STATEMENT OF RESEARCH INTERESTS

Hyogi Sim (hyogi@vt.edu)

The advent of NAND flash-based Solid State Drive (SSD), has introduced a new availability in the storage hierarchy. Although most SSDs provide the same interface with traditional hard disk drive, their sophisticated characteristics make it intricate to deploy them in existing systems [1][2]. Moreover, their relatively high price and limited lifespan place a vast decision space where thoughtful insight into storage systems is required, to achieve a both cost- and performance- efficient solution.

My past research was predominantly focused on building an high performance SSD and deployment of it in uni-core environments. And now my research interest is being expanded and my primary goal is to establish a efficient storage system exploiting SSD, particularly in distributed computing environments.

Current Research Agenda

Alleviating Checkpoint Bottleneck via Introducing SSD Aggregators

Periodical checkpointing is the most widely adopted solution to provide fault tolerance in High Performance Computing (HPC) clusters. Although the checkpointing becomes crucial as the HPC clusters get larger, it incurs significant extra overhead where computing processors become idle. My current research aims at alleviating the checkpoint overhead by optimally placing SSDs in a HPC cluster. Our work is dissimilar to existing approaches [3][4], suggesting to place SSDs by intuition. We are building a theoretical model, based on in-depth topology analysis of a real HPC cluster, Jaguar [5], which allows us to estimate performance impact of any given SSD placement decision accurately.

Past Research Agendas

The works listed here are the projects in which I was involved during my Master's studies.

Managing Multimedia Data for Contents Server with Hybrid Storage

The goal of this project was to build a hybrid storage system for a commercial video streaming service. We adopted hybrid storage system architecture utilizing both HDDs and SSDs. We prioritized all the contents in the server according to their popularity and cached most popular contents in SSDs. In addition to that we prototype a file system for non-volatile memory devices.

Design of a Flash File System Supporting Transaction and Record Structure

This project aimed at developing a database management system running at a set-top box, which was equipped with a raw flash storage. For backend of the database we adopted Wisconsin Storage System [6] and MiniSQL for the frontend [7]. The biggest challenge was that the performance of the database system did not fit with the system requirement, which led us to re-implement the level 0 of Wisconsin Storage System to manage the raw flash directly. And we proposed an algorithm to improve the Flash Translation Layer by introducing a small NVRAM write buffer [8]. We also presented how the buffer management layer of the database can conform to flash memory devices [9].

Future Research Agendas

I intend to continue working in the agenda of deploying SSDs in distributed systems. Especially, my perspective is that eventually SSD will become closer to a non-volatile memory device, speed of which is comparable to the dram, rather than being a fast hard disk drive. Meanwhile, data-intensive web applications, such as Facebook, are migrating their massive amount data from disk to memory for providing fast user experiences [10]. I am confident that we can build a disk IO-free and reliable distributed storage system by exploiting SSDs. More specifically, my future research topics include the followings, but not limited to.

1. Distributed Cache Management with Heterogeneous Cache Devices
2. Distributed File System for Heterogeneous Devices
3. Realizing a Durable Data Storage in Memory

References

[1] Chen, Feng, David A Koufaty, and Xiaodong Zhang. "Understanding intrinsic characteristics and system implications of flash memory based solid state drives." *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems* 15 Jun. 2009: 181-192.

[2] Agrawal, Nitin et al. "Design tradeoffs for SSD performance." *USENIX 2008 Annual Technical Conference on Annual Technical Conference* 22 Jun. 2008: 57-70.

[3] Vishwanath, Venkatram et al. "Topology-aware data movement and staging for I/O acceleration on Blue Gene/P supercomputing systems." *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis* 12 Nov. 2011: 19.

[4] Liu, Ning et al. "On the role of burst buffers in leadership-class storage systems." *Mass Storage Systems and Technologies (MSST), 2012 IEEE 28th Symposium on* 16 Apr. 2012: 1-11.

[5] Bland, Arthur S et al. "Jaguar: The worlds most powerful computer." *Memory (TB)* 300.62 (2009): 362.

[6] Chou, HT et al. "Design and implementation of the Wisconsin storage system." *Software: Practice and Experience* 15.10 (2006): 943-962.

[7] http://www.hughes.com.au/products/msql/

[8] Kang, Sooyong et al. "Performance trade-offs in using nvram write buffer for flash memory-based storage devices." *Computers, IEEE Transactions on* 58.6 (2009): 744-758.

[9] Jung, Hoyoung et al. "LRU-WSR: integration of LRU and writes sequence reordering for flash memory." *Consumer Electronics, IEEE Transactions on* 54.3 (2008): 1215-1223.

[10] Beaver, Doug et al. "Finding a needle in Haystack: Facebooks photo storage." *Proc. 9th USENIX OSDI* (2010).