

STORAGE SWITZERLAND REPORT

DRIVING DOWN STORAGE COMPLEXITY WITH SSD



George Crump, Senior Analyst

The mechanical hard disk has been rotating at about the same 15K RPM maximum speed for roughly a decade now. Since rotational speed is the primary determinant of disk drive read/write speed, the industry has had to develop other methods to deliver acceptable performance. The challenge is that these workarounds are now buckling under their own weight with dozens of hard drives being chained together, consuming power and floor space, while still not meeting the demands of the applications. An alternative may be to drive down storage complexity with solid state disk (SSD). When time spent tuning and retuning the mechanical storage environment is factored into its total cost, along with power and cooling, SSD may be a far more cost effective and simpler approach than trying to wring every last ounce of performance out of mechanical based technology.

Storage systems made up of mechanical hard drives certainly have their place in the modern data center, but the case can now readily be made that this place should not include performance oriented environments. If the applications, or the servers hosting those applications, cannot exceed the demands of a basic RAID configuration, then mechanical hard drive storage is the right choice. However, if *any* storage I/O tuning is needed, especially if tuning will be an ongoing task as the application matures, then SSD easily becomes the more cost effective and least complex storage performance solution.

The Complexity of Tuning Mechanical Based Storage Systems

The only real improvement in storage performance over the last decade has been from increased density per platter, not from an increase in rotation speed. Because of this density more data can now be collected with each rotation of the platter; but it can't spin faster. As a result the aforementioned workarounds needed to be developed. The first of which was to add more and more drives to an array group. In theory, with each drive added, RAID group performance should improve, until queue depth runs out. Queue depth, as explained in this [Visual SSD white paper](#), is essentially the number of I/O requests that are waiting on the array. Once that depth is depleted adding drives to the RAID group has little to no impact on performance. Even if enough queue depth can be created there comes a point where the amount of drives required to provide optimal performance is no longer cost effective.

If queue depth is depleted, but more performance is still needed, the only course of action is to increase the speed with which the drive responds to I/O requests. This is also known as reducing latency. Once the maximum RPM drive is deployed the only way to reduce latency is to make sure that data is only placed on the fastest outer sections (or cylinders) of each platter. To make sure that this happens, only the outer section of the platters within the drive can be formatted. This technique, known as "short stroking", will reduce latency to some degree but at the cost of 50% or more of the capacity of the drive. The result is the fastest drive, which is also the most expensive, formatted to half of its capacity.

The techniques of adding drive count and potentially short stroking those drives greatly increases the complexity of the environment. Also high drive count RAID groups have a greater probability of failure of a single drive, reducing reliability and performance when a RAID rebuild occurs and that data is put at risk. Initial design and layout of these configurations is critical to their performance. For example, best practice from most manufacturers is to have one drive per RAID group, per shelf. This protects against a shelf failure resulting in data loss; but in a 50-drive RAID configuration, this is practically impossible. As a result performance is either contained to the maximum number of available shelves or reliability is put in jeopardy by having multiple drives from the same RAID group on a shelf.

Once the layout and drive count is determined then a RAID level or RAID type must be selected. The default is to use RAID 5, but as stated earlier, there's a higher probability of a single drive failure in large drive-count RAIDs. A single drive failure, at a minimum, will impact performance, but a second drive failure can cause total data loss. Dual parity RAID 6 is an alternative, but brings with it an immediate write performance penalty. The only other option is to mirror drives, which of course increases the solution cost by creating a 1:1 copy of data.

It is also more complex to scale these high drive count configurations. Each increase in capacity requires that RAID stripes be carefully planned, the drive locations within shelves be understood and that impact to other systems sharing those shelves be considered.

Finally, there is the issue that after all of this planning and the initial implementation phase is complete, performance demands are not still not met. It's very possible that the application can still require more storage I/O than what even this finely tuned mechanical drive configuration is able to deliver, or what the data center can afford to provide. Even if the configuration is acceptable initially, performance driven applications have a tendency to require more performance as they mature and become more widely utilized. Applications that generate this type of storage I/O

demand are often the 'mission critical' types of applications that have high visibility throughout the organization. Having to pause or worse, shut down the application to do performance tuning is not lightly considered.

The alternative is to address the solution, potentially permanently, with solid state storage (SSS). While SSS has the perception of being more expensive than mechanical storage it has the ability to eliminate performance problems, at the current moment and far into the future. This allows the storage performance problem to be addressed once, with no future impact to application performance for tuning.

One of the biggest issues is the impact on storage administration time that these complex mechanical RAID systems need. The system has so many components that it needs almost constant management attention. Tuning of performance needs to be revisited time and again to make sure that every possible I/O is delivered. Compare that with the SSD reality. Planning does need to take place to make sure SSD is implemented correctly, but once that task is complete, for the most part it's done and doesn't need to be revisited.

SSD is as close to a 'set it and forget it' option that storage I/O performance tuning has. While the implementation needs to be planned to account for integration into the overall data protection strategy, once done the tuning work is over.

When the issues with trying to manage the workarounds for improving performance on mechanical based arrays are factored into the overall cost of the system, SSS becomes a far more economical alternative. Dozens, if not hundreds, of drives can be reduced to a single SSD storage system, lowering power and footprint requirements at the same time. By eliminating large, complex RAID array configurations the infrastructure is simplified. Time spent managing storage is reduced significantly and performance is fixed and rarely needs to be revisited.

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>.