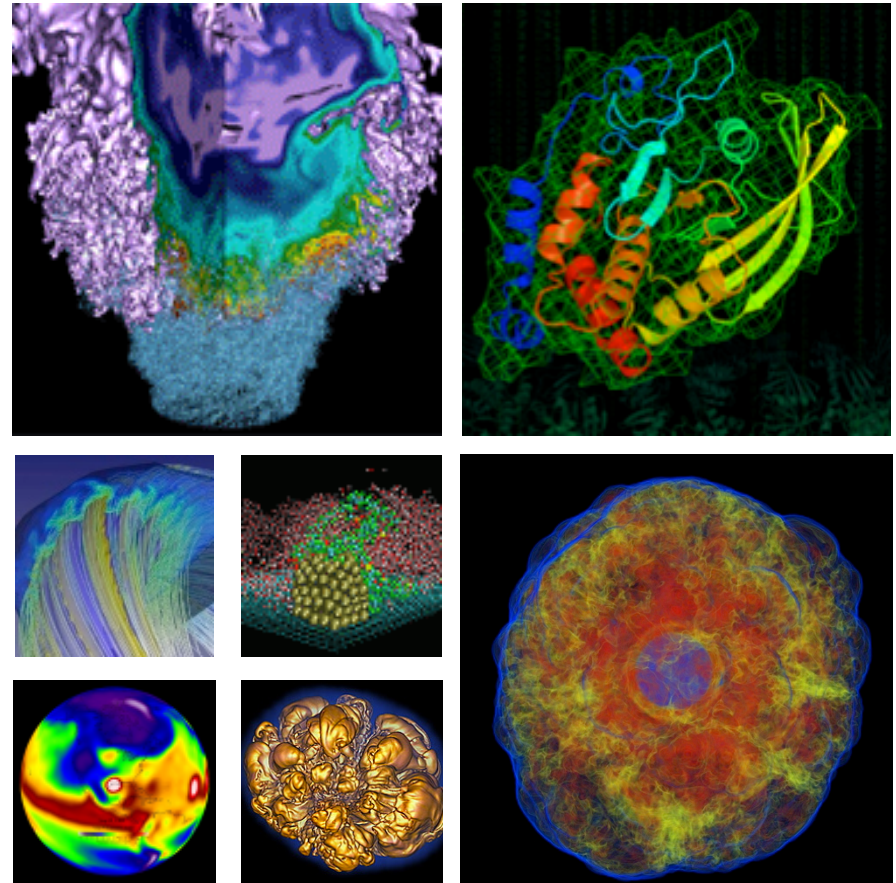


GPFS at LBNL/ NERSC



Jason Hick
Storage Systems Group

January 8, 2013

The HPC Production Facility for DOE SC



- **Our mission is to accelerate scientific discovery through High Performance Production Computing**
- **We provide computation and data storage for approximately 4,500 users, 500 projects, and 400 codes**

Science Discovery with Data Analysis

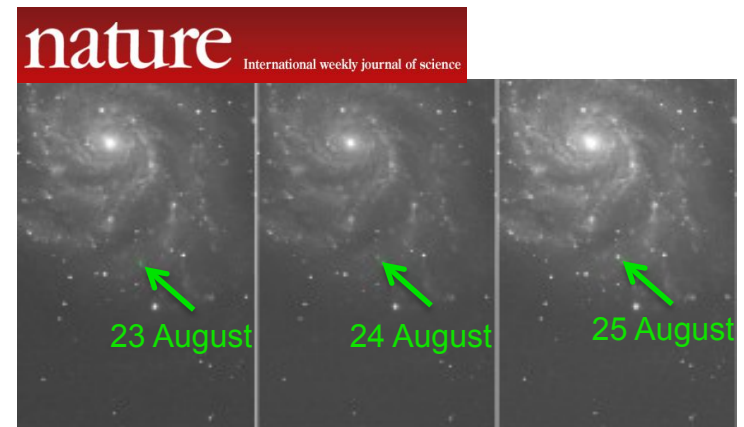


Astrophysics discover early nearby supernova

- Palomar Transient Factory runs machine learning algorithms on ~300GB/night delivered by ESnet “science network”
- Rare glimpse of a supernova within 11 hours of explosion, 20M light years away
- Telescopes world-wide redirected within 1 hour

Data systems essential to science success

- GPFS /project file system mounted on resources centerwide, brings broad range of resources to the data
- Data Transfer Nodes and Science Gateway Nodes improve data acquisition, access and processing capabilities

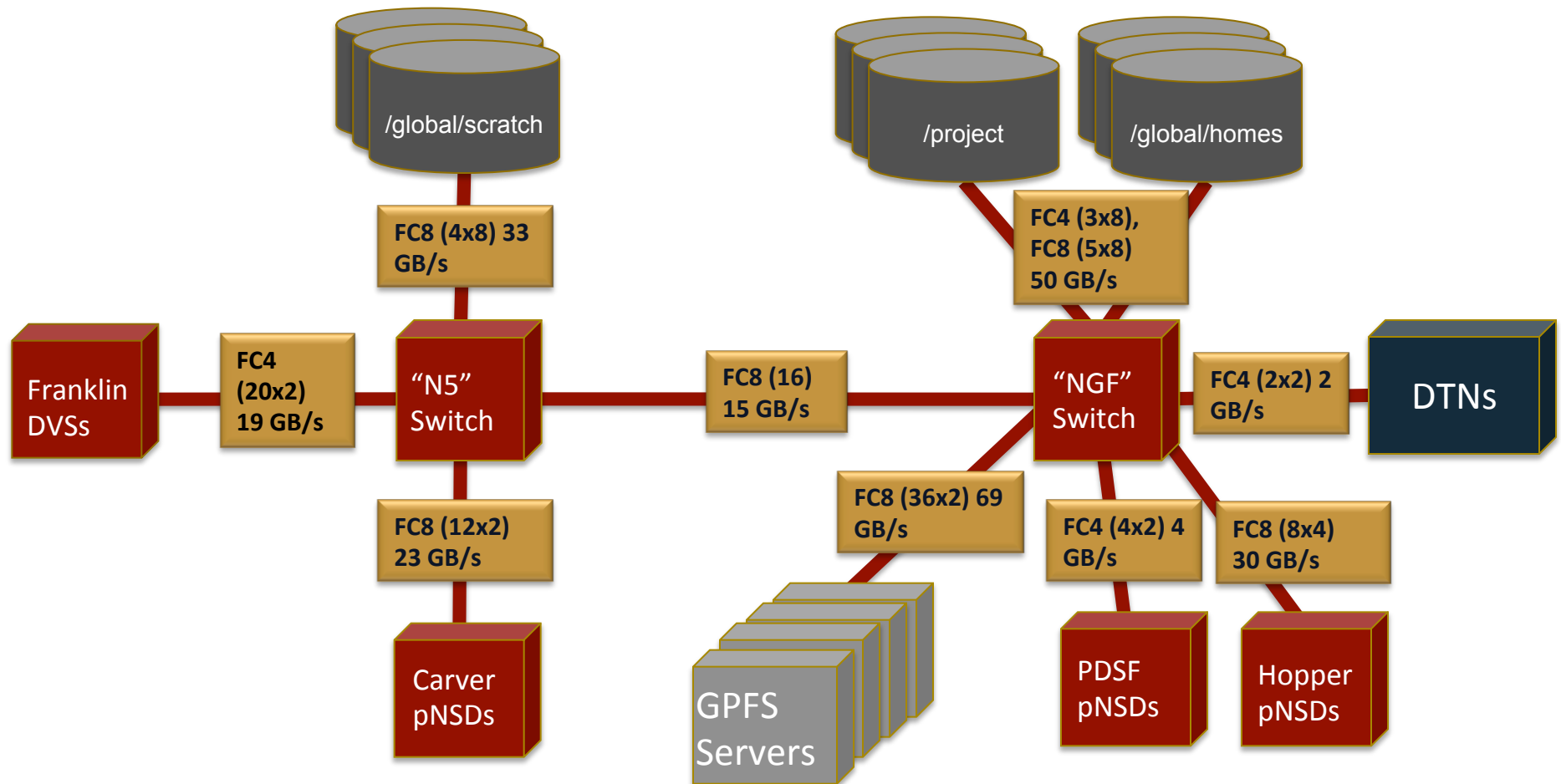


GPFS resources

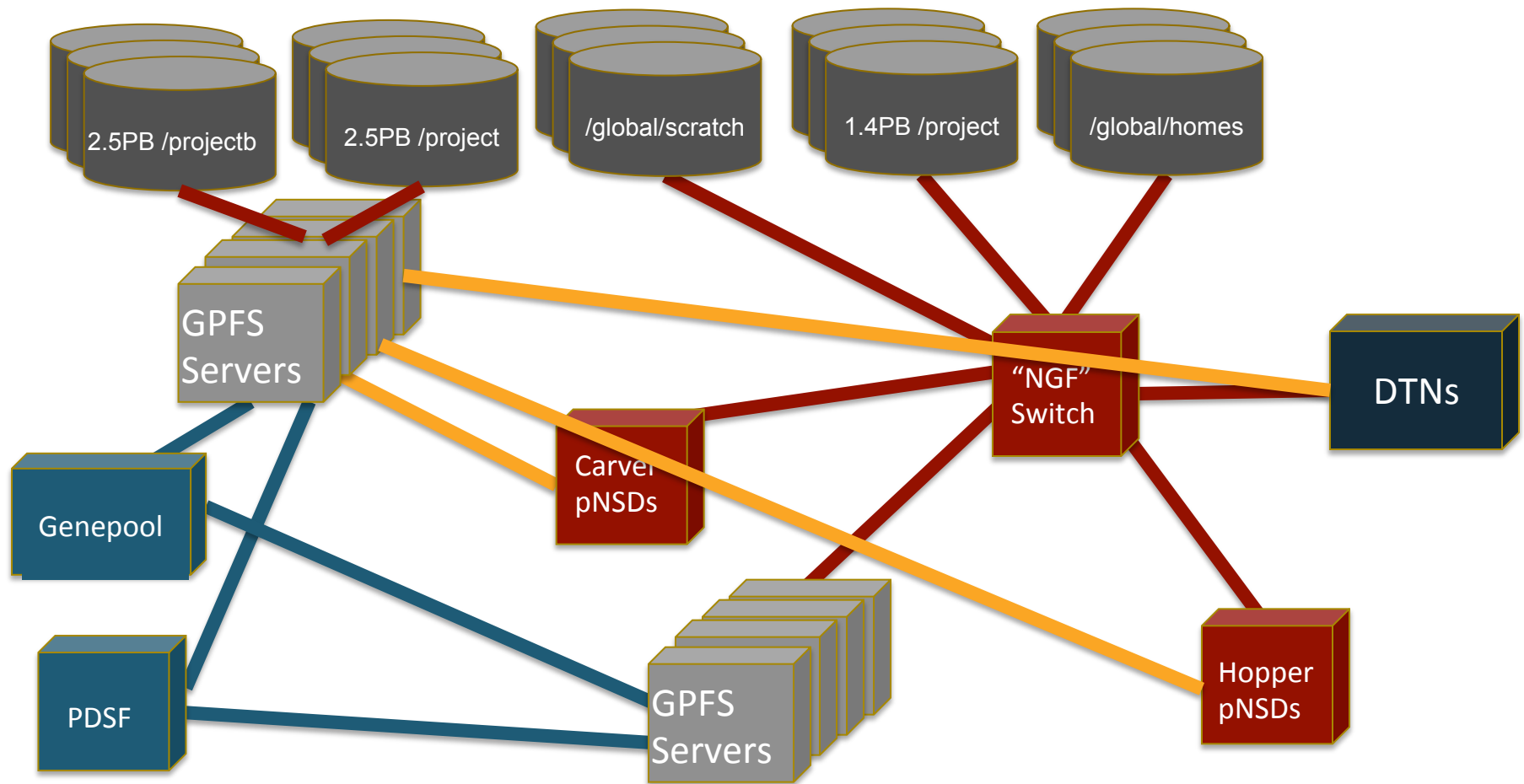


- **/project is for sharing and long-term residence of data on all NERSC computational systems.**
 - 4% monthly growth, 60% growth per year
 - Not purged, quota enforced (4TB default per project), projects under 5TB backed up daily
 - Serves 337 projects over FC8, QDR/FDR IB, and 10Gb ethernet
 - 3.8 PB total capacity
 - ~10TB average daily IO
- **/global/homes provides a common login environment for users across systems.**
 - Not purged but archived, quota enforced (40GB per user), backed up daily
 - Serves 4500 users, 400 per day over 10Gb Ethernet
 - 250TB total capacity
 - 100's of GBs average daily IO
- **/global/common provides a common installed software environment across systems.**
 - 5TB total capacity
 - Provides software packages common across platforms
- **/global/scratch provides high bandwidth and capacity data across systems.**
 - Purged, quota enforced (20TB per user), not backed up
 - Serves 4500 users over FC8 primarily, 10Gb ethernet alternatively
 - 15GB/sec and 1PB total capacity
 - Increasing to 80GB/sec and 4PB total by Jun 2013

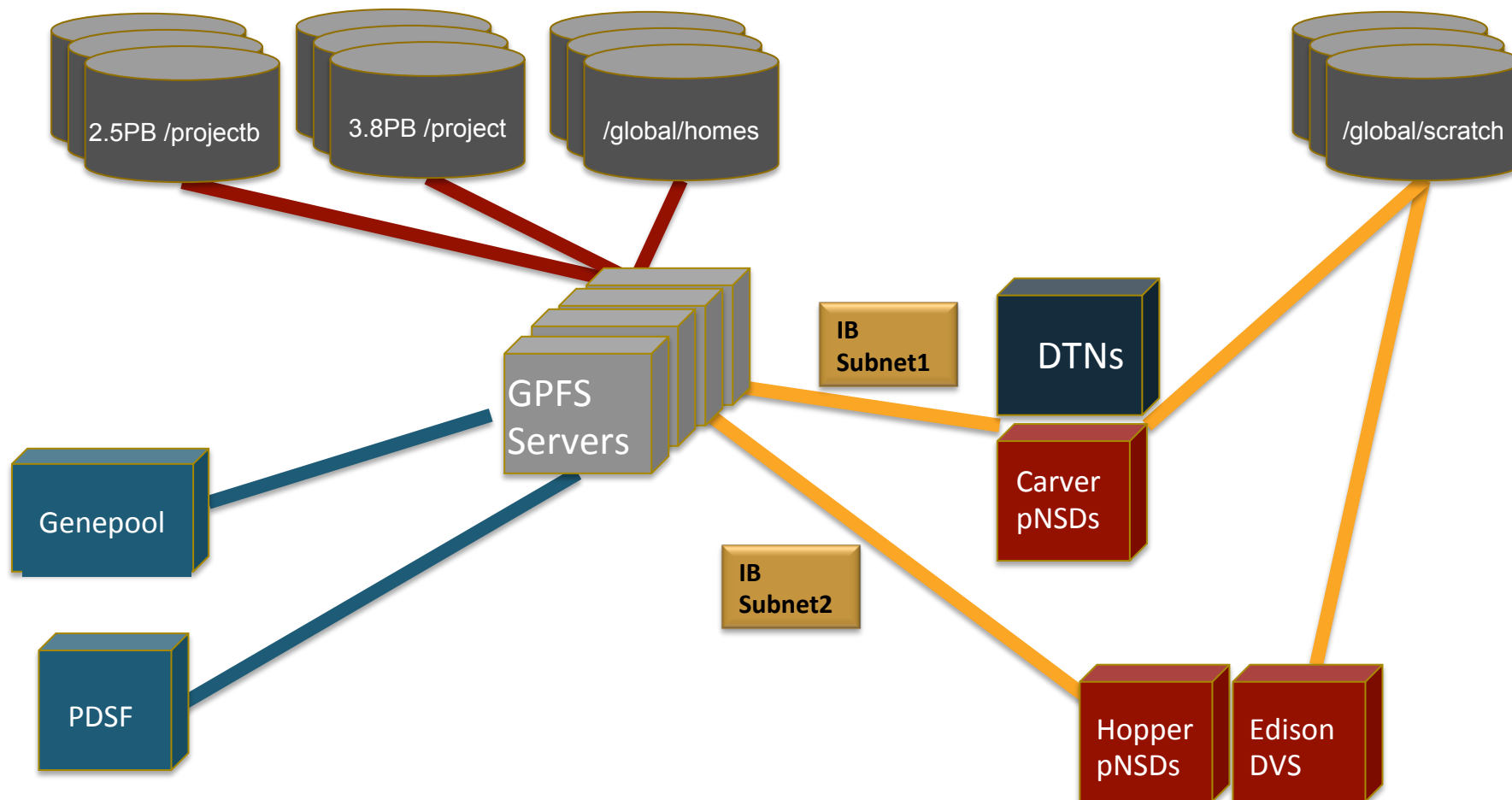
GPFS Storage Network 2011



GPFS Storage Network 2012



GPFS Storage Network 2013



U.S. DEPARTMENT OF
ENERGY

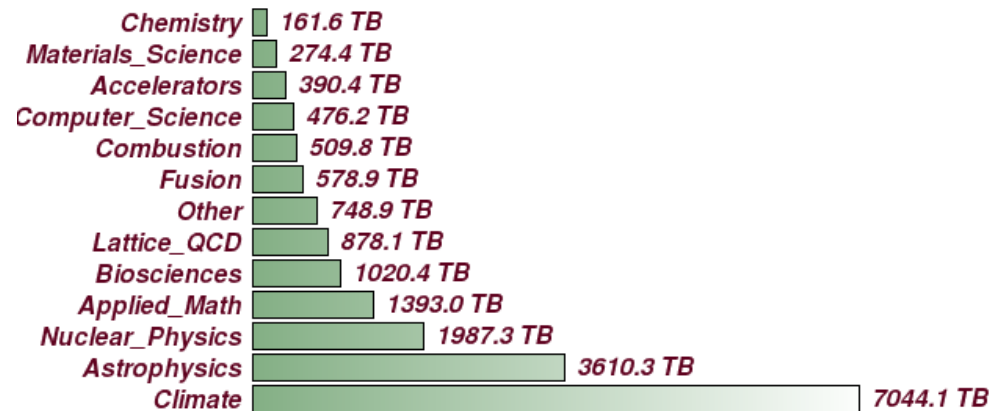
Office of
Science



Statistics involving data



Storage Utilized by Discipline (2012/12)



- Lately, the HPSS archive grows at about 1PB of data each month
- Import more data to the Center than export
- Support 337 of our 500 projects with at least 4TB allocations on /project
- Incremental backups of multi-PB GPFS file systems daily (~10TB per day) and successfully transitioned backups to support users (26 user data restore operations in 2012)
- Big users of GlobusOnline (JGI, GPFS, and HPSS endpoints)

Accomplishments 2012



- **Expanded HPSS bandwidth and capacity at least doubling both**
 - Production introduction of TS3500 Library with TS1140 drives (4TB tapes) and more T10KC drives (5TB tapes) enabling us to meet exponential growth needs
 - Deployed 3 new disk arrays and new HPSS p750 movers
- **Renewed GPFS contract**
 - Supporting our file system exponential growth demands through 2019
- **Expanded /project to enable Data Intensive Pilot awards**
 - Capacity more than doubled and we were able to deliver that capacity to science projects with high data demands rapidly
- **Support of JGI**
 - Oracle databases continue to serve JGI production
 - Deployed new /projectb file system to alleviate /house issues
- **Work to improve GPFS availability**
 - Replace problematic hardware
 - Lots of software upgrades (firmware, drivers, OS, GPFS server/clients)
 - New deployment strategy from GPFS server to clients (networking, server consolidation & specialization, direct attached storage)
 - Disk vendor engineering working on design/firmware improvements (~8 new firmware releases from 2 vendors)
 - GPFS features/bugs (14 PMRs in last 12 months, 1 DCR accepted)

Goals 2012-2013



- **/projectb expansion for JGI (~1PB)**
- **/global/scratch replacement**
 - 80GB/s with 4PB capacity using DDN SFA12KE (embedded software on controllers)
- **Continue work toward single architecture of GPFS (need IBM GPFS support/features)**
 - Consolidation of servers
 - Move storage network from Ethernet + multiple IB to single IB storage network
- **NIM integration of storage services**
 - /project new directories, renames, archiving, quota changes, backup notifications
- **Initiate sunsetting of 9840D and T10KB tape drives and media**
 - Hardware, planning
 - 24x7 work from Operations and SSG to migrate data
- **HPSS software upgrade to v7.3 and implementation of new features**
 - Small file improvements
 - File aggregation for migration
 - Small file creation rate improved (~2x)
 - Checksumming for clients (HSI)
 - IBM tape library into full production
 - Improved SCSI control library



National Energy Research Scientific Computing Center