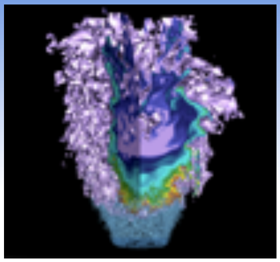
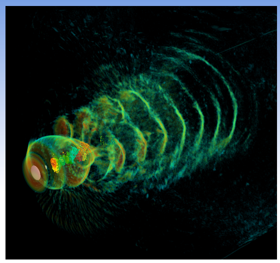
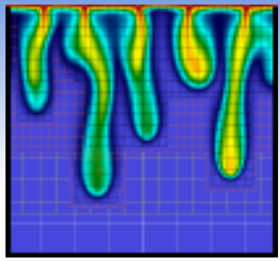
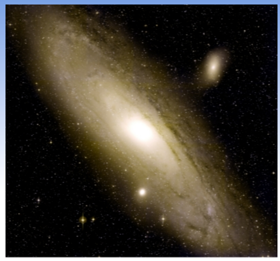
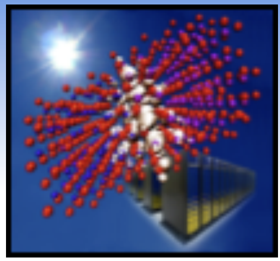
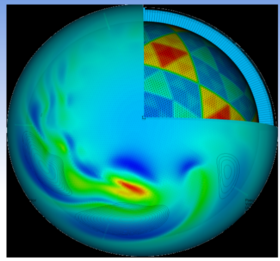







NERSC/Magellan Overview

Exploring Cloud Computing for Science

Shane Canon

Discovery 2015
June 20, 2011

U.S. DEPARTMENT OF ENERGY | Office of Science

Berkeley Lab 75 Years of World-Class Science 1931-2006

NERSC Facility Leads DOE in Scientific Computing Productivity

NERSC computing for science

- 4000 users, 500 projects
- From 48 states; 65% from universities
- Hundreds of users each day
- ***1500 publications per year***

Systems designed for science

- 1.3PF Petaflop Cray system, Hopper
  - 2nd Fastest computer in US
  - Fastest open Cray XE6 system
  - Additional .5 PF in Franklin system and smaller clusters

U.S. DEPARTMENT OF ENERGY

Office of Science

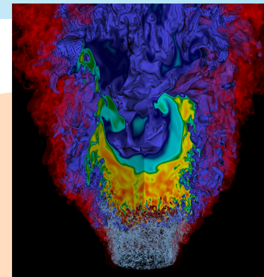
BERKELEY LAB

75 Years of World-Class Science 1931-2006

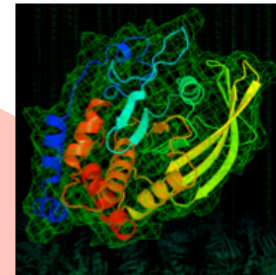
NERSC Serves the Computing and Data Needs of Science

- **NERSC provides computing, data, and consulting services for science**
- **Allocations managed by DOE based on mission priorities**

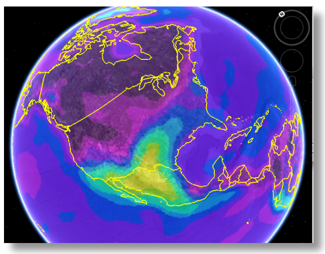
ASCR 5%
NP 11%
BER 19%
FES 19%
BES 31%



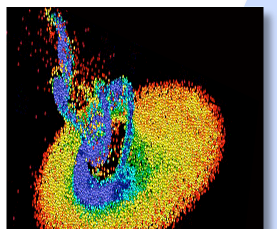
Burning structure in hydrogen leads to pockets of emissions



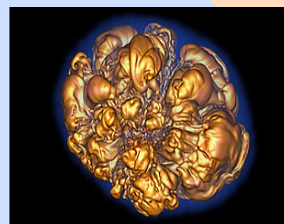
11,000 protein foldings, show common feature in amyloid development,



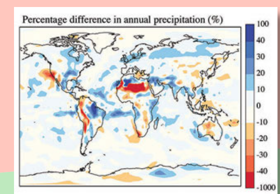
20th Century 3D climate maps reconstructed and in public database



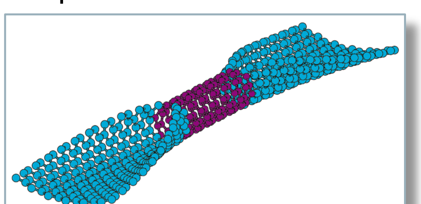
Location of dark companion to Milky Way found



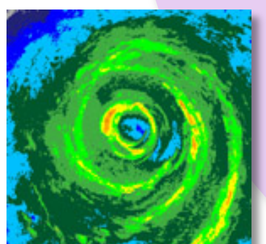
Supernova ignition depends on dimensionality of neutrino heating



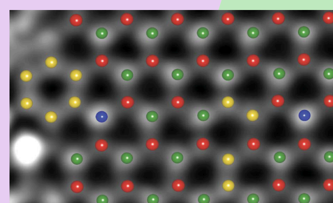
Flowering plants cool the earth



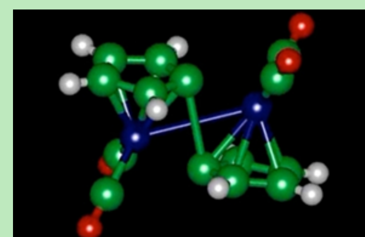
Carbon-based transistor junction created



Higher temperatures in Pliocene era linked to cyclones

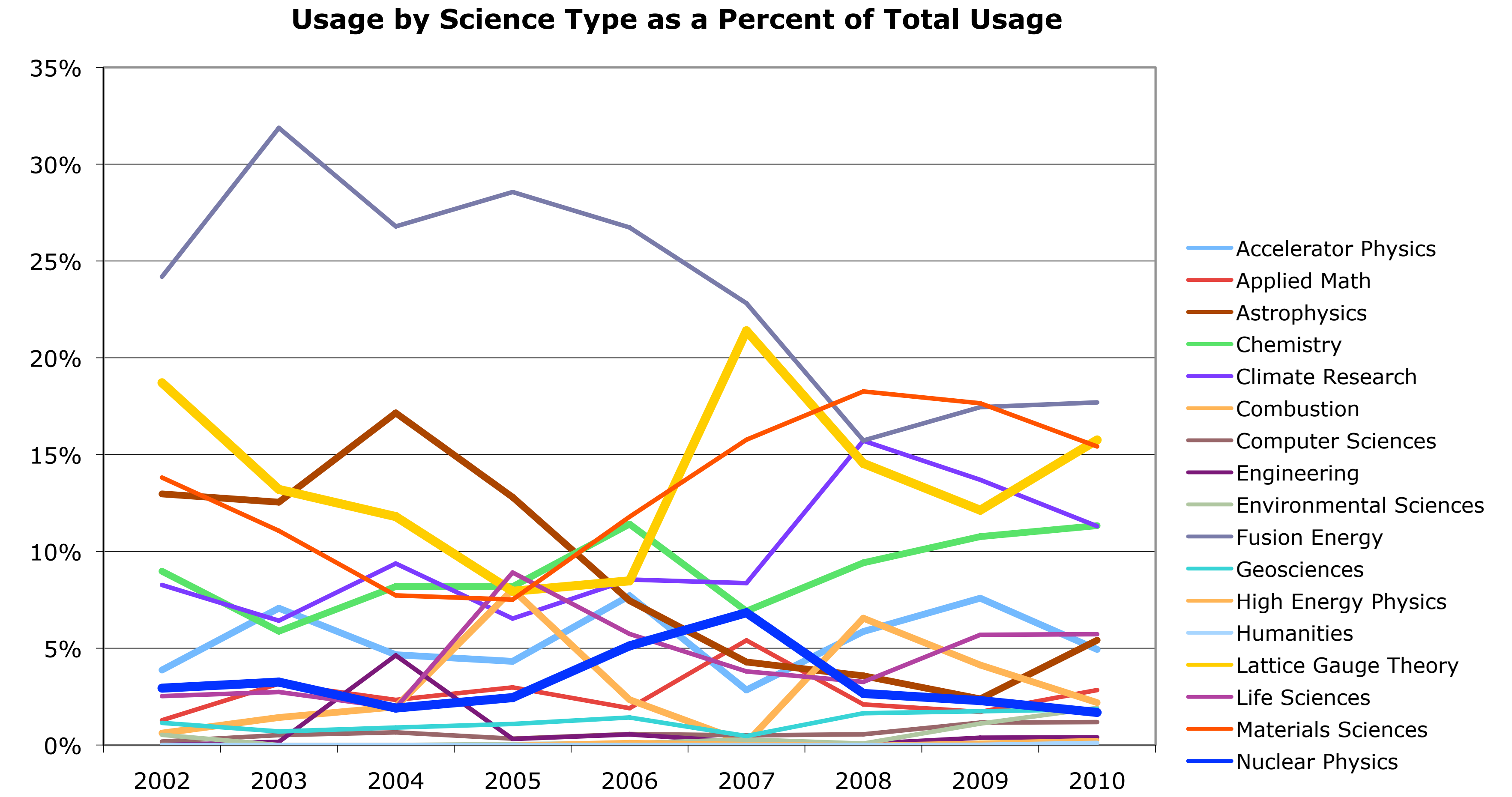


Experiments+simulations “show” individual atoms of boron, carbon, & nitrogen.



Candidate molecule for reversible storage of solar energy identified

DOE Priorities for NERSC Change Over Time



U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB 75 Years of World-Class Science 1931-2006

NERSC ASCR's Computing Facilities

NERSC at LBNL

- 1000+ users,100+ projects
- Allocations:
  - 80% DOE program manager control
  - 10% ASCR Leadership Computing Challenge*
  - 10% NERSC reserve
- Science includes all of DOE Office of Science
- Machines procured competitively

LCFs at ORNL and ANL

- 100+ users 10+ projects
- Allocations:
  - 60% ANL/ORNL managed INCITE process
  - 30% ACSR Leadership Computing Challenge*
  - 10% LCF reserve
- Science limited to largest scale; no limit to DOE/SC
- Machines procured through partnerships

NERSC Strategy: *Science First*

- **Response to scientific needs**
  - **Requirements setting activities**
- **Support computational science:**
  - **Provide effective machines that support fast algorithms**
  - **Deploy with flexible software**
  - **Help users with expert services**
- **NERSC future priorities are driven by science:**
  - **Increase application capability: "usable Exascale"**
  - **For simulation and data analysis**

NERSC Systems

Large-Scale Computing Systems


Franklin (NERSC-5): Cray XT4

- 9,532 compute nodes; 38,128 cores
- ~25 Tflop/s on applications; 356 Tflop/s peak


Hopper (NERSC-6): Cray XE6

- 6,384 compute nodes, 153,216 cores
- 120 Tflop/s on applications; 1.3 Pflop/s peak

Clusters

140 Tflops total

Carver

- IBM iDataplex cluster

PDSF (HEP/NP)

- ~1K core cluster

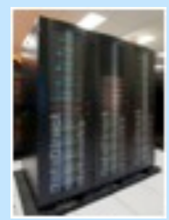
Magellan Cloud testbed

- IBM iDataplex cluster

GenePool (JGI)

- ~5K core cluster

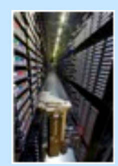
NERSC Global Filesystem (NGF)



Uses IBM's GPFS

- 1.5 PB capacity
- 5.5 GB/s of bandwidth

HPSS Archival Storage



- 40 PB capacity
- 4 Tape libraries
- 150 TB disk cache

Analytics

Euclid

(512 GB shared memory)

Dirac GPU testbed (48 nodes)

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB | 75 Years of World-Class Science 1931-2006

What is a Cloud? Definition

According to the National Institute of Standards & Technology (NIST)…

- *Resource pooling.* Computing resources are pooled to serve multiple consumers.
- *Broad network access.* Capabilities are available over the network.
- *Measured Service.* Resource usage is monitored and reported for transparency.
- *Rapid elasticity.* Capabilities can be rapidly scaled out and in (pay-as-you-go)
- *On-demand self-service.* Consumers can provision capabilities automatically.

What is a cloud? Cloud Models

Hardware focus

Applicaiton focus

Infrastructure as a Service (IaaS)

Provisions processing, storage, networks, and other fundamental computing resources. Consumer can deploy and run arbitrary software, including OS.

- Amazon EC2
- RackSpace

Platform as a Service (PaaS)

Provides programming languages and tools. Consumer applications created with provider's tools.

- Microsoft Azure
- Google AppEngine

Software as a Service (SaaS)

Provides applications on a cloud infrastructure. Consumer provides data.

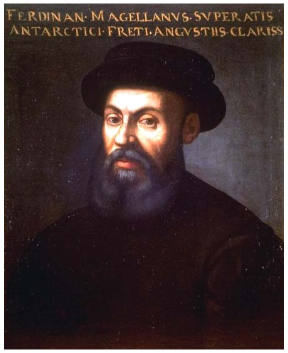
- Salesforce.com
- Google Docs
- Application Portals

Magellan

Exploring Cloud Computing

Co-located at two DOE-SC Facilities

- **Argonne Leadership Computing Facility (ALCF)**
- **National Energy Research Scientific Computing Center (NERSC)**
- **Funded by DOE under the American Recovery and Reinvestment Act (ARRA)**

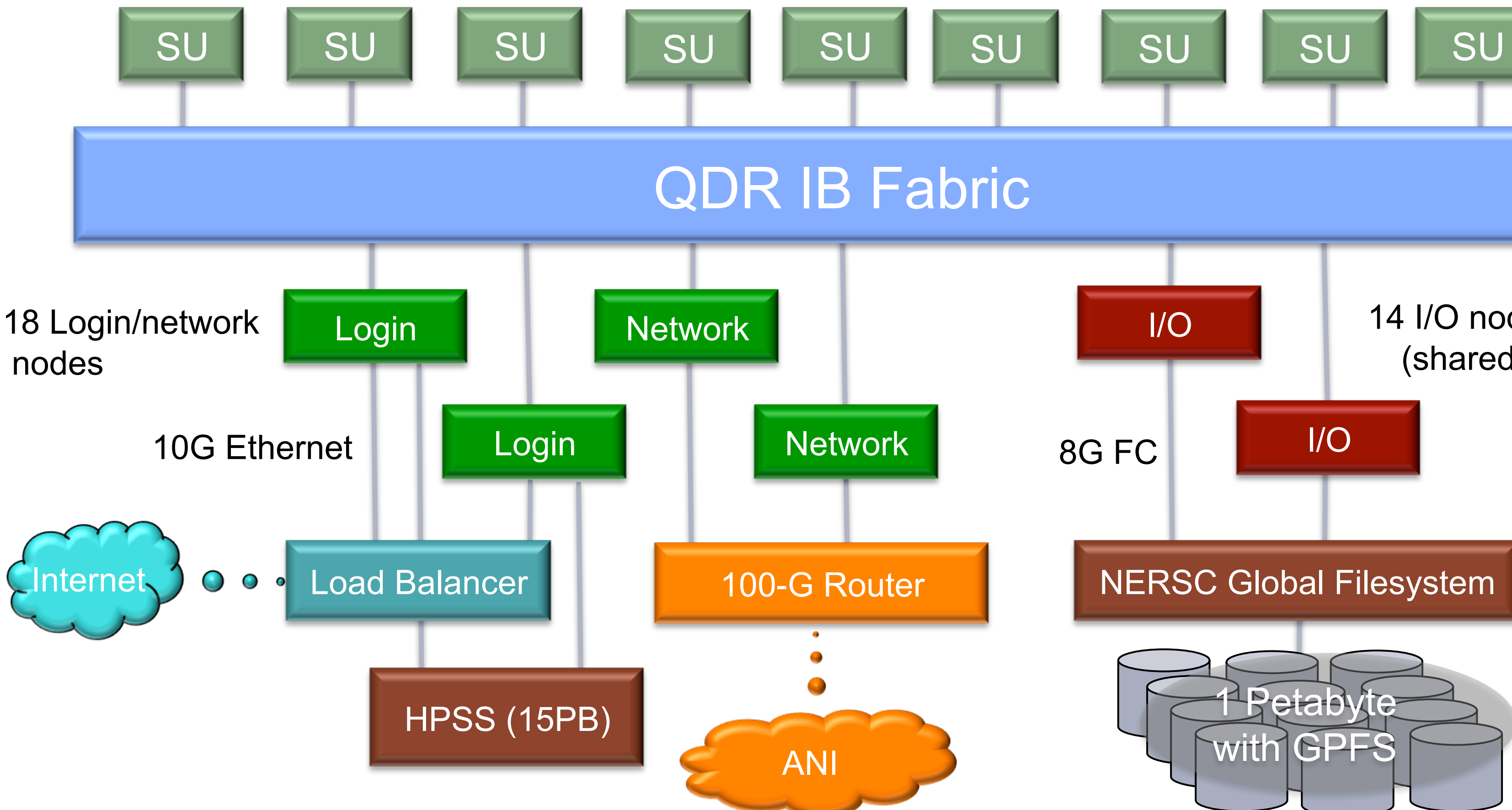


Magellan Scope

- **Mission**
  - Determine the appropriate role for private cloud computing for DOE/SC midrange workloads
- **Approach**
  - Deploy a test bed to investigate the use of cloud computing for mid-range scientific computing
  - Evaluate the effectiveness of cloud computing models for a wide spectrum of DOE/SC applications

Magellan Test Bed at NERSC
Purpose-built for Science Applications

720 nodes, 5760 cores in 9 Scalable Units (SUs) ➔ 61.9 Teraflops
SU = IBM iDataplex rack with 640 Intel Nehalem cores



Magellan Computing Models

Purpose	Comments
CLUSTER RESOURCES INC.	Mix of node types and queues. Future: Dynamic provisioning, VMs, and virtual private clusters
Eucalyptus Systems	Can expand based on demand. Supports: VMs, block storage
hadoop	MapReduce. Both configured with HDFS

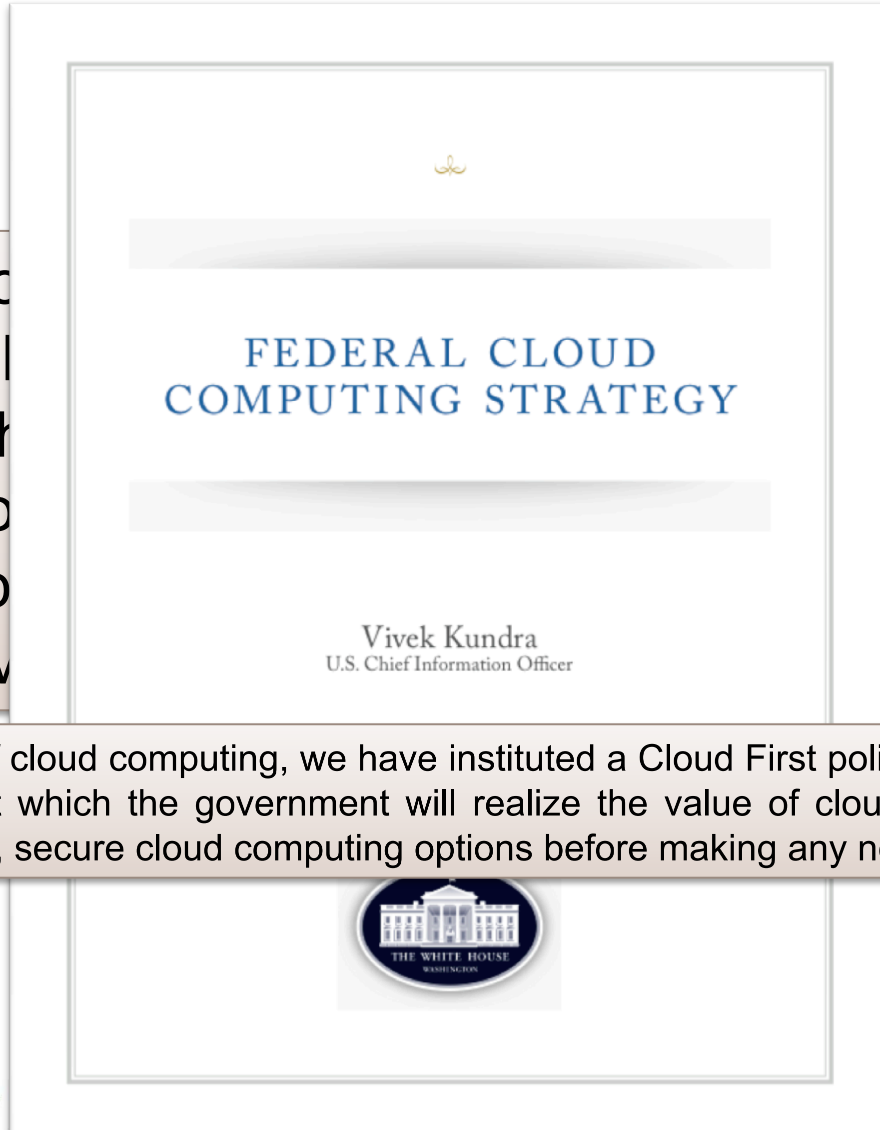
Magellan Research Agenda and Lines of Inquiry

- Are the *open source* cloud software stacks ready for DOE HPC science?
- Can DOE cyber security requirements be met within a cloud?
- Are the new cloud programming models useful for scientific computing?
- Can DOE HPC applications run efficiently in the cloud? What applications are suitable for clouds?
- How usable are cloud environments for scientific applications?
- When is it cost effective to run DOE HPC science in a cloud?
- What are the ramifications for data intensive computing?

The Push towards Clouds

“39
on
--S
Re

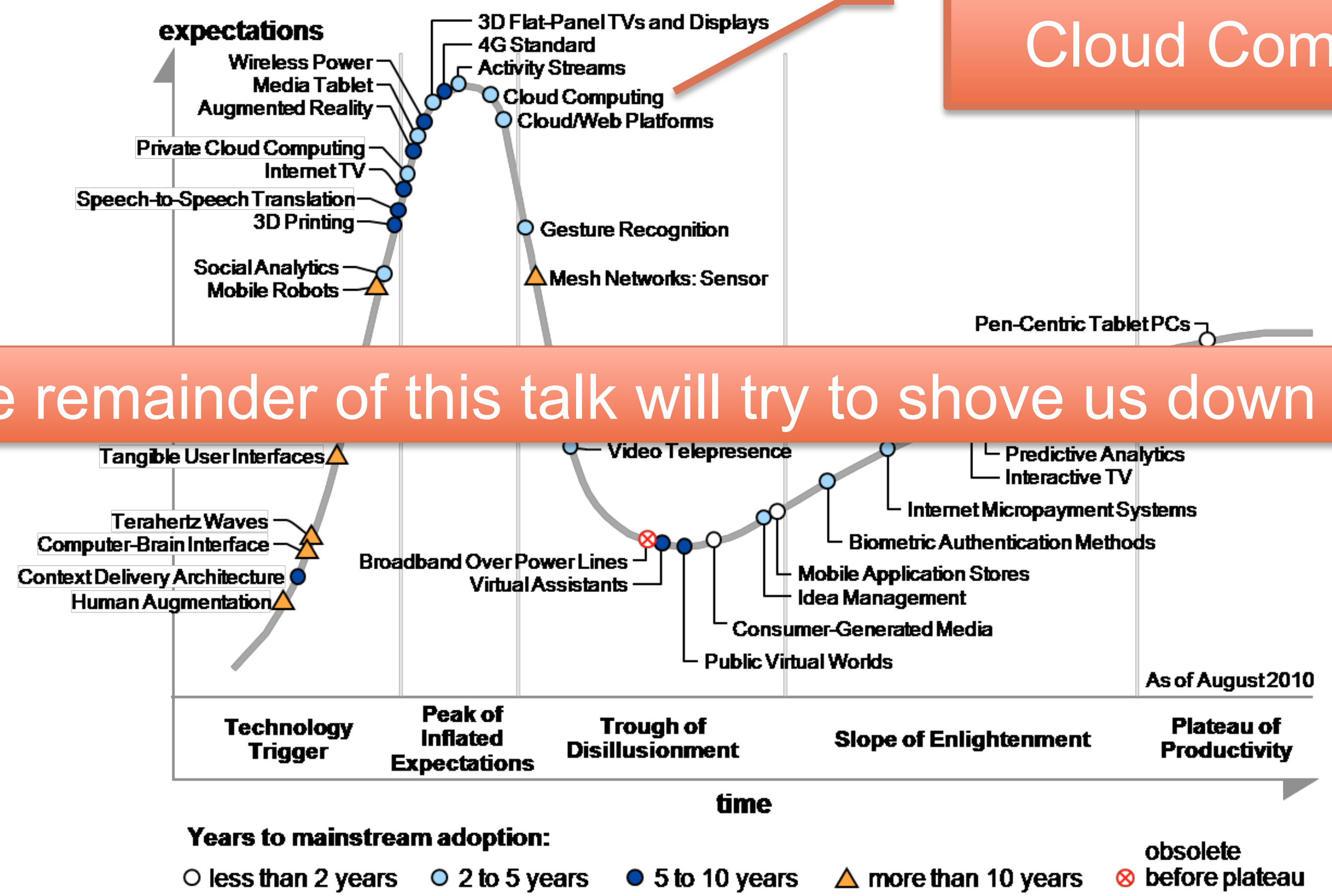
A survey c … anies across
the world l … ercent on
average) h … -party-
hosted clo … with last
year (28 p
--Ovum (M



To harness the benefits of cloud computing, we have instituted a Cloud First policy. This policy is intended to accelerate the pace at which the government will realize the value of cloud computing by requiring agencies to evaluate safe, secure cloud computing options before making any new investments

The Hype around Clouds

Gartner's 2010 Emerging Technologies Hype Cycle



Cloud Computing

The remainder of this talk will try to shove us down the hill

U.S. DEPARTMENT OF ENERGY | Office of Science

75 Years of World-Class Science 1931-2006 BERKELEY LAB

Cloud Misconceptions

- Clouds are simple to use and don’t require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore’s Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.

Are Clouds Easy to Use?

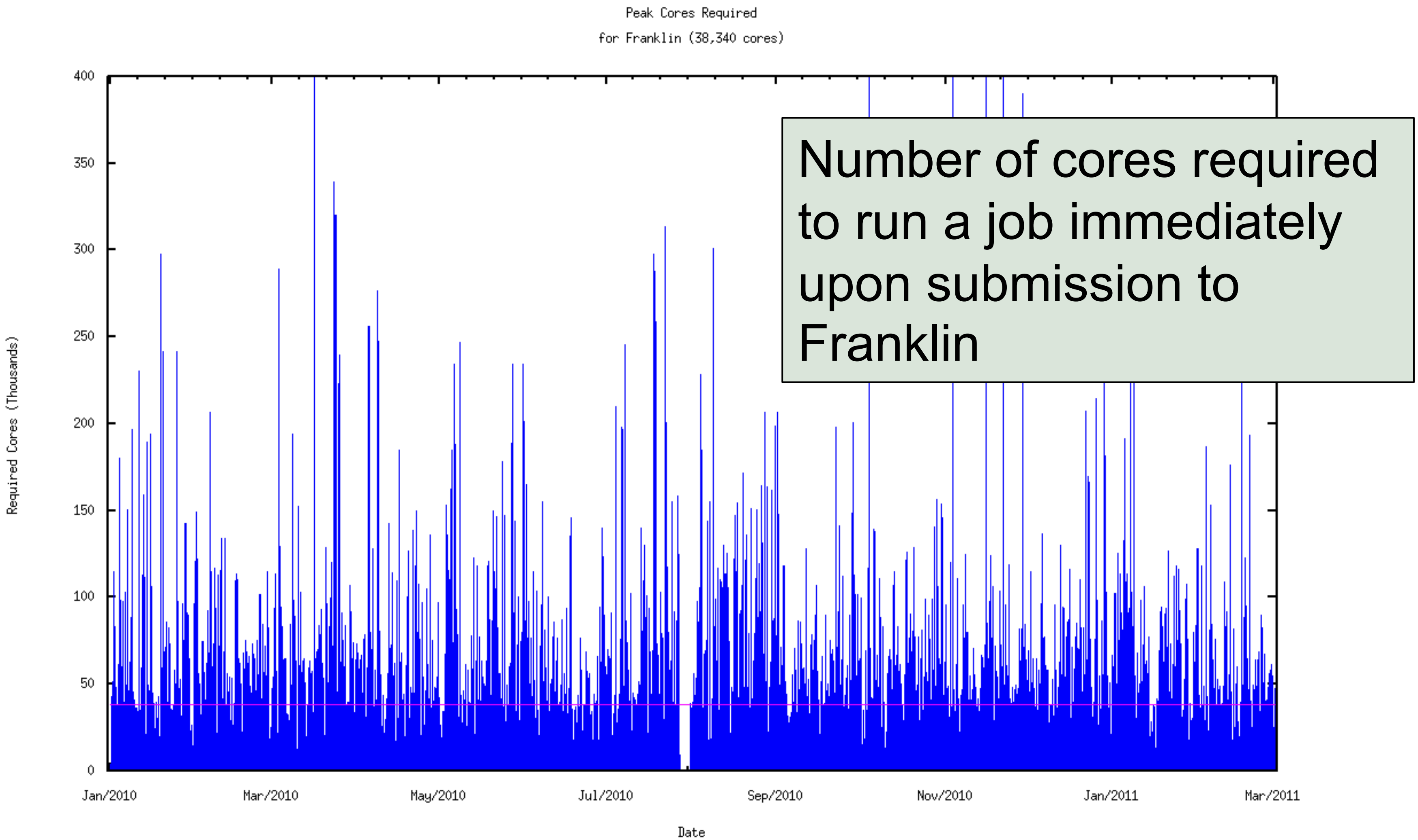
From Experience with Magellan we have Learned

- IaaS Clouds can require significant amounts of system administration expertise
- Images must be customized for the application
- No batch environment. No global file system.
- Users must properly secure and protect their images and instances.
- Do we want to turn scientists into system administrators?

Cloud Misconceptions

- Clouds are simple to use and don’t require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore’s Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.

Is the Cloud Elastic enough for HPC?



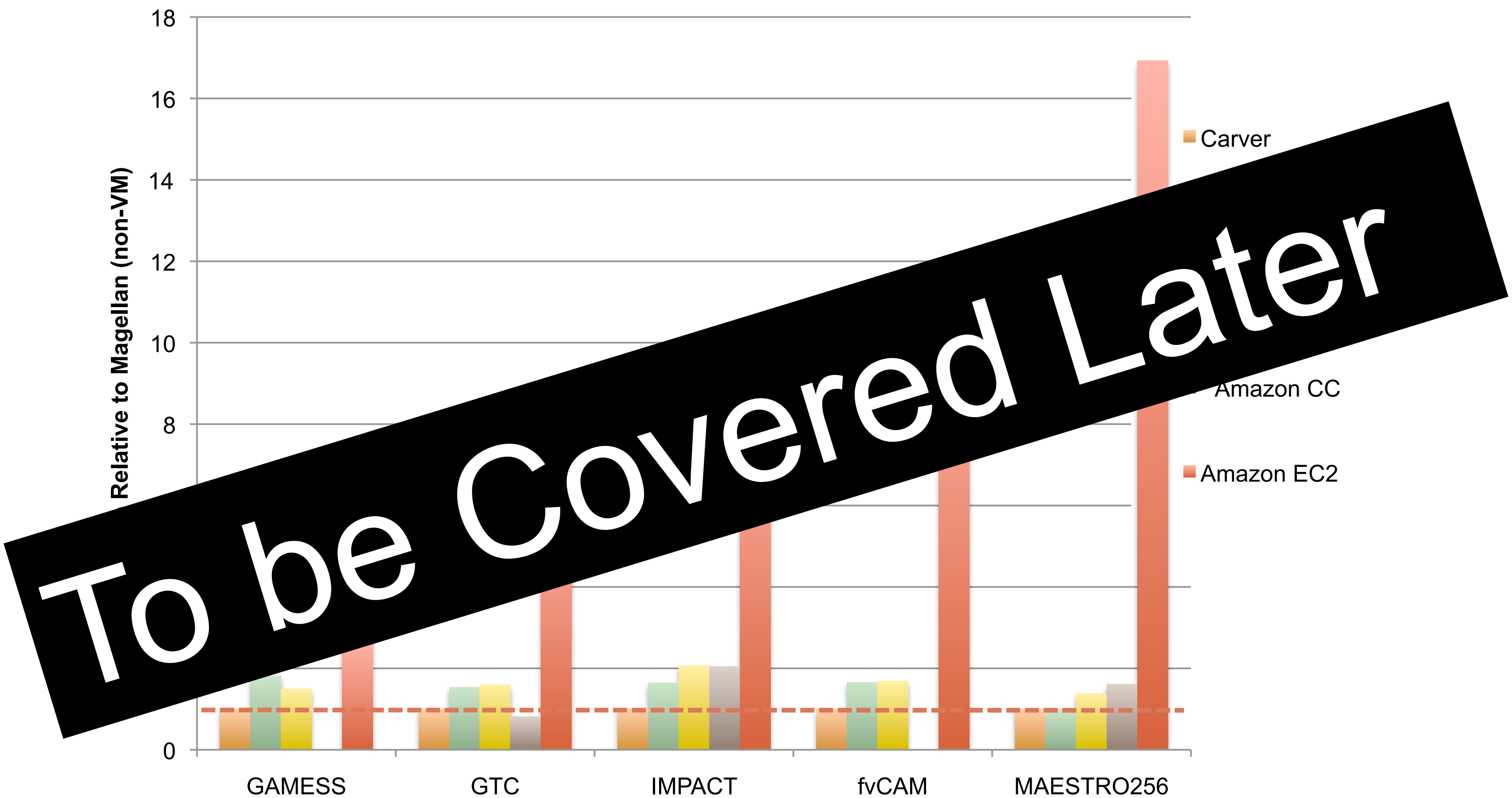
U.S. DEPARTMENT OF ENERGY | Office of Science

Berkeley Lab — 75 Years of World-Class Science 1931-2006

Cloud Misconceptions

- Clouds are simple to use and don’t require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore’s Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.

Application Performance
Application Benchmarks



Assumptions for cost saving from Clouds aren't true for HPC Centers.

EFFICIENCY	
Cloud Benefits	**Current Environment**
• Improved asset utilization (server utilization > 60-70%) • Aggregated demand and accelerated system consolidation (e.g., Federal Data Center Consolidation Initiative) • Improved productivity in application development, application management, network, and end-user	• Low asset utilization (server utilization < 30% typical) • Fragmented demand and duplicative systems • Difficult-to-manage systems
AGILITY	
Cloud Benefits	**Current Environment**
• Purchase "as-a-service" from trusted cloud providers • Near-instantaneous increases and reductions in capacity • More responsive to urgent agency needs	• Years required to build data centers for new services • Months required to increase capacity of existing services
INNOVATION	
Cloud Benefits	**Current Environment**
• Shift focus from asset ownership to service management • Tap into private sector innovation • Encourages entrepreneurial culture • Better linked to emerging technologies (e.g., devices)	• Burdened by asset management • De-coupled from private sector innovation engines • Risk-adverse culture

- HPC Centers run at >90% CPU utilization and >90% scheduled utilization.
- HPC Centers partner with Vendors to field cutting edge systems
- HPC more aggressive with technical risks

From the Federal Cloud Computing Strategy

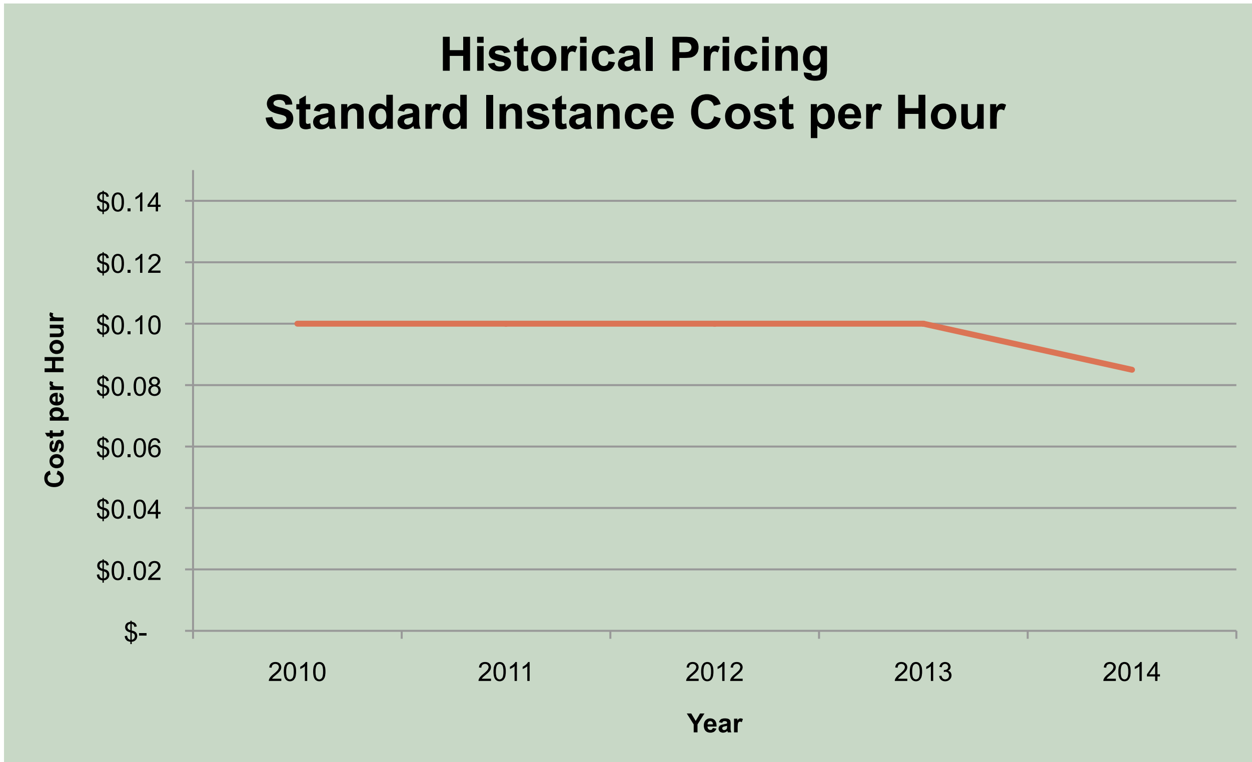
Enterprise IT versus HPC

	Traditional Enterprise IT	HPC Centers
Typical Load Average	30% *	90%
Computational Needs	Bounded computing requirements – Sufficient to meet customer demand or transaction rates. (i.e. If you gave a typical business free computing, would they suddenly be able to take advantage of it?)	Virtually unbounded requirements – Scientist always have larger, more complicated problems to simulate or analyze.
Scaling Approach	Scale-in. Emphasis on consolidating in a node using virtualization	Scale-Out Applications run in parallel across multiple nodes.

Cloud Misconceptions

- Clouds are simple to use and don't require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore's Law without additional investment.
- Commercial clouds are much cheaper than operating your own system.

Cloud Pricing Trends



The cost of a standard cloud instance has dropped 18% over 5 years. Meanwhile, cores per socket have increased 2x-5x per socket in the same time-frame at roughly constant cost.

Cloud Misconceptions

- Clouds are simple to use and don't require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore's Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.

Cost of NERSC in the Cloud

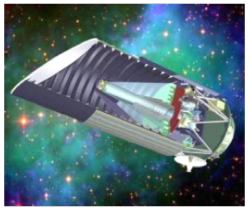
Component	Cost
Compute Systems (1.38B hours)	$180,900,000
HPSS (17 PB)	$12,200,000
File Systems (2 PB)	$2,500,000
Total (Annual Cost)	**$195,600,000**

Assumes 85% utilization and zero growth in HPSS and File System data.
Doesn't include the 2x-10x performance impact that has been measured.
This still only captures about 65% of NERSC's $55M annual budget.
No consulting staff, no administration, no support.

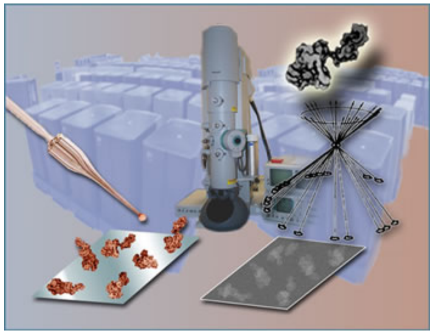
Where are (commercial) clouds effective?

- **Individual projects with high-burst needs.**
  - **Avoid paying for idle hardware**
  - **Access to larger scale (elasticity)**
  - **Alternative: Pool with other users (condo model)**
- **High-Throughput Applications with modest data needs**
  - **Bioinformatics**
  - **Monte-Carlo simulations**
- **Infrastructure Challenged Sites**
  - **Facilities cost >> IT costs**
  - **Consider the long-term costs**
- **Undetermined or Volatile Needs**
  - **Use Clouds to baseline requirements and build in-house**

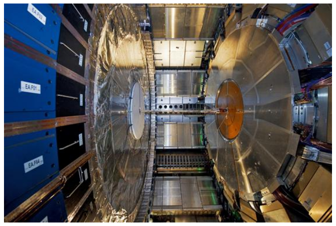
What HPC Can Learn from Clouds



- **Need to support surge computing**
  - Predictable: monthly processing of genome data; nightly processing of telescope data
  - Unpredictable: computing for disaster recovery; response to facility outage



- **Support for tailored software stack**
- **Different levels of service**
  - Virtual private cluster: guaranteed service
  - Regular: low average wait time
  - Scavenger mode, including preemption



What should an HPC Cloud Solution Look Like?

- High-performance interconnect (high bandwidth, low latency) with fast access from the application
- Fast access to a high-performance file system
- No penalty to gather resources
- Non-Virtualized/bare-metal?

Is an HPC Center a Cloud?

HPC Centers ?

- *Resource pooling.* ✓
- *Broad network access.* ✓
- *Measured Service.* ✓
- *Rapid elasticity.*
  - Usage can grow/shrink; pay-as-you-go. ✓
- *On-demand self-service.*
  - Users cannot demand (or pay for) more service than their allocation allows X
  - Jobs often wait for hours or days in queues

It's All Business

- Cloud computing is a business model
- It can be applied to HPC systems as well as traditional clouds (ethernet clusters)
- Can get on-demand elasticity through:
  - Idle hardware (at ownership cost)
  - Sharing cores/nodes (at performance cost)
  - Scheduling policies (pre-emption)

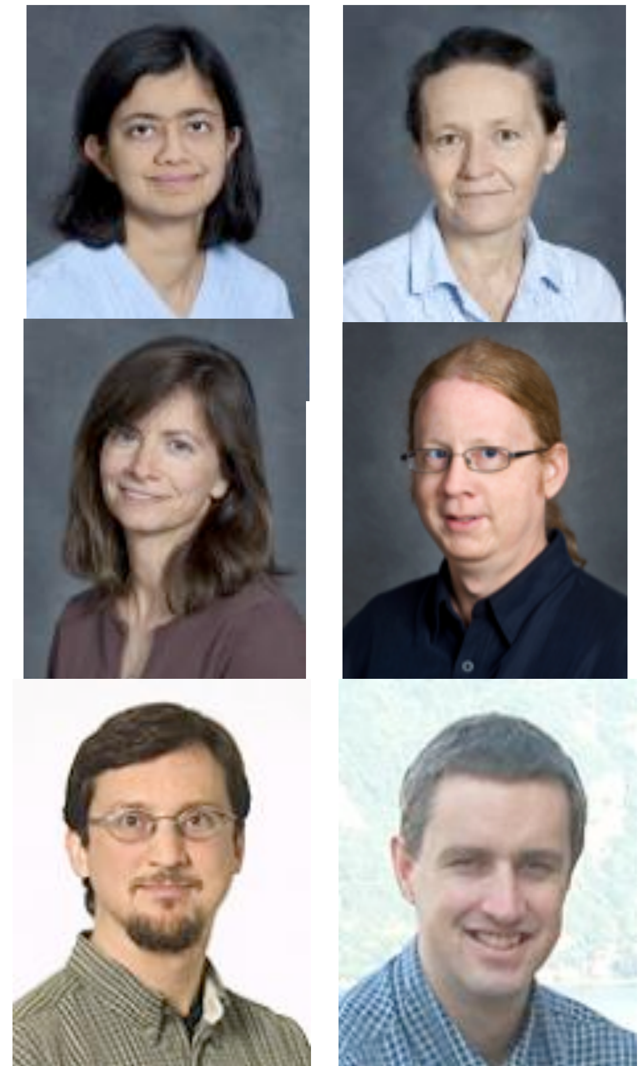
Closing Remarks

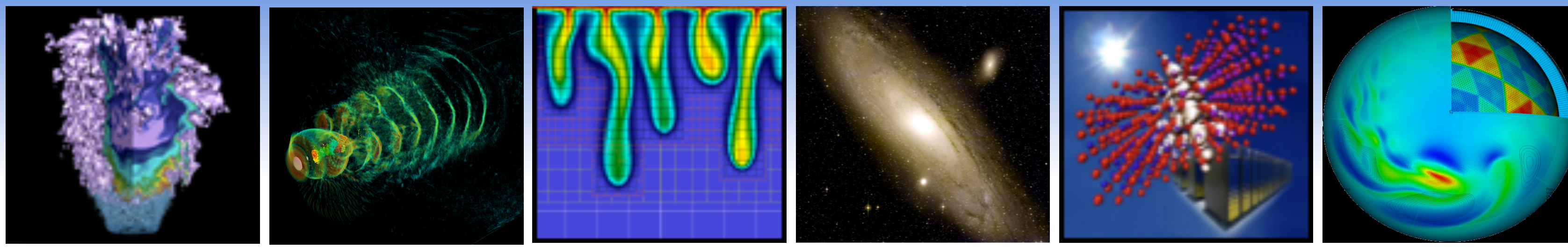
- Cloud Computing is changing the face of computing
- HPC Centers can learn new tricks from the Cloud Computing space
- Cloud Computing is not a pancea: they can be more difficult to use, slower and more expensive than in-house solutions.

NERSC

Acknowledgements

- Lavanya Ramakrishnan
- Iwona Sakrejda
- Tina Declerck
- Others
  - Keith Jackson
  - Nick Wright
  - John Shalf
  - Krishna Muriki (not picture)





Thank you!

Contact Info:
Shane Canon
Scanon@lbl.gov
magellan.nersc.gov