Designing and Evaluating Hybrid Storage for High Performance Cloud Computing

Abstract

- The need for reliable and fast storage systems is increasingly critical in various fields including artificial intelligence and data analytics. This paper proposes a new architecture for large-scale data storage systems, focusing on comparing performance of software and hardware storage technologies that effectively reduce the computational latency and improve performance.
- The main contributions include: (1) the combination of Singled Magnetic Recording (SMR) for storing data and Solid State Devices (SSD) for storing metadata is a viable solution for implementing large data storage systems; and (2) the combination of Conventional Magnetic Recording (CMR) for storing data and SSD for storing metadata shows the highest performance for high performance computing.
- Our experiments are carried out in multiple settings, demonstrating that the proposed architecture successfully improves performance for sequential and random reads/writes.
- The prototypes are evaluated with a set of workloads, showing the superiority of the proposed data storage configurations. This work provides new opportunities for efficiently processing and storing data and metadata in largescaledata analysis systems.

Existing

- A critical challenge faced by data storage community exists in how effectively storing data without storing the same data again and again in different locations and storage devices.
- Hardware companies are trying to invent cost-effective solutions; and software companies are developing solutions that make Big Data Storage (BDS) easier to manage and analyze.

Proposed

- This paper proposes a new architecture for large-scale data storage systems, focusing on comparing performance of software and hardware storage technologies that effectively reduce the computational latency and improve performance.
- The main contributions include: (1) the combination of Singled Magnetic Recording (SMR) for storing data and Solid State Devices (SSD) for storing metadata is a viable solution for implementing large data storage systems;
- and (2) the combination of Conventional Magnetic Recording (CMR) for storing data and SSD for storing metadata shows the highest performance for high performance computing.

HARDWARE REQUIREMENTS

- Processor – Pentium –III
- Speed – 1.1 Ghz
- RAM – 256 MB(min)
- Hard Disk – 20 GB
- Floppy Drive – 1.44 MB
- Key Board – Standard Windows Keyboard
- Mouse – Two or Three Button Mouse
- Monitor – SVGA

SOFTWARE REQUIREMENTS

- Operating System : Windows 8
- Front End : Java /DOTNET
- Database : Mysql/HEIDISQL

Conclusion

- In this paper, we propose and evaluate cloud-based big data storage architectures working with the Ceph file sys-tem. Selecting a particular Ceph server architecture should be based on the tradeoff between speed (IOPS) and drive density (Tb/sq.).
- The results show that the combination of CMR + Metadata on SSD gives us the optimum read/write performance, but these CMR drives have very low drive density.
- This CMR + Metadata on SSD architecture can be best utilized when a user needs high response times but not high storage density.
- The combination of SMR + Metadata on SSD gives us very high drive density though this system has moderately lower sequential read speed (IOPS) compared to the CMR + Metadata on SSD server.
- SMR + Metadata on SSD architecture can be best utilized when a user needs high storage capacity for big data applications.

Reference

[1] [Online]. Available: http://www.openstack.org/

[2] A. Amer, J. Holliday, D. D. E. Long, E. L. Miller, J. F. Paris, and T. Schwarz, "Data management and layout for shingled magnetic recording," IEEE Transactions on Magnetics, vol. 47, no. 10, Oct 2011.

[3] P. S. Andrei, M. Florica, M. Alin, A. Victor, and C. M. Claudiu, "Hardware acceleration in ceph distributed file system," in IEEE International Symposium on Parallel and Distributed Computing, 2013.

[4] R. Fontana, S. Hetzler, and G. Decad, "Technology roadmap comparisons for tape, hdd and nand flash: Impli-cations for data storage applications," IEEE Transactions on Magnetics, vol. 48, no. 5, 2012.

[5] D. Gudu, M. Hardt, and A. Streit, "Evaluating the per-formance and scalability of the ceph distributed storagesystem," in IEEE International