrastructure of roads, bridges, power grids, telephone lines, and water systems that support modern society, **“Cyberinfrastructure"** refers to the distributed computer, information and communication technologies combined with the personnel and integrating components that provide a long-term platform to empower the modern scientific research endeavor” ‡

US NSF investments in cyberinfrastructure (CI)

- **Atkins Report created change in NSF**
- **US Congress authorization for 1 G$US initiative. After a false start much of that investment has been realized.**
  - **Office of Cyberinfrastructure (OCI) created**
  - **1,100M$US CI investment in OCI**
  - **New agency-wide initiatives (CDI)**



List of OCI investments: HW ($US)

- **TeraGrid partners: 05-10, 30M$/yr**
- **HPC ops 28M$/yr 07-09**
- **Track 2**
  - **A: 2006 Ranger/TACC 30M$ const. 4x8M$ ops**
  - **B: 2007 Kracken/NICS 30M$ const. 4x8M$ ops**
  - **C: 2008 <SGI-Altix>/PSC 30M$ const. 4x8M$ ops**
  - **D: 2009 – proposals due 11/18/08 26M$ const. 4x11M$ ops**
- **Track 1: NCSA/<IBM PERCS?> 200M$ const, 5x25M$? Ops**
- **TGXD : 10-15 32M$/yr 5 yrs (pre-proposals due tomorrow!)**

List of OCI CI investments: SW ($US)

- **SDCI: 08 12M$ (2-3 yr projects, total 12M$)**
- **STCI 12M$(?)**
- **CI-Team 8 M$ in 2006 (?)**
- **PetaApps (OCI led, agency-wide)**
  - **2008: 21.5M$**
  - **2009: 18M$**
- **DataNet: 09-14 100M$ over 5 yrs/5 projects**

NSF Agency-wide CI investments ($US) Summary

- **CI HW and facility Ops: 947M$**
- **SW dev. and Science engagement: 171.5M$**
- **Total OCI: 1,118M$**
- **Agency-wide: (CDI, …) 52M$, ….**
- **In the research directorates:**
  - **Some clearly CI**
  - **Some CI components (Ex Danse as a 7M$ IMR 06-11)**
  - **Some not clearly identified**
  - **Some are part of MRE's**
- **These cover many years**
- **Est: Agency-wide: annual 1G$US/yr spent on CI today.**

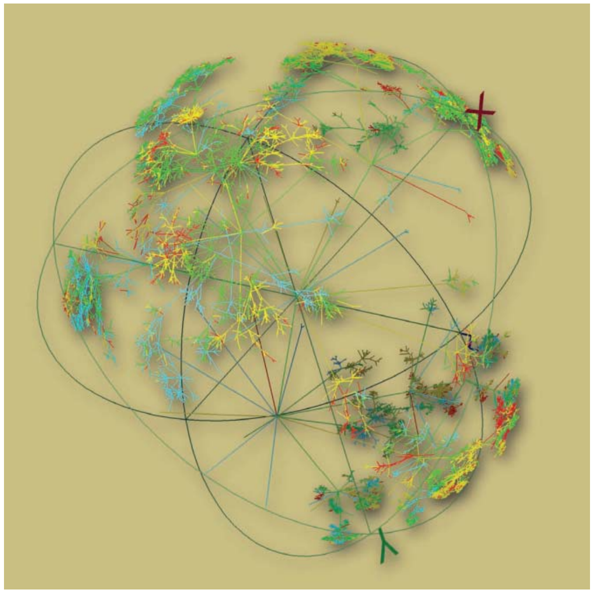
Outline

1. The "Other" facility report: HPC Cyberinfrastructure
2. Cyberinfrastructure
3. TeraGrid and other grids
4. Science Gateways and the Neutron Science TeraGrid Gateway
5. Issues in engagement: Building the Bridge
6. What's Next?

TeraGrid overview

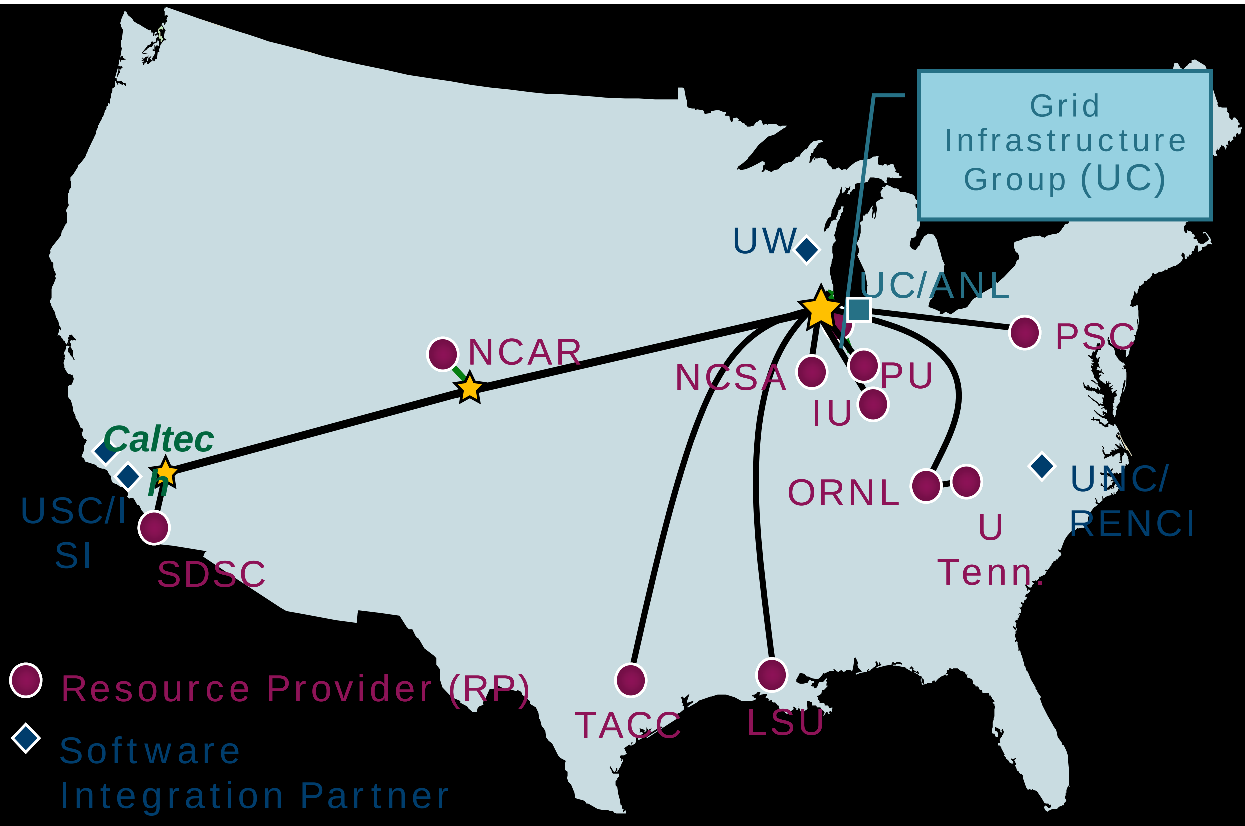
> **"The TeraGrid is the world's largest, most powerful and comprehensive distributed cyberinfrastructure for open scientific research. It currently supports more than 1,000 projects and over 4,000 researchers geographically spanning the entire United States."**
>
> *National Science Foundation in press release 07-09*

- **Noun: Cyberinfrastructure**
- **Adjectives**
  - **Most powerful**
  - **Comprehensive**
  - **Distributed**
  - **For open research**



TeraGrid resource providers

- **Dedicated Links**
- **Resource to Resource**
- **Bandwidth unit: 10 Gbps**
- **Substrate for high performance high quality data movement and communications**



 Managed by UT-Battelle
for the Department of Energy

TeraGrid resources

TeraGrid Systems Monitor

HPC Systems | Storage Systems | Advanced Vis Systems | Special Purpose Systems

High Performance Systems

Name	Institution	System	*Peak TFlops*	Memory TBytes	Status	Load	Running Jobs	Queued Jobs	Other Jobs
Ranger	TACC	Sun Constellation	579.40	123.00	Up		392	3	53
Kraken	NICS	Cray XT4	166.00	18.04	Up		28	7	19
Abe	NCSA	Dell Intel 64 Linux Cluster	89.47	9.38	Up		162	23	47
Lonestar	TACC	Dell PowerEdge Linux Cluster	62.16	11.60	Up		142	79	4
Steele	Purdue	Dell Intel 64 Linux Cluster	60.00	12.40	Up		597	52	38
Queen Bee	LONI	Dell Intel 64 Linux Cluster	50.70	5.31	Up		90	39	0
Big Red	IU	IBM e1350	30.60	6.00	Up		310	156	791
BigBen	PSC	Cray XT3	21.50	4.04	Up		26	25	41
TeraGrid Cluster	NCSA	IBM Itanium2 Cluster	10.23	4.47	Up		134	72	0
Cobalt	NCSA	SGI Altix	6.55	3.00	Up		78	256	0
Frost	NCAR	IBM BlueGene/L	5.73	0.51	Up		10	0	1
Pople	PSC	SGI Altix 4700	5.00	1.54	Up		26	39	30
TeraGrid Cluster	SDSC	IBM Itanium2 Cluster	3.10	1.02	Up		8	0	6
TeraGrid Cluster	UC/ANL	IBM Itanium2 Cluster	0.61	0.24	Up		54	14	0
NSTG	ORNL	IBM IA-32 Cluster	0.34	0.07	Up		1	0	1
		Total:	**1091.39**	**200.62**			**2058**	**765**	**1031**

**Indicates failure of one or more status test.*
Hover mouse pointer over Resource Name, Resource Status, and headings to see additional information.

- **11 distributed centers**
- **> 1 Petaflop**
  On path to double every year
- **2.9 PetaByte disk**
- **6 PetaByte archive**
- **Different resource types**
  - **Capability**
  - **Capacity**
  - **High throughput**
  - **Visualization**
  - **Specialized use**
- **Common user environments**
- **Tuned data transfer over dedicated 10+ gbs network**

TeraGrid as a Grid

- **TeraGrid has dual (multiple) personalities**
  - **Traditional HPC supercomputers**
  - **Grid**
  - **Special purpose computers**
- **TeraGrid Grid services**
  - **CTSS: Common TeraGrid Software and Services**
  - **"KITS" groupings of software packages**
    - **Remote computing**
    - **File transfer**
    - **…**
- **Most standard Globus Toolkit offerings**

TeraGrid features

- **20 resources for the price of 1**
- **User portal to manage accounts/allocations**
- **Unified processes: allocations, account creation, helpdesk, user support, V&V,**
- **TeraGrid is a user facility**
  - **Access via refereed allocations proposals (judged on soundness of computational plan)**
  - **4 allocation meetings /yr.**
  - **~ 200 proposals per meeting**
  - **Difference from Neutron sources: scheduling via batch rather than Holiday inn**

TeraGrid features

- **All TeraGrid sites are connect by dedicated 10 gbs**
- **TeraGrid provides tools to assist in data movement**
  - **GridFTP**
  - **Wide Area File systems (GPFS, Lustre, …)**
- **TeraGrid seeks to grow from 4,000 to 40,000 to 400,000 users. This presents scaling issues and leads to the Gateways program**

TeraGrid innovations

- **Cyber-security HPC use of GSI.**
  - **PKI certs for the science world**
  - **Enables Single sign-on**
  - **TG recognizes a set of Certificate Authorities**
    - **Annual CA audit/review**
    - **Accept certificate from CA's for authentication**
    - **Some TeraGrid sites have eliminated passwords**
  - **Distributes identity management**
- **Advanced User Support (inherited from NSF Supercomputing Centers programs**
- **Science Gateway Concept**

Many other grid efforts exist (some at similar scale to TG)

Outline

1. **The "Other" facility report: HPC Cyberinfrastructure**
2. **Cyberinfrastructure**
3. **TeraGrid and other grids**
4. **Science Gateways and the Neutron Science TeraGrid Gateway**
5. **Issues in engagement: Building the Bridge**
6. **What's Next?**

Crticism: Grids are hard to use

- **Accepted. But why? Difficulty is inherent in the problem**
  - **Many many clusters with many cores**
  - **Manage many projects**
  - **Manage many user**
  - **Cop-out: I just want my own cluster.**
    - **Denial of real issues**
    - **Delay of addressing the problems**
- **Middle ground: Hire programmers to program the grid. Between computer science and discipline science**
- **How to scale?**
- **Solution: TeraGrid Science Gateways program**

Science gateways

- **TeraGrid size: How can we manage it?**
  - **Today: 4000+ user accounts across 20 resources**
  - **Target: 10–100× increase in user base; increase in number of resources**
- **TeraGrid Science Gateways**
  1. **Web portal with users in front and TeraGrid services in back**
  2. **Grid-bridging gateways: Extending the reach of a community grid (devoted to a single area of science) so it can use TeraGrid resources**
  3. **Application programs running on users' machines (e.g., workstations and desktops) that access TeraGrid (and other) services**
- **The NSTG is one of 32 existing gateways**

Charting new territory for wide cyberinfrastructure services:

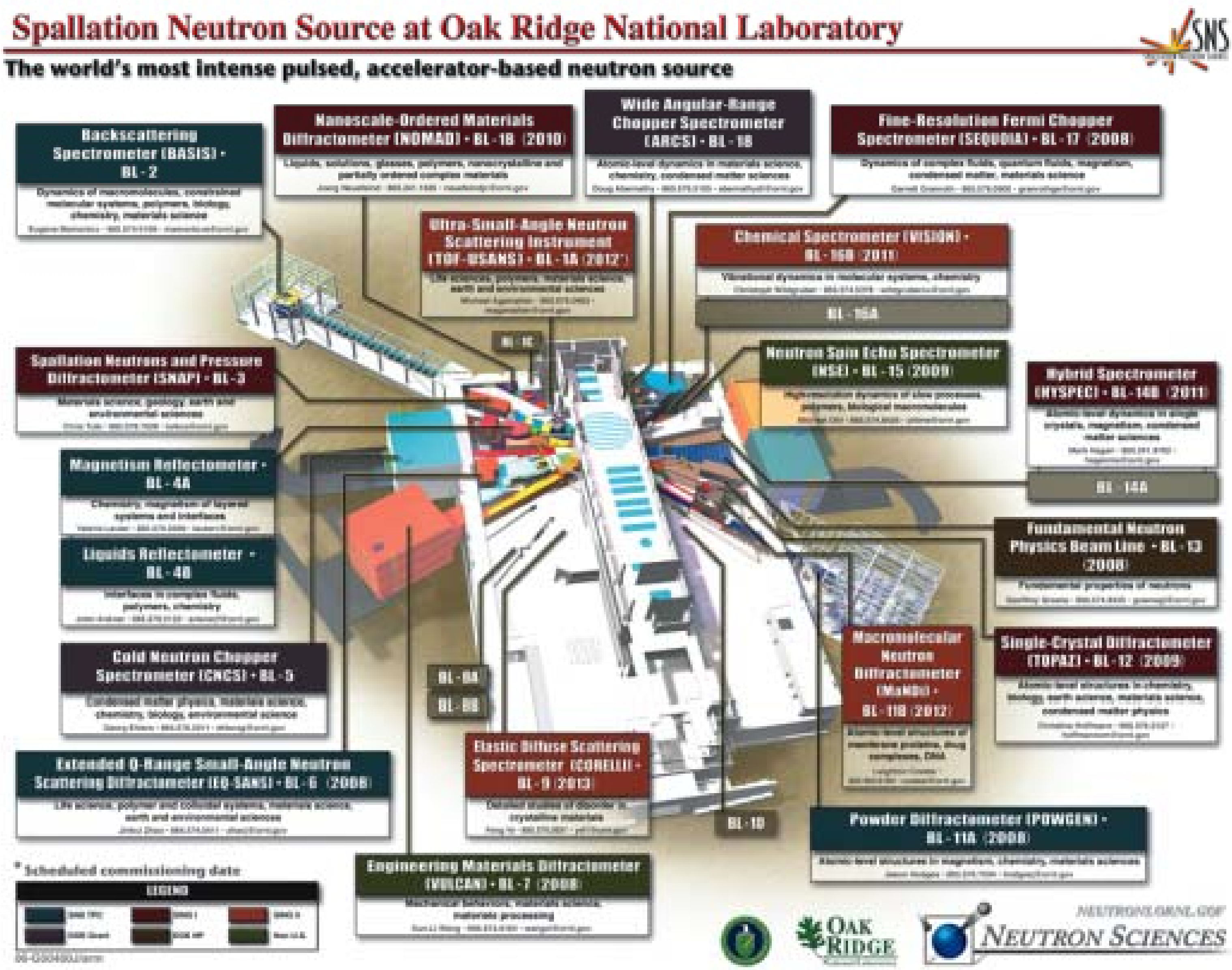
- **Federated user management**
- **Callback end-user identification with automated service auditing across federated identity space "gateway community accounts"**
- **Grid interoperability**
- **Participation in attribute-based authorization**

Neutron Science TeraGrid Gateway (NSTG)

One of 11 TeraGrid partner Resource Providers

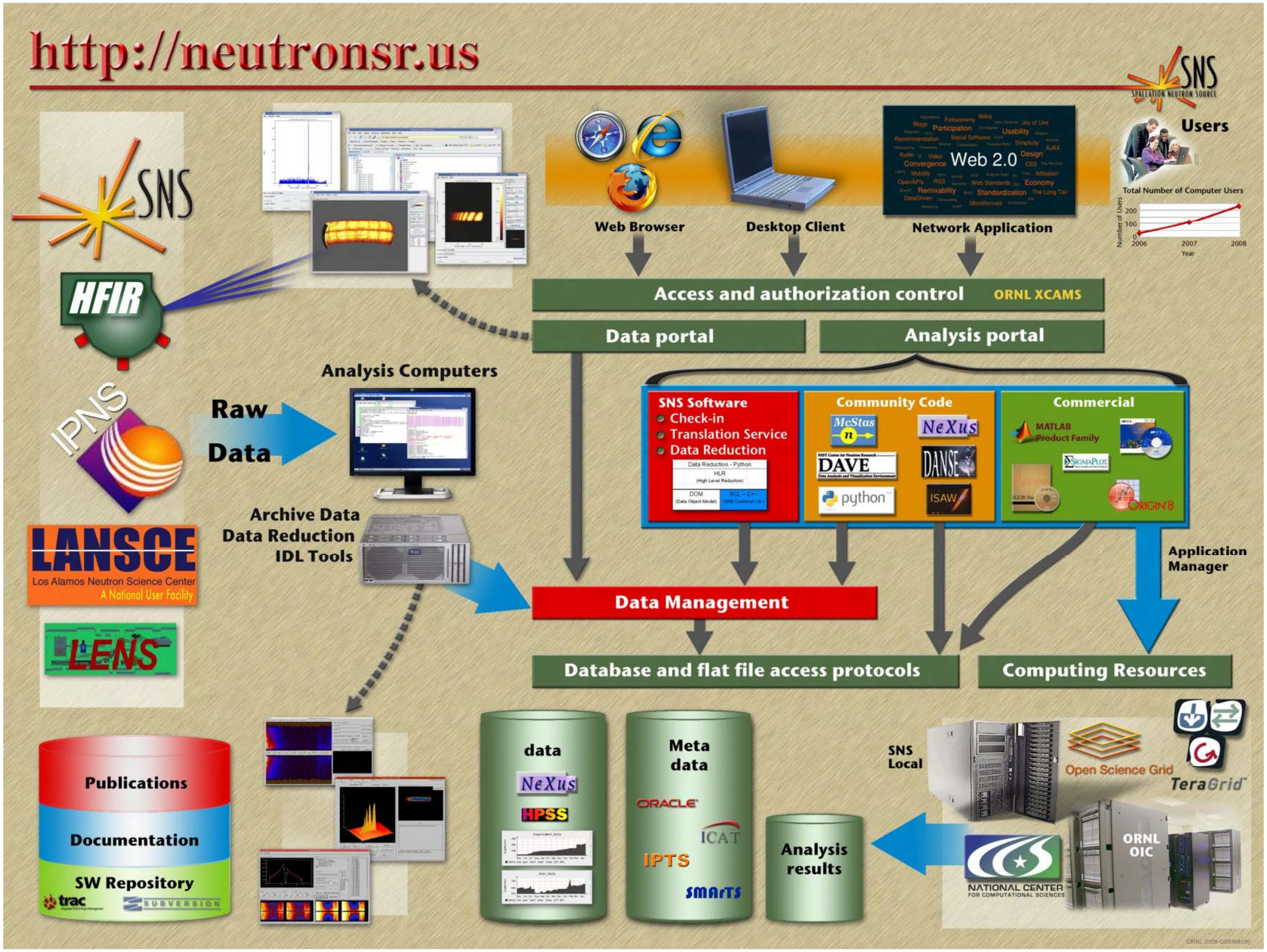
Focus areas	Resources provided
• Neutron science • Connecting facilities with cyberinfrastructure • Bridging cyberinfrastructures • Data movement within and across TeraGrid 	• Outreach to a specific science community (neutron science) – Expertise – Technology transfer – Education, in a broad sense – User outreach • A science gateway interface for neutron science • Exploration of large science facility integration with national-scale cyberinfrastructure • Combine TeraGrid computational resources with available neutron scattering datasets for analysis and simulation comparison

SNS instrument suite



 Managed by UT-Battelle
for the Department of Energy

Neutron science portal view



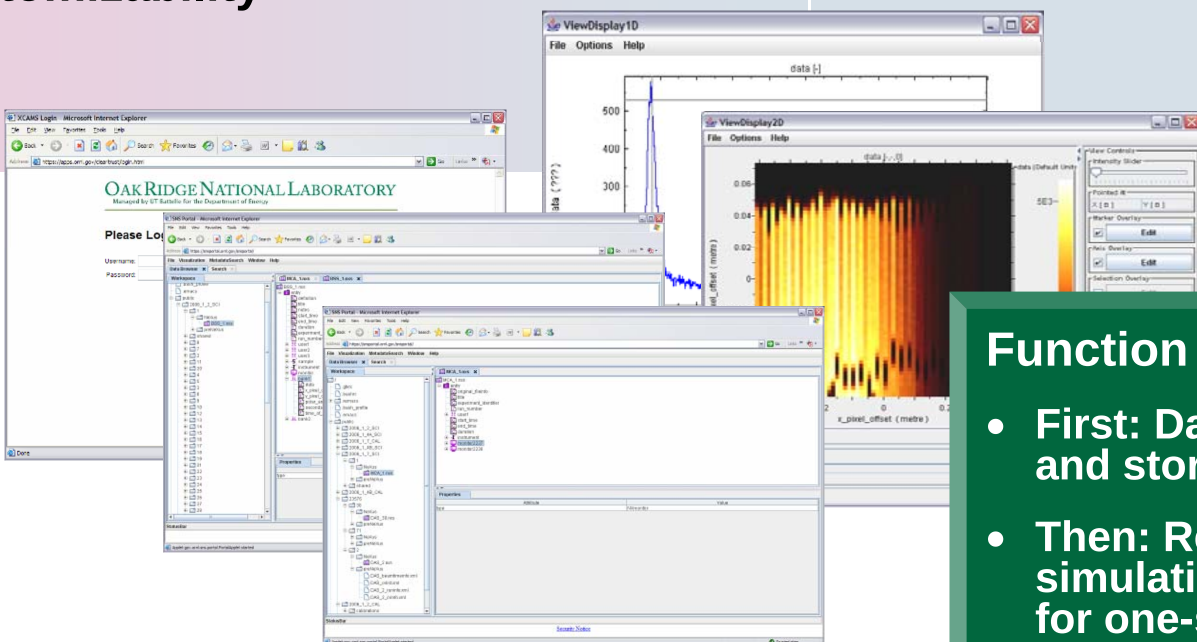
Currently hosts data from 5 different facilities:
- SNS
- HFIR
- IPNS
- LANSCE
- LENS

Ack: From SNS ASG

SNS data access via portal
(Reduction, analysis, and simulation as well)

Turnkey solutions are being requested
by users and user groups

Problem	SNS approach
Providing advanced tools in an easy fashion while retaining customizability	• **Neutron Science User Portal** – **Totally in Java**

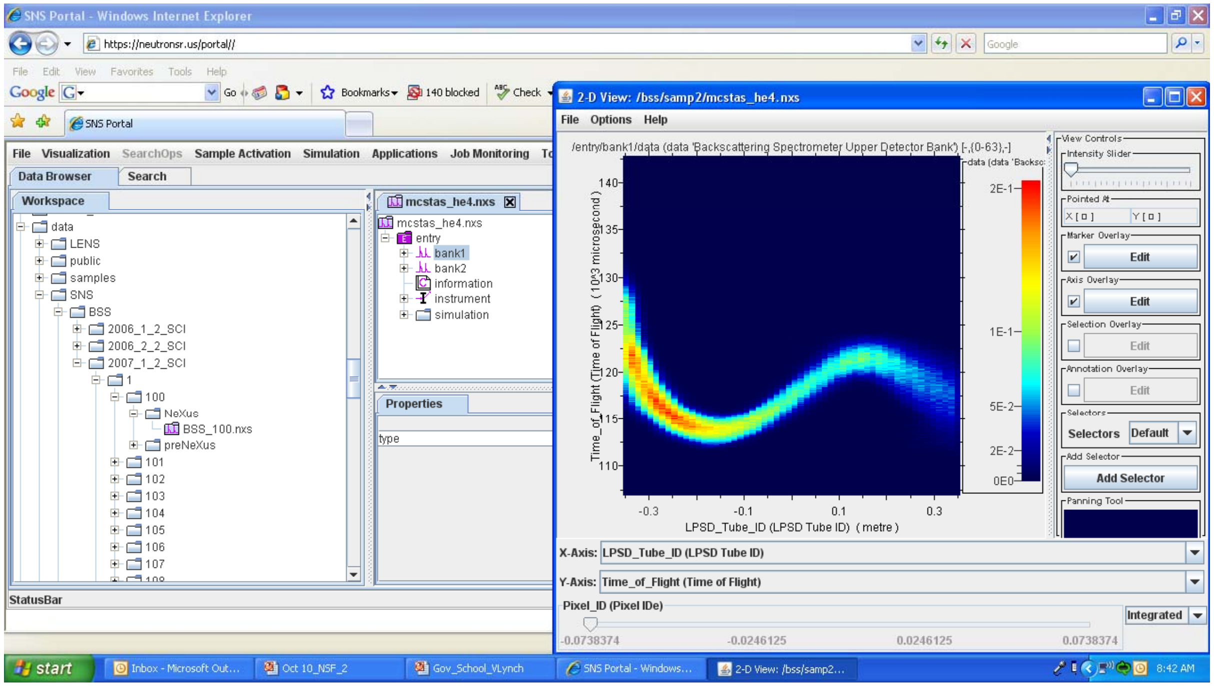


Function deployment order

- **First: Data acquisition and storage/stewardship**
- **Then: Reduction, analysis, simulation, proposals for one-stop access**

Acknowledgement: S. Miller and SNS Advisory SW group;
J. Kohl, S. Vazhkudai and ORNL/CSM division)

Simulation available from portal



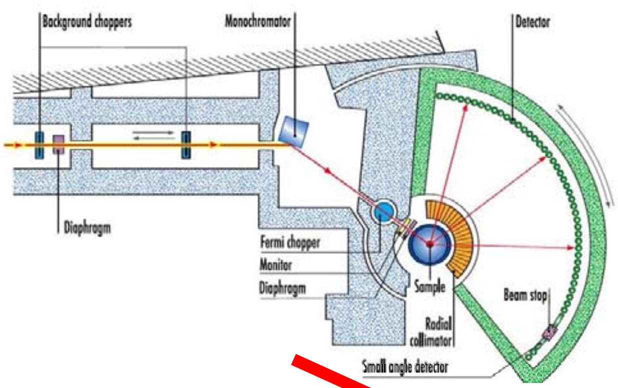
Ack: V. E. Lynch, M. Hagen

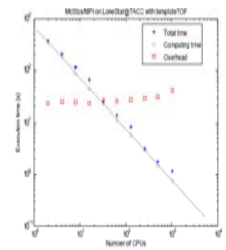
Simulation of inelastic scattering off of He4 sample in Backscattering Spectrometer (BASIS)

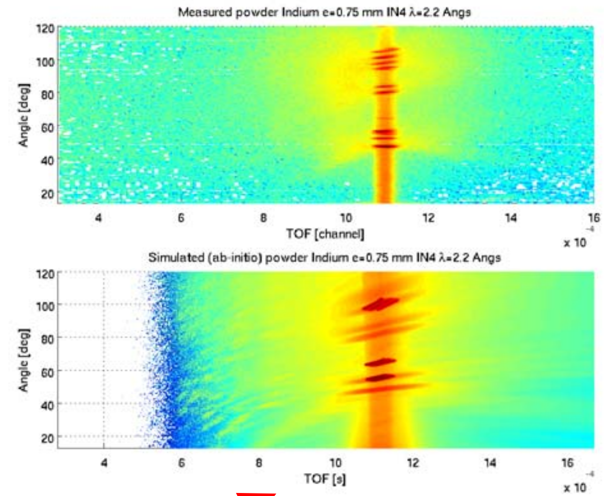
- Using McStas
- 10^{10} Simulation neutrons
- Performed on NSTG cluster

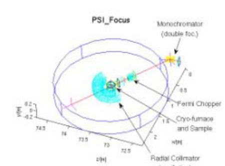
McStas

- **Widely used Monte Carlo neutron ray tracing neutron instrument simulation package for instrument design and experiment planning**
- **Developed at Risø Nat. Lab. (Den.) and Institut Laue-Langevin (France)**
- **Simulations from moderator through sample to detectors agree with experimental results**
- **Use cases:**
  - **Instrument design and construction**
  - **Experiment interpretation**
  - **Experiment planning**
  - **Experiment proposal evaluation**
- **E. Farhi has completed a preliminary study to computationally compute scattering kernels using VASPE.**
- **McStas scalable: 1024+ cores**









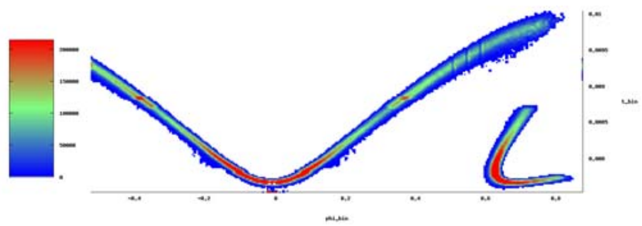
McStas – A neutron ray−trace simulation package RISØ

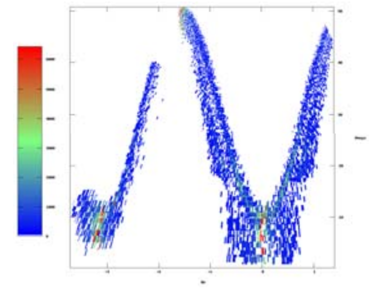
Neutron science simulations

Next step: Provide analysis and relevant simulations in support of neutron science

Instrument design optimization:

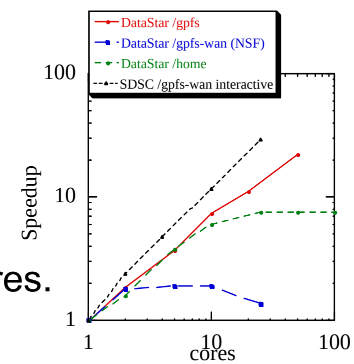
Moving simulations to TeraGrid allows larger runs and faster turnaround. Assisted in effort to redesign instrument to lower cost. Simulations provided confidence that resolution would not be sacrificed.

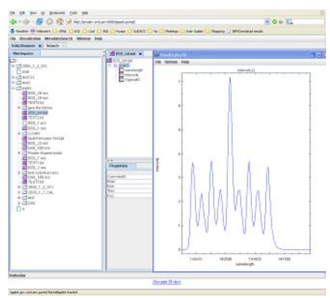


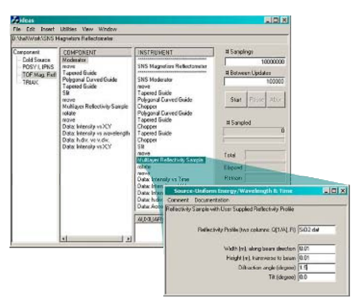


Porting and deployment of simulation tools to TeraGrid cyber-infrastructure: McStas

Many codes currently I/O limited. Next step is to modify I/O strategy in codes to take advantage of I/O architectures.

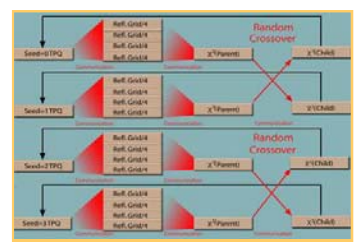


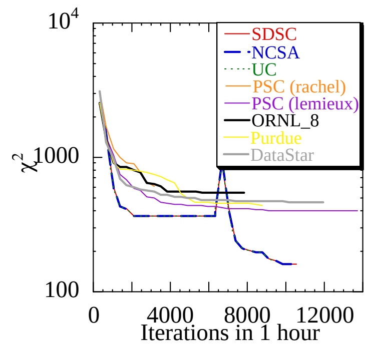




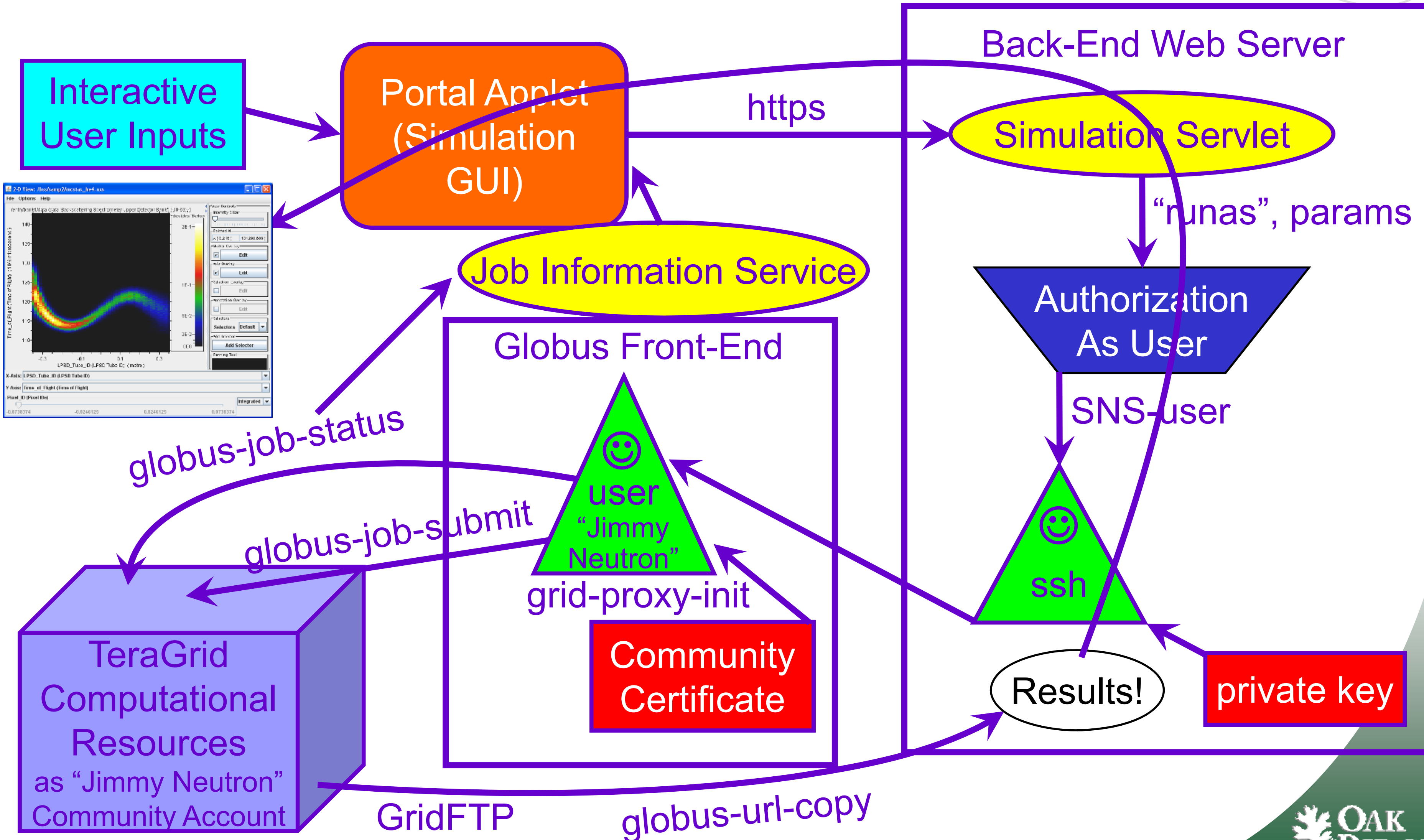
Improved and faster data analysis: Reflectometry

Faster and better χ^2 minimization.





Portal-initiated simulations . . . under the covers!



NSTG results: Portal deployment reduction to practice

- **In development: A generalized fitting service for experimental data integrated into neutron science portal and using TeraGrid computational resources**
- **Outreach: mentored RAMS summer Student: Jessica Travierso (Austin Peay St. U.) Implemented XML driven fitting service GUI. 1st prize undergraduate research at TG'08**
- **Possible follow-on: Can an undergrad build a portal? Winter 2009, stay tuned**

NSTG lessons:

- **Gateway user concept important and takes time to "sink in"**
- **Batch processing and live data needs impedance matching**
- **Reservations**
- **Batch queue prediction**

How long will it take my job to begin running?

Enter your job requirements below and *wait time prediction* can help you answer this question with a 95% confidence.

Predictions will be made on all available resources using each system's default queue. If you would like to limit the prediction to specific systems and/or specific queues, please click on Select System & Queue Details.

Job Information

CPUs: 24

Runtime: 12 Hour(s)

Quantile: 50% What is this?

Predict Wait Time

[Select System & Queue Details]

[Predictions provided by Network Weather Service]

A **24** node, **12 hour** job will exit the queue in less than:

System	Queue	Results
NCSA TeraGrid Cluster	dque	1 minute(s)
ORNL NSTG	dque	1 minute(s)
UC/ANL TeraGrid Cluster	dque	2 minute(s)
TACC Lonestar	normal	3 minute(s)
LONI QueenBee	workq	45 minute(s)
NCSA Abe	dque	54 minute(s)
PSC Bigben	batch	1.7 hour(s)
IU Big Red	NORMAL	2.1 hour(s)
SDSC TeraGrid Cluster	dque	4.5 hour(s)
NCSA Cobalt	standard	60 hour(s)

NSTG was and is "postured as a pathfinder"

- **Proposed a Science Gateway as part of ORNL RP in 2003, before creation of TG Science Gateways Program**
- **Included a focus on cyber-enabled science from the outset: Enabling neutron science**
- **One of the first intergrid production operations**
- **Early concentration on intra-TeraGrid high-performance wide area data transfer (Scxy bandwidth challenge teams)**
- **Concentration on high-performance transfer between TeraGrid and external endpoints**
- **Early production implementation of gateway user concept and simulation-portal integration**
- **Early adopter and pathfinder postures fit with NSTG RP role**

Outline

1. **The “Other” facility report: HPC Cyberinfrastructure**
2. **Cyberinfrastructure**
3. **TeraGrid and other grids**
4. **Science Gateways and the Neutron Science TeraGrid Gateway**
5. **Issues in engagement: Building the Bridge**
6. **What’s Next?**

How to create a productive working relationship?

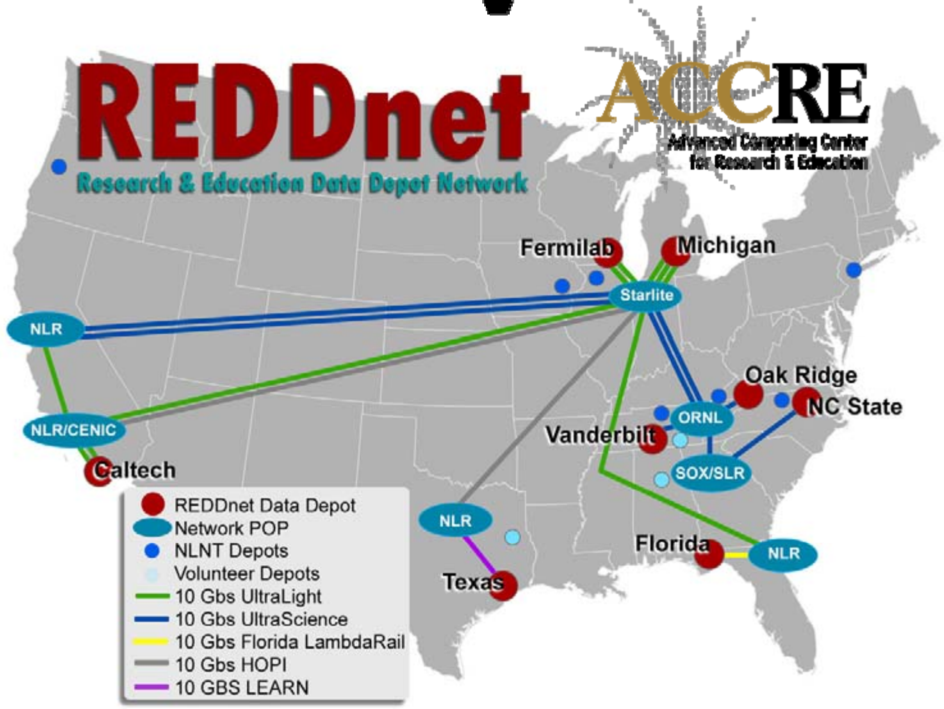
- **Common Goals**
- **Grassroots commitment to collaboration**
- **Collaborative *LEADERSHIP***
  - **Must finesse financing issue**
  - **Leaders (at appropriate levels) must be committed**
  - **Leaders must have appropriate "EQ" Prima Donnas, Control Freaks, the self-absorbed, and credit grabbers are not suitable.**

Outline

1. **The "Other" facility report: HPC Cyberinfrastructure**
2. **Cyberinfrastructure**
3. **TeraGrid and other grids**
4. **Science Gateways and the Neutron Science Terag Grid Gateway**
5. **Issues in engagement: Building the Bridge**
6. **What's Next?**

REDDnet
Research and Education Data Depot Network

VANDERBILT UNIVERSITY



ORNL REDDnet nodes in NSTG and LCF

See Vanderbilt research booth for SC'08 demo, bandwidth challenge and more details

http://www.reddnet.org/

- National-scale logistical storage network
- Distributed **(but acts local!)**
- **Funding from NSF and Library of Congress**
- **>700 TB (500 disk, 200 tape)**
- **Multiple application domains**
  - **Satellite imagery (AmericaView)**
  - **High energy physics: LHC**
  - **Terascale Supernova Initiative**
  - **Structural biology**
  - **Vanderbilt TV news archive**
  - **National Geospatial Data Archive**
- Data everywhere
  (under the cover tools to manage movement, replication, …)

Send in the Clouds (apologies to Stephen Sondheim)

- Grids have grown up and are now commercially viable compute and storage services
  - Amazon: S3 for Storage, EC2 for computing.
  - Google: Google FS, Google Apps
  - Univa-UD
  - IBM
  - Recently Microsoft Azure (what is it? Is it real?)
- Enabling technologies:
  - Virtualization technologies (VmWare, Xen, Hypervisor, …)
  - Global distributed file systems: (soon ready)
    - Hadoop, GPFS, Lustre-WAN – pNFS
  - User community acceptance of “life in the cloud”

Also See: ISIS Prod. Grid-MP implementation

OAK RIDGE National Laboratory

More clouds (no storms)

- Reliability?
  - Untested
  - Recent “burps” imply < 100%
  - In fairness, compare to your local IT support?

  “How many 9’s do you have?”
- Redundancy: a plus and a must
- Control: upper management must come to terms with the reality of their lack of control.
- Innovation: moving to clouds ties projects to hotbeds of innovation, including social computing trends

Cost Model -1

- Ex: Amazon cost: (US prices and $US)
  - EC2 Elastic Cloud Computing
    - Executes a virtual image uploaded to Amazon
    - Based on "instances" - # cores, memory footprint
    - $0.10/CPU-hour for small instance
    - $0.04/core-hr for High CPU instance
  - S3 storage
    - Storage $0.17 / GB / mon.
    - (Bulk price falls to $0.10/GB-mon. for sizes > 150TB)
    - Data-In $0.10 / GB
    - Data Out $0.16/GB

Cost Model -2

- Case: SNS
  - 1.3 TB/yr
  - write once/yr
  - Read 5/yr
  - Use 100,000 CPU-hrs (small instance)

Amazon S3 (US)	Storage	$ 195.00	
	Data Transfer	$ 104.00	
	Requests	$ 1.01	
	Amazon S3 (US) Bill:		$ 300.01
Amazon S3 (EUR)	Storage	$ 0.00	
	Data Transfer	$ 0.00	
	Requests	$ 0.00	
	Amazon S3 (EUR) Bill:		$ 0.00
Amazon EC2	Compute	$ 5,000.00	
	Data Transfer	$ 970.00	
	EBS Volumes	$ 0.00	
	EBS Snapshots	$ 0.00	
	Amazon EC2 Bill:		$ 5,970.00
Amazon SQS	Messaging	$ 0.00	
	Data Transfer	$ 0.00	
	Amazon SQS Bill:		$ 0.00
Total Monthly Payment:			**$** 6,270.01

Cost Model -3

- Case: SKA
  - 1 PB/day
  - write once
  - Read once
  - Exaflop: 167M 6 Gflop cores: 8,350,000 Large CPU instances
- Generic Moore's law predicts 256X improvement or $63,635/mon
- Additional problem: Network
- Sorry Tom. I tried

Amazon S3 (US)	Storage	$ 3,606,656.00	
	Data Transfer	$ 6,002,969.60	
	Requests	$ 1.01	
	Amazon S3 (US) Bill:		$ 9,609,626.61
Amazon S3 (EUR)	Storage	$ 0.00	
	Data Transfer	$ 0.00	
	Requests	$ 0.00	
	Amazon S3 (EUR) Bill:		$ 0.00
Amazon EC2	Compute	$ 6,680,000.00	
	Data Transfer	$ 970.00	
	EBS Volumes	$ 0.00	
	EBS Snapshots	$ 0.00	
	Amazon EC2 Bill:		$ 6,680,970.00
Amazon SQS	Messaging	$ 0.00	
	Data Transfer	$ 0.00	
	Amazon SQS Bill:		$ 0.00
Total Monthly Payment:		**$**	16,290,596.61

Observations on clouds

- For large HPC, go with Supercomputer Centers
- For less demanding cyberinfrastructure clouds offer real promise
  - Low cost
  - Well defined interface to support service level arrangements
  - This is assisted by standard technologies: data transfer, virtualization, …
- The real facility and project cost and complexity is in the Software

It's the Software Stupid!
So you better keep it simple!

 Managed by UT-Battelle
for the Department of Energy

Causes of McGreevy's Telescope?

Caveat: A Personal View:

1. Data and Software are at the pinnacle of the CI pyramid. They are irrelevant if the facility is not completed in the same sense that happiness is not the main pursuit of a hungry and homeless person.
2. Tradition: software (and even instruments) are often viewed as project contingency
3. Scope/Resources. Software projects are then doomed. Resources insufficient for goals and contingency is not available. Failure leads to lack of trust of software efforts (got to #1 and repeat at a lower level)

Software lessons learned - 1

- Eliciting requirements is elusive. Users not engaged
  - Waterfall models are often not an option.
  - Instead develop a use cases and (internally) begin to develop a design methodology

- “All is Flux” – Heraclitus
  - Is Python the PERL of our decade?
  - 30 years ago Fortran was unquestioned.
  - Neutron facilities have a useful life of ~30 years (HFIR built in 1950’s)
  - Response: Plan for upgrades; drop-in replacements; rip-out replacements; fork-lift upgrades
  - Compatibility: The metric that matters is backward compatibility as seen from the future – future proofing.

Software lessons learned - 2

- Beware computer scientists.
  - Are goals compatible? (WE both have publish, but not in the same journals)
  - Insure that project objectives are commonly shared ?
    (Ex. NEES)
  - Are working arrangements mutually beneficial (recall "other facility" slide)
  - If someone comes to you and says: "I've looked at your problem and I know what to do. I propose building a framework to …. " – RUN!
- Look for re-use: Cobb's Rule: "If I can identify and need and imagine a solution, someone has probably already done it" deviate rarely.
- The value of open source is in licensing, maintenance, and operations, not development
- Do not re-invent the wheel for only modest improvements. The maintenance cost outweighs the benefit. "The KISS rule is true, deeply true"
- Software collaboration occurs at the spec/documentation level. At CVS level – its just forks.

Software lessons learned – 3 Collab

- The VO concept is deep and important
  - Virtual organization (VO) is NOT a distributed organization.
  - DO's can be command and control, not VO's
  - Facility software collaborations will be VO's
  - There successful (and unsuccessful) VO's lessons. Learn from them.
- User scaling is important and difficult
- Federated identity is needed. No one wants to take on the burden of identity management for thousands of users. Let's don't do it multiple times.
- Example of successful collaborative, cross –facility software development: EPICS (Considering our familiarity with EPICS, why has this not been mentioned earlier?)
- CEI effort should be engaged. (NSF, CIMA, McMullin, et. al)

Contact

John W. Cobb, Ph.D.

Computer Science and Mathematics
(865) 576-5439
cobbjw@ornl.gov