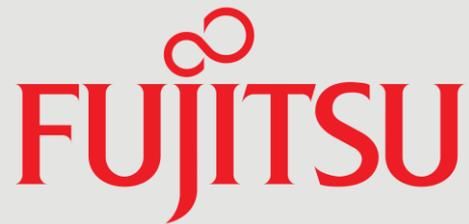




Inside Track Research Note

In association with



Data Protection and RAID

Modern business needs new
storage protection

February 2015

About this Inside Track

The insights presented in this document are derived from independent research conducted by Freeform Dynamics. Inputs include in-depth discussions on the latest technology developments with IT vendors and service providers, along with intelligence gathered from mainstream enterprises during broader market studies.

Users nowadays consider pretty much all forms of data to be 'business critical'.

Traditional approaches to data protection may not be serving you as well as they used to, and RAID technology in particular is increasingly falling into this category.

In a nutshell

Rapid data growth, a greater emphasis on operational risk and developments in storage hardware are conspiring to question the suitability of some traditional data protection solutions to deal with evolving needs. Against this background, a blend of approaches is required to effectively manage data-related risks going forward.

Increased pressure on data protection and recovery

You probably don't need reminding of the ever-increasing amount of data that's currently generated across your organisation on a daily basis. This includes structured data from core business applications, along with documents and objects that are more unstructured in nature arising from the use of office tools, communication and collaboration systems, creative solutions, and so on. Throw cloud services into the mix and the chances are that your storage systems are starting to feel the pressure.

To make matters worse, expectations across your business are also undoubtedly rising. Users nowadays consider pretty much all forms of data to be 'business critical'. They expect you to look after it and make sure it's available on a 24 x 7 basis. In the unfortunate event of it going off-line for some reason, they want their data back and accessible in the shortest possible time.

As a result of all this, within IT you are almost certainly having to place a higher degree of emphasis on availability and recovery today than you have done in the past. At the same time, traditional approaches to data protection may not be serving you as well as they used to, and RAID technology in particular is increasingly falling into this category. In the remainder of this document we will explore why this is the case, and look at ways to overcome the challenges.

Traditional RAID is falling short

Almost all organisations use some form of RAID somewhere in their infrastructure to protect data held on hard disk drives. It can be implemented in various ways that each have their pros and cons, but the basic idea underpinning all RAID systems is the spreading of redundant data across multiple drives to prevent data loss and ensure continuity of service if a disk fails. When this happens, protection is reinstated automatically while the system is still running once a replacement disk is installed.

Over the years, RAID configurations have become an integral part of the way we protect data at a fundamental level, and it has generally proved to be both effective and convenient. So what's changed?

Higher capacity disks mean lengthier rebuilds

The storage capacity of individual disks continues to increase steadily as technology companies find ways of fitting more data onto a single drive. This is generally good news from a cost perspective as this trend brings with it a reduced cost per unit of storage. The need for fewer moving parts can also simplify systems and lead to a corresponding reduction in administrative overhead.

The traditional method by which RAID rebuilds occur following a failure can take a significant amount of time, and the higher the capacity of the disk(s) involved, the longer it takes.

But the news isn't so good when it comes to the impact of increased drive capacity on RAID configurations. This is because the traditional method by which RAID rebuilds occur following a failure can take a significant amount of time, and the higher the capacity of the disk(s) involved, the longer it takes.

Furthermore, while rebuilds are in progress, performance of the overall storage system, and therefore service levels experienced by users, can be negatively impacted. In addition, and perhaps more significantly, the storage system is potentially exposed until a rebuild completes. A second disk failure before the first one is fully recovered could compromise the integrity of the entire RAID array, leading to both downtime and potential loss of data.

In summary, higher capacity disks translate to longer RAID rebuild times, which in turn leads to greater business risk.

But this isn't the only way in which some of the trends we have been discussing are leading to RAID-related challenges.

More disks and aging kit mean higher risk

While the capacity of individual disks is increasing year-on-year, the rate of change is nowhere near the growth rates you are likely to be experiencing in terms of overall data volumes. There are then situations in which it is preferable to use a higher number of lower capacity disks (rather than fewer larger ones) for performance reasons – e.g. to maximise the degree of parallel activity.

Pull these factors together and the reality is that as data volumes grow, the number of physical hard disks required in your storage systems invariably increases too, despite the fact that each device can hold more. And this matters because the probability of experiencing a disk failure increases linearly with the number of physical disks deployed.

Aggravating the situation, the average age of equipment is escalating. You would not be unusual if you were expecting the storage equipment you currently have in place to remain in service for longer than in the past, with a life expectancy of 4, 5 or even 6 years and upwards. Given that the likelihood of failure increases as devices get older, this multiplies the chances of problems occurring, and therefore the frequency at which RAID rebuilds become necessary.

All of this means more risk of both data loss and service interruption or degradation.

Acknowledging the RAID challenge

If you are unsure of how much you should be taking some of these challenges seriously, the results of a recent Freeform Dynamics research study might help. More than a third of the 403 IT professionals taking part considered RAID rebuild times to be a major risk to the business, with a further quarter regarding the issue as a significant distraction.

While the remainder didn't see any problems in this area, we have to bear in mind that a proportion of these probably hadn't thought about it enough to acknowledge the challenges. If you live with constraints and issues for long enough, the tendency is to just accept them as normal to the point where you have stopped even looking for better ways of doing things.

If you live with constraints and issues for long enough, the tendency is to just accept them as normal.

Data protection is an area that often doesn't receive the attention it deserves.

It may be time to pause and assess where you are.

Despite the traditional challenges, RAID still has a role to play.

The fact that you have read this far suggests that you are aware of the need for review and action. So, if you have experienced problems already, or see the potential for challenges to emerge as data volumes and variety continue to grow, then stay with us as we look at a few different ways of approaching resilience and recovery.

Options to improve data protection

Our research at Freeform Dynamics tells us time and time again that data protection is an area that often doesn't receive the attention it deserves. The tendency is to put solutions in place then consider the problem to be solved. Not only does this mean that protection measures drift out of line with business requirements, especially given the pace of change we have mentioned, but that it's all too easy to fail to keep up with the way in which the solution side of the equation is developing.

If it has been a while since you performed an objective review of how you protect data in your organisation, then it may be time to pause and assess where you are.

A good place to start is with data classification and requirements definition. While users might claim that everything is critical, the reality is that needs will vary in relation to factors like performance, resilience, tolerance of data loss and required recovery times in the event of a failure. Once you understand the requirements of different data sets, you can start to look at appropriate protection measures more objectively. Let's look at some of the options you might consider while doing this.

More effective use of RAID

Despite the traditional challenges, RAID still has a role to play. As you review its use, however, it's important to address the exposure represented by elongated rebuild times. Here are some approaches now available to help:

- Selective recovery, in which only the 'active' data elements of a disk are rebuilt, rather than including unused 'free' space in the recovery process. Recovery times are reduced simply because less work needs to be done.
- Distributed parallel recovery, in which redundant data generated during a rebuild is spread across all disks within a RAID array, rather than being targeted at a single new or spare disk. The idea here is to reduce recovery times by spreading the I/O load and avoiding bottlenecks.
- Use of high quality disks and storage controllers to minimise failures. This is about appreciating that effective protection should not be regarded as a license to use commodity equipment without thinking.
- Matching RAID formats with requirements, i.e. appreciating that a different type of RAID may be appropriate depending on what's important in relation to the data, application or workload at hand.

The following advice from a participant in our research serves as a good reminder:

"Always look at the applications using RAID. For example, mission critical applications are at their most vulnerable during a RAID rebuild, so rather than use RAID 5, it would be better to switch to RAID 6. I always find it best to refer back to the basic principles of RAID before making these kinds of decisions."

But RAID alone is never going to meet all requirements.

'High Availability' (HA) solutions allow you to create a 'hot standby' of your storage environment so that applications can failover automatically.

Storage Cluster offerings may be supplied as ready to use solutions designed to provide DR / HA capabilities without the need for complex additional implementation services.

But RAID alone is never going to meet all requirements, so let's take a look at some of the options when it comes to a system or site failure.

Data availability during a system or site failure

Recovering from the failure of a single disk can be challenging, but there are data sets that are so critical to your business operations that it is essential the data is available even in the case of the failure of a storage system or the loss of your site. In the past setting up systems and networks to give you DR data protection capabilities was complex and expensive. As a consequence such solutions were only available to one or two business critical applications in the largest organisations.

But some vendors can now offer you storage arrays with functional capabilities designed to ensure that in the event of a disaster your storage can failover automatically to a second system running at another location. This could be in another office, computer room or data centre belonging either to your organisation or one run by your chosen service provider.

The second option is particularly appealing if you work in a smaller business where your IT runs only in one place. Indeed, the availability of storage systems able to provide remote DR failover capabilities at an affordable price without the need for highly specialised skills is one that has rarely been seen in many small and mid-sized organisations.

There are now several options available to provide automated storage failover in the event of a disaster, and some are offered as ready to buy solutions from a number of vendors. It is important that you ensure any solutions proffered are able to operate as seamlessly as possible with your IT infrastructure.

Technologies that allow data to be mirrored, snapshotted or replicated between systems, or even locations (e.g. across multiple data centres, or between your data centre and the cloud) are now widely available. Within this space, so-called 'High Availability' (HA) solutions allow you to create a 'hot standby' of your storage environment so that applications can failover automatically with little or no downtime should a serious failure occur.

DR / HA storage approaches include:

Storage Clustering

Storage Clustering is similar to that which has long been available in servers and provides capabilities that automatically synchronise data between two storage arrays in different locations. Clearly good network connectivity is essential, but after the initial data load and synchronisation has taken place, only changes to data are replicated between sites. Some suppliers include management tools that can operate either fully automatically or in manual mode within these systems. Other solutions may make use of 'wizards' to simplify both the setup and routine operation of the storage cluster. Top of the range systems also offer HA or even 'continuous availability' to data in the event of a failure. Storage Cluster offerings may be supplied as ready to use solutions designed to provide DR / HA capabilities without the need for complex additional implementation services.

It is now possible for organisations of all sizes to consider implementing storage solutions that provide effective DR capabilities in the event of a systems or site failure, without breaking the bank.

Larger players who offer a broad range of storage options are more likely to give you balanced advice.

Scale-out redundant storage

Born out of the open source and service communities, 'hyper-scale' software defined storage (SDS) environments provide an alternative way of handling the relentless growth in data volumes. Some simultaneously solve the data protection problem into the bargain, e.g. solutions based on open source projects like Ceph allow you to specify the number of copies of data that are created on-the-fly when a file or object is stored. Such solutions need not involve working through all the complexity and uncertainty often associated with the use of open source. Pre-integrated and fully supported systems are available from mainstream vendors in appliance or software package form.

Storage Virtualisation

Storage virtualisation solutions are built using special software hypervisors that layer on top of storage arrays. Such systems can come with functionality to replicate data around the storage pool created, allowing you to failover in the event of one array failing to data held in other parts of the pool. While such systems can be effective they also introduce another technology to be managed which can place additional burdens on IT staff and may also carry the financial cost of the storage virtualisation software itself.

It is now becoming possible for organisations of all sizes to consider implementing storage solutions that provide effective DR capabilities in the event of a systems or site failure, without breaking the bank or needing dedicated specialists to implement.

The bottom line

It will be clear from our discussion that a blend of approaches is required to strengthen, streamline and future-proof your storage systems. As you explore the options, however, beware of niche suppliers who sometimes over-position point products. It's not that you should necessarily avoid them, but larger players who offer a broad range of storage options are more likely to give you balanced advice.

The overriding imperative, however, is to appreciate that the world has moved on in terms of requirements and technology, so it's critical to review your current setup and make sure you are in a good position to handle what the future has in store.

About Freeform Dynamics

Freeform Dynamics is an IT industry analyst firm. Through our research and insights, we aim to help busy IT and business professionals get up to speed on the latest technology developments, and make better informed investment decisions.

For more information, and access to our library of free research, please visit www.freeformdynamics.com.

About Fujitsu

Fujitsu is the leading Japanese information and communication technology (ICT) company offering a full range of technology products, solutions and services. Approximately 162,000 Fujitsu people support customers in more than 100 countries. We use our experience and the power of ICT to shape the future of society with our customers.

Business-centric storage

Under the ETERNUS storage solution brand, Fujitsu provides disk systems and data protection appliances, which allow customers to flexibly manage their increasing data volumes at a lower cost of growth, and to benefit from a very reliable architecture and radically simplified operation.

Storage Management Software

The ETERNUS SF storage management software is a uniform management solution that works across the storage portfolio, from entry-level systems to high-end configurations, including virtualised server infrastructures. All essential operations for resource management, monitoring, reporting, tiering, performance management, disaster resilience and business continuity are integrated.

For more information, please see <http://www.fujitsu.com/eternus>.

Terms of Use

This document is Copyright 2015 Freeform Dynamics Ltd. It may be freely duplicated and distributed in its entirety on an individual one to one basis, either electronically or in hard copy form. It may not, however, be disassembled or modified in any way as part of the duplication process. Hosting of the entire report for download and/or mass distribution by any means is prohibited unless express permission is obtained from Freeform Dynamics Ltd or Fujitsu. The contents contained herein are provided for your general information and use only, and neither Freeform Dynamics Ltd nor any third party provide any warranty or guarantee as to its suitability for any particular purpose.