National Energy Research Scientific Computing Center (NERSC)

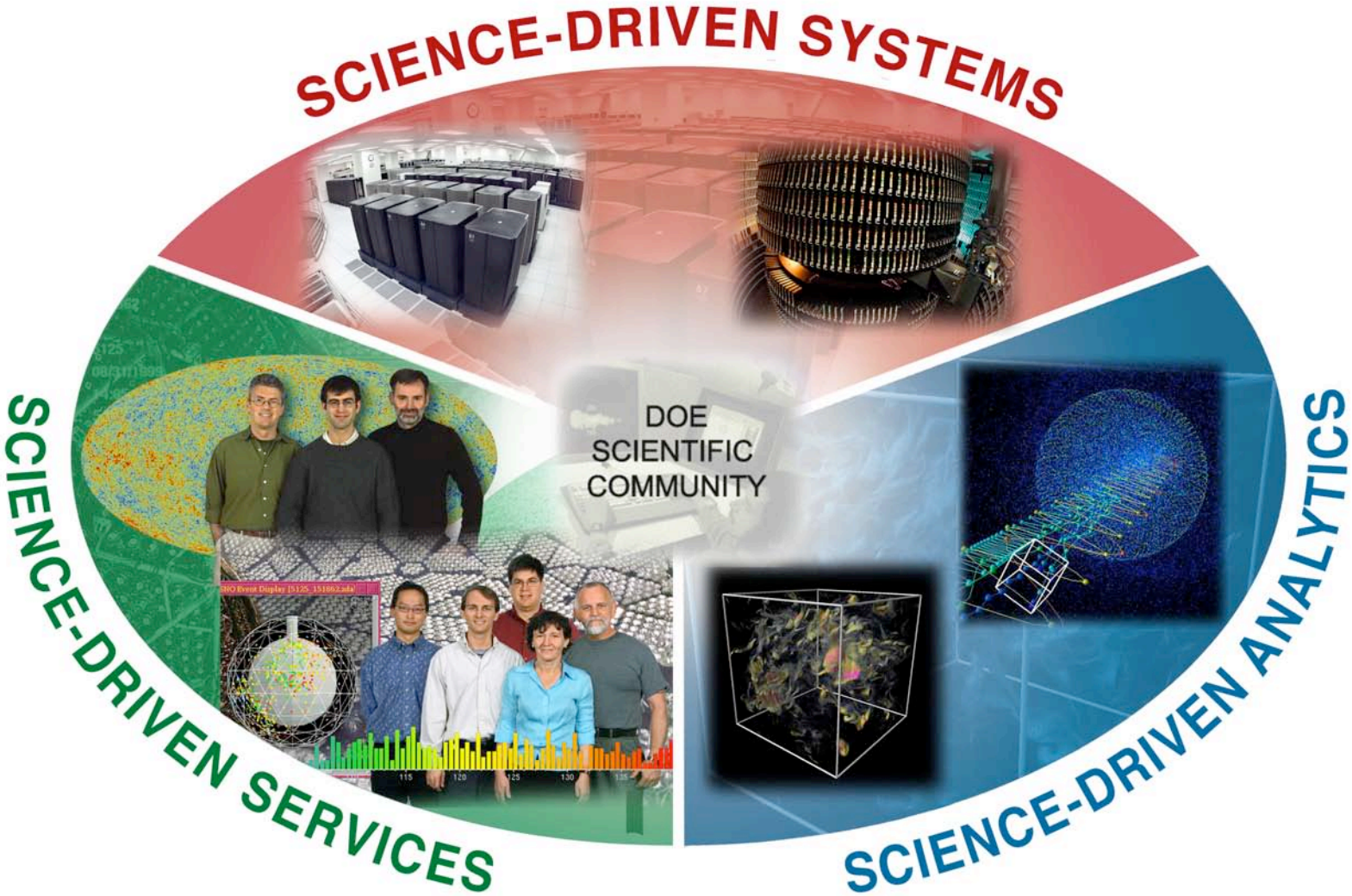
Jonathan Carter

Scientific Impacts and Opportunities in Computing
January 10, 2008

NERSC Mission

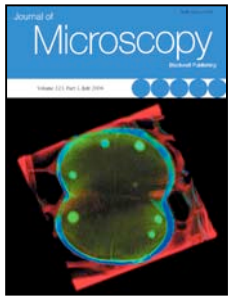
The mission of the National Energy Research Scientific Computing Center (NERSC) is to accelerate the pace of scientific discovery by providing high performance computing, information, data, and communications services for research sponsored by the DOE Office of Science (SC).

Science-Driven Computing Strategy 2006 -2010



Office of Science
U.S. DEPARTMENT OF ENERGY

Science-Driven Computing



National Energy Research Scientific Computer Center (NERSC)

- computational facility for open science
- international user community
- 2500 users
- 300 projects, including 5-10 INCITE awards

NERSC is enabling new science

Office of Science
U.S. DEPARTMENT OF ENERGY

Science-Driven Services

Office of Science
U.S. DEPARTMENT OF ENERGY

Science-Driven Services

- Provide the entire range of services from high-quality operations to direct scientific support
- Enable a broad range of scientists to effectively use NERSC in their research
- Concentrate on resources for scaling to large numbers of processors/cores, and for supporting multidisciplinary computational science teams

NERSC User Services

- Consulting
  - Increase scientific productivity of users
- Web Content/Training
  - Distill and augment vendor and community technical information
  - Develop customized teaching materials
- Software Installation and Support
  - Efficient and robust algorithms
  - Choice and coverage, over 120 software packages installed
  - Flexible model to allow contributed software
  - Track use for commercial software
  - Testing new vendor software
- Account Support
  - Account creation
  - Passwords
  - Charging questions
  - ERCAP
- Trouble Ticket System
  - Track user inquiries, requests, across NERSC groups
- User advocate in system management decisions
  - Monitoring performance, advise on upgrades
- Procurement teams
  - Focus on usability and programmability
- User Survey

Consulting

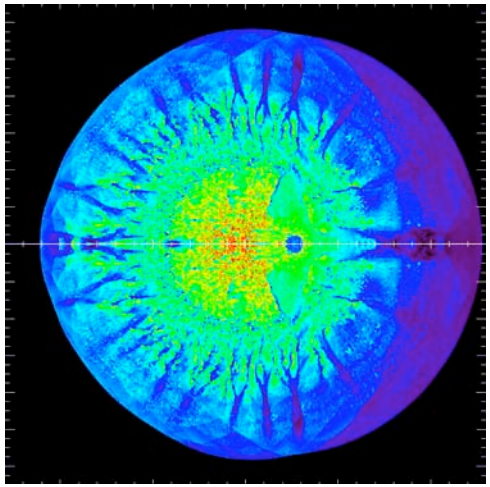
- Contact via web, mail, or phone 9x5
- All tickets are owned by a USG member who monitors through to resolution (almost 4000 incidents last year)
- Respond within 4 hours
- Resolve or "set plan" for 80% of incidents within 3 days
- Advice on building, tuning, and running on NERSC systems
- How to select the most appropriate resource: compiler flags, libraries, data storage & movement
- Scaling to higher concurrencies
  - Scaling incentives
- INCITE and SciDAC support

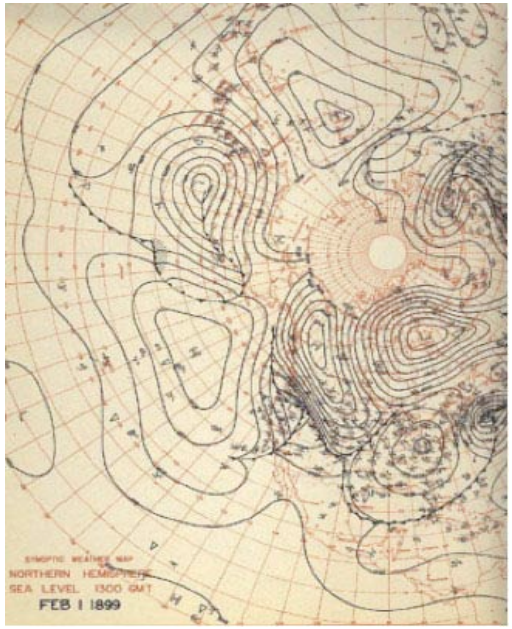
Consulting (cont)

- Consulting is not a static skill, algorithms, programming paradigms, etc. evolve over time
  - USG contributed to 25 papers and technical reports in physical sciences, computer science, and other areas
  - Attended numerous conferences and presented 13 talks and posters

Successes for INCITE 2007

- Tuning the FLASH code for memory use to correct an error condition
  - *"We could not have asked for better or more support than we got from the folks at NERSC, in helping us to get on the NERSC machines quickly, in giving the job special status, and in helping us meet the challenges of running a large job on Bassi."*
    —Don Lamb
- Tuning and debugging the global tropospheric circulation analysis code
  - Adding multi-level parallelism to bundle several associated parallel jobs

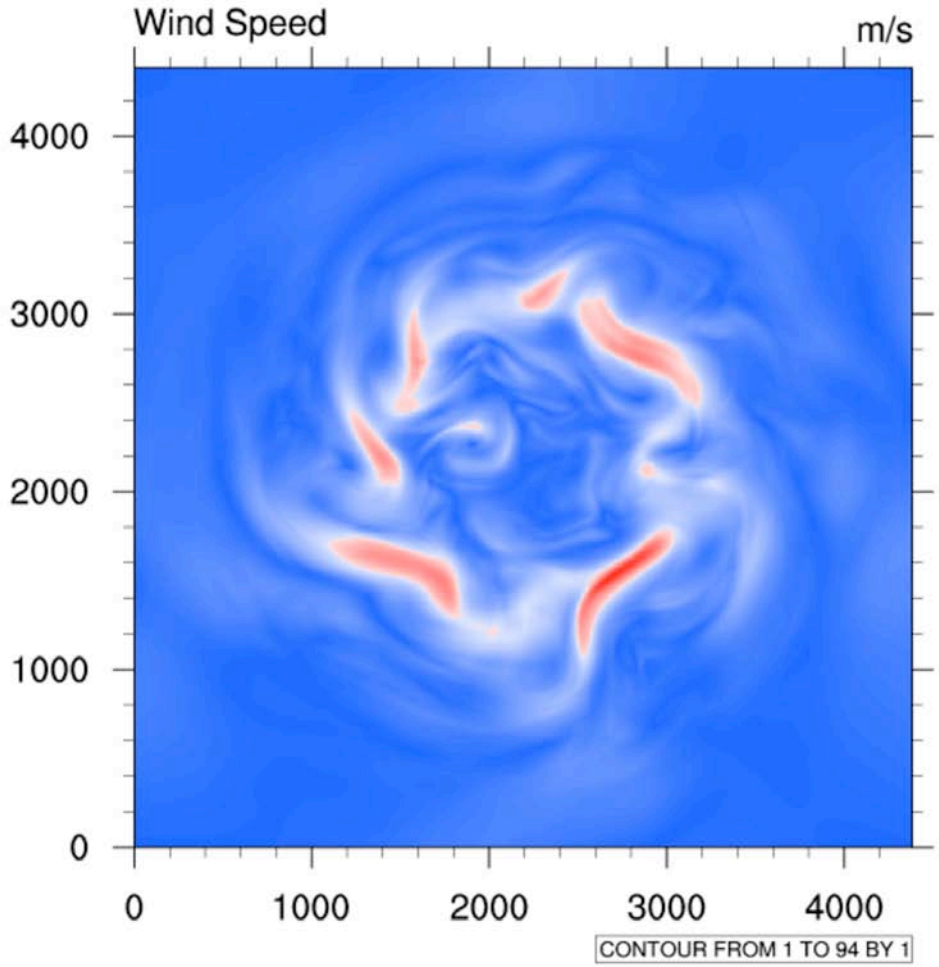


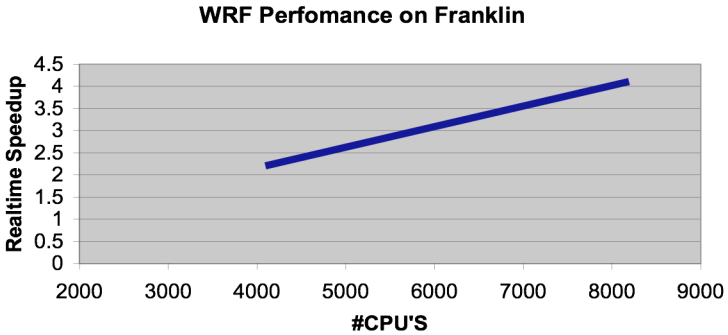


Office of Science
U.S. DEPARTMENT OF ENERGY

WRF Nature Run on Franklin

John Michalakes[a], Josh Hacker[a], Richard Loft[a], Michael O.McCracken[b], Allan Snavely[b], Nicholas J. Wright[b]
a) University Corporation for Atmospheric Research (UCAR), Boulder, CO.
b) Performance Modeling and Characterization Lab San Diego Supercomputer Center, La Jolla, CA





- WRF is a model of the atmosphere for mesoscale research and operational numerical weather prediction.
- An idealized high resolution simulation at a size and resolution never before attempted – 2 billion cells @ 5km resolution
- Science Goal: It provides a very high-resolution "truth" against which more coarse simulations or perturbation runs may be compared
- Huge volume of data! – 200GB input and 40GB per simulated hour output.
- 6.27 TF @ 8192 cores
- 3.39 TF @ 4096 cores
- Code tuning for I/O and MPI, job scheduling by NERSC

Office of Science
U.S. DEPARTMENT OF ENERGY

Science-Driven Systems

Office of Science
U.S. DEPARTMENT OF ENERGY

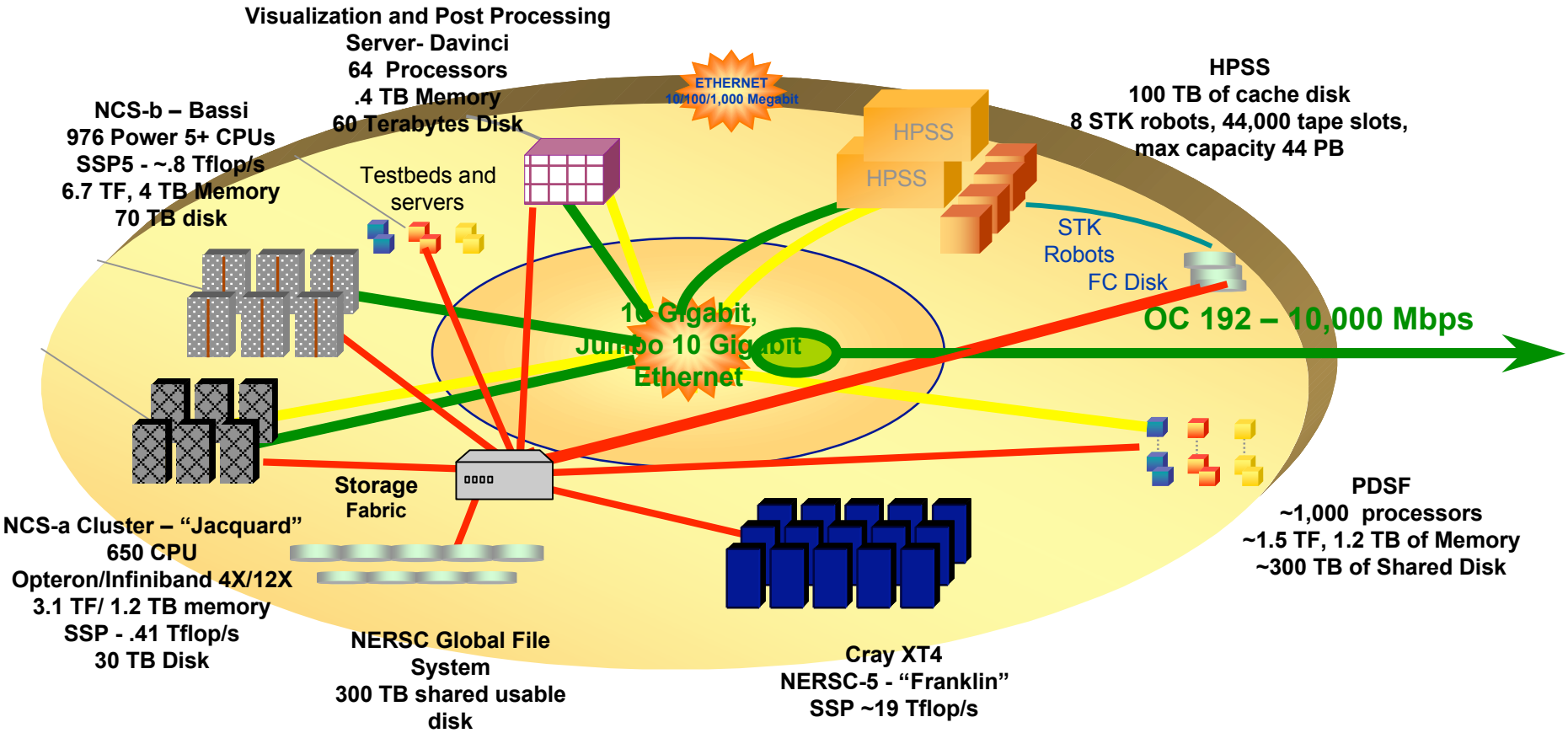
Science-Driven Systems

- Balanced and timely introduction of best new technology for complete computational systems (computing, storage, networking, analytics)
- Engage and work directly with vendors in addressing the SC requirements in their roadmaps
- Collaborate with DOE labs and other sites in technology evaluation and introduction

Computational System Strategy

- Balanced and timely introduction of the best new technologies for complete systems
  - Often the first and/or largest of its kind
- Computational systems address the widest breadth of DOE capability science
  - NERSC has at least one major computational systems in place at a time – called NERSC-n
    - NERSC's major computational systems arrive every three to four years
      - The oldest-generation system is replaced with the latest-generation system.
  - Modest-sized systems (NCSy) arrive between the major systems as funding and technology allows
    - Typically focus on a subset of the workload

NERSC Systems 2008



Office of Science
U.S. DEPARTMENT OF ENERGY

Franklin Cray XT-4 System

Newest NERSC System is a Cray XT-4

- Parallelism: 9,740 nodes with 19,480 cores
  - 102 Cabinets
  - AMD Dual core Opteron 2.6 GHz
- 39.5 TBs Aggregate Memory
- Peak performance: >100 Tflop/s
- Sustained performance: >19 Tflop/s
- Interconnect: Cray SeaStar2, 3D Torus
  - >6 TB/s Bisection Bandwidth
  - >7 GB/s Link Bandwidth
- Shared Disk: 345 TBs
- Network Connections
  - 4 x 10 Gbps + 16 x 1 Gbps

Benjamin Franklin, one of America's first scientists, performed ground breaking work in energy efficiency, electricity, materials, climate, ocean currents, transportation, health, medicine, acoustics and heat transfer.

Franklin arrives, January 16, 2007

Timeline

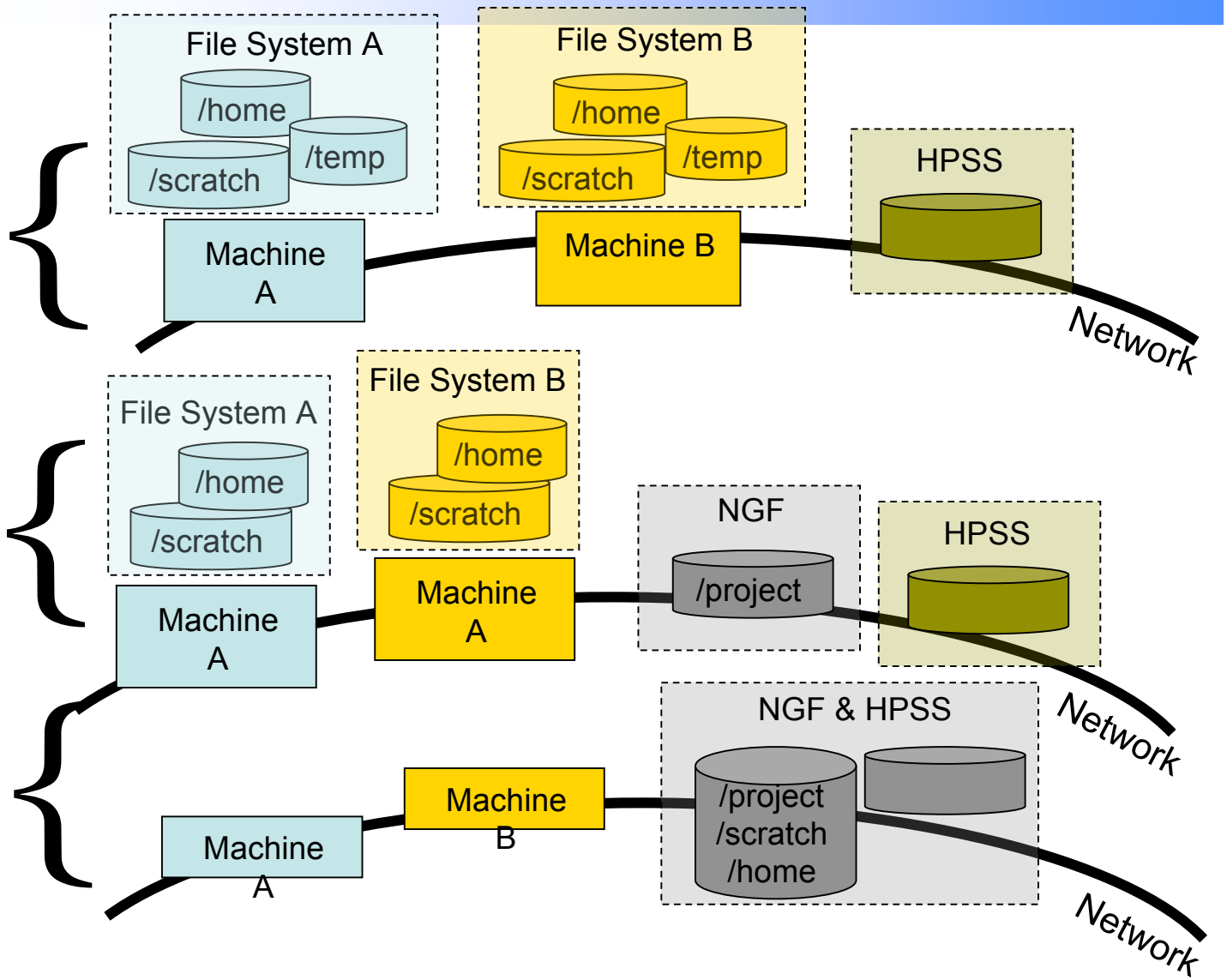
- Phase I arrives Jan 2007
- Phase 2 arrives Feb 2007
- Initial CVN testing Mar 2007
  - Memory replacement
  - File system instability (MDS, Lustre, dual-core)
  - Over 100 reported problems
- Initial CNL testing June 2007
  - UME rates
  - Batch scheduler
- Early user program July 2007
- Acceptance Test begins Aug 2007
- Acceptance concludes Oct 2007
- Enters production service Jan 2008

NERSC Global Filesystem (NGF) in Full Production

- After thorough evaluation and testing phase in production
- Based on IBM GPFS
- Seamless data access from all of NERSC's computational and analysis resources
- Single unified namespace makes it easier for users to manage their data across multiple system
- First production global filesystem spanning four platforms, three architectures, and four different vendors

NERSC Storage Roadmap

- **Past**
  - **Local Disk**
    - **/scratch**
    - **/home**
    - **/tmp**
  - **HPSS**
- **Now**
  - **Local Disk**
    - **/scratch**
    - **/home**
  - **NGF**
    - **/project**
  - **HPSS**
- **Future**
  - **Local Disk**
    - **/home**
  - **NGF-HPSS**
    - **/scratch**
    - **/project**



Office of Science
U.S. DEPARTMENT OF ENERGY

Future Systems Plans

- Analytics server to be installed 2008 to replace Davinci
- Possible AMD quad-core upgrade to Franklin, based on benchmarking, program feedback, user experience at ORNL
- Finalize NERSC6 RFP based on budget, program budget

Science-Driven Analytics

Office of Science
U.S. DEPARTMENT OF ENERGY

What is the Analytics Program at NERSC?

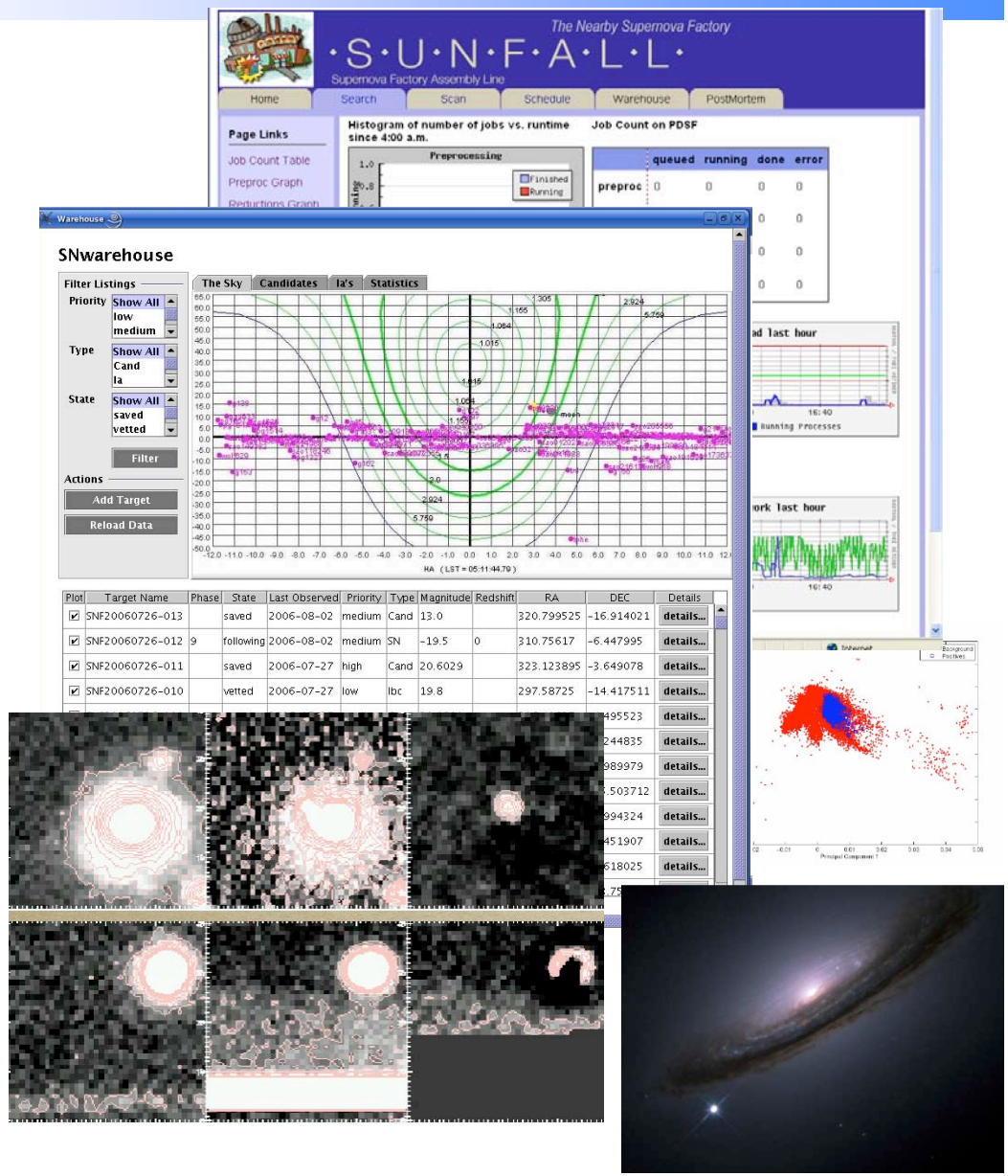
- Analytics is the *"science of analysis"*.
- At NERSC, the Analytics Program is the confluence of several key technologies:
  - Data management
    - Data storage/retrieval/sharing/movement, data indexing/querying
  - Data analysis and data mining
    - Compare datasets, find features within a dataset.
  - Data visualization of data
    - Primary method of data exploration.
  - Workflow management
    - Systematic approach to "data processing pipelines," especially those that use multiple distributed resources and automate scientific data processing activities

NERSC Analytics Services

- A dedicated interactive analysis platform that is well integrated with the rest of the Center
- Software applications, libraries, etc. that are the building blocks for Analytics solutions
- In-depth, collaborative help to science stakeholders to implement effective analytics solutions
  - There is no “one-size-fits-all” solution due to the diversity of stakeholder problems
- Serving shared licenses to NERSC users
- Technology path-finding to stay abreast of latest developments in the field, including interactions with the CS research community

Production Analytics – SNFactory

- New supernova data analysis and workflow visualization tools (Sunfall and SNwarehouse) have improved usability and situational awareness, and enabled faster and easier access to data for supernova scientists worldwide
- Advanced image processing (Fourier contour analysis) and machine learning techniques running on NERSC platforms have achieved a >40% decrease in human workload in nightly supernova search (>75% FTE)



Summary

- NERSC supports a diverse range of requirements and science
- NERSC systems, services and analytics are providing high-impact for the DOE science community
- NERSC is helping to create highly successful systems of the future