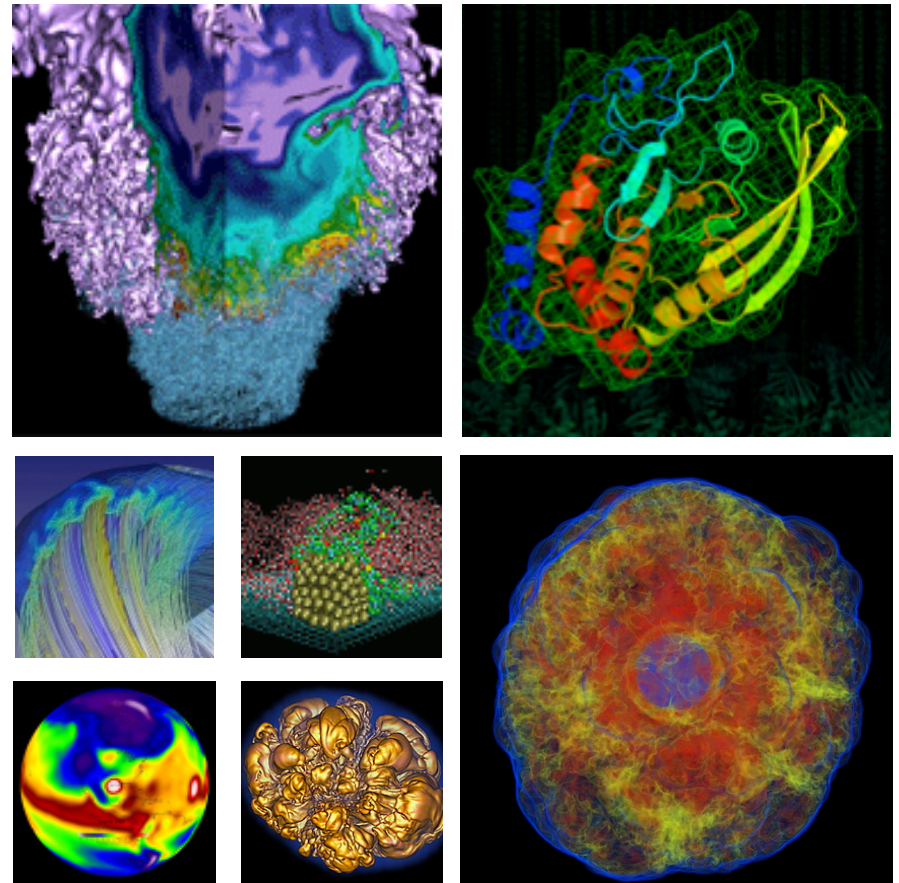


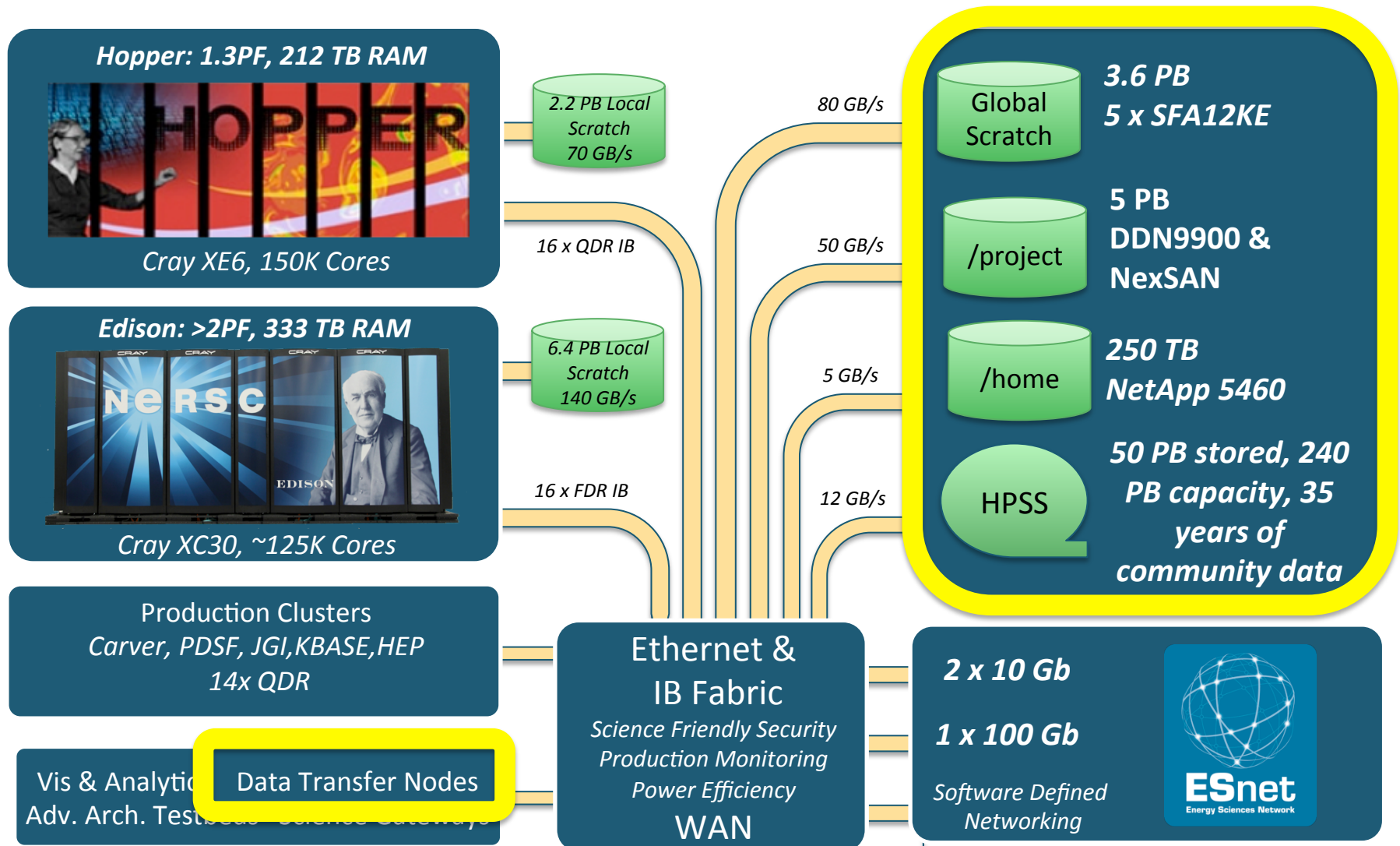
Storage Systems: 2014 and beyond



Jason Hick
Storage Systems Group

NERSC User Group Meeting
February 6, 2014

The compute and storage systems 2013



What each storage system/service is best at

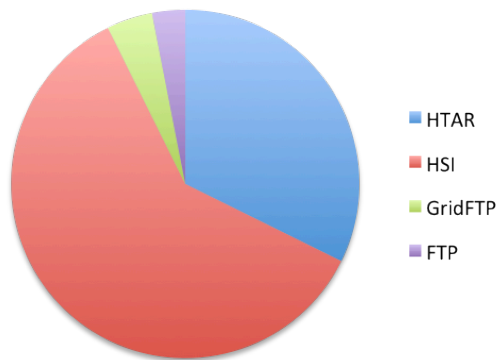


- **Global scratch**
 - Available across all NERSC platforms (avoids extra data movement)
 - High bandwidth large file I/O
 - Large individual allocation of 20TB
- **Project**
 - PI/user-managed storage, not purged, 5TB and under backed-up
 - Sharing with collaborators or across NERSC platforms
 - Different flavors of allocations/directories
 - Default project directories – 1TB for every repository in NERSC
 - Designer project directories – by special request up to 40TB
 - Sponsored storage – by special arrangement, over 40TB
- **Home**
 - Small file I/O, lower latency
 - Not for job output, or data transfer source/sink!!
 - Common login environment, source code
- **Archive**
 - Long-term retention of data
 - Stability and resiliency of system

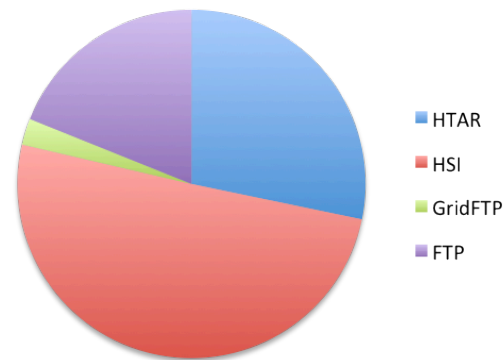
Some numbers for usage of storage

- **38% of the 930 project directories are designer project directories**
 - Designer project directories are user requested project directories, normally between 1TB and 40TB in size
- **571 new project directories created with 1TB allocation so that every NERSC repo has a project directory**
 - Default project directories are automatic for each repository, 1TB in size
- **The HPSS Archive system is growing by about 1.2PB per month, and processed over 50M transfers in 2013**
- **HPSS transfers in 2013:**

Number of Users



Amount of Data



Some more numbers



- **We backup about 500TB of data from NGF file systems every month**
 - We backup project directories 5TB and less in size, and each user's global/home directory
 - We restore a user's data about once every other week
- **Reads represent 30% of the I/O to HPSS, and users read most (>90%) of their data within 90 days of archiving**

Archive what's new?



- **Improving Globus Online NERSC#HPSS endpoint**
 - Administration features enabling us deal with transfers/ users while keeping endpoint in production
 - Better testing to validate compatibility of new Globus Online features
 - Provider contract with GO for development of new features
- **Migrating data to new technology**
 - Rewriting data to new tape media/technology

Archive what's coming?

- **New larger, better performing disk cache**
 - Goal is to retain 30 days of data on disk
 - Working to improve transfer bandwidth from NERSC client systems
- **Preparation for move to CRT**
 - Complete rewriting data to new tape media/technology
 - Replace some old equipment
 - ~3-week outage to move equipment

NGF what's new?



- **/project capacity upgrade ~1PB**
 - Requests for over 2PB of allocations in /project Nov-Dec 2013, we awarded ~1PB
- **Data pilot awards**
 - Supported data pilot (generally large allocations) with 1PB of capacity
- **/global/scratch upgraded**
 - New hardware at 3.6PB and 80GB/sec

NGF what's coming?



- **Sponsored storage option for /project allocations above 40TB**
- **Mounting local scratch file systems across center**
 - Enables us to redeploy global scratch hardware in /project
- **SSDs for global home file system metadata**
 - Expect biggest performance gain will be for nightly scans for backups
 - Less contention on metadata for user operations
- **Upgrade of DTNs to more capable systems**
 - Currently capable of 1GB/sec transfers to HPSS, and 3GB/s transfers to NGF
 - Planning to double those
- **Preparation for move to CRT**
 - Make redundant file systems, full backups, lots of new hardware
 - Performance cut in half for January 2015 – April 2015 during replication/move, with benefit of availability throughout the move



Thank you.