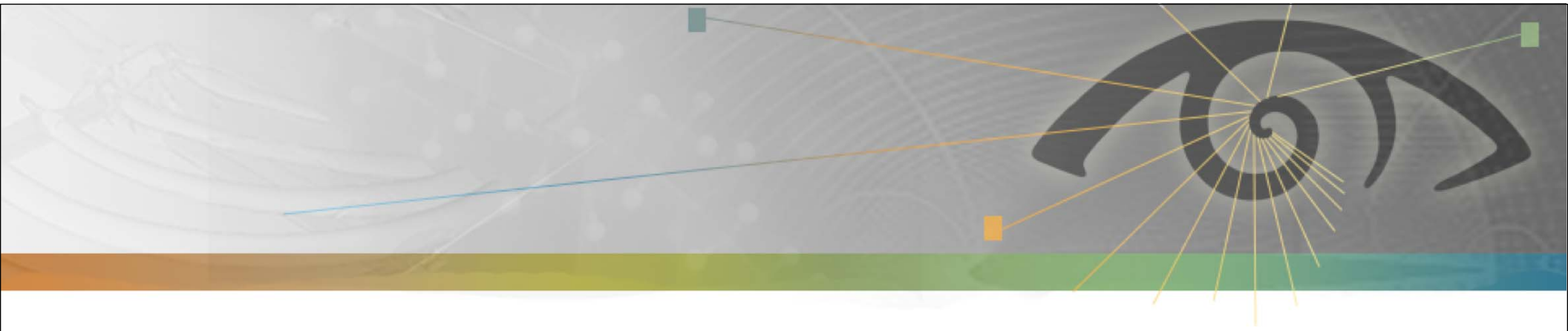


The University of Tennessee Center for Remote Data Analysis and Visualization (RDAV)

Sean Ahern, PI – University of Tennessee
Jian Huang, co-I – University of Tennessee
Wes Bethel, co-I – Lawrence Berkeley National Laboratory
Scott Klasky, co-I – Oak Ridge National Laboratory
Dave Semeraro, co-I – National Center for Supercomputing Applications (NCSA)
George Ostrouchov, Senior staff – Oak Ridge National Laboratory
Miron Livny, Senior staff – University of Wisconsin

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation

Co-located at National Institute for Computational Sciences (NICS)

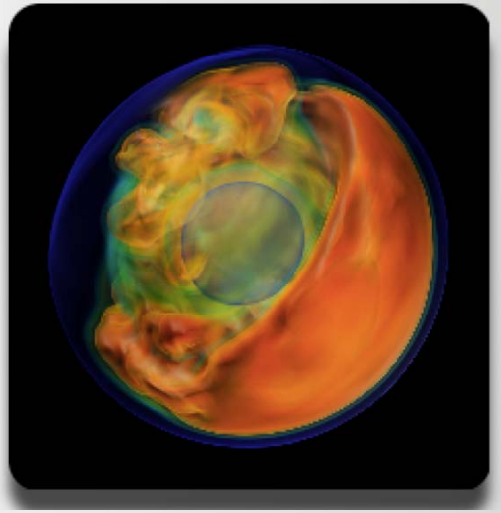
- NICS is a collaboration between UT and ORNL
- Awarded the NSF Track 2B ($65M)
- Phased deployment of Cray XT systems
- Home of Kraken, used to be #3 on Top 500

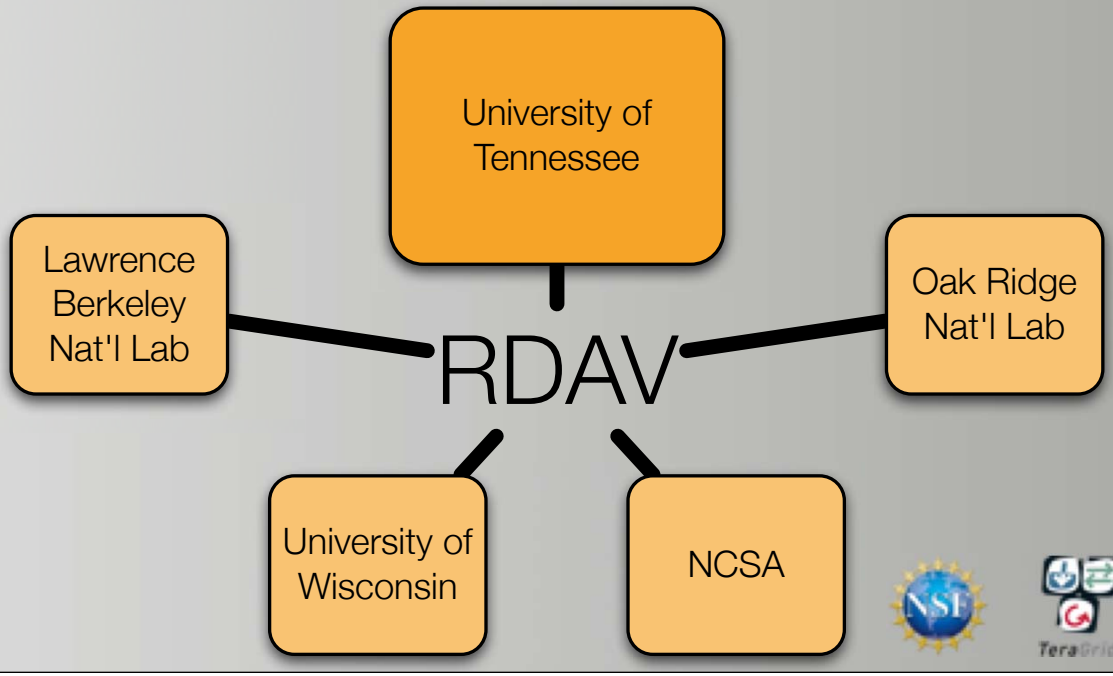
NSF Teragrid

- Under the auspices of NSF OCI (Office of Cyberinfrastructure)
- The world's largest distributed cyberinfrastructure for open science research
- 11 partner sites of integrated, persistent computational resources
- 2.5 petaflops, 50 petabytes storage, 100+ discipline specific databases -- connected through high speed network
- Entering its new era of XD: eXtreme Digital (2011 - 2016)

RDAV - Eyes of the Teragrid

- Provide remote and shared resources for the purpose of exploring/analyzing/ visualizing large scale data.
- Provide the ability to easily take advantage of remote and shared computing/ data storage infrastructure.
- Provide unique architecture for data analysis and visualization
- Leverage large amount of existing experience in deploying similar capabilities.
- Allocated through TRAC

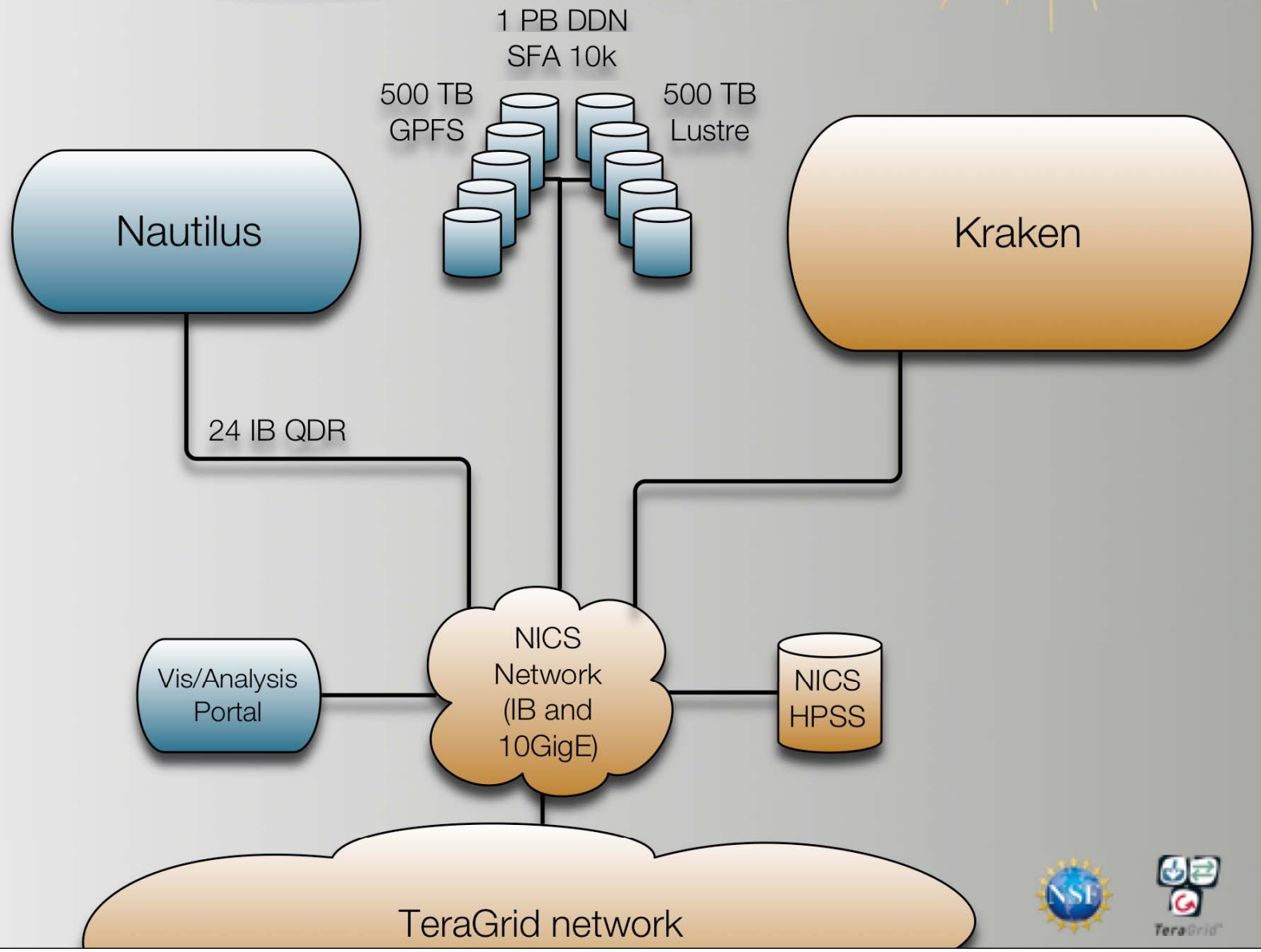




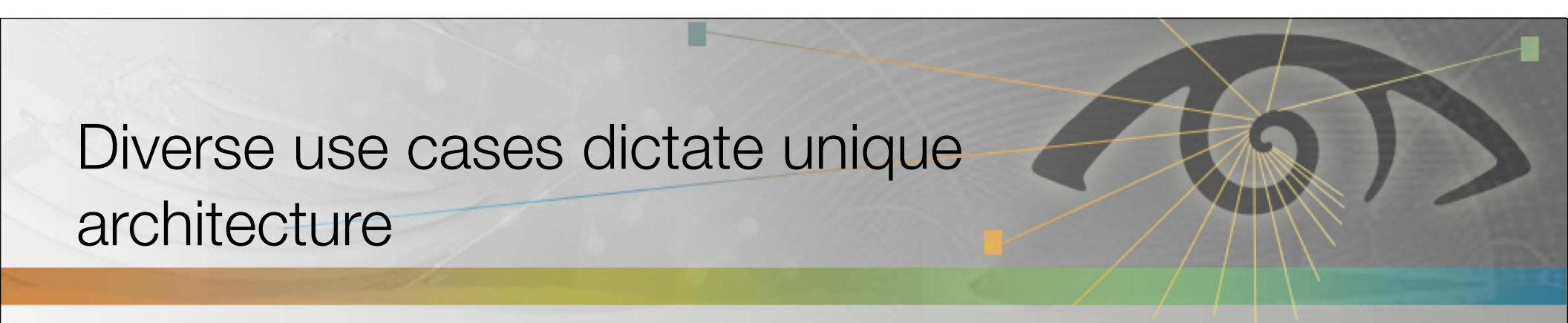
RDAV's Central Hardware: Nautilus - SGI Ultraviolet SMP

- Nautilus in stable production
- System configuration:
  - 1024 Nehalem cores
  - 128x blades/sockets
    - 4 TB memory
    - SLES 11
  - Infiniband 24x QDR
  - 4x 10 Gigabit Ethernet
  - 8-16x Nvidia Fermi Tesla
  - ~1 PB parallel file system

How Nautilus Fits Into NICS



Diverse use cases dictate unique architecture



- Many HPC users can use distributed memory analysis
  - data parallel, time parallel
- However, many general and statistical analysis algorithms favor large shared memory
  - Document clustering/searching
  - Generalized graph structures
  - Bioinformatics, genomics
  - …
- *Large shared memory* is the only reasonable way to address all of these needs
- SGI UltraViolet architecture provides:
  - Large memory single-system image through NUMA
  - A “better” cluster architecture, accelerating distributed memory MPI

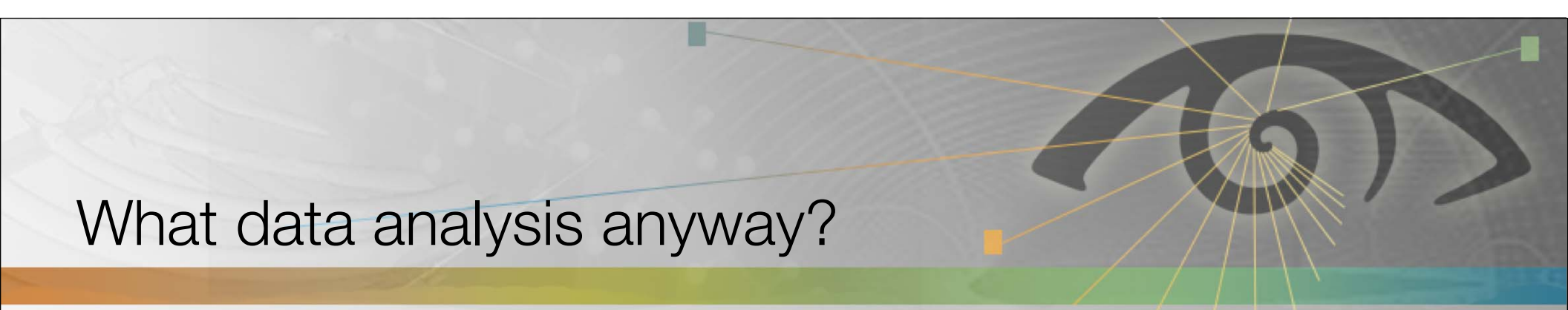
RDAV Users

- RDAV user accounts are managed by Teragrid
- RDAV is open to all science users with data analysis needs
- So far, 261 total users (minus training, student, staff, vendor accounts)
  - 81 projects from about 37 universities and laboratories
  - 30+ disciplines (ranked by # of users): astronomy and cosmology, astrophysics, climate dynamics, magnetospheric physics, meteorology, hydraulic systems, fluid systems, systematic and population biology, computational art, blood flow, hazard mitigation, ecology, solar research, condensed matter physics, molecular bioscience, biophysics, genetics, structures and building systems, ocean sciences ...
  - Use case by % of users: data analysis (74%), vis (53%), and computation (32%)
  - Desired features: accessing large amounts of memory from a single thread/process (61%), running their existing software using large number of cores (49%), shared file system (46%), ... 20% expressed needs for GPU

RDAV provides User Services

- Routine user services staff – Connected to NICS
  - Routine user services
  - Ticket triage and routing to specialist
- Specialized staff for Advanced Support for TeraGrid Applications (ASTA)
  - Specific for remote visualization, data analysis, workflow services, portal
  - Educates on effective use of existing tools or on custom development
  - Provides individualized assistance for center-wide software
- Dedicated staff for education, outreach, and training
- Dedicated staff for tool discovery and certification

What data analysis anyway?



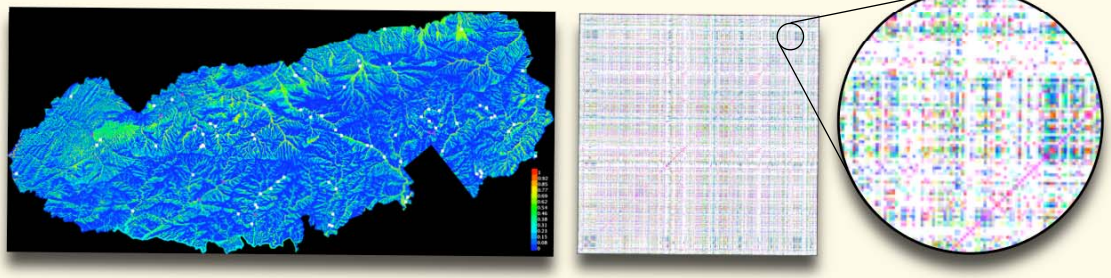
- By popularity in tools: IDL, GNUplot, Paraview, Matlab, R, VisIt, Yt, NCL, VMD, numpy, scipy, GrADS, Mathematica, VAPOR, AVS, ARCGIS, and other in-house custom tools
- By programming language for in-house user codes: C/C++/Fortran, Python, Matlab, R, IDL, Perl, Java, CUDA, ...
- By source of data (each user can name multiple choices): HPC simulation (93%), observational (17%), experiments and others (7%)
- By structure of data: structured and unstructured grids; XML text files; images; plain text; temporal layers of grids, 3D meshes; and particle sets;
- Range of data size: 1GB to 40TB
- Range of file counts: 1 to 1000
- Challenges: not all tools can be rewritten for parallelism; many tools are actually black box code; need to study simulation together with observation/experiment data

RDAV provides a range of software services

- Analysis applications: to be dictated by user needs and technology needed to solve user problems.
  "Whatever it takes!"
- Write any glue software needed
  - Eden
  - Custom scripts in python, etc.
- Remote visualization and image generation
  - Provide interactive and batch image generation tools. (gnuplot, ImageMagick, etc.)
  - Remote parallel visualization (VisIt, ParaView, etc.)
  - Tools for custom application development
- Data analysis and statistical analysis
  - Octave, Parallel R, Matlab, etc.
- Workflow systems
  - DAGMan system automates batch actions on behalf of users
  - Infrequent current use, however, value is increasing and many users wish to explore.
- Dashboard delivery
  - Leverage DoE funding for eSimMon dashboard system.
- Portal system
  - Builds upon standard Liferay platform
  - Provides SAS services for analysis and visualization

Quotes and re

- NIMBioS species modeling of National Park at ecosystem scale
- Performed comparative analysis of MaxEnt output (pairwise comparisons of 6000+ grids ***facilitated by large shared memory***)
- Few had the capability to evaluate the feasibility and the proper protocol to apply MaxEnt type of models to an entire ecosystem
- We used RDAV tools to manage running many parallel instances of MaxEnt (***what did take 6-7 days can now be done in under an hour***)



"The ***large amount of s*** and the ability to run exist standard Linux kernel to access massive amounts memory from a single thread has been a ***crit*** ***enabler*** of my research. In fact, my allocation Nautilus has enabled me ***for the first time*** to ans fundamental "whole data" questions from my datas rather than being limited to the small 1-5% samples I have been using to this point. Being able to move a 5% sample to working with the entire dataset ***fundamentally altered the kinds of questions able to tackle***. In particular, the algorithms I use continual random strides across the entire datase thus do not lend themselves well to message p architectures, so SMP is the only kind of system th really support my research."

– Kalev Leetaru (NCSA)

I want to share my happiness about nautilus with you! I received my advance last week and decided to try it out this weekend. I was able to ***reduce 1 week***'s worth of computation ***to about 30 minutes*** on nautilus! Literally! It used to take me 4-5 days total to transfer all the data to our local machines, then another 2-3 days to generate all the movies (we have very capable machines but this amount of data is simply overwhelming to my raid array). I was able to generate the exact same data on nautilus in 30 minutes. Hooray!!!"

–Amy McGovern (Oklahoma)