

Analysis and Optimization of an OrangeFS Block Driver for QEMU/KVM

Timothy Scott, Walt Ligon, Clemson University
{tscott2, walt} @clemson.edu

1 Introduction

Virtualization in the data center is a powerful tool for increasing system utilization and efficiency. Expanding the infrastructure, increasing efficiency, and creating better tooling can greatly improve ease and cost of development for big systems. OpenStack hopes to become a major part of the developing supercomputing ecosystem through its datacenter and virtual system management technology. QEMU is an integral piece of the OpenStack framework and offers powerful virtualization technology for processor and virtual machine creation and management. OrangeFS is the new branch of the Parallel Virtual File System (PVFS). OrangeFS works to bring the traditional scientific computing paradigms employed in PVFS and extend them to a wider application space. The goal of this project is to create a proof-of-concept application of OrangeFS and its user tools as a back end image storage and service platform for deploying and snapshotting QEMU images. We were able to develop a new block driver protocol as part of QEMU and provide some preliminary results showing performance gains for OrangeFS systems using the specialized block driver.

2 Background

As previously stated, OrangeFS is a new branch of PVFS that seeks to offer support for a wider application space than PVFS. OrangeFS is primarily developed at Clemson University and offers support contracts through an official agency, Omnibond. OrangeFS is the first commercial development of PVFS and will be the primary development branch in the future. A new feature in OrangeFS that enables this shift to more broad based applicability is *libofs*. This library sits on top of *stdio* and POSIX libraries to enable seamless integration with out-of-kernel OrangeFS operations. This enables more efficient data transfers and the use of higher level OrangeFS mechanisms. Clemson University is working on integrating OpenStack. This partnership and combined interest offered the perfect motivation for exploring integrating OrangeFS and OpenStack.

3 Design

The QEMU block driver software is written to the QEMU block driver API. The block driver reuses code

from the RAW/POSIX block driver. The block driver builds against *libofs* to use kernel module-bypassing procedures to improve efficiency. This allows for higher level constructs and more advanced configuration in OrangeFS client and server settings. Future designs will better leverage file system configuration and integration with virtual machine nodes to produce greater speedups. The initial design uses a naive approach that does maximize speedups for certain areas, specifically asynchronous I/O operations.

3.1 Preliminary Performance Analysis and Methods

The preliminary results from testing the block driver are encouraging. An across the board speedup of 1.08 running QEMU against an OrangeFS installation configured for scientific workloads shows the raw performance gains. This test suite used is a synthetic suite designed to test data validity for QEMU block drivers. The tests focus on very small I/O operations and do not necessarily reflect real world conditions. We expect that with additional tuning the performance results will improve further. All testing was performed on the Palmetto Cluster [1]. Specifically, these tests leveraged *newsratch*, a 218TB OrangeFS installation, and were conducted on a node featuring an Intel Xeon E5-2665, 60GB ram, and 240GB local disk space.

4 Discussion and Future Work

Generally, we are encouraged by the results. With more careful tuning and a more targeted OrangeFS installation, we believe these speedups can be dramatically improved. As OrangeFS continues to evolve to have a broader application focus, we expect that new features and *libofs* maturity will further improve the performance increases seen with the block driver. Providing a more full feature set is critical to improving adoption of OrangeFS within and without the scientific community. This case study hopefully shows that OrangeFS can provide a useful backend storage component for scientific and nonscientific communities.

References

- [1] Palmetto cluster user guide.
- [2] M. Great and U. Advisor. Dummy Paper. In *Proceedings of ISCA*, 2011.