OBJECTIVE ANALYSIS
Semiconductor Market Research

USING FLASH AS MEMORY
A More Straightforward Way to Boost Performance

OBJECTIVE ANALYSIS WHITE PAPER

The SSD has become a standard part of computing architecture – in the enterprise only the most cautious managers avoid using SSDs in speed-sensitive applications, while in personal computing some end users purchase an all-SSD system as a costly way to boost their performance, but those on a tighter budget will often purchase systems that use a small amount of flash or SSD to boost the performance of a standard HDD-based system.

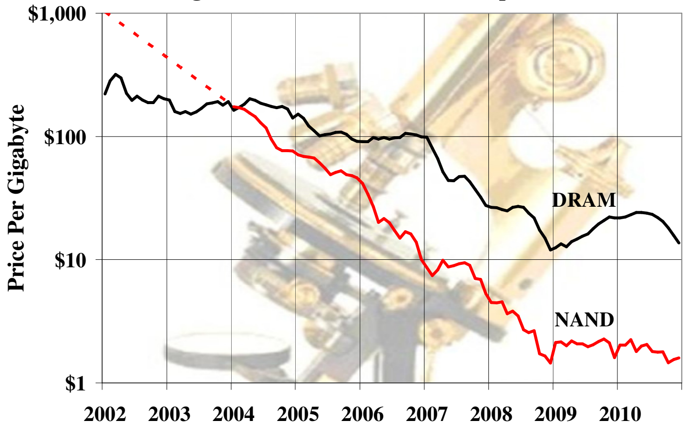
Few would question whether adding flash to a system (in the form of an SSD) is a good way to improve performance, but almost nobody considers whether that is the *best* way to add the flash to the system. The common wisdom dictates that since NAND flash is persistent and HDDs are persistent then flash should be added to the system as an HDD look-alike. In this white paper we will explore some reasonable alternatives to that perspective.

To put things into perspective, though, we need to understand why flash is just now finding its way into computing even though the technology has been available since the early 1990s.

NAND is Cheaper than DRAM

How did we get to this point? The graph in Figure 1 shows a history of DRAM and NAND flash price per gigabyte from 1998 to 2011. While NAND started out more expensive than DRAM, it gained rapid adoption in markets for which it was a cheaper alternative than the established medium, displacing photographic film, floppy disks, and eventually the compact disc. As the market grew NAND producers took several significant steps to reduce production costs, to the point that NAND's price per gigabyte fell below that of DRAM in 2004 and today is more than an order of magnitude less expensive than DRAM.

Figure 1.) NAND vs DRAM Price per GB



This throws a new light on how the computing memory/storage hierarchy should be designed. Today it makes sense for a flash layer to be added to any and all computing systems. NAND flash fits into the memory/storage hierarchy for two very good reasons:

- NAND is cheaper than DRAM but more costly than HDD
- NAND is slower than DRAM but faster than HDD

Since the memory/storage hierarchy is built on such relationships from processor cache to DRAM to HDD to tape, this same set of relationships argues in favor of adding NAND flash to any computing system.

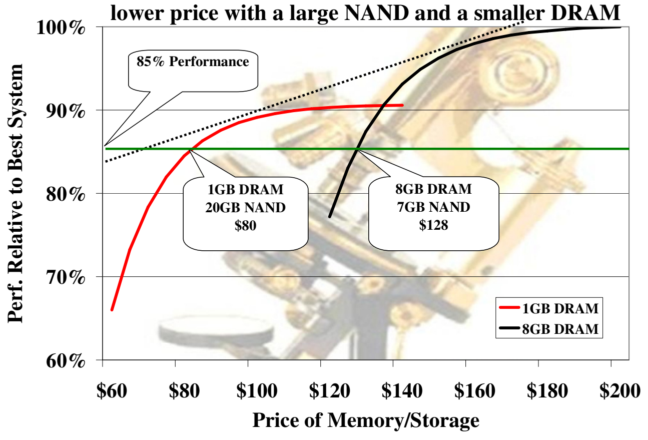
NAND vs. DRAM Performance

NAND flash is a very slow medium, and nobody would argue that it can be compared to DRAM when it comes to either write or read speed. It's not even close! The difference between a NAND write and a DRAM write is a few orders of magnitude. Even so, NAND can significantly improve system performance as a buffer between HDD and DRAM simply because of the fact that an HDD is about six orders of magnitude slower than DRAM. This phenomenal gap can be partially filled by adding a NAND layer. Not only does NAND improve the speed of a system when added to an existing DRAM+HDD complement, but it can actually help reduce the amount of DRAM required in that system, reducing capital costs and improving power and cooling, along with other critical cost factors.

In a 2010 report: ***How PC NAND Will Undermine DRAM***, Objective Analysis compared nearly 300 benchmarks performed on a standard Windows PC. These benchmarks, which included standard tests like PCmark and SysMark, were run on three levels of NAND: Zero, 4GB (within a Seagate Momentus XT Hybrid Hard Disk Drive or HHDD) and a 120GB OCZ VERTEX 2 SSD. For each of these levels the benchmarks were run using six DRAM sizes: 1, 2, 3, 4, 6, and 8GB.

In brief, we found that, after a certain minimum DRAM threshold was achieved, a greater performance boost was attained for each dollar's worth of added NAND than for the same dollar's worth of DRAM.

Figure 2.) The same performance can be reached at a lower price with a large NAND and a smaller DRAM



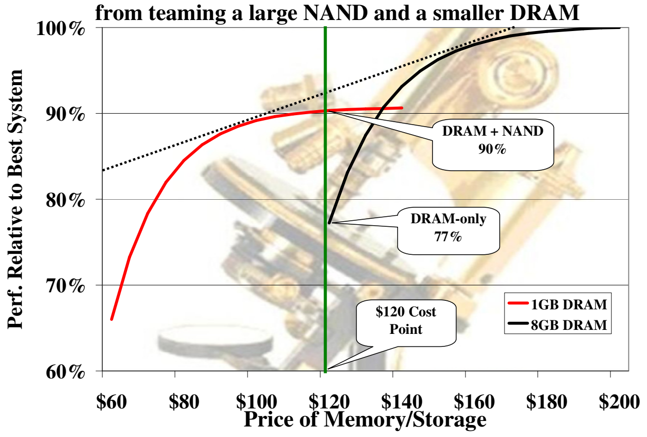
Some of the results of these benchmarks are illustrated in Figures 2 & 3. These figures show averages of all of the benchmarks, with the exception of the HDD-specific benchmarks. For the sake of clarity we only show the two extremes of the DRAM spectrum: 1GB and 8GB, with continuously-varying amounts of NAND flash. Since only three NAND sizes were used the lines interpolate between the measured levels.

The charts compare throughput against the cost of combined NAND and DRAM in the system. The cost of the DRAM+storage complement (HDD, HHDD, or SSD) runs along the x-axis, and the performance (relative to the costliest system with 8GB of DRAM and the 128GB SSD) is measured on the y-axis.

In Figure 2 a horizontal line shows the 85% performance level to ask the question: "What does it take to achieve 85% of the performance of the most expensive system?" This performance level can be achieved using 1GB of DRAM and 20GB of NAND, at a cost of $80, or with 8GB of DRAM and 7GB of NAND for a cost of $128. Naturally, the best value is obtained with the 1GB DRAM system with a generous amount of NAND flash.

Figure 3 compares the performance of the same two DRAM extremes with a vertical green line to indicate a fixed memory/storage complement cost of \$120. The system using 8GB of DRAM and no NAND achieves 77% of the performance of the highest-performing system, but the system with 1GB of DRAM and 60GB of NAND reaches a full 90% of the performance of that system at the same cost. Once again this underscores how performance can be maximized through the addition of flash instead of DRAM.

Figure 3.) Better performance for a fixed price results from teaming a large NAND and a smaller DRAM



The Best Way to Add NAND

We mentioned before that convention has dictated that NAND flash should be added to a system in the form of an HDD. This argument is based upon the rationale that NAND is persistent, so it must be used as persistent storage. Although this may be the simplest way to add NAND to existing computer architectures, HDD interfaces severely degrade the performance level of the flash.

Several "fixes" have been developed to address this problem. An early one was to add NAND through the PCIe interface, making it appear somewhat similar to a PCIe-based RAID controller card with several HDDs attached. Another has been to accelerate standard HDD interfaces to speeds far beyond anything useful to an actual HDD. SATA and SAS interfaces were pushed beyond 6Gb/s, a speed that is far faster than is useable by HDDs, and are now moving to 12Gb/s. While this is a step in the right direction, important delays still result from putting memory behind an HDD interface.

But the best performance would be attained through the addition of NAND flash directly onto the system memory bus. This was attempted by Intel with its Braidwood initiative, which was aborted by the company prior to actual shipments. Braidwood added a NAND-only bus to the PC that was hierarchically equal to the main memory bus but was based on the ONFI, rather than the current DDR2, interface. Another equally-good approach is to add NAND in the form of a DRAM module or DIMM. A few companies have tried this approach, and it appears to be gaining some momentum. Let's look at those designs.

Three of the systems use a similar approach – this is the NV-DIMM sold by Cypress Semiconductor's AGIGA, Viking Technology, and Micron Technology. These companies have taken the approach of adding NAND to a standard DRAM DIMM as a means of retaining critical data when power is lost. The DIMM contains a standard amount of DRAM plus one or two NAND flash chips capable of copying the entire contents of the DRAM. An external supercapacitor keeps the module alive after the system's power has been lost. When power is interrupted, the processor (through a special utility) dumps its register contents into the DRAM, then allows the DRAM to copy its entire contents into the NAND flash chip. Once power is restored, the data moves in the opposite direction and the system returns to the state it was in immediately prior to the power loss.

This approach increases the cost of the memory module, but gives it persistence, a trade-off that is favored in certain applications like SSD arrays, in which it is used to store critical metadata. It does not avail the system of the performance/cost benefit detailed in the prior section of this white paper.

A completely different approach is used by Diablo Technologies. Diablo's Memory Channel Storage (MCS) technology uses a NAND-only DIMM to provide a much larger memory space, achieving very high densities at a low cost. The NAND is managed by an ASIC to perform at DRAM speeds without requiring the module to contain any DRAM.

This MCS architecture also allows nonvolatile storage to be added directly to the memory channel, rather than behind an HDD interface. However, rather than performing solely as a nonvolatile DRAM or storage device, the Memory Channel Storage is meant to provide the system with a very large memory complement at a modest price. This approach allows servers to have a significantly larger memory size than the server can support using standard DRAM modules through NAND's higher density and lower power consumption.

Since NAND can't perform at DRAM speeds, the MCS architecture has critical logic in the form of an ASIC, as well as special software installed in the host system to manage the NAND as a separate layer in the memory hierarchy. NAND writes are buffered in a combination of the system DRAM and special ASIC, which coalesces writes and presents them to the NAND as fast as possible. The software manages conflicts to prevent NAND's slow writes and erase-before-write sequence from creating difficulties. The net result is NAND latencies of 5µs, rivaling those of DRAM.

Summary

Objective Analysis has consistently predicted for the past four or more years that a flash layer will become a universal feature in computers of all sorts. We have demonstrated its cost/performance advantage, and have argued the logic of connecting the technology to the system in a way that circumvents interface and software stack delays.

It is refreshing to see that manufacturers are starting to provide products that support this approach. After a frustrated effort by Intel to attach NAND to the PC through the ONFI interface, Diablo has finally found a way to connect NAND flash to a server's main memory bus that makes sense.

Jim Handy, July 2013