ESSENTIAL ELEMENTS OF A SERVER SAN SOLUTION FOR VIRTUALIZED ENVIRONMENTS

By: Trevor Pott, Systems Administrator, Columnist The Register

Being one of the three fundamental components of the modern datacenter – alongside networking and compute - storage is a huge sector of IT. There are innumerable players all vying for the crown, each with their own unique twist on technology, pricing and customer service. With so many entrants, I have been frequently asked why I use Maxta's technology; the answer is as complicated as the space.

SIMPLICITY

Should storage be easy to use? On the face of it, this question would seem to have an obvious answer: yes! All other things being equal, technologies should be as easy to use as possible. Despite this, debates over storage technologies and implementations continue to rage, with many promoting difficult-to-use technologies and implementations. Some storage administrators prefer to have a vast array of knobs to twiddle; they are uncomfortable letting storage software developers make the decisions about what's best. The problem with this approach is that storage is a huge field; truly understanding the consequences of twiddling a given knob takes not only training, but a great deal of real world experience.

Storage administrators who have a deep understanding of storage are rare and thus expensive. These folks are more expensive than most companies can afford to hire and even more than most companies can afford to bring on as consultants; how then is any company supposed to purchase and design proper storage?

What is needed is a "fire and forget" implementation that works, works well and works without requiring intervention. We, as systems administrators and business owners, pay storage vendors not for an incomprehensible multi-headed hydra, but to provide us storage that "just works".

The experts design the storage software, we buy it and implement it, the software decides what the settings on all the various knobs should be for us. This is true software-defined storage and it is where true value for dollar is found.

RESILIENCY

The most important aspect of storage is resiliency. Resiliency has three primary aspects: the ability to withstand loss or damage to the primary storage medium, high availability and disaster recovery.

The ability to withstand loss or damage to the primary storage medium has been handled for decades through the use of a Redundant Array of Inexpensive Disks (RAID). Sadly, this tried and true technology is reaching the end of its usefulness. Bulk storage drives are increasing rapidly in size, but both disk reliability and disk speed have remained roughly static.

Rebuild maths for RAID 5 (single disk parity) has been known to be an issue for some time. Even without a total disk failure,

Unrecoverable Read Errors are a fact of life. Increased disk sizes without increased disk speed means longer rebuild times. Already we are at the point that the chances of a second disk failure during a RAID 5 rebuild are so high that RAID 5 cannot be considered safe for production use. Even RAID 6 (dual disk parity) is increasingly questionable.

High availability is all about the ability to withstand the loss of an entire storage node. It isn't just the disks that can fail, controller cards, motherboards, RAM, CPUs, network cards…any and all of it can die without warning. Car versus utility pole can cause a power surge, and sometimes the UPS just doesn't take the hit like it's supposed to.

The solution to both problems is to move away from a single storage system running RAID and to turn to a Redundant Array of Inexpensive Nodes (RAIN) architecture. RAIN keeps a copy of the data on another storage node entirely, synchronising writes across the network. If one node is rendered unusable or inaccessible for any reason, the data is available elsewhere.

Capacity Optimization (Thin Provisioning, Compression and Deduplication)

Capacity optimization features have been around for some time. As technologies they offer a straightforward compromise: more efficient use of storage space in exchange for some CPU power as required to perform the calculations. The three technologies ready for mainstream consumption in today's storage are thin provisioning, compression and deduplication.

At one time, each of these technologies was a product unto themselves; companies that offered only one of these technologies. Storage is so expensive for large enterprises that the ability to delay refreshes by a year or two made many a start-up their fortune.

Though these technologies have been around for years, the trade-off for their mainstream use hasn't always made practical sense. Moore's law, however, made their adoption inevitable: processing power doubles roughly every 18 months and the cost of each cycle has dropped dramatically with time.

Today, storage is far more expensive than processing power, and expending a little bit of CPU in order to see a dramatic reduction in storage usage is a rational exchange to make.

Flash changed everything. Not only is it far more costly per gigabyte than spinning rust, but flash has a write limit. Making optimal use of flash is all about minimising writes, and here compression and deduplication become must-have technologies. Instead of writing the same block umpteen times across the available storage it is written only once.

This helps preserve flash lifetimes, increases practical storage. Highly de-duplicatable workloads (such as persistent VDI) can see a marked increase in performance from the utilisation of deduplication and compression in hybrid arrays as it allows more hot blocks to be stored in the SSD tier than would otherwise be possible.

Some storage vendors are charging a premium for these features. With the cost of storage growing so rapidly, compression, deduplication and thin provisioning absolutely need to be included as standard with all storage. It's high time that the

technologies needed to ensure we are getting the maximum possible usage out of our storage are made standard, and that the complexity of their implementation is handled without need for administrator intervention.

PERFORMANCE

The performance of storage – and what is needed to optimise it – is a hotly debated topic. Different workloads have different requirements; however, virtualisation blends up all the I/O of various workloads into a single stream. Server SANs and host-based caches add additional factors to consider as they replicated between them and can serve up storage to workloads directly from the compute node on which they operate.

The underlying storage technology in use – flash versus spinning rust – makes a big difference. So does proper software. Far from being a drag on performance, capacity optimisation technologies can in fact be a boon, if the storage software is up to the challenge of doing things right.

Spinning rust is just not enough. Meeting the demands of today's workloads requires flash. Sadly, the cost of all-flash arrays is realistically beyond most companies. They look to stay that way, as storage volumes are increasing faster than the cost of flash is decreasing.

The middle ground solution to this the hybrid array. Bulk data storage lives on spinning rust with hot blocks moved up to SSDs. If the majority of data can be served off the SSD then the requests that do go to the spinning rust disks shouldn't overwhelm their limited performance. If the amount of flash is properly balanced, then the system as a whole will be lightning-fast, even when fetching data from spinning rust.

In a perfect world, you feed the cluster more flash and watch it go faster, but this only works if the software knows what to do with it. The most important aspect is to make sure that the storage solution is designed to make the best use of the resources available.

Capacity optimisation technologies are a performance hit to any storage system. Turning them into a performance gain requires clever software designed by clever people. In a hybrid storage solution their use allows us to put more data in flash, thus the whole system goes faster, turning a drag on performance into a remarkable gain.

Consider also snapshots and clones. Taking a snapshot or making a clone through the hypervisor alone can take quite a lot of time, especially if the attached virtual disks are over a terabyte.
The solution to this is to hand off creation of snapshots and clones to the storage array. Each vendor has their own behind-the-scenes wizardry to accelerate such tasks, with varying degrees of success. A successful implementation will enable the creation of snapshots and clones in seconds, not minutes or hours.
The key to enable the creation of these efficient snapshots and clones is for the solution is not have to make a copy of the data to make a snapshot or clone. It should simply update some metadata so that there exist two pointers to the same blocks. This is convenient with small VMs. As the size of the virtual disk increases, it moves into the realm of absolute necessity.

Just as with capacity optimization features storage solutions should not be charging extra for the convenience; storage should be simple and this functionality should be just part of the offering.

SCALABILITY

Among the most dominant storage issues is the issue of scalability. Getting more storage – or faster storage – frequently requires a forklift upgrade. Move all data off the old unit and over to a new one, with the attendant capital costs and service disruptions that entails.

The alternative is sharing; increasing the number of arrays in play and moving workloads from one array to another. This is either a tediously manual process requiring both monitoring and vigilance, or it involves the purchase of some rather expensive management software to go along with your already painfully expensive storage arrays.

Storage solutions utilizing RAIN architecture allows for storage to grow simply by adding nodes to the cluster while providing for data resiliency. Your ability to choose the storage mix in your servers means you determine how must storage each node has as well as how much flash is made available for performance acceleration. Capacity optimization ensures that you make optimal use of the storage you purchase while also getting lightning fast snapshot and clone creation.

WHY I CHOOSE MAXTA

Maxta's software is simple, resilient and it gives me access to the latest capacity optimization technologies. It offers both performance and scalability in a converged infrastructure offering while letting me choose the balance between compute and storage by designing my own infrastructure to suit the needs of my workloads.

Maxta Storage Platform (MxSP) delivers high availability by ensuring that all data exists on at least two nodes without any expensive proprietary hardware. The lack of configuration and implementation complexity keep the costs of managing and maintaining storage to a manageable level.

With Maxta there's no more over purchasing storage to meet compute requirements, or buying too many converged nodes just to get enough raw storage. If I need more IOPS, I can add SSDs. If I need more storage, I can add some spinning rust. One size does not fit all, and Maxta helps me deliver exactly what my clients need.

Additionally, MxSP bundles all enterprise-class features such as snapshots, clones and capacity optimization features into a single offering with no complex licencing to worry about. Storage should be simple. With Maxta's MxSP, it is.