STORAGE SWITZERLAND

DECIDING ON THE RIGHT FLASH STORAGE FOR HIGH DENSITY VIRTUAL INFRASTRUCTURES

Storage Switzerland, LLC

George Crump, Lead Analyst

The modern data center consists of mission critical application and virtual server/desktop environments that are clustered for scalability and reliability. Each of the managers of these environments is trying to increase their density (the number of users or virtual machines supported on each physical host) in order to fully exploit the processing power available to them. If high density can be achieved, it drives down the cost to deliver IT services, reduces the data center footprint and minimizes power and cooling requirements.

But the denser and more scalable these virtualized environments become, the greater the strain it puts on the storage infrastructure. Most notably, legacy hard drive based storage systems simply can't keep up. Disk drives have been stuck at the same performance levels for over a decade. The only way to increase storage performance is to increase the number of hard disk drives. This is an unsustainable model from a hard cost perspective as well as the ongoing costs of power, cooling and floor space.

As a result, IT planners are looking to flash based storage to allow the quest for scale and density to continue. This has lead to an almost overwhelming number of flash implementation options including Hybrid Flash Systems, All-Flash Arrays and a variety of server side flash solutions. Depending on the data center each of these can have a role to play in increasing the density of the environment.

Flash, The Performance Sledgehammer

While there are performance differences between various flash systems and different implementation methods, most flash systems will provide performance beyond what a highly dense virtual infrastructure can require. The key is for that performance to be delivered consistently and reliably across the virtual cluster, in a way that is also easy to use.

Shared Flash vs. Server Flash

The shortcut to flash performance at first glance seems to be one of the various server side solutions. These solutions leverage either a solid state disk (SSD) drive in the server or a PCIe based flash card. While they do gain a latency advantage by being in the server and avoiding the network, these solutions also complicate matters in a clustered virtual environment.

The clustered environment needs shared access to a common pool of storage to enable functions like virtual machine migration, automated resource balancing and storage migration. The server side flash approach inhibits that sharing by creating data silos on each physical host.

Server side flash does work well when combined with read caching software which forces all new or changed data to the shared storage system. This means that only reads are accelerated but, considering that most virtual environments are at least 50% reads, it can be a worthwhile compliment to accelerating storage performance.

When the focus turns to the other half of the I/O puzzle, data writes, these server side solutions often fall short. Write caching acknowledges the write back to the application prior to it being safely store on permanent storage. This means that the write cache had to be highly available and in the case of server side flash, the second copy should be outside of the server. As a result many vendors are creating a series of workarounds that often involve building their own private networks between the servers. This also involves building custom drivers for the hypervisor.

Typically the software that aggregates this flash capacity is still in its infancy and does not offer the enterprise class features that traditional shared arrays do. 'Missing in action' are capabilities like snapshots and clones. While these can be provided to some extent by the hypervisor, data efficiency techniques like deduplication and compression are also not typically provided. Data efficiency is critical for bringing down the cost of implementing flash on a broad scale.

In the end, when server side flash solutions require building a private network, it is an admission that a network is indeed needed. If that's the case it seems more logical to leverage the storage network that is probably in place already that has been more thoroughly vetted than a new network established by a start-up server side flash vendor.

A shared flash array, be it All-Flash or Hybrid, allows the various virtualization functions that require cluster-like access to work as they would in hard drive based systems, albeit significantly faster. After all, all hypervisors are built assuming a shared storage system on a storage network.

Hybrid Flash vs. All-Flash

There are two types of flash arrays available that target the mainstream data center. The first is commonly called the "Hybrid" Flash Array. This is a storage system that leverages flash for performance and hard drives for capacity. The focus of these systems is still to achieve an attractive price per gigabyte with an acceptable IOPS specification to represent the performance potential.

The key problem with hybrid systems is that their performance reporting is often an arbitrary or average performance number. The problem is there is no average in these systems. Data is either accessed from fast flash or slow disk and as a result, there is an unpredictable nature to what the exact performance expectation should be.

The number one concern IT professionals express when considering a Hybrid Array is what performance will be like when there is a cache miss, a situation where the requested data is not on flash and needs to be retrieved from hard disk. While cost is always important, consistency is more so. It is the unpredictable drop in performance that will get an IT department into trouble, not a consistently bad performance number.

Another challenge with Hybrid systems is the way in which the flash tier is utilized. For example, many implement their flash tier as a read-only or write-through cache. This means that all new writes have to go to the hard disk tier first, although some hybrid system write to hard drive and flash simultaneously. In either case the acknowledgement of the received write must come from the hard disk. The result is essentially that all writes are at hard disk speed.

Consider that most virtual server environments are 40-60% writes and that most virtual desktop environments are 60-70% writes during the workday. And the need to have all those writes confirmed by a slower hard disk tier can eliminate much of the performance gain of a flash system.

This write acknowledgement problem as well as the cache miss problem is further compounded in Hybrid Arrays because of the hard disk configuration. As mentioned earlier, the only way to increase the performance of HDDs is to have more of them. The more spindles, the better the performance. The problem is that this costs money, which with their focus on dollar per GB is something that the Hybrid Systems would like to avoid.

As a result, these systems are configured with a finite number of high capacity HDDs. This means that in the case of a write acknowledgement or a cache miss not only does the I/O occur at a slower tier of storage (hard disk drives), it may occur at the very slowest type of storage for that tier: a high capacity, low performance hard disk.

Hybrid systems, like server side flash have developed workarounds for these challenges. First, most hybrid systems provide deduplication and compression on the flash storage area to get maximum capacity out of the tier. Some even extend that optimization to the hard disk tier. Second, a few of these systems can pin certain volumes to the flash tier, or specific VMs to the flash tier.

The problem with cache pinning is that it requires the ability to monitor performance on a per VM basis, something that many systems do not have, and it consumes more flash resources to accommodate the entire VM. Essentially the flash is no longer used as a temporal cache but as a tier and much of the perceived cost advantage versus an All-Flash system is lost.

A few Hybrid vendors have also worked in appropriate HA functions on the flash tier to accommodate writes to flash first. This requires that the flash tier either be RAIDed or Mirrored as well, and requires that the software have the sophistication to manage through a failure. Once again, while this does lead to more consistent performance it does so at increased cost.

All-Flash Systems like Pure Storage's FlashArray, on the other hand, take a different approach. They provide 100% consistent performance. There is no chance of a cache miss. All data is fast all the time. Reads and writes are also treated equally and no special workarounds are needed.

This leads to a system that is significantly easier to administer since volumes are created with no future tuning typically needed. In fact, at Pure Storage's user conference "Pure Evolve", the number one comment about life after implementing a Pure Storage Array was how simple storage management had become.

The theoretical downside to All-Flash systems is the relative cost per GB. Thanks to the decreasing lithography of the physical NAND flash plus the effectiveness of data efficiency technologies like compression and deduplication, the delta between the cost of an All-Flash storage system and a performance HDD system largely evaporates.

All-Flash Systems unlock the potential to create very dense virtual server and virtual desktop systems. This density leads to reduced server hardware cost since fewer hosts need to be purchased. This reality, when factored with the aforementioned simplicity, makes the total cost of ownership of an All-Flash system very attractive.

Recommendations

With all these flash options available, it can be confusing to design a storage strategy that will stand the test of time. Ideally, if it is time for a storage refresh, then an All-Flash system should be strongly considered. This allows for the creation of a high performance baseline without concern for some of the predictability and reliability issues of the other two flash categories. It should allow for the reduction in time spent managing storage and a reduction in server hardware costs.

After this baseline is established, server SSDs should be considered as boot drives so that hypervisors can rapidly boot. These drives could also be used for virtual memory paging but the All-Flash Array should also deliver excellent virtual memory performance. Finally, PCIe SSD with read-caching software should be considered for hosts where there is high read I/O. With an All-Flash Array, there is no need to cache writes locally.

Selecting the Right All-Flash Vendor

The key for All-Flash Systems is to provide consistent, high performance at a cost-effective price. This means extracting maximum data efficiency from the data being stored on the All-Flash Array which requires a combination of technologies, including deduplication, compression, snapshots and thin provisioning. Each of these can lower the amount of upfront capacity that needs to be purchased. It is important that systems are evaluated with ALL the features enabled and in-place. Ideally, these capabilities are in an "always-on" state, which simplifies management and measurement of any performance impact.

Systems like Pure Storage's All-Flash Array have all these features enabled and consistently report a 12X or greater efficiency in virtualized environments. That means that a 5TB investment in All-Flash can lead to as much as 60TBs of net capacity. This increase in effective capacity should overcome any price delta when compared to a high performance HDD system.

Conclusion

With any new technology, users are typically flooded with options as the market matures. When designing a new storage architecture it makes sense to establish a baseline first by selecting a system that addresses the broadest range of performance challenge at one time. All-Flash Arrays are a good example of a product that does just that. Server side flash solutions can back fill performance hot spots, if any, on an as needed basis.

About Storage Switzerland

Storage Switzerland is an analyst firm focused on the virtualization and storage marketplaces. For more information please visit our web site: http://www.storage-switzerland.com

Copyright © 2013 Storage Switzerland, Inc. - All rights reserved