



OpenSFS Benchmarking Work Group (BWG)

Scalable parallel file systems I/O characterization 2015 survey

The OpenSFS (<http://www.opensfs.org/>) Benchmarking Work Group (http://wiki.opensfs.org/Benchmarking_Working_Group) aims to provide an open source file system benchmark suite and guidance on file system benchmarking best practices to satisfy the benchmarking requirements of the scalable parallel file system users and facilities. Towards this end, BWG aims to characterize the file system workloads deployed at various high-performance and parallel computing facilities and institutions. Using these gathered characteristics, the workgroup will identify and build the required I/O benchmarks to emulate these workloads and provide required documentation about the benchmark suite. An initial version of this effort has been published in 2014: ([OpenSFS BWG 2014 file system characterization survey](#)).

Your assistance is needed once more to provide a better characterization of the file system workload at your facility by completing the survey below.

Collected data will not be used for marketing or sales purposes. Gathered data and results will be shared with OpenSFS members and public, after the data is anonymized, if requested.

Please contact bwg-survey-data@opensfs.org for questions or comments or completed survey data.

Target release date for the community: November 2014

Deadline to gather the answers from the community: February 2015

On behalf of the OpenSFS BWG community,

Devesh Tiwari

BWG File System Characterization
Task Lead

Sarp Oral and Andrew Uselton

BWG Co-chairs

Center/Facility Information:

Name:

Affiliation or funding agency/division:

Primary File System Contact:

Telephone number:

Email address:

1. How many different HPC systems/platforms operated concurrently?

- Number of users that use compute or storage resources
- Are there users who exclusively use only one of the resources?

2. Major site activities/functions for your facility

- Bio-science research and industry
- National research laboratory
- National security
- Academic (not for profit)
- Energy
- Media/entertainment
- Other (please specify)

3. Are there any restrictions on releasing your data to the public after identifying information (name of the center and name of any applications) have been removed?

4. File System Overview

- i. Name and version of the deployed scalable parallel file system
- ii. Connectivity diagram of file system, servers and clients, if there is technical document that describes your file system characteristics, please share with us (an URL link or PDF).
- iii. Number and type of scalable storage units (SSU), if details of your SSU are not easily available online, please describe briefly.
- iv. What is the composition of their storage back-end (disks, SSDs, tapes)? How did the center come up with that composition?
- v. Number and type of I/O controllers (possibly within each SSU)
- vi. Total number of disks
- vii. Disk types and interfaces (FC, SAS, SATA, NVRAM)
- viii. Redundancy configuration of disk groups in file system
- ix. Number and type of metadata servers
- x. How often and in what fashion (incrementally vs all at once) storage hardware is upgraded? Similarly, how often client and server side software are upgraded (not patched)?
- xi. How much data is transferred between different kinds of file systems (for example, scratch and archival). Both estimates and actual measurements are acceptable.
- xii. How much data is transferred to-and-from wide area network to the file system? Both estimates and actual measurements are acceptable.
- xiii. What tools are used to monitor the health of file system: performance and reliability? How often are these tools run and what is the estimated overhead?
- xiv. Typically, what is the distribution of number of files within a directory?
- xv. How much file system activity changes over time (day, night, weekends)?

Both estimates and actual measurements are acceptable.

- xvi. How has the amount of I/O activity changed over past several years? What is the data growth rate at your center?
- xvii. Can the data growth rate be tied to a particular set of applications?
- xviii. What fraction of all applications are the most I/O intensive? As per our definition, if an application performs more than 10% of the I/O performed by all applications, then it can be classified as an I/O intensive application.
- xix. What benchmarks and tools are being used for performance debugging?
- xx. What are some of the file system related policies (w.r.t. user quota, purge policy, maintenance window)?
- xxi. How many failure events are observed due to the parallel file system or the hardware?
- xxii. Rank the failures in the order they are most frequent: file system bugs, disk failures, I/O controller failures, disk enclosure failure, I/O module failure, Fan failure, power supply failure?
- xxiii. Rank the failures in the order they are expected to impact the applications most: file system bugs, disk failures, I/O controller failures, disk enclosure failure, I/O module failure, Fan failure, power supply failure?
- xxiv. For example, if there are multiple redundant I/O controllers then I/O controller failure may not impact the application as much as less frequent event such as power supply failure.
- xxv. What are the perceived causes of storage system performance problems, if any? For example, slow metadata server, random I/O workload, unpredictable disk performance, slow disks, slow controllers.
- xxvi. Is it possible to share the failure log of the storage system?
- xxvii. What are some of the most common mistakes or problems found in the user application codes?

- xxviii. How can users change their code to improve the overall efficiency of the storage system (of course, other than by avoiding random writes)?
- xxix. What are some of the most common operational lessons or mistakes?
- xxx. Does the center collect the user I/O patterns (e.g., sequential, random, stripe count etc.)?
- xxxi. Please share a couple of other interesting file system detail or operational experience/anecdote

5. What other question(s) do you think should be in subsequent surveys?