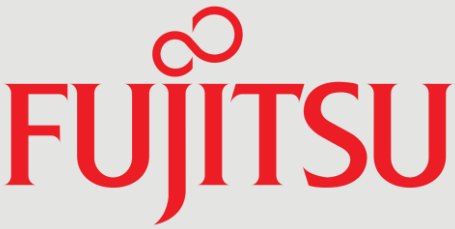
Inside Track Research Note

In association with



Data Protection and Management in a Box

Appliances that take the pain out of backup and archiving

August 2015

In a nutshell

In today's fast-moving, information-intensive business environment, data management is more of a challenge than ever. Relying on manual processes and scripts, or ad hoc piecemeal automation, is not sustainable over the longer term. But implementing comprehensive software-based data protection and archiving can be quite daunting. Could the answer be to explore appliance-based options?

About this Inside Track

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

That was then, this is now

Data protection arrangements are often implemented, then forgotten. Whether at the level of specific systems or your infrastructure as a whole, you may have backup and recovery mechanisms in place that haven't been reviewed or even tested for a while. It also wouldn't be unusual if you were relying on multiple tools to protect your data from loss or corruption, each needing to be separately administered in a different manner. These are some of the realities resulting from the historical practice of dealing with data protection requirements on a case-by-case basis as they arise.

Added to this, the world changes, and so does your business and systems landscape. Tools and processes that were appropriate in the past may no longer be fully up to the job. Looked at in the context of today's higher data volumes and heightened expectations of availability and recovery, they may be judged to be inefficient, costly to run, and even inadequate to meet your organisation's current and evolving needs. Meanwhile, data protection technology has advanced significantly over recent years, so even if you have properly working solutions in place, there may still be more efficient and effective ways of doing things.

Tools and processes that were appropriate in the past may no longer be fully up to the job.

Assessing your current position

If it's been a while since you stood back and reviewed your data protection arrangements, then consider the list of issues and challenges below that frequently come up in our research. Do you recognise any of these?

- Business data has become fragmented across application data stores, file shares, PC hard drives, your email system, and other locations, and you can't be certain everything is backed up that needs to be.
- As users take more control, documents and datasets get downloaded, copied, forwarded, sync'd and otherwise replicated to the extent that the exact same information is often backed up many times from multiple locations.
- Static, historical data that seldom or never gets accessed clogs up live systems, occupying much-needed space on expensive premium storage devices, while you scrabble around for budget to buy more capacity.
- That same historical data is backed up nightly, regularly snapshotted, and/or continuously mirrored, even though it never changes and the last time it was accessed by a user was months or even years ago.
- Backup jobs are taking longer and longer to run as data volumes continue to grow, to the point where fitting everything into the available backup window is becoming more of a challenge.

Business data has become fragmented across application data stores, file shares, PC hard drives, your email system, and other locations.

Keeping track of all the mechanisms and jobs that your backup and archiving processes depend on is hard, and some of the measures in place require ninja-like scripting skills.

- You periodically archive historical data, but according to relatively crude policies based primarily on age, then users complain when they need something that's been taken offline because it takes so long to retrieve.
- Keeping track of all the mechanisms and jobs that your backup and archiving processes depend on is hard, and some of the measures in place require ninja-like scripting skills if modification or troubleshooting becomes necessary.
- You know the organisation is becoming more and more dependent on its virtualised infrastructure, but you've never quite figured out how best to handle the backup of virtual machines (VMs) and the data they might contain.
- While you have disaster recovery (DR) policies and procedures in place, you know that a lot has changed since they were devised, yet the thought of working through recovery time/point requirements and testing against these is daunting.

It could be time for you to reconsider your options.

We could go on to speak about security and access implications of the way data is backed up and archived, the degree to which off-site storage is used, whether you have measures in place to deal with data held in the cloud, how easy it is to deal with regulatory requirements around data retention and disposal, and the performance implications of running backup jobs on live application servers. If three or more of the points we have mentioned sound familiar, it could be time to reconsider your options.

In an ideal world…

So let's imagine we have a magic wand, and could wave it to conjure up the perfect data protection environment. What would this look like?

The ideal solution would enable everything to be managed from a central point – file stores, application databases, email repositories and virtual environments.

Well the ideal solution would enable everything to be managed from a central point – file stores, application databases, email repositories and virtual environments. It would allow flexible policies to be set up through a simple point-and-click interface to deal with information backup (take a copy for safe keeping), archiving (move it from the live system to long term storage) and disposal (purge it). Such policies could vary depending on the type of data, the source system and/or the owner or department involved (which might, for example, dictate regulatory requirements).

Our ideal data protection system would obviously know what's needed to take valid backups, e.g. quiescing databases before starting a job to assure backup integrity. It would also compress data to save space, and be clever enough not to back up the same stuff more than once. When it came across the second, third or fourth copy of a document, for example, it would simply replace it with a pointer to the original material (unless policy dictated otherwise). And to aid subsequent retrieval, our data protection system would create comprehensive meta-data so files and objects could be quickly located on backup disks or tapes, or in the information archive.

Designing, let alone implementing, a comprehensive data protection and archiving environment is not an easy task.

From an operational perspective, the architecture of the solution would be based on a dedicated server and storage environment so backup and archiving jobs could be run against live systems without an undue impact on their performance.

Now you might be thinking that you have come across promises like these before from various software vendors, but when you have looked at how to make it work in practice, the integration effort required has been prohibitive. Designing, let alone implementing, a comprehensive data protection and archiving environment is not an easy task.

Making it easy – or at least a lot easier

Most of the complexity and specialist skills requirements associated with a comprehensive solution can be avoided through the use of a backup and archive appliance.

Most of the complexity and specialist skills requirements associated with a comprehensive solution can be avoided through the use of a backup and archive appliance. This will most commonly come in one of two forms:

Integrated Appliance: The idea here is to bring the right software together with the right mix of hardware and management tooling to create a self-contained single product that arrives in your datacentre or computer room in a fully functional and pre-optimised state.

Target Appliance: This option provides a sophisticated, pre-optimised storage platform to underpin your existing backup and archiving software, and is particularly useful if you are using more than one software suite. Replication and deduplication features are usually included to reduce capacity demands and to ensure protection against system failures.

The storage layer typically includes a full policy-driven management environment to minimise administration overhead.

Apart from the speed of initial provisioning, one of the main advantages of the appliance-based approach is that all of the components included have been selected and integrated so they work together in a seamless and optimal manner. Furthermore, in both types of appliance, the storage layer typically includes a full policy-driven management environment to minimise administration overhead.

You can then expect at least a couple of additional advanced features. One example here is data replication to boost the overall level of protection and reduce recovery times in the event of failure.

Another important feature to look out for is storage-level deduplication. This is designed to complement higher level deduplication mechanisms implemented in the backup and archiving software layer. When integrated into the underlying platform, such functionality is extremely efficient, and can further condense data volumes to save even more time, space and money.

Assessing the options available

When procuring either an integrated or target appliance, the first step is to review your requirements and existing landscape.

When procuring either type of solution, the first step is to review your requirements and existing landscape with the supplier to determine the appropriate size of appliance (bearing growth needs in mind), and the options to be included (e.g. mix of disks, incorporation of tape drives, etc).

Manufacturers tend to offer a range of solutions at different entry levels and price points. What they all have in common is an ability to populate the appliance with the amount of storage you need for the short to medium term, then add more capacity later. As an example, your appliance may be initially delivered with 30% of the available device slots populated. As your requirements grow, you can then add more disk (of whatever type makes sense) or swap out smaller capacity units for larger ones.

Offerings in this space have been designed both technically and commercially so you only pay for what you initially need.

The whole point is that offerings in this space have been designed both technically and commercially so you only pay for the capability you initially need, but have the peace of mind that you can expand later as your needs evolve. The trick is therefore to select a point in the manufacturer's product range that provides room for growth, without paying a premium for expansion capacity that you may never exploit.

Some specifics to consider

It's important to do your due diligence and check that the box contains all the functionality you are likely to need or that relevant features can be added.

While solutions are generally built with capacity expansion in mind, the same is not true when it comes to the scope of their functionality, especially if you are considering an integrated appliance. As you research this area, you will come across options at all levels of capability, and it can be tempting just to focus on what you absolutely need in the short term.

Remember though, that unlike buying solutions as individual components that you integrate yourself (or get a reseller or integrator to put together for you), you may not be able to make big changes to the functional scope down the line. It's important to do your due diligence and check that either the box contains all the functionality you are likely to need to meet longer term requirements, or that relevant features can be added at a future date cost-effectively.

As a simple example, if your appliance provides straightforward file and email backup and recovery capability based on an open source software component, but does not deal with virtual machine environments or application databases intelligently, then you may not be able to add this later. Similarly, if you decide in the future that you need flexible policy-based archiving capability, then this too may be difficult to add in.

If the appliance doesn't include self-management capability such as auto-migration of data between disks based on real-life access statistics, then you probably won't be able to retrofit this later.

And the same is true of the storage layer. If the appliance doesn't include self-management capability such as auto-migration of data between disks based on real-life access statistics, then you probably won't be able to retrofit this later if, for example, you decide to go down the online rather than offline archiving route. Again, the flexibility to add different 'grades' (or even form factors) of disks, e.g. cheaper slower devices to handle longer-term online storage cost-effectively, is important from a future-proofing perspective.

Fortunately, some manufacturers have constructed appliances based on state-of-the-art commercial software packages. This means you aren't being asked to compromise from a software perspective; indeed it's sometimes the opposite, in that the appliance manufacturer has made sure the right hardware configuration is in place for you to take full advantage of all the software features available.

Having said all that, there are situations in which it can be difficult to pin down your longer-term requirements for specific backup, recovery and archiving features. You may also have perfectly legitimate reasons to start simple but leave the door open to enhance functionality over time. If either of these apply to you then the answer is probably to go down the target appliance route. You can then upgrade or extend your software when the time is right.

Moving from a storage layer made up of generic disk and/or tape technology to a modern, integrated all-disk based platform can simplify management and allow you to keep more data online.

Target appliances are also useful to strengthen and enhance your existing software environment. Moving from a storage layer made up of generic disk and/or tape technology to a modern, integrated all-disk based platform can simplify management and allow you to keep more data online to speed access and recovery. It will also generally make everything a lot more robust and efficient. If you are using multiple software suites, for example, the consolidated data reduction process will lead to higher deduplication rates, further lowering storage capacity demands.

The right solution will even emulate aspects of your old platform, e.g. by presenting virtual tape devices so existing software and processes are not unduly disrupted as you initially migrate from your current environment.

But the technology only gets you so far

Implementing an effective data protection and archiving strategy needs to be driven as a business initiative.

While an appropriately selected modern appliance removes much of the pain and provides many options from a technology perspective, implementing an effective data protection and archiving strategy needs to be driven as a business initiative. It's beyond the scope of this paper to go into details, but as a minimum you will need to:

- Review your current application portfolio and associated databases or file stores and categorise documents and data sets according to business requirements. Considerations here include disaster recovery basics such as recovery point objectives (RPOs) and recovery time objectives (RTOs), along with retention, archiving and disposal policies in line with both regulatory and business needs. It is perfectly legitimate to focus on key applications and data here, setting default policies for the remainder, as trying to be exhaustive risks your project stalling due to excessive demands on time and resources.

Focus on key applications and data, setting default policies for the remainder.

- Don't just focus on the initial setup and implementation of the solution. Once a policy is established, whether at the document lifecycle or physical storage level, it will generally be applied automatically, but policies are often not static. The right data protection and archiving setup therefore will allow you to create a 'living' policy environment, i.e. one in which you can optimise storage, retention, migration and disposal on an ongoing basis. For this reason, you need to think about who is responsible for which policies, and make sure procedures are in place to deal with reviews, escalations, exception handling, etc.

This last point illustrates that in some ways, modern technology in this area can actually create work – but importantly, it's the right kind of work. Rather than crossing your fingers and hoping on the data protection front, and dealing with storage and retention policy on an ad hoc, manual basis, you now have the visibility and an automation capability to deal with such matters in a robust and business-like manner.

The bottom line

Requirements for effective and efficient data protection and management are only going to increase in the future.

Requirements for effective and efficient data protection and management are only going to increase in the future. Transaction systems are producing higher volumes of data at faster rates as a result of business process automation and digital business. Meanwhile, the seemingly limitless capacity for users to create more and different types of content as they collaborate electronically creates additional pressures. And if these factors weren't enough, regulators continuously ask for more in the area of compliance, whether to do with data privacy, financial integrity or industrial safety.

Reduce risks immediately, save money over the longer term, and generally sleep more soundly.

Against this background, a piecemeal approach relying on old, fragmented and often manual solutions to protect and manage data is not sustainable. Whether you go down the appliance route we have been discussing or not, one way or another you need to get yourself onto a firmer footing. Do this well and you'll reduce risks immediately, save money over the longer term, and generally sleep more soundly.

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.

Fujitsu offers a rich portfolio of backup appliances for different usage scenarios and enterprise sizes helping customers to drive efficiency in relation to backup processes and to increase service levels while reducing costs.

For more information, please see www.fujitsu.com/fts/products/computing/storage/data-protection/

Terms of Use