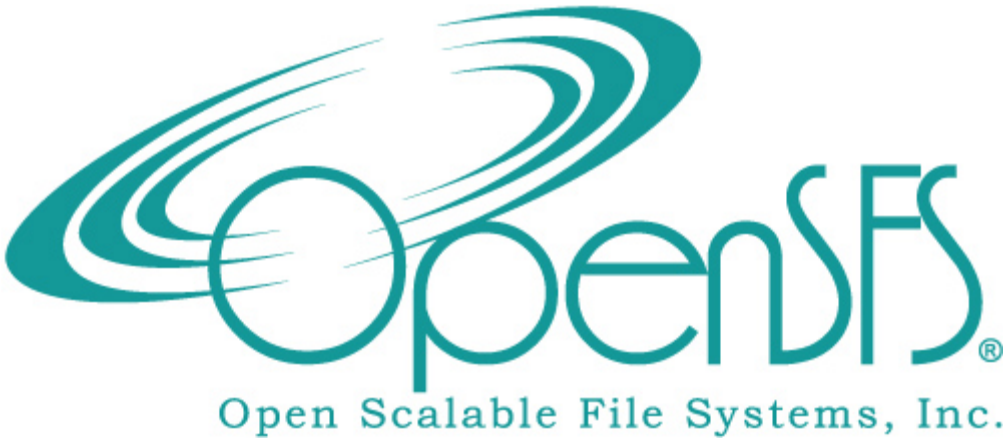

OpenSFS

Parallel File System Tools, Addressing Lustre Technical Debt, and Parallel File System Incubators

Technical Specification

Version 1.0

OpenSFS Solicitation W4570

February 20, 2013

Background

The work described herein is in support of three broad areas of open source parallel file system advancement.

1. User-space tools for parallel file systems
2. Addressing Lustre Technical Debt
3. Open source parallel file system incubators

This Technical Specification describes the potential areas of work in these three broad categories that will advance the open source parallel file system ecosystem. OpenSFS is seeking proposals to develop or productize user-space tools for parallel file systems, documentation of the Lustre source code and overall architecture, as well as incubator projects to further the advancement of other open source parallel file systems.

1.0 User-space tools for parallel file systems

While parallel file system technologies continue to mature, the ecosystem of tools required to productively use these technologies has remained significantly less mature. OpenSFS seeks proposals to develop new and/or harden existing tools that improve productive use of parallel file systems. Areas of interest are listed below but Offeror's may propose work on tools in other areas as well.

Data movement, transformation, and validation tools

Many sites require tools for data movement and data archiving that leverage the performance capabilities of parallel file systems. Examples of these tools include parallel copy, parallel tar utilities, and tools to validate data integrity. While these tools have been developed by multiple organizations in the past none have matured to a level of hardened and supported software. OpenSFS seeks proposals that will develop and/or mature these tools into hardened and supported products.

Storage monitoring and management

Open source parallel file systems such as Lustre, continue to be difficult to manage and monitor in even small-scale environments. While some open source technologies have improved monitoring and management few if any of these tools are production hardened and fully supported. OpenSFS seeks proposals that will develop and/or mature these tools into hardened and supported products.

2.0 Addressing Lustre Technical Debt

The Lustre codebase has, over its decade-plus of development, accumulated a great deal of "technical debt". Technical debt has many causes, best summarized at http://en.wikipedia.org/wiki/Technical_debt. The Lustre code base has suffered from each of these causes of technical debt at one time or another. This technical debt directly impacts the

sustainability of the Lustre code, making Lustre feature development difficult, increasing the incidents of new bugs, raising the difficulty to repair new bugs, and contributing to an imposing learning curve for new Lustre software developers.

In OpenSFS contracts for new Lustre features, we require that the contractor address some level of technical debt that is directly touched by the new code. This tends to improve the code over a long period of time, but there are large areas of code that are not directly connected with new features that continue to be overlooked. Our intent with this request is to focus directly on technical debt, without the added complexity of also modifying the code to add new functionality.

OpenSFS recognizes that it is very important to address technical debt in order for Lustre to continue to meet the requirements of our participants and the broader Lustre community. OpenSFS therefore seeks proposals that will address one or more areas of technical debt in the Lustre codebase. General categories of technical debt repair include:

- Code refactoring
- Code documentation
- Lustre internals and architecture documentation
- Writing Lustre tests

Our preference is for proposals addressing the first three categories, but proposals for additions to the Lustre test suite will also be considered.

3.0 Open source parallel file system incubators

OpenSFS seeks to broaden the set of solutions available to the community of users of open source parallel file systems for High Performance and Data Intensive Computing. To this end, OpenSFS seeks proposals to further develop and mature existing open source parallel file systems that are based on technologies other than Lustre. The intent of these incubator projects is to continue to mature open source parallel file systems towards meeting the requirements of the OpenSFS community[1].

A variety of activities are appropriate for incubator projects such as improving efficiency, scalability, and resilience of the software to better meet the requirements of the community. Offerors may choose to propose an initial scoping and design activity to be funded by OpenSFS or may choose to propose activities to include development, testing, and delivery of the solution for consideration.

[1]See the first appendix of the OpenSFS TWG Requirements report: http://wiki.opensfs.org/images/f/f9/OpenSFSTWGRequirements2012.pdf