Inside Track Research Note

In association with

Protecting data in your hyper-converged environment

Keep it safe, keep it simple

Freeform Dynamics, 2016

In a nutshell

Hyper-converged infrastructure (HCI) has been gaining popularity as a way of simplifying IT environments. Solutions in this area are already finding their place in the context of scale out storage, desktop virtualisation and the creation of flexible cloud-like platforms to support the 'service provider' IT delivery model.

Whatever the use case, a frequent question is whether the data protection functionality built into HCI systems is adequate to meet business needs. Most often the answer is 'no', so the question then becomes how best to manage data-related risks. Protecting an appliance with another appliance is an option to consider.

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.

Hyper-converged essentials

Hyper-converged infrastructure, 'HCI' for short, is the latest manifestation of a trend that has been unfolding in the IT industry for a while. Over recent years we have seen suppliers pre-integrating collections of hardware and software into appliances aimed at serving a whole range of specific needs. Whether it's security monitoring, relational database management, in-memory databases, or even general-purpose application serving, options are now available to buy everything you need in a single box for a single price supported and maintained under a single contract.

The emergence of convergence

As part of this trend, we have also seen so-called 'converged infrastructure' (CI) offerings emerge. These bring together compute, storage and networking components into self-contained boxes designed for more general purpose use. Entry-level CI systems have effectively provided a 'data centre in a box' for smaller businesses, and the building blocks for conveniently deploying chunks of standard resource in a larger enterprise data centre environment.

In a hyper-converged system, compute, storage and sometimes even networking resources are 'fused' tightly, and are generally only accessible through a virtualisation layer.

Going hyper

Against this background, HCI at first glance appears to meet a similar set of needs, but if you look under the covers you will find some important architectural differences. It is beyond the scope of this paper to go into technical detail, but suffice it to say that convergence is implemented at a lower level. In a hyper-converged system, compute, storage and sometimes even networking resources are 'fused' more tightly, and are generally only accessible through a virtualisation layer. The approach has many benefits in terms of increased abstraction, simplification, and the smooth scaling of systems using a building block approach as demands increase.

You'll find lots of literature out there on the Web if you want to get up to speed on the nitty-gritty of HCI architecture and benefits, but for the purposes of our discussion here, there's one general attribute of this kind of system that's particularly important.

Black box thinking

HCI systems effectively operate as 'black boxes'.

HCI systems effectively operate as 'black boxes'. This makes working with them quite different to the 'white box' CI approach, in which physical storage, compute and other resources are still managed separately and explicitly (albeit more conveniently through a single pane of glass) using traditional methods. In an HCI solution, the emphasis is on the system rather than the administrator deciding how physical

From an application perspective, HCI systems offer up 'logical' pools of resources.

resources are allocated and exploited. As a simple example, when you give the system a new unit of combined resource (commonly referred to as a 'node'), or replace an existing one, it immediately assumes control and integrates it with minimal administrator intervention.

The interface to the outside world then becomes very simple. From an application perspective, HCI systems offer up 'logical' pools of resources that are accessed strictly through the aforementioned virtualisation layer. This is why HCI solutions invariably include one of the industry standard hypervisors.

Taking care of your data too?

HCI solutions usually include functionality such as automated data replication and snapshotting.

In line with the self-managing black box spirit, HCI solutions usually also include functionality such as automated data replication and snapshotting. This brings us to the main purpose of our discussion in this paper, as such capabilities can create the impression that nothing more is needed to keep your data safe in an HCI environment. But is this really the case?

The truth about embedded protection

Lack of separation

While embedded data management and protection tools can be very sophisticated, they almost always place replicated data on the same platform as the source data. Doing so makes copying existing information for use in other contexts, perhaps for testing or staging purposes, very swift. It also means that data snapshots can be recovered very rapidly and easily should the need arise.

You can get the impression that nothing more than the embedded capabilities are needed to keep your data safe in an HCI environment, but this is rarely the case.

But the problem is that, unless another copy of the data is taken and moved to a separate system situated in a completely different location, there is a risk of potential data loss if an entire rack or possibly even a data centre should fail. And such occurrences in the shape of floods and other catastrophic weather events or site failures can and do occur, irrespective of how much redundancy has been introduced into data centre facilities.

But these scenarios also pose another question: is data replication alone enough to safeguard your corporate information?

Replication is useful, but isn't enough

Replicating data inside systems is now a widely deployed technique and can be very helpful when used appropriately. But in isolation, it is unlikely to provide the total protection required for many business and IT scenarios.

Replication alone is unlikely to provide the total level of protection needed.

Not only will the local nature of the protection be inadequate, but some data replication and snapshot solutions may also be unable to recover information with sufficient granularity. For example, it can be problematic if replication software only enables entire VMs to be recovered should only a small subset of the data need to be salvaged. Replication solutions also have a nasty habit of propagating corrupt data, mistaken deletions, and so on.

It is therefore likely that many organisations will find the most effective means of tackling these issues is to bolster the built-in data replication capabilities of many of

today's HCI platforms with additional backup, archive and other management functionality provided by specialist solutions. Let's take a closer look.

A more holistic view of protection

Even if protection capabilities are present, they may not be strong enough or comprehensive enough to deal with your business requirements.

Sometimes it's a case of the protection functionality you need simply not being included in the HCI system. When you consider your specific environment, you may also conclude that even if capabilities are present, they may not be strong enough or comprehensive enough to deal with your business requirements. So what are we talking about here? Well, this will become clear if we step back and take a more holistic view.

What do we mean by data protection anyway?

The phrase 'data protection' covers many interlinked but different capabilities that IT systems may need in order to ensure that business information remains available, regardless of mistakes, data corruption or disasters. These include:

- Fast replication of data sets
- Rapid recovery of entire data sets/VMs/VDI instances
- Short-term versioning
- Long-term versioning
- Granular recovery of individual data files/email
- Creation of data protection 'air gaps'
- Long-term archiving
- Audits/reporting

This is a pretty long list, but it needs to be.

If you take an objective business view, it quickly becomes clear that protection requirements often vary considerably between data-sets.

Horses for courses

If you take an objective business view, it quickly becomes clear that protection requirements often vary considerably between data-sets. Recovery time and recovery point objectives (RTO and RPO) obviously depend on the criticality of data and the speed with which it is created or updated. These will in turn dictate the protection approach (or approaches) used on a case-by-case basis to balance costs and risks.

And the same is true when you look at retention requirements. For transient data, you may need to do little or nothing other than periodically clean things out. For highly regulated and/or data that has long-term business value, full-blown policy-driven archiving capability may be a mandatory requirement to deal with what's expected from a retention, disposal, discovery and access perspective.

HCI systems are increasingly likely to be used as general-purpose platforms over time. The need for comprehensive data protection capabilities is therefore clear.

Versatile platforms need comprehensive protection

Only you can work out what your precise mix of requirements are in relation to the data you hold. However, coming back to HCI, one of the big advantages of such platforms is that the scalability and operational simplicity that comes with them means they are increasingly likely to be used as general-purpose platforms over time, supporting an increasing range of different applications, workloads and data-sets.

As you look forward, the need for a comprehensive set of capabilities to protect and manage data held in your HCI environment becomes clear.

Implementation practicalities

At the time of writing, the majority of HCI offerings come in the form of hardware appliances. Software solutions are available to allow you to construct your own HCI system from standard hardware components, but whether you buy it prebuilt or build it yourself, the black box, self-managing spirit of the end result from an operational perspective will be the same.

If you're not careful, you could end up with a backup and archiving stack that is more complex to implement, maintain and support than the production environment itself.

Given this, constructing a data protection environment in the traditional manner by buying in and integrating the necessary hardware and software can seem like a backward step. If you're not careful, you could end up with a backup and archiving stack that is more complex to implement, maintain and support than the production application and storage environment itself.

Protecting an appliance with another appliance

Fortunately, new options are now available in the shape of backup appliances. The full benefits of such systems are explored in a recent paper by Freeform Dynamics entitled 'Data Protection and Management in a Box'. Suffice it to say, for the purposes of our discussion here, appliances in this area typically deliver state-of-the-art backup, recovery, archiving and other information management capabilities, pre-integrated onto a suitably configured hardware platform. In doing this, they offer the same kind of simplicity as HCI from an implementation, management and support perspective.

And this simplicity and convenience also really matters when it comes to exploiting the core capabilities of this kind of solution, as it means you are more likely to implement the right kind of protection measures and keep them up to date. The importance of this should not be overlooked when considering the difference between a backup appliance and a traditional DIY data protection approach.

Given that most HCI environments will over time end up supporting a range of application requirements with differing data protection needs, appliances in this area represent a cost effective and future proof way of bringing the required depth of breadth of functionality to bear.

Not just data in the traditional sense

As the world moves increasingly to more rapid deployment and turnover of applications - DevOps-style - comprehensive data protection capabilities can help.

Beyond user and application data, there is another factor to consider when evaluating your protection needs – the fact that HCI systems will usually be hosting virtual machines and storing the associated image and configuration files. Indeed, as we said at the outset, HCI is already finding its place in the implementation of private clouds.

As we consider needs in this area, it's not just backup and recovery of data that's relevant. Regulated applications implemented through VMs may well be subject to lifecycle management for compliance purposes. There is then the data that's held in virtual desktops and servers themselves.

And as the world moves increasingly to more rapid deployment and turnover of applications - DevOps-style - comprehensive data protection capabilities can help here too. For example, consider that when deploying new virtual machines unexpected issues may arise. In such circumstances the use of a dedicated backup appliance can be used to protect older versions of the VMs, providing fast rollback capabilities should the new version cause problems. Similarly, it would also be possible to archive

virtual machines that are only required at certain times during the year or for longer term discovery / reporting purposes.

It all comes down to the fact that many data sets, VMs included, must be held securely for long periods of time, often at least three to seven years. It may also be necessary to retain certain data sets for even longer periods, possibly many decades. In some business scenarios, there may even be a legal mandate to ensure that specific data sets are permanently erased at different points in time - another feature for which built-in data replication functionality has not been designed.

Data protection appliances are potentially a good fit not only for HCI systems, but to cover the backup and archive requirements of the entire 'traditional' IT infrastructure.

The role of data protection appliances beyond HCI

Data protection appliances are now very mature technologies and such solutions have the potential to simplify operational processes with which many organisations still struggle. Fully functional backup appliances often combine traditional backup and archiving software from specialist data protection vendors in a simple-to-manage box. This potentially makes such appliances a good fit not only for HCI systems, but to cover the backup and archive requirements of the entire 'traditional' IT infrastructure.

In some circumstances a good backup appliance could even help facilitate the transition to an HCI from a traditional infrastructure by capturing applications, data and VMs as part of a migration process.

There is clear potential for backup and recovery appliances to become the solution of choice for providing essential data protection capabilities.

Pulling it all together

The use of HCI solutions is rising, and the fact that they are being employed to support an expanding range of enterprise workloads will inevitably generate exacting data protection requirements. This means that various forms of specialist data protection functionality will be needed to complement the built-in replication tools.

Given that these systems are built on the premise of being easy to acquire, simple to get working and easy to manage in a daily operational sense, it is also safe to assume that similar expectations will exist for the backup, recovery and archiving systems that support them. As a result, there is clear potential for backup and recovery appliances to become the solution of choice for providing essential data protection capabilities.

And don't forget the basics

Retro-fitting adequate data protection can end up being more complex and expensive than designing it in from the outset.

As a final word, in order to ensure adequate data protection when using HCI systems, it is essential to consider a few key issues before implementation:

- Know what applications are likely to be moved onto the systems
- Think about recovery point objectives (RPOs) and recovery time objectives (RTOs)
- Consider long-term data lifecycle management needs and the role of archiving
- Work out the DR requirements in the event of a system or site failover
- Design data protection capabilities in light of the above.

Experience tells us that data protection can easily be overlooked in the rush to implement the IT solutions that business users want yesterday. This means that data protection requirements often only become apparent after something has gone wrong. But retro-fitting them can end up being more complex and expensive than designing them in from the outset. The lesson is always to think ahead, and we hope our discussion has helped you to identify what to consider as you do this.

The lesson is always to think ahead.

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.

For more information, please see www.fujitsu.com/fts/products/computing/storage/data-protection/cs200c/index.html

Fujitsu partners with Commvault to build easy-to use backup appliances helping customers to radically simplify their backup environments.

About Commvault

Commvault's data protection and information management solutions provide mid- and enterprise-level organizations worldwide with a significantly better way to get value from their data. Commvault can help companies protect, access and use all of their data, anywhere and anytime, turning data into a powerful strategic asset.

For more information, please see www.commvault.com.

Terms of Use