

NERSC Systems Reliability

NERSC - PDSI

Bill Kramer

Akbar Mokhtarani

HEC FSIO Workshop

Aug. 7, 2007



NERSC Tasks for 2007-2008

- Gather supercomputer, networking, and I/O and file storage usage data, including job length, size, processor usage and other usage profile data for analysis.
- Gather supercomputer, high performance networking, and I/O and file storage system reliability data including machine and environment configuration information, and system logs and failure data.
- Evaluate different approaches to proactively assess reliability issues.
- Evaluate the suitability of the existing archive and backup and storage usage practice on NERSC platforms and the NERSC Global File-system and characterize the archive/backup extension needed to support a facility-wide file system.

Production Systems at NERSC

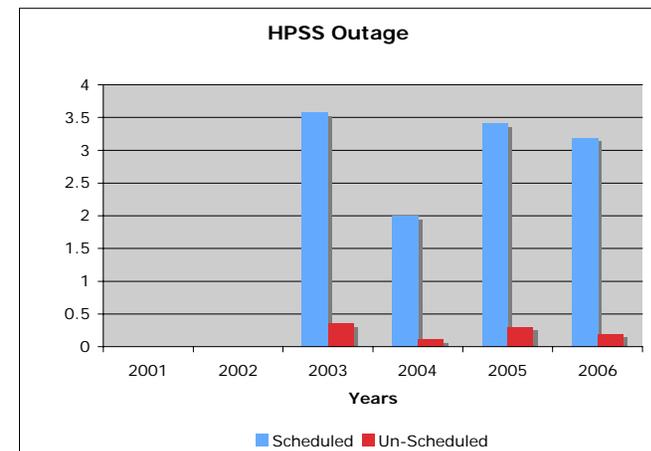
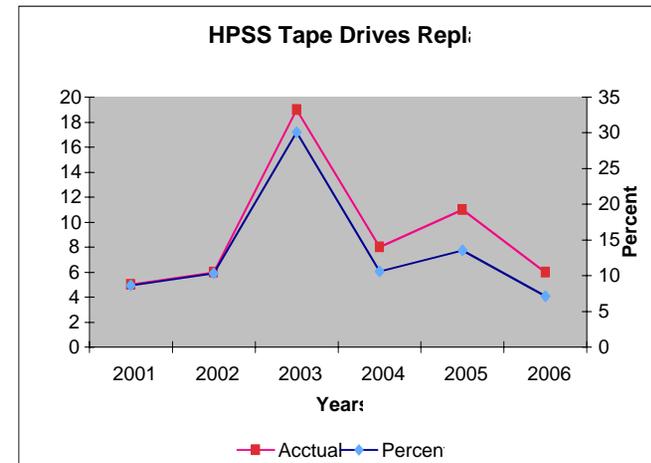
- **HPSS:** 2 High Performance Storage Systems
- **Seaborg:** IBM SP RS/6000, AIX, 416 nodes (380 compute)
- **Bassi:** IBM p575 POWER 5, AIX, 122 nodes
- **DaVinci:** SGI Altrix 350 (SGI PropPack 4 64-bit Linux)
- **Jacquard:** Opteron Cluster, Linux, 356 nodes
- **PDSF:** Networked distributed computing, Linux
- **NERSC Global File-system:** Shared file-system based on IBM's GPFS

Data Sets

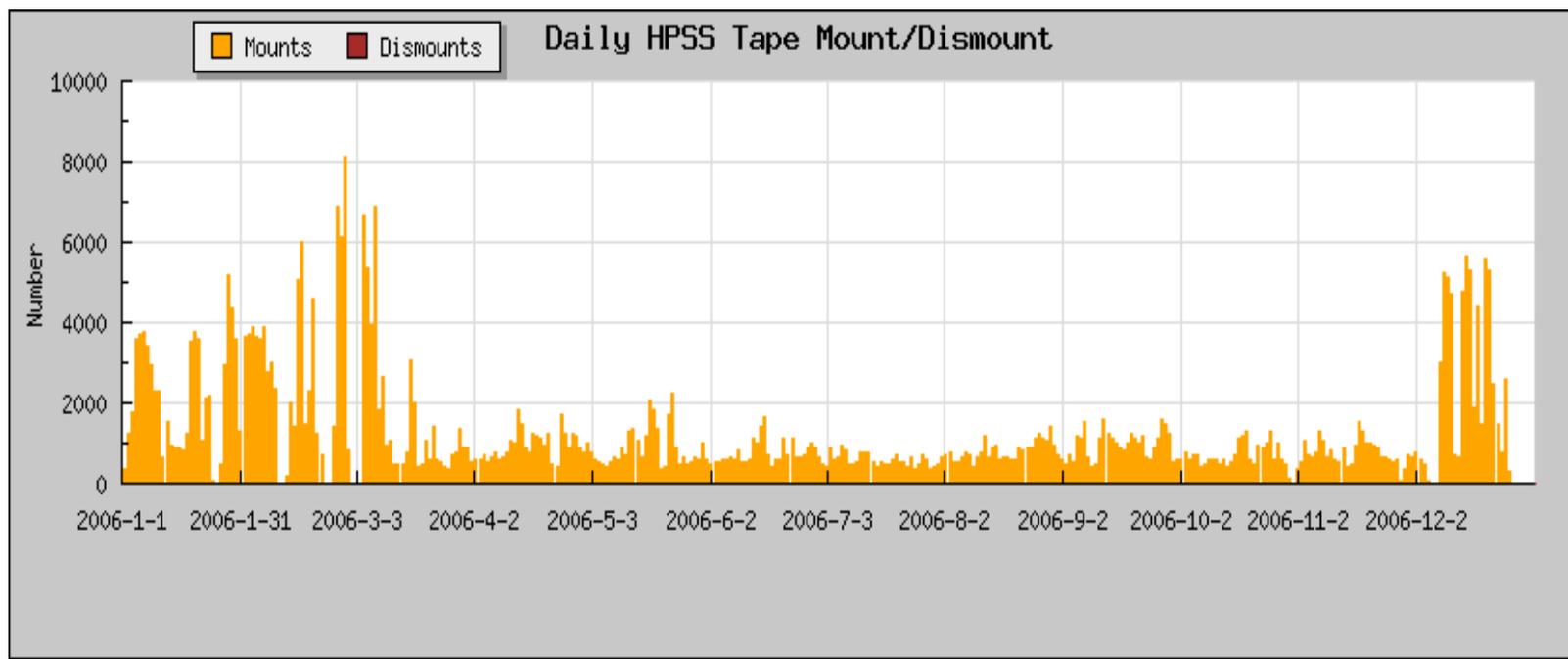
- Data were extracted from the following sources:
 - Remedy, a problem tracking database used by operation
 - Paper records kept by the operations staff
 - Vendor's repair records
 - Operating system error reports
- Coverage is from 2001 - 2006, or whatever was available
- Data collected into a MySQL database
- Web site allows interactive queries, charting, and exporting of data to CSV formatted files. <http://pdsi.nerisc.gov>

HPSS Data

- Two HPSS systems available at NERSC
- Eight tape silos with 100 tape drives attached
- 650 disk drives providing 100 Terabytes of buffer cache
- Theoretical capacity: 22 Petabytes.
- Theoretical maximum throughput: 3.2 GB/sec.
- ~1500 tape mounts/dismounts per day

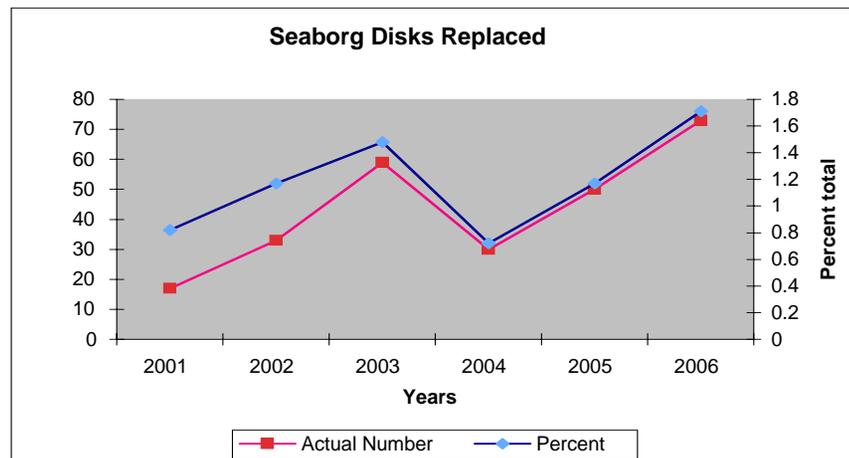
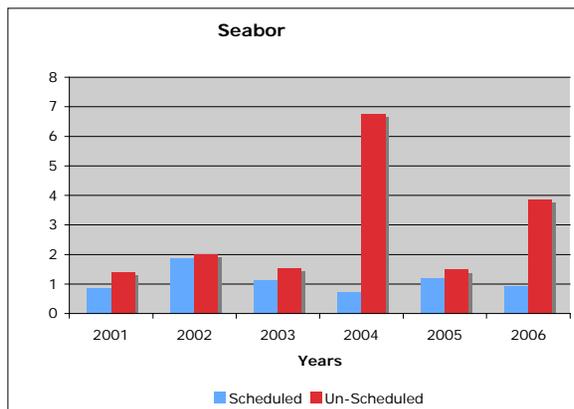
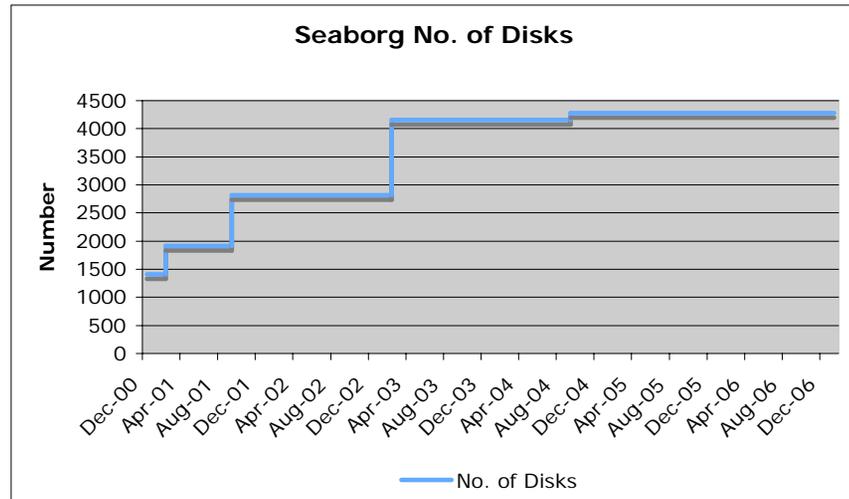


HPSS Data

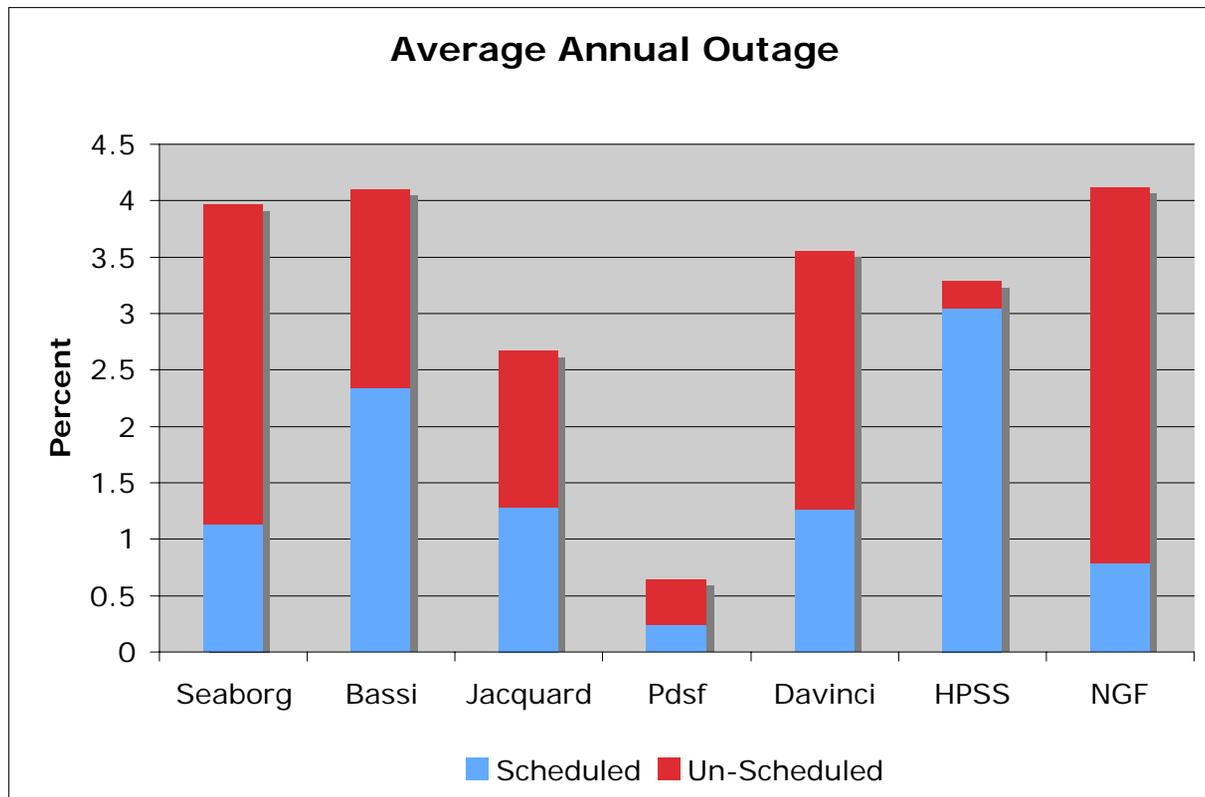


Seaborg Data

- IBM SP RS/6000, AIX 5.2
- 416 nodes; 380 compute nodes
- 16 CPUs per node
- 16-64 GB shared memory
- Total disk space: 44 TeraBytes

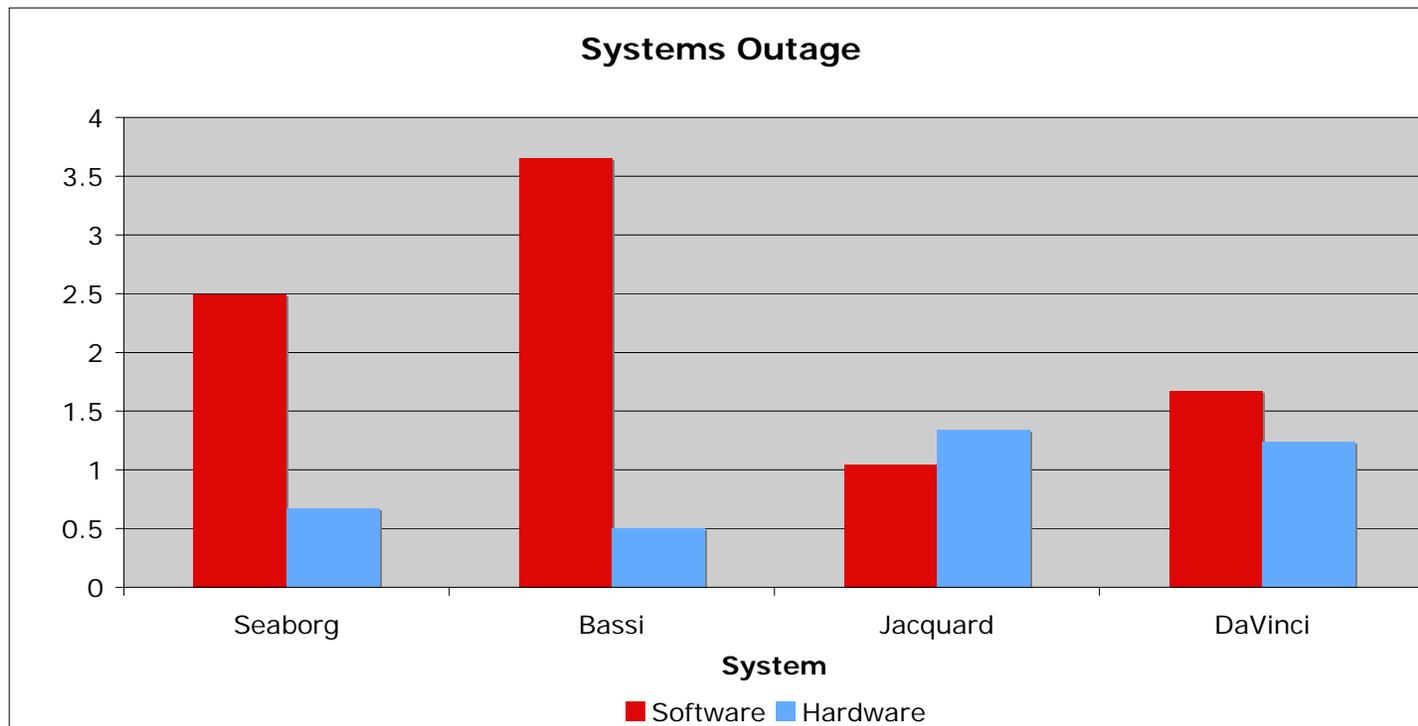


Average Annual Outage



Data since: Seaborg(2001), Bassi(Dec. 2005), Jacquard(July 2005), DaVinci(Sept. 2005), HPSS(2003), NGF(Oct. 2005), PDSF(2001)

HW vs SW Outage



Summary and Future Work

- Lots of data available but not complete
- Some initial records required interpretation
- Will continue to improve data set to make it easier for others to extract the data.
- Will try to calculate failure rates (MTBF, ...)
- Need better tracking and monitoring systems
- Right Now Web (RNW) and NAGIOS are being deployed
- We have started producing trace data and profiling of scientific applications at NERSC