# It's Time for Disk Drives 2.0

**Date:** March, 2008

**Author:** Mark Peters, Analyst

**Abstract:** Disk drives are mechanical, subject to breakage and adversely impact IT operations. After decades of addressing the symptoms and masking the issue, it is time to apply new thought and technology to address the cause and deliver Disk Drives 2.0—drives that do not negatively impact IT.

## What If…?

There's a well-worn saying in the storage business—disks do three things: read, write and break. But what if they could be limited to just the first two? It's not talked about much because it's so much an accepted way of IT life—one of 'the rules,' if you like. But what could change if disks and disk systems didn't break like they do now? It would be a whole new fact to deal with—and there would be implications in terms of service, reliability, personnel support, management skills and so forth. Would RAID as we know it be relevant? Could we manage with fewer copies of data? Indeed, do we even need to change? After all, there isn't a general, audible market clamor for more reliable disks, but that is most probably explained by the mind-set that a) no-one really believes it's possible, and b) in any case, the industry has perfected many 'band aids' to overcome the vagaries of existing disk systems.

## The Way Things Are – Disks '101'

Primary data lives on HDDs—'round, brown, spinning things'—as it has for years. That hasn't changed and neither have the underlying physics, despite the multiple generations of size and protocol. For instance, contention has always been an issue, and it's well known that areal densities (capacity) growing far more rapidly than the RPM rate (performance) leads to challenges when managing the retrieval of ever-growing volumes of data. So, a slew of work-arounds have been delivered over the years, such as multi-pathing, caching and striping. The number one problem with disks, however, has always been lack of reliability. Although MTBF (Mean Time Between Failure) numbers have generally improved, so has the number of HDDs installed and the amount of data on each drive; such that the impact of disk breakage and replacement remains a significant factor in just about all IT environments. The storage industry has been most creative in developing methods to mask, manipulate and manage this issue as well—both via straightforward backups and the many elegant levels of RAID architecture. However, all the methods share one aspect: *they address the symptom rather than the cause*.

Before turning to the possibilities that could be enjoyed IF highly reliable and truly self-healing disks existed, it is worth restating the main impacts of disk breakage—impacts to which data managers are inured.

a) **Management impact** – both to architect around the inevitable breakages (these days usually done with RAID of some sort) and also by handling the business-risk and user-distress following the invariable performance degradation associated with disk failures. Depending on the business and security sensitivities, there can also be issues with the disposal and destruction of the devices.

b) **Operational impact** – failed drives, RAID protected or not, need to be physically replaced as fast as practicable. The risk of actual data loss increases whenever a drive is lost; even if the system is designed to protect data availability. A further, very practical, risk is that either the wrong drive might be removed (fairly easy with today's small and tightly packed systems) or the right drive might be incorrectly removed. Drive replacement is no small deal—in large data centers, there are often engineers, as well as dedicated spares, on site 7/24.

It's reasonable to wonder why the band-aid approach continues. Partly it is due to natural human and business inertia (very few organizations want to 'start from scratch') and partly it is because *this is the way it is.* The pertinent questions are:

- Can band-aids continue to work as the nature of IT and its storage needs change?
- What if a better, virtually 100% reliable disk system existed?

## The Way Things Should Be

What would 'reliable' mean in practice for disk systems? Certainly, there are improvements that can be made in component packaging, cooling efficiencies and vibration control. However, there's no serious assertion that disks, being mechanical, will ever reach the stage where breakage is totally eliminated. What is possible, however, is to move the *impact* of any breakage—that is, the maintainability and reliability features—*into* the disk subsystem.

A self-healing, self managing disk subsystem would preclude both the management and operational impacts described previously. Addressing the cause, rather than the symptom makes perfect sense, since it could:

- Reduce the need for ever-more-complex RAID implementations
- Remove the data integrity exposure during drive replacements (whether from lower data protection while RAID is impacted, or from human error)
- Eliminate the need to choose between data exposure and performance during RAID rebuilds (since these can take hours, a half-day or more, often the only way to retain system performance is to 'dial down' the rebuild priority, which extends the period of elevated data risk)

What does this mean? Today, disk breakages are all about managerial and operational impact. By both limiting disk breakages and keeping those that remain within an intelligent, self-healing system, IT managers and users would be insulated from any adverse impact. It's true that current RAID systems help to mitigate the issues, but they are complex and expensive in terms of overhead (for example, the 'unit of failure' is an entire HDD, even if only one platter surface has gone bad; an enclosed self-managing system could be far less wasteful and require less spare capacity).

Now also, in addition, no reasonable IT manager would likely be prepared to trade other key attributes in order to get more reliable disks; in other words, factors such as performance, availability and cost would have to be at least as good as now—and preferably better of course! However, whereas the disk buying decision has traditionally been 'you can pick any two from reliability, performance and capacity,' it would be replaced with 'it's reliable, with performance and capacity scaling as needed.'

## The Business Value of 'Disks 2.0'

So, imagine a world with disk systems that never cause a service event. Today, there are frequent disk replacements being done by service personnel (trained engineers)—this would stop. Moreover, with the gradual shift to disks being CRUs (Customer Replaceable Units) there are also many replacements being handled by relatively inexperienced personnel—this would also stop. Both of these changes represent a security improvement, as well as an operational one. Having no service events would mean significantly lower service costs, fewer engineer journeys and an environmental benefit in terms of better usage of fewer components (less to manufacture and less landfill).

More importantly, such reliable disk systems would be a means and not an end—what happens to applications, performance and business costs is what's crucial. The reason such systems could be so important is simply because storage capacity requirements continue to rise. Today's demands around unstructured data, rich media, compliance, backup-to-disk and archive are only accelerating the storage capacity trend. This in turn means that configuring and managing storage keeps getting harder, with the impacts of unreliable disk becoming continually more painful and simultaneously less acceptable. The overall umbrella movement referred to as 'Web 2.0' compounds the issue of massive growth by adding unknown elements of varied scalability into the demand mix. In plain English, this further drives the need for improved ways to run our storage… and indeed, preferably for it to run itself. Rather than continually adding band aids that mask the true cause, it is time to address the underlying issue.

## The Bottom Line

*"To see what is in front of one's nose requires constant struggle" – George Orwell*

Contrary to what might be thought, there is relatively little change in the basics of storage from year to year. And what does change is often introduced to deal with the problems created to date. Disk hardware has followed a similar path, yet there is clearly room in the market for an improvement to the basics of the technology. The current approach works only because we wrap unreliable HDDs in complex layers of expensive managerial and operational protection. That said, any vendor proclaiming victory in having a fully reliable disk system faces a number of stiff challenges, as does anyone seeking to change the status quo. Proof will be vital. As will an understanding that no market changes overnight. Such a move is unlikely to come from any of the current large players as it's not in their interests, meaning the 'new' Disk 2.0 vendor faces even more of an uphill credibility battle. Large incumbent suppliers and manufactures would 'FUD' the new technology, as it is better for them to supply ever increasing amounts of disks that fail, and thus generate additional revenue for themselves as well as support a whole service industry (which would be threatened by the new model).

However, a new approach makes sense and for security, cost and ease (that is, lack of impact), it is easy to imagine users adopting such an offering. The more flexible any such implementation can be, the better—it should be capable of being configured as anything from DAS to SAN to GRID and of integrating new storage components into its architecture. In terms of changes in storage, some have been revealed as more "marchitecture" than miracle (think of ILM). Some never made it as big as maybe they could have (optical for example). Some have been, gone and have re-appeared (SSD and virtualization). But some seem ripe to be addressed. Why keep managing around the lack of disk system reliability if it can actually be fixed?