

Memory in the Data Center – The Series

Originally published at www.thoughtput.com

Thoughtput

Gear6 accelerates storage and delivers real time application performance with centralized storage caching. Deployed in the data center, Gear6 scalable caching appliances provide scalable and transparent acceleration for existing storage infrastructures to speed up applications, eliminate peak load disruptions, and simplify enterprise configurations.

Memory in the Data Center - The Series

Memory in the Data Center - part I - Media

I have been talking with several industry folks recently about the push for more memory in the data center. A friend asked the following question, and it led to this series of posts.

Question: The battle lines on the storage pyramid seem to be in flux. Flash on disks. Flash drives in arrays. Ever larger server memory capacity. D2D backup replacing tape. Disk is encroaching on tape. Flash is encroaching on disk. And DRAM is encroaching on both in some cases. Can you shed some light on this?

Sure. Let's break this down into four sections.

- 1) Media
- 2) Architecture and Implementation
- 3) Installation and Operation
- 4) Overall Data Center Impact

Let's start with the media and get that out of the way first. Ultimately this discussion is less about media and more about the architecture, specifically how do you architect a solution that matches the way you use your data.

There are several developments happening across different media types and price points that are particularly exciting.

Flash is one of the hottest topics because it is a truly groundbreaking shift in the price/performance spectrum. This is not to say that flash will rule the world. Rather, we are going to see a variety of ways to make use of this media.

As consumers, we have already benefited from flash in our iPods, camera memory cards, USB keys and more. The next wave for flash is in the enterprise.

As an industry, we're racing to figure out how to extract the most value from different types of flash media based on flash-level specs. This analysis is important, but we need to measure media types at an application level as well.

Application metrics. Metrics such as overall system latency, total application operations per second, and application performance per dollar of infrastructure will become more important than what one flash based drive can do.

Staying on the media discussion, the prices of DRAM continue to drop significantly with CPU and motherboard suppliers creating denser memory

configurations on industry standard platforms. For many industry applications the cost of DRAM, particularly when measured in a cost-per-IOP basis, is extremely favorable.

Then there are spinning mechanical disks. The big changes in the disk drive market are the amazing increases in capacity, but unfortunately with declining performance. But the availability of new memory-based solutions completely reshapes our ability to make use of this valuable drive-based media.

SATA and SAS drives can provide a near infinite amount of capacity for common applications. Now users can enhance the "cheap and deep" capacity with performance-based memory solutions. In and of themselves, the low-cost, high-capacity disks were only good for archiving types of applications. But with the addition of memory-based storage enhancements, the picture changes completely, and "cheap and deep" now has instant performance.

We will forever be presented with new and interesting media choices. Just a couple of months ago, Intel and ST Microelectronics announced major achievements with phase-change memory, a technology that some claim may eventually replace flash.

More recently, at the FAST 08 Conference, IBM introduced a tutorial on Storage Class Memory, and finding the right applications for the right technologies within a triangle of Speed, Write Endurance, and Cost per bit. In addition to expectations for improved flash, other candidates to optimize this triangle include Ferroelectric RAM (FeRAM), Magnetic RAM (MRAM), Resistive RAM (RRAM), Solid Electrolyte, and Phase-change RAM (PC-RAM).

Rather than assuming this media discussion will ever be over, we need to focus on the overall architecture of memory-based solutions and how one makes effective use of currently available media. More on that next.

Memory in the Data Center - part II - Architecture and Implementation

We all want higher throughput and lower latency data centers.

That translates to adding more memory. The question is how best to do that for maximum application performance at the least overall infrastructure cost?

Flash or DRAM drives? One approach is to take the new memory-based solutions and make them look like disks. That is to make flash or DRAM emulate the characteristics of a disk drive using some flavor of SCSI, FC or SATA disk-level operations to communicate with the overall system.

Examples of this implementation include both flash and DRAM based solid state drives, solid state arrays, PCI-based memory repositories, and static memory appliances.

This approach can benefit very specific data sets that are not likely to grow because the memory is being presented as a fixed LUN to the system, and resizing is difficult. It is tricky to resize a group of SSDs in an array when there are no more drive slots left, or add another PCI card to a full bus.

This memory-as-disk approach also applies to small data sets because adopting it forces you to forgo all of the benefits of low-cost, high-capacity storage. Granted you can manually move data around, but that is generally regarded as troublesome and time consuming maintenance that could be better spent on new initiatives.

One further complication with memory-as-disk approach is the difficulty in extracting the active data set from the entire file or volume. This can be near impossible to assess manually and therefore results in over-provisioning memory resources to accommodate the entire data file or LUN as opposed to the smaller percentage of active or “hot” data.

Finally, memory-as-a-disk assumes that the memory has disk level retention characteristics. So in the case of DRAM that requires a robust and highly available battery backup system, along with all of the storage management responsibilities of persistence. And in the case of flash some concerns remain about wear levels and reliability, all of which will improve and possibly be resolved over time, but still may cause near-term concern for enterprise IT departments.

An interesting alternative to deploying memory as a disk, is to use memory as a cache. This concept itself is not new as cache has been deployed at nearly every level of data center systems from L1 and L2 CPU based caches, to motherboard-attached memory, to storage subsystem caches, and even cache at the drive level.

The CPU-level and drive-level caching should and will continue for a long time to come. What I'd like to focus on now is the differences between client side caching (at the application server), storage side caching (at the subsystem), and caching in between in the network.

The end-node server and subsystem solutions have worked well in the past, particularly in a world where single servers connected to single storage systems, because in that case having memory on one side of the connection or the other was guaranteed to provide assistance. But in an increasingly networked world where many application servers are connecting to multiple storage systems,

including the deployment of clustered file systems, having the memory and caching resources in the network makes a lot more sense.

No doubt that this line of thinking will stir some debate. And I agree that there will always be room for caching at the end-node server and storage systems. But the reality is that once you are in a multi-device world, the most efficient and effective use of a memory-based resource is to apply it to the maximum number of servers and storage systems. This is analogous to the migration from direct-attached to network storage many years ago.

With a memory-based caching resource in the network any application server requesting data from any storage system can benefit from the ability to cache frequently accessed data in high-speed memory compared to slower mechanical disk. This guarantees that in cases of shifting hot-spots or hot-systems the benefit of caching applies across the board.

Alternatives would be to put the maximum amount of cache memory within each server and storage system which falsely assumes that each system is under the exact same utilization all of the time. It is simply not an option for most customers.

Another major reason to cache in the network is for scalability. Caching at an end node devices means that each increment in cache requires another end-node device. Need more cache in your subsystem? Buy another subsystem. If the cache is in the network, that subsystem might last a lot longer, and in some cases customers might actually reduce the overall number of storage systems required to deliver a set performance level.

Over and over, we've seen valued system level technologies be accessible within the network. Servers have long been there, storage more recently, and caching with high-speed, high-performance memory is another step in the network direction.

Memory in the Data Center - part III - Installation and Operation

So far we've covered a bit about media, and also architecture. Now let's talk about installation, use, and administration across these different approaches.

Perhaps it's best to step back for a second and look at the differences between the memory-as-disk approach compared to the memory-as-cache approach.

In the memory-as-disk approach administrators must decide what size memory footprint they need. This is far easier said than done because few tools exist to figure out the "active data set" within an existing LUN. So inevitably

manual intervention in this process leads to gross over provisioning of memory resources for the entire data set as compared to the active, frequently used data.

The next task is to create and provision LUNs, and then migrate data to the new LUN(s). Once in place, the LUN must be protected with backup and recovery procedures, overall storage management, and the ongoing monitoring to determine if that data set is growing beyond the LUN size. If the data set exceeds the size of the LUN, administrators must be able to rapidly provision additional space (of the same speed and performance) while not disrupting the application.

In the memory-as-caching approach, the goal is to enhance the existing disk infrastructure as a seamless complement that delivers performance without creating additional maintenance and management items.

With caching, particularly when deployed as a network resource, applications request data from any amount of storage capacity, but will “view” that infinite capacity through a caching appliance. The network-based cache only retains the actively used data, dynamically populating the cache based on application requests and letting all unused (but important) data remain on disk-based, protected persistent storage.

By enhancing a traditional or clustered file system with a network-based cache, customers get the capacity depth of a their chosen file system, and the optimized performance of a cache sized perfectly to their active data set.

As workloads shift and change, the intelligence of a network-based caching appliance automatically adjusts to the data I/O patterns. This continuously alleviates hot-spots on and across systems.

Using Technology Efficiently

An important trigger for market growth and adoption is how to use more memory, more effectively in the data center. This memory can be in any shape or form but we need to turn our attention to the system level implementation and specifically **how do we use memory to remove action items from the administrative to-do list, not add to it.**

Using memory as a persistent storage device will tend to add way more to-do items than it might alleviate. Using memory as an intelligent cache will remove action items and migrate data centers to more automated, dynamic operation.

Memory in the Data Center - part IV - Overall Data Center Impact

The final area to explore is the overall impact these new memory based approaches will have on the data center. In particular we want to keep the major themes in mind of spindle reduction, consolidation, and a need to reduce power, space, and cooling.

Memory in all of its various shapes and forms will be responsible for keeping our data centers from spiraling out of control with disk drive proliferation. Coupling high-speed memory with low-cost, high-capacity drives delivers both the performance and capacity requirements for modern data centers in an efficient footprint.

But as with all data center transformations, this will not happen overnight. And there will be stages of deployment as data center managers find the best way to enhance their current infrastructure with new memory-based solutions.

The most visible aspect of the data transformation will be the spread of silicon (in terms of processing cores and memory) to complement the storage layer. Of the three primary layers in the data center – servers, networks, and storage – the storage component is the last to rely on physical moving parts. Some believe this as “rotating rust” but I prefer to think of it as the most cost effective means of retaining high-capacity persistent data. But today’s robust application servers demand more in terms of IOPS and bandwidth than what a typical disk-based storage system can provide.

Disks are not going away, and neither is tape. But we will see memory advance to complement, and eventually displace, some disk-based systems that have relied on spindles for performance.

The big wins in adding memory to the data center will not come by trying to replace disk, but rather by effectively enhancing it. We often mistake the decline in memory prices as justification that disks will go away sooner, but I see it differently. The decline of memory prices makes the use of both technologies more applicable....performance when needed through memory, and the benefits of a high-capacity, low-cost persistent storage layer for a never-ending amount of content.

Over time (and I use these words carefully because it may be five years, or it might be ten) memory as a persistent storage media will become a more effective means in and of itself. But this will require more in terms of reconfiguring data centers than might be fully realized. New persistent storage systems that rely on memory for persistent storage not only need the basic

components in place, but also years of maturity to where all of the exceptions and error codes can be easily handled in a standards-based manner understood by multiple vendors. That day is a bit further off than the current headlines might indicate.

In the meantime, we have new ways to make use of memory that did not exist a few years ago giving us the ability to:

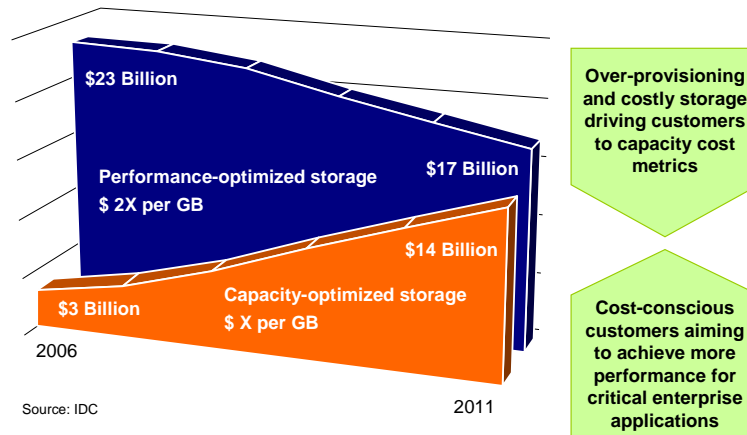
1. complement and enhance our existing storage systems
2. right-size memory-based caching solutions in terms of IOPS, bandwidth, low latency, and capacity
3. deploy memory as a network resource that is addressable by any application server accessing any storage system
4. retain existing applications without modification
5. reduce the amount of active administration and eliminate the need for manual data movement to optimize performance
6. intelligently improve our ability to rapidly access large file repositories

Individually, these capabilities might seem straightforward. Taken together, they expand and open up the architectural options for data center managers. But most importantly, they represent an immediate opportunity for IT professionals to dramatically improve current data center performance without causing unnecessary and premature overhauls of existing data center equipment.

By freeing data center managers from having to buy disks for performance, they will now be able to change the way they purchase and configure storage. The impact will reshape our map of modern data centers.

Final Conclusions

The world is moving away from configuring for performance with disk. As the following chart shows, the spending on performance-optimized storage is declining because customers are tired of paying twice the amount per Gigabyte. And there is an overwhelming trend to move more things to capacity-optimized storage.



Why isn't this happening faster?

1. Performance-oriented customers are still afraid that they will not be able to achieve their performance needs on capacity-optimized storage. But with the introduction of more memory in the data center in new and unique ways like scalable caching appliances, that changes completely. Now customers can have a performance insurance policy to move more data and more applications to low-cost, high-capacity storage. The result...performance-optimized storage spending declines faster than anticipated.
2. Customers looking to increase their capacity-optimized storage to encompass more enterprise applications often have push back about performance. But now they can achieve the same performance or greater than traditional disk-oriented performance configurations. Watch out, the capacity optimized storage systems are about to get significantly more interesting. The end result...faster adoption of single-tier, capacity-optimized systems that are complemented by sophisticated clustered caching solutions to dynamically improve performance.



© 2008 Gear6. All rights reserved. All other trademarks and copyrights are property of their respective owners. Information provided in this paper may be subject to change.

For more information, visit www.gear6.com