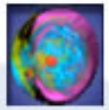
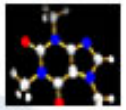
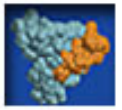
SciDAC

Scientific Discovery through Advanced Computing



File System Statistics

Shobhit Dayal, CMU

Garth Gibson, CMU

Marc Unangst, Panasas

Project under PDL and PDSI

Designing better File Systems

- Need better understanding of file systems files
- Many types of understanding
  - Content specific understanding
  - Dynamic traffic usage
  - Static file and directory attribute sizing.
- Static file tree attributes
  - Least invasive
  - An important source of understanding

FSstats

- A tool for collecting summaries of static file attributes.
- Perl code - works on UNIX based file systems, licensed under GPL
- Walks file trees gathers anonymous attributes
- Stop and restart from checkpoint
  - Useful for long captures
- CSV and ASCII output

Our Goals

- Encourage contributions from many sources
  - Distribute tool widely
  - Tool is available at: - http://www.pdsi-scidac.org/fsstats
- Upload statistics to our website
- Fill a form that gives us more information about the source of data
- Repeated runs over time on the same tree
  - longitudinal stats

Website view- uploaded results

Uploaded File	Organization	Date	Data Size	System Name	Form Questions	Formatted Result	fssta Vers
PhilRoth_fsstats.csv	ORNL	Oct102007	305GB	Panasas	Form	Result	1.4
EvanFelix_mpp2dtemp.csv	PNNL	Oct102007	12TB	ext3	Form	Result	1.4
EvanFelix_nwfs.csv	PNNL	Oct102007	233TB	ext3	Form	Result	1.4
EvanFelix_mpp2home.csv	PNNL	Oct102007	4TB	advfs	Form	Result	1.4

Snapshot of a report

```
"skipped special files",1
"skipped duplicate hardlinks",9271
"skipped snapshot dirs",0
"total capacity used",304922368.00,KB
"total user data",212806201.21,KB
"percent overhead",0.302087

file size:
count=2160683 average=97.991729
min=0 max=9715350
```

File Stats

[0- 2 KB):	1014550	(46.96%)	(46.96%	cumulative)
[2- 4 KB):	357372	(16.54%)	(63.49%	cumulative)
[4- 8 KB):	283323	(13.11%)	(76.61%	cumulative)
[8- 16 KB):	200451	(9.28%)	(85.88%	cumulative)
[16- 32 KB):	124820	(5.78%)	(91.66%	cumulative)
[32- 64 KB):	72150	(3.34%)	(95.00%	cumulative)
[64- 128 KB):	43722	(2.02%)	(97.02%	cumulative)
[128- 256 KB):	34768	(1.61%)	(98.63%	cumulative)
[256- 512 KB):	12236	(0.57%)	(99.20%	cumulative)
[512- 1024 KB):	7314	(0.34%)	(99.54%	cumulative)
[1024- 2048 KB):	3679	(0.17%)	(99.71%	cumulative)
[2048- 4096 KB):	2726	(0.13%)	(99.83%	cumulative)
[4096- 8192 KB):	1472	(0.07%)	(99.90%	cumulative)
[8192- 16384 KB):	1056	(0.05%)	(99.95%	cumulative)
[16384- 32768 KB):	467	(0.02%)	(99.97%	cumulative)
[32768- 65536 KB):	322	(0.01%)	(99.99%	cumulative)
[65536- 131072 KB):	95	(0.00%)	(99.99%	cumulative)
[131072- 262144 KB):	67	(0.00%)	(100.00%	cumulative)
[262144- 524288 KB):	40	(0.00%)	(100.00%	cumulative)
[524288- 1048576 KB):	30	(0.00%)	(100.00%	cumulative)
[1048576- 2097152 KB):	15	(0.00%)	(100.00%	cumulative)
[2097152- 4194304 KB):	4	(0.00%)	(100.00%	cumulative)

Space Stats

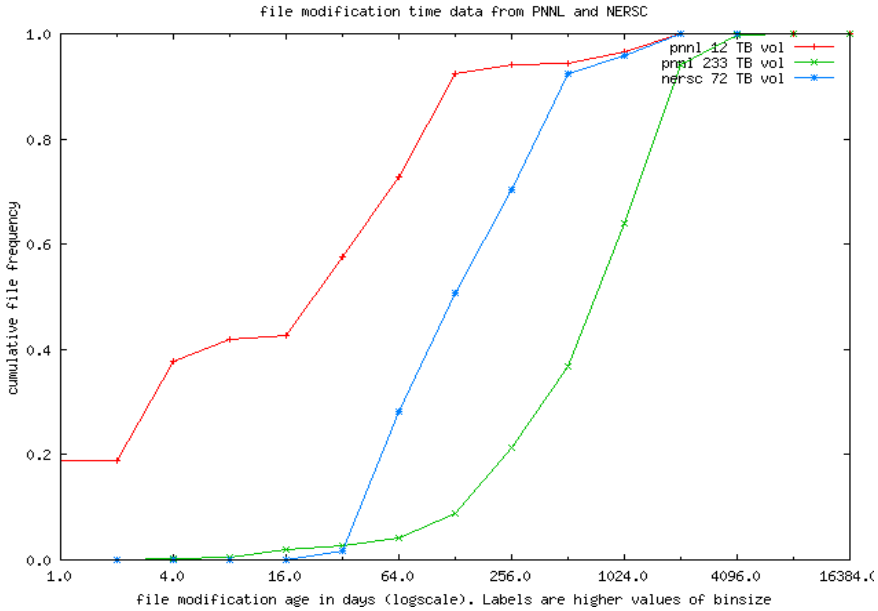
650790.78 KB	(0.31%)	(0.31%	cumulative)	
1028250.04 KB	(0.49%)	(0.79%	cumulative)	
1598852.48 KB	(0.76%)	(1.55%	cumulative)	
2263306.97 KB	(1.07%)	(2.62%	cumulative)	
2781341.97 KB	(1.31%)	(3.93%	cumulative)	
3236567.59 KB	(1.53%)	(5.46%	cumulative)	
3853971.42 KB	(1.82%)	(7.28%	cumulative)	
6197173.07 KB	(2.93%)	(10.21%	cumulative)	
4416223.82 KB	(2.09%)	(12.29%	cumulative)	
5223425.88 KB	(2.47%)	(14.76%	cumulative)	
5262715.81 KB	(2.49%)	(17.24%	cumulative)	
7698948.13 KB	(3.64%)	(20.88%	cumulative)	
8777611.05 KB	(4.15%)	(25.03%	cumulative)	
11752915.19 KB	(5.55%)	(30.58%	cumulative)	
10972048.80 KB	(5.18%)	(35.76%	cumulative)	
14274067.17 KB	(6.74%)	(42.50%	cumulative)	
8626297.36 KB	(4.07%)	(46.58%	cumulative)	
12182783.76 KB	(5.75%)	(52.33%	cumulative)	
13638314.66 KB	(6.44%)	(58.77%	cumulative)	
22571606.96 KB	(10.66%)	(69.43%	cumulative)	
22388984.26 KB	(10.57%)	(80.01%	cumulative)	
10965937.76 KB	(5.18%)	(85.19%	cumulative)	

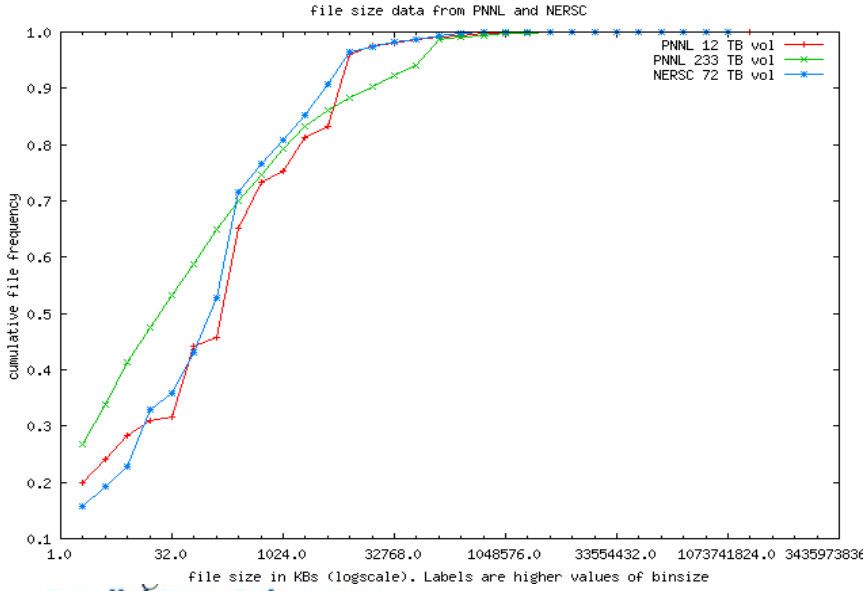
What we have so far

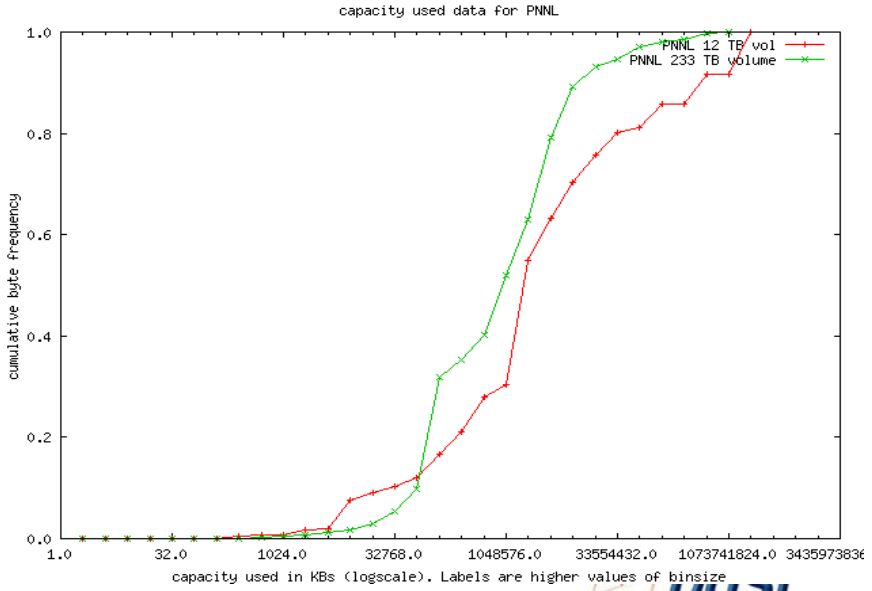
- Data from Large supercomputing sites (LANL, ORNL, PNNL, PSC, NERSC etc)
  - Some using their own tools in conjunction with fsstats
  - Integration with the histogram module is easy
- Our own big servers (PDL server)
- Workstation data form LANL
  - User data in a scientific computing environment, Engineers, scientists, Sys Admins.
  - Some preliminary results from over 200 volumes
  - Expecting data from over 2000 volumes
  - With node, volumeId to enable repeated stats

Early Data ….

- 50% files < 64KB, but
  50% space >1 GB files
- Ages vary more
  - 20% 1 day old vs.
    < 5% 64 days old







Going forward….

- Analysis of contributed data
  - Focus on scientific computing data
  - Compare results with previously published results for workstation data
- Creating a database with statistics
  - Will be made open for querying
  - View statistics summaries and graphs
  - Can be used to view interesting statistics

Looking for your help

- Encourage more folks to contribute !
  - We'd like to get as much data as we can get our hands on
  - Suggestions to improve form questions
    - What are the right question to ask submitters about data they are giving us
    - Capture as much context information as possible without making it too tedious or data too vaugue