Lustre Distributed File System Graduate Research Project Proposal

Sponsor

The Department of Electrical, Computer, Software & Systems Engineering at Embry-Riddle Aeronautical University

Released December 4, 2014

	Name	Signature	Date
Student	Mr. Justin Albano		
GRP Advisor	Dr. Remzi Seker		
MSE Coordinator / Department Chair	Dr. Remzi Seker		

Proposal

Objective

To study and implement portions of the solution architecture and solution high level design proposed in [1] and [2], respectively, for the Layout Enhancement (LE) established in the Technical Proposal by High Performance Data Division of Intel for OpenSFS Contract SFS-DEV-003 as signed on Friday 23rd August, 2013. Through the research and application of these solutions, the intricacies of the design can be realized and vetted, ultimately leading to improvements in the design that may not have been foreseen except upon implementation. Likewise, this research may eventually lead to an alternative solution design that incorporates the improvements found during application of the LE solution architecture and LE high level design.

Background

The Lustre File System is a high-performance computing (HPC), POSIX-compliant distributed file system that is used on over 60% of the TOP100 sites, as recorded by Alexa [3]. While there are many open source distributed file systems available for use today, the architecture and design of the Lustre file system is particularly suited for extreme-throughput environments and platforms, and is capable of not only storing petabytes of data on its file system, but is also capable of providing terabytes per second of aggregate input/output (I/O) bandwidth across the file system. Compounding the interest of many of the TOP100 sites with the capabilities provided, Lustre is quickly becoming the *de facto* standard for HPC file systems and, in turn, has gained increased support from companies such as Intel, providing a financial and technical means for improving the file system [3].

Problem Statement

While Lustre has proven to be a capable file system, there are many improvements that can be made to increase both the efficiency and simplicity of the system. The Lustre distributed file system is based on the concept of dividing a file into objects deposited on various storage nodes and in a network, and storing an accompanying manifest data structure (referred to as a layout in the Lustre nomenclature), that contains the location of the file objects within the network. Currently, there are four main areas of improvement that Lustre is seeking to develop: (1) file replication, (2) Redundant Array of Independent/Inexpensive Disks (RAID) support, (3) compaction for widely stripped files, and (4) handling large layouts.

In order to develop solutions to these outstanding problems, the High Performance Data Division of Intel signed a contract (SFS-DEV-003) with Open Scalable File Systems, Inc. (OpenSFS) to create a solution architecture, solution high level design, and implementation assessment for enhancements to the layouts used in Lustre (see [4] for the scope of these enhancements). At the time of writing, the solution architecture ([1]) and solution high level design ([2]) have been released, but an implementation assessment has yet to be created. This implementation assessment presents an opportunity for improvement to both the proposed architecture and high level design, and provides an avenue for a possible alternative design, leveraging the knowledge gained from the implementation of the proposed design as the basis for additions and modifications to the solution architecture and design.

References

[1] "Layout Enhancement Solution Architecture." *OpenSFS: The Lustre File System Community*. Open Scalable File Systems, Inc., 20 Dec. 2013. Web. 30 Nov. 2014.

[2] Hammond, John. "Layout Enhancement High Level Design." Ed. Richard Henwood. *OpenSFS: The Lustre File System Community*. Open Scalable File Systems, Inc., 7 Feb. 2014. Web. 30 Nov. 2014.

[3] "Lustre® File System." OpenSFS: The Lustre File System Community. Open Scalable File Systems, Inc., n.d. Web. 04 Dec. 2014.

[4] "Layout Enhancement Scope Statement." *OpenSFS: The Lustre File System Community*. Open Scalable File Systems, Inc., 10 Oct. 2013. Web. 30 Nov. 2014.