White Paper

NRAM as Rotating Media Cache

NRAM Cache: A Fundamental Improvement in Hard Drive Controller Design

By Bill Gervasi

Abstract: Data caches for hard drives have always been a requirement for buffering data from the computer system host in order to compensate for the slow speed of the rotating media. These caches also provide performance enhancements such as gathering sequential read and write operations to optimize performance. These caches, however, become a weak link in data reliability when power fails. This white paper will discuss the compelling value proposition of storage class memory in hard drive controllers and discuss why NRAM® – based on Nantero's carbon nanotube technology (CNT) - is well suited to provide such value.

About the Author: Bill Gervasi is Principal Systems Architect for Nantero, and a chairman of the JEDEC international standards organization coordinating the development of memory and storage solutions for the computer industry.

The DRAM cache inside a hard disk drive controller is a key determinant of the performance of the hard drive, yet its use also restricts many aspects of hard drive controller design. Hard disk drives (HDDs) incorporate a DRAM cache for data and commands for many reasons. Keeping up with the data rate of the interface (SATA, etc.) requires that data be stored in read and write buffers. Re-ordering commands helps improve performance and power efficiency by stringing reads and writes together. A larger cache means that more data can be buffered and more commands may be reordered, resulting in a fairly direct relationship between increased cache size and increased performance—however cache size is limited by the energy required to maintain its contents and therefore the ideal cache size cannot be achieved today

This cache, typically implemented using DRAM or less typically, SRAM, may also be used for program execution and data for the HDD controller device, thereby keeping the cost of the controller down since it requires fewer on-chip resources that consume die area. Data tables, bitmaps, compression intermediate results, etc. can be stored in the DRAM if the interface speed is high enough to prevent performance bottlenecks.

Because power can fail with very little advance notice, and because data stored in the HDD cache must be preserved to avoid corruption of the rotating media, each HDD must incorporate an energy source that keeps power stable until a flush operation can commit all the data in the cache to the rotating media.

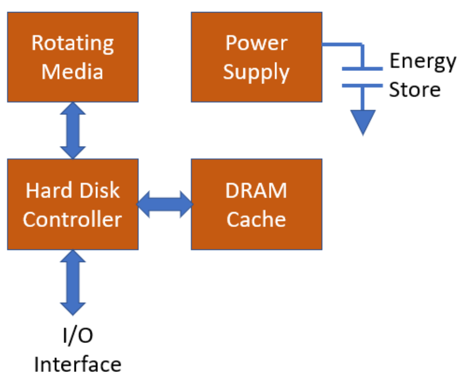

Figure 1: Block Diagram of a Typical HDD Controller

This energy source requirement is one of the most difficult aspects of HDD design. The amount of energy required to flush a cache to rotating media is quite high since the DRAM, the controller, and the rotating media all need to be powered continuously until the flush is complete. While the total amount of energy is pretty dependent on the design, it can take multiple Farads of capacitance to maintain power during a flush of 128MB of cache to rotating media.

This also explains why caches are kept relatively small in HDDs despite the performance gains that could be realized from a larger cache. If you increase cache size, you must also increase the energy store needed for flushing the cache to the rotating media.

Capacitance is a real pain for systems designers. The main options are wet dielectric supercapacitors or dry dielectric metallic types (such as tantalum). The cost and size of a supercapacitor solution can be quite attractive, around $1 and 400mm2 of board space, or 7% of available area. However, the long term reliability of a supercapacitor is not very attractive for high end solutions. The wet dielectric tends to evaporate over time, especially in the hot environments most HDDs are used in. This forces greater overprovisioning, and largely restricts supercapacitors to lower end HDD solutions.

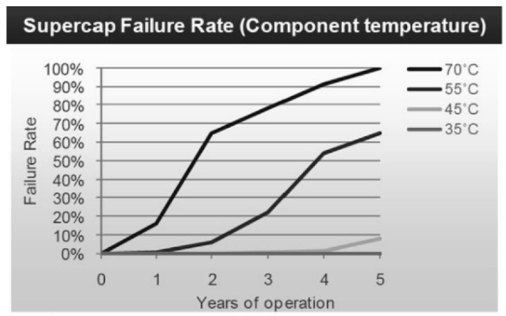

Figure 1: Long Term Reliability of Supercapacitors

The alternative, metallic capacitors, for the equivalent storage capacity as a supercapacitor is a painful tradeoff. The solution cost can jump to $4 and the footprint increase to 1000mm2, or 18% of available space, and often require cutouts in the circuit board to fit the radial bodies. This severely impacts not only the layout and routing on the PCBs, it also becomes a mechanical issue to prevent the radial components from vibrating and the solder joints from cracking.

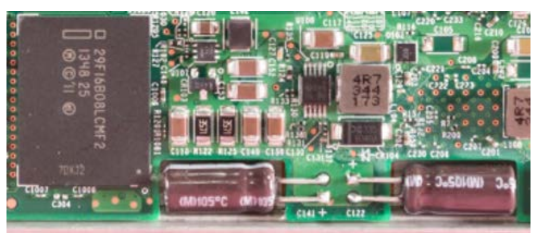

Figure 2: Tantalum Capacitors for Cache Flushing

Nantero NRAM is a non-volatile memory with DRAM performance and capacity. It connects to the HDD controller exactly like a DRAM, providing the necessary throughput to handle data and command buffering as well as controller data access. Based on resilient carbon nanotube technology, NRAM does not suffer from leakage, dielectric fractures, heat sensitivity, or other problems associated with other memory devices. When power fail is detected, the controller need only complete a data burst in process then shut down and wait for power to be restored. This eliminates the need for expensive and space consuming energy sources.

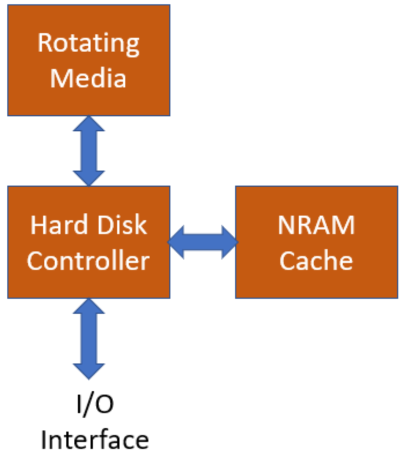

Figure 3: HDD Block Diagram Using NRAM

NRAM offers other features that can help simplify HDD controller design as well. Today's HDD controllers using DRAM for a cache must support the needs of the DRAM for going offline regularly to implement the required refresh cycles of a DRAM. If data comes in on the host interface during a refresh cycle, the HDD controller must buffer all such data until the refresh cycle completes and all buffered data can be written to the DRAM. This additional on-chip buffering increases HDD controller die size and cost.

NRAM does not require any refresh cycles, therefore 100% of the bus bandwidth is available for data traffic. This means that future HDD controllers designed to take advantage of NRAM can reduce the required internal data buffering size, further reducing cost, complexity, and increasing reliability.

The DRAM compatible speed of the NRAM is a related benefit. Controllers that want to use the cache for controller code or data may simply access the NRAM at DRAM speeds at any time, without the required times for refresh cycles where the memory is unavailable.

As long as we're tallying the advantages of NRAM over DRAM, perhaps one of the biggest advantages is the temperature insensitivity of the NRAM. Due to the capacitive nature of the storage cell in a DRAM, they tend to be very sensitive to heat, and in fact will fail above 95℃. HDDs must work extra hard to deal with this in many ways including increasing refresh rates, throttling accesses, or eventually shutting down if the temperature exceeds the DRAM specification.

NRAM's carbon nanotube memory cells do not share this sensitivity to heat, operating in ranges far beyond the other components in the system with the possible exception of the metal screws that hold the HDD together. NRAM eliminates these thermally related problems inherent in DRAM-based designs.

Though it may be obvious from the concepts described above, a little indulgence is to point out that the elimination of the energy source requirements decouples the cache size calculation from system design. Doubling, quadrupling, even growing the cache by 16X or 256X has absolutely no impact on energy requirements for power fail cycles. When power fails, the same procedure is followed: complete any single access in process then shut down and wait for the juice to come back. This opens the way for significant increases in cache size that would lead to dramatic improvements in overall HDD performance.

In summary, NRAM offers a number of advantages for HDD design, from eliminating energy source requirements to non-volatility, and from thermal insensitivity to improving data availability. The combination of these advantages can lead to an HDD with a value proposition that is simply not possible today, and one that the users of the HDDs will immediately notice and appreciate. By decoupling the cache size from concerns about an energy store, HDD designers can provide users with the performance levels expected of tomorrow's computing systems.