

Research Directions in Parallel I/O for Clusters

Walter B. Ligon III
Clemson University

Abstract

Parallel I/O remains a critical problem for cluster computing. A significant number of important applications need high performance parallel I/O and most cluster systems provide enough hardware to deliver the required performance. System software for achieving the desired goals remains in the research and development stage. A number of parallel file systems have achieved remarkable goals in one or more of several key areas related to parallel I/O, but there is still great reluctance to commit to any file system currently available. This is mostly due to the fact that these file systems do not address enough issues at once in a package that is robust enough for widespread use.

Critical goals in the development of an operation parallel file system for clusters include:

- *High performance with scalability*
- *Reliability/fault tolerance*
- *Flexible and efficient integration with parallel codes*
- *Portability*

These issues give rise to problems with interfaces and semantics, in addition to specific technical problems such as distributed locking, caching, and redundancy. The next generation of parallel file systems must look beyond traditional interfaces, semantics, and implementation methods in order to achieve the desired goals. Of equal importance is the issue of knowing to what extent a given file system achieves these goals. Given that no file system is likely to address ALL of these goals equally well, it is important to be able to measure a given file system's utility in these areas through benchmarking or other evaluation methods.

In this talk we will explore a few of these issues. Included are specific examples and a case study of the PVFS V2 team's approach to these issues. This talk is intended to point out areas where we feel parallel file systems must grow beyond the confines of traditional file systems or traditional implementation methods in hopes of generating discussion of these views. While we will present details of PVFS V2 design and implementation this talk focuses more on future development directions that might apply to any parallel file system project.