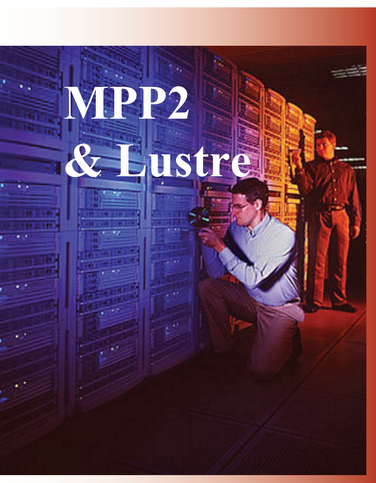
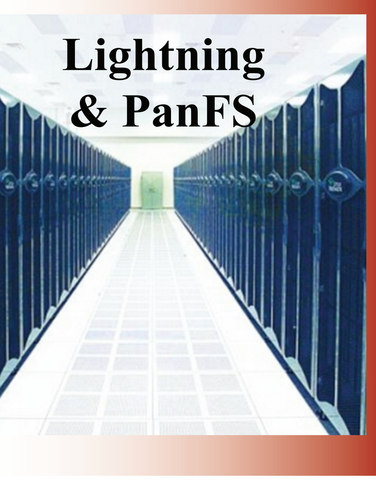
PDSI Data Releases and Repository

Office of Science
U.S. DEPARTMENT OF ENERGY

SciDAC
Scientific Discovery through Advanced Computing





petascale data storage institute
www.pdsi-scidac.org/

pdsi
www.pdsi-scidac.org

MEMBER ORGANIZATIONS

- Los Alamos National Laboratory – institute.lanl.gov/pdsi/
- Oak Ridge National Laboratory – www.csm.ornl.gov/
- National Energy Research Scientific Computing Center pdsi.nersc.gov/
- Pacific Northwest National Laboratory – www.pnl.gov/
- Parallel Data Lab, Carnegie Mellon University – www.pdl.cmu.edu/
- Sandia National Laboratories – www.sandia.gov/
- Center for Information Technology Integration, U. of Michigan www.citi.umich.edu/projects/pdsi/
- University of California at Santa Cruz – www.pdsi.ucsc.edu/

Community Contributions

THE COMPUTER FAILURE DATA REPOSITORY

- The Computer Failure Data Repository (CFDR) initiated at CMU in 2006
- Motivated by the fact that hardly any failure data from real, large-scale production systems is available to researchers
- Goal: to collect and make available failure data from a large variety of sites
  - Better understanding of the characteristics of failures in the real world
- Become reality when Los Alamos National Laboratory (LANL) released a large set of failure data collected at LANL's HPC systems.
- Now maintained by USENIX at cfdr.usenix.org/

CFDR

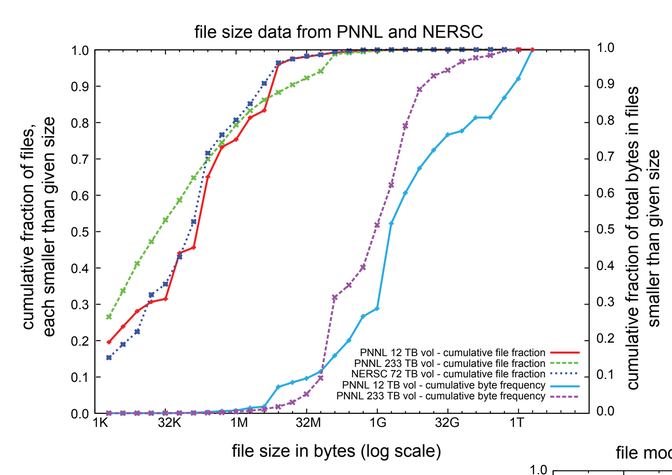
NAME	SYSTEM TYPE	SYSTEM SIZE	TIME PERIOD	TYPE OF DATA
Los Alamos National Laboratory	22 HPC clusters	5000 nodes	9 years	Any node outage
Pittsburgh Supercomputing Center	1 HPC cluster	765 nodes 3,400 disks	5 years	Hardware/ disk drive replacements
1 Internet service, Various HPC sites	3 storage, many HPC clusters	>10,000 nodes >100,000 disks	1 mth - 5 yrs	
Pacific Northwest National Laboratory	MPP2 system HPC cluster	980 nodes	4 yrs	Hardware failures
NERSC	HPC cluster	A number of production systems	5 years	I/O specific failures
COM 1	Internet services cluster	Multiple distributed sites	1 mth	Hardware failures
COM 2	Internet services cluster	Multiple distributed sites	20 mths	Warranty service log of hardware failures
COM 3	Internet services cluster	Large external storage system	1 yr	Aggregate quarterly stats of disk failures

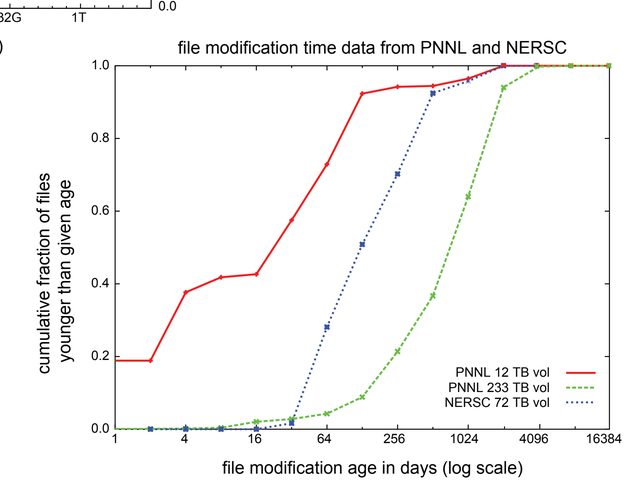
SUPERCOMPUTING FILE SYSTEMS STATISTICS DATABASE

- To facilitate worldwide data collection of static file tree attributes
- Aggregate collected data into a large, public, shared database
- We offer a small Perl tool to walk a file tree and record only aggregate statistics, not file contents or names
- Generates periodic checkpoint files to allow collection kill-and-restart
- Upload text file outputs to shared database
- Partners developing more specialized collection tools with same upload data format
- Available at www.pdsi-scidac.org/fsstats
- The following graphs show early data collections from large file systems from NERSC and PNNL
  - For example:
    - Approximately 50% of files are smaller than 64 KB but over 50% of space is in files bigger than 1 GB
    - The 90% smallest files contain only 10% of the space
  - Ages vary more in this data:
    - In one case 20% of files are 1 day old and 80% are <2 months old
    - In another case the median file age is 3 years and only 5% of files are younger than 2 months

SOFTWARE – BENCHMARKS, TRACE AND STATS COLLECTIONS

- LANL High Performance Computing (HPC-5) – institutes.lanl.gov/data/
  - Open Source Software
  - Operational Data to Support and Enable Computer Science Research
  - Trace Data to Support and Enable Computer Science Research
  - High End Computing Interagency Working Group sponsored File Systems and I/O (FSIO) research
- NERSC sources – pdsi.nersc.gov/
  - Data for Storage System Failure and Network Outages
  - Global File System failure and statistics
  - Systems Diskfailure
  - Application I/O Benchmarking and Characterization
  - Workload Profiles – workload characterization data of scientific apps
- PNNL PDSI SciDAC Debian Distribution Repository – www.pdsi-scidac.org/repository/debian/index.htm
  - PNNL has also released statistics and failure data in CFDR and stats data in the PDSI FSSTATS database
- PDL, Carnegie Mellon Univ. – FSSTATS code release – www.pdsi-scidac.org/fsstats/index.html
- Univ. of Michigan – Parallel NFS (pNFS) Linux implementation – www.citi.umich.edu/projects/asci/pnfs/linux/
- UCSC Ceph petabyte-scale object-based storage – www.pdsi.ucsc.edu/proj/ceph.html and ceph.sourceforge.net
- Sandia is supporting LLNL's IOR software, for benchmarking parallel file systems using POSIX, MPIIO, or HDF5 interfaces – sourceforge.net/projects/ior-sio





PDSI Event at FAST '08 Today!

PETASCALE DATA STORAGE BOF

- www.usenix.org/events/fast08/bofs.html
- "Gold" Room at 7-9 p.m. on Wednesday, Feb. 27
- No registration required to attend – please join us
- PDSI speakers will discuss released code and data sets
- John Shalf of LBNL will present "A User Perspective on HPC I/O Requirements"

RECURRING COMMUNITY EVENTS

- Petascale Data Storage BOF at FAST (TODAY!) – February (www.usenix.org/events/fast08/bofs.html)
- HECIWG Sponsored HEC FSIO Workshop – August (institute.lanl.gov/hec-fsio/workshops/2008/)
- Petascale Data Storage Workshop at Supercomputing – November
- Talks, papers and posters from PDSW '07 at www.pdsi-scidac.org/SC07/index.html

Los Alamos National Laboratory EST. 1943 · Oak Ridge National Laboratory · NERSC National Energy Research Scientific Computing Center · Pacific Northwest National Laboratory Operated by Battelle for the U.S. Department of Energy · Parallel Data Laboratory Carnegie Mellon University · Sandia National Laboratories · center for information technology integration University of Michigan · The University of California