THE GORILLA GUIDE TO...™

Hyperconverged Infrastructure for Data Protection

Scott D. Lowe
Consultant & Industry Veteran

Brought to you by
Hewlett Packard Enterprise

Helping you navigate the technology jungle

THE GORILLA GUIDE TO...

Hyperconverged Infrastructure for Data Protection

Scott D. Lowe • **David M. Davis**

Author: Scott D. Lowe, ActualTech Media
Editors: David M. Davis, ActualTech Media
Hilary Kerchner, Dream Write Creative

Book Design: Braeden Black, Avalon Media Productions
Geordie Carswell, ActualTech Media

Layout: Braeden Black, Avalon Media Productions

Printed in the United States of America

First Printing, 2015

ISBN 978-1-943952-02-1

ActualTech Media
Okatie Village Ste 103-157
Bluffton, SC 29909
www.actualtechmedia.com

Table of Contents

1

Introduction to Hyperconverged Infrastructure

In recent years, it seems like technology is changing faster than it used to in decades past. As employees devour newer technologies such as smartphones, tablets, wearables, and other devices, and as they become more comfortable with solutions such as Dropbox and Skype, their demands on enterprise IT intensify. Plus, management and other decision makers are also increasing their demands on enterprise IT to provide more infrastructure with less cost and time. Unfortunately, enterprise IT organizations often don't see much, if any, associated increases in funding to accomodate these demands.

These demands have resulted in the need for IT organizations to attempt to mimic NASA's much-heralded "Faster, Better, Cheaper" operational campaign. As the name suggests, NASA made great attempts to build new missions far more quickly than was possible in the past, with greater levels of success, and with costs that were dramatically lower than previous missions. NASA was largely successful in their efforts, but the new missions tended to look very different from the ones in the past. For example, the early missions were big and

complicated with a ton of moving parts, while modern missions have been much smaller in scale with far more focused mission deliverables.

What is NASA?

NASA is the United States National Aeronautical and Space Administration and has been responsible for helping the U.S. achieve success in its space programs, from the moon landing to recent high quality photographs of Pluto. NASA has faced serious budget cuts in recent years, but has been able to retool itself around smaller, more focused missions that cost less and have achieved incredible results.

The same "faster, better, cheaper" challenge is hitting enterprise IT, although even the hardest working IT pros don't usually have to make robots rove the surface of an inhospitable planet! Today's IT departments must meet a growing list of business needs while, at the same time, appeasing the decision makers who demand far more positive economic outcomes (either by cutting costs overall or doing more work within the existing budget).

Unfortunately, most of today's data center architectures actively work against these goals, because with increasing complexity comes increased costs — and things have definitely become more complex. Virtualization has been a fantastic opportunity for companies, but with virtualization has come some new challenges, including major issues with storage. With virtualization, enterprise IT has moved from physical servers, where storage services could be configured on a per-server basis, to shared storage systems. These shared storage systems, while offering plenty of capacity, have often not been able to keep up in terms of performance, forcing IT departments to take corrective actions that don't always align with good economic practices.

For example, it's common for IT pros to add entire shelves of disks, not because they need the capacity, but because they need the spindles to increase overall storage performance. There are, of course, other ways to combat storage performance issues, such as through the use of solid state disk (SSD) caching systems, but these also add complexity to what is already a complex situation.

There are other challenges that administrators of legacy data centers need to consider as well:

- **Hardware sprawl**. Data centers are littered with separate infrastructure silos that are all painstakingly cobbled together to form a complete solution. This hardware sprawl results in a data center that is increasingly complex, decreasing flexibility, and expensive to maintain.

- **Policy sprawl**. The more variety of solutions in the data center, the more touch points that exist when it comes to applying consistent policies across all workloads.

- **Scaling challenges**. Predictability is becoming really important. That is, being able to predict ongoing budgetary costs and how well a solution will perform after purchase are important. Legacy infrastructure and its lack of inherent feature-like scaling capability make both predictability metrics very difficult to achieve.

- **Desire for less technical overhead**. Businesses want analysts and employees that can help drive top line revenue growth. Purely technical staff are often considered expenses that must be minimized. Businesses today are looking for ways to make the IT function easier to manage overall so that they can redeploy technical personnel to more business-facing needs. Legacy data centers are a major hurdle in this transition.

So, with all of this in mind, what are you to do?

Hyperconverged Infrastructure from 30,000 Feet

An emerging data center architectural option, dubbed *hyperconverged infrastructure*, is a new way to reduce your costs and better align enterprise IT with business needs. At its most basic, hyperconverged infrastructure is the conglomeration of the servers and storage devices that comprise the data center. These systems are wrapped in comprehensive and easy-to-use management tools designed to help shield the administrator from much of the underlying architectural complexity.

Why are these two resources, storage and compute, at the core of hyperconverged infrastructure? Simply put, storage has become an incredible challenge for many companies. It's one of— if not *the* — most expensive resources in the data center and often requires a highly skilled person or team to keep it running. Moreover, for many companies, it's a single point of failure. When storage fails, swaths of services are negatively impacted.

Combining storage with compute is in many ways a return to the past, but this time many new technologies have been wrapped around it. Before virtualization and before SANs, many companies ran physical servers with directly attached storage systems, and they tailored these storage systems to meet the unique needs for whatever applications might have been running on the physical servers. The problem with this approach was it created numerous "islands" of storage and compute resources. Virtualization solved this resource-sharing problem but introduced its own problems previously described.

Hyperconverged infrastructure distributes the storage resource among the various nodes that comprise a cluster. Often built using commodity server chasses and hardware, hyperconverged infrastructure nodes and appliances are bound together via Ethernet and a powerful software

layer. The software layer often includes a *virtual storage appliance* (VSA) that runs on each cluster node. Each VSA then communicates with all of the other VSAs in the cluster over an Ethernet link, thus forming a distributed file system across which virtual machines are run.

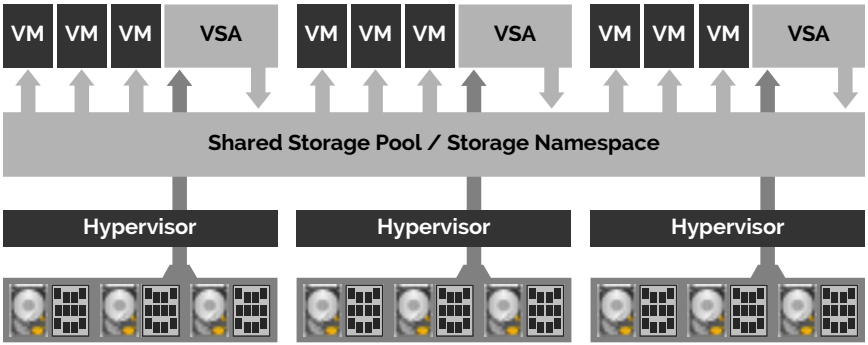

Figure 1-1: An overview of a Virtual Storage Appliance

The fact that these systems leverage commodity hardware is critical. The power behind hyperconverged infrastructure lies in its ability to coral resources – RAM, compute, and data storage – from hardware that doesn't all have to be custom-engineered. This is the basis for hyperconverged infrastructure's ability to scale granularly and the beginnings of cost reduction processes.

The basics behind hyperconverged infrastructure should be well understood before proceeding with the remainder of this book. If you're new to hyperconverged infrastructure or are unfamiliar with the basics, please read *Hyperconverged Infrastructure for Dummies*, available now for free from www.hyperconverged.org.

Resources to Consolidate

The basic combination of storage and servers is a good start, but once one looks beyond the confines of this baseline definition, hyperconverged infrastructure begins to reveal its true power. The more hardware devices and software systems that can be collapsed into a hyperconverged solution, the easier it becomes to manage the solution and the less expensive it becomes to operate.

Here are some data center elements that can be integrated in a hyperconverged infrastructure.

Deduplication Appliances

In order to achieve the most storage capacity, deduplication technologies are common in today's data center. Dedicated appliances are now available which handle complex and CPU-intensive deduplication tasks, ultimately reducing the amount of data that has to be housed on primary storage. Deduplication services are also included with storage arrays in many cases. However, deduplication in both cases is not as comprehensive as it could be. As data moves around the organization, data is rehydrated into its original form and may or may not be reduced via deduplication as it moves between services.

SSD Caches/All-Flash Array

To address storage performance issues, companies sometimes deploy either solid state disk (SSD)-based caching systems or full SSD/flash-based storage arrays. However, both solutions have the potential to increase complexity as well as cost. When server-side PCI-e SSD cards are deployed, there also has to be a third-party software layer that allows them to act as a cache, if that is the desire. With all-flash arrays or flash-based stand-alone caching systems, administrators are asked to support new hardware in addition to everything else in the data center.

Backup Software

Data protection in the form of backup and recovery remains a critical task for IT and is one that's often not meeting organizational needs. Recovery time objectives (RTO) and recovery point objectives (RPO) — both described in the deep dive section entitled "The Ins and Outs of Backup and Recovery" — are both shrinking metrics that IT needs to improve upon. Using traditional hardware and software solutions to meet this need has been increasingly challenging. As RPO and RTO needs get shorter, costs get higher with traditional solutions.

With the right hyperconverged infrastructure solution, the picture changes a bit. In fact, included in some baseline solutions is a comprehensive backup and recovery capability that can enable extremely short RTO windows while also featuring very small RPO metrics.

The Ins & Outs of Backup & Recovery

There are critical recovery metrics – known as Recovery Time Objective (RTO) and Recovery Point Objective (RTO) that must be considered in your data protection plans. You can learn a lot more about these two metrics in Chapter 4.

Data Replication

Data protection is about far more than just backup and recovery. What happens if the primary data center is lost? This is where replicated data comes into play. By making copies of data and replicating that data to remote sites, companies can rest assured that critical data won't be lost.

To enable these data replication services, companies implement a variety of other data center services. For example, to minimize replication impact on bandwidth, companies deploy WAN accel-

eration devices intended to reduce the volume of data traversing the Internet to a secondary site. WAN accelerators are yet another device that needs to be managed, monitored, and maintained. There are acquisition costs to procure these devices; there are costs to operate these devices in the form of staff time and training; and there are annual maintenance costs to make sure that these devices remain supported by the vendor.

2

Ensuring Availability, Data Protection and Disaster Recovery

Even the smallest of small businesses today depend on their IT resources being available on a 24/7 basis. Even short periods of downtime can wreak havoc, impact the bottom line, and mean having to cancel going out to lunch. Maintaining an agreed-upon level of infrastructure availability is critically important. On top of that, outages or other events resulting in loss of data can be a death knell for the business. Many businesses that suffer major data loss fail to recover in the long-term and eventually make their way down the drain. Data protection is one of IT's core services. Unfortunately, it's also a hard service to provide at times, or at least, it was. There are now some hyperconverged infrastructure solutions that are uniquely positioned to solve, once and for all, the challenges across the entire data protection spectrum.

The Ins & Outs of Backup & Recovery

There are two primary metrics to consider when it comes to disaster recovery.

Recovery Point Objective (RPO)

If you're using a nightly backup system, you're implicitly adhering to a *24-hour Recovery Point Objective (RPO)*. You're basically saying that losing up to 24 hours worth of data is acceptable to the business. RPO is the metric that defines how much data your organization is willing to lose in the event of a failure that has the potential to result in data loss. To reduce RPO, you need to back data up more often.

Recovery Time Objective (RTO)

RPO is critically important as it defines just how much data you're willing to lose. Once you've suffered a data loss, the critical metric shifts. Now, you're more interested in how long it takes you to recover from that failure. How long is your organization willing to be without data while you work to recover it from backup systems? This metric is often used to support such statements as, "For every minute we're down, the company loses $X."

The Recovery Time Objective (RTO) is the formal name for this metric and is one that companies will go to great lengths to minimize. As is the case with RPO, the closer to zero that you attempt to get to RTO — that is, the less time that you're willing to be down — the more it costs to support.

To achieve very low RTO values, companies will often implement multi-pronged solutions, such as disaster recovery sites, fault tolerant virtual machines, clustered systems, and more.

The Data Protection and Disaster Recovery Spectrum

Let's talk a bit about data protection as a whole. When you really look at it, data protection is a spectrum of features and services. If you assume that data protection means "ensuring that data is available when it's needed," the spectrum also includes high availability for individual workloads. **Figure 2-1** provides you with a look at this spectrum.

Figure 2-1: The Data Protection Spectrum

RAID

Yes, RAID is a part of your availability strategy, but it's also a big part of your data protection strategy. IT pros have been using RAID for decades. For the most part, it has proven to be very reliable and has enabled companies to deploy servers without much fear of negative consequences in the event of a hard drive or two failing. Over the years, companies have changed their default RAID levels as the business needs have changed, but the fact is that RAID remains a key component in even the most modern arrays.

The RAID level you choose is really important, but you shouldn't have to worry about it. The solution should do it for you. That said, don't forget that it's pretty well-known that today's really large hard drives have made traditional RAID systems really tough to support. When drives fail in a traditional RAID array, it can take hours or even days to fully rebuild that drive. Don't forget this as you read on; we'll be back to this shortly.

RAID is also leveraged in some hyperconverged infrastructure systems; however, with these systems, administrators are shielded from some of the complexity and configuration options that they used to work with on stand-alone storage arrays. Bear in mind that one of the tenets of hyperconverged infrastructure is simplicity. As such, you don't have to go through a lot of effort to manage RAID in a hyperconverged system. It's simply leveraged behind the scenes by the system itself. In **Figure 2-2**, you get a look at how RAID protects data.

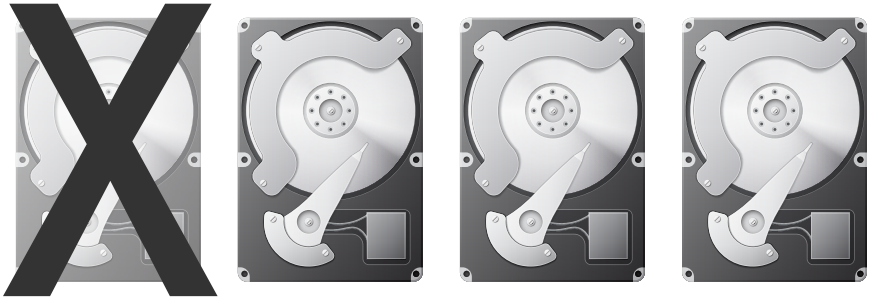

Figure 2-2: Key Takeaway: On the data protection spectrum, RAID helps you survive the loss of a drive or two.

Replication/RAIN/Disaster Recovery

RAID means you can lose a drive and still continue to operate, but what happens if you happen to lose an entire node in a hyperconverged infrastructure cluster? That's where replication jumps in to save the day. Many hyperconverged infrastructure solutions on the market leverage replication as a method for ensuring ongoing availability and data protection in the event that something takes down a node, such as a hardware failure or an administrator accidentally pulling the wrong power cord.

This is possible because *replication* means "making multiple copies of data and storing them on different nodes in the cluster."

Therefore, if a node is wiped off the face of the earth, there are one or more copies of that data stored on other cluster nodes.

Two kinds of replication

There are two different kinds of replication to keep in mind. One is called *local* and the other is called *remote*. Local replication generally serves to maintain availability in the event of a hardware failure. Data is replicated to other nodes in the cluster in the same data center. Remote replication is leveraged in more robust disaster recovery scenarios and enables organizations to withstand the loss of an entire site.

In some hyperconverged infrastructure solutions, like those shown in **Figure 2-3**, you can configure what is known as the *replication factor* (RF). The replication factor is just a fancy way of telling the system how many copies of your data you'd like to have. For example, if you specify a replication factor of 3 (RF3), there will be 3 copies of your data created and stored across disparate nodes. You will sometimes see replication-based availability mechanisms referred to as RAIN, which stands for Redundant Arrays of Independent Nodes.

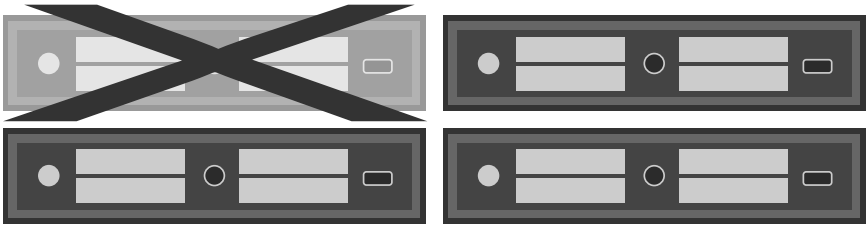

Figure 2-3: Lost a node? Can't find it? Don't worry! Replication will save the day!

Besides helping you to make sure that your services remain available, replication goes way beyond just allowing you with withstand the loss of a node, too. When replication is taken beyond the data center to other sites, you suddenly gain disaster recovery capability, too. In fact, in some hyperconverged systems that leverage inline deduplication across primary and secondary storage tiers, that's exactly what happens. After deduplication, data is replicated to other nodes and to other data centers, forming the basis for incredibly efficient availability and disaster recovery.

How About Both – RAID and RAIN Combined

Let's go a little deeper into the RAID/RAIN discussion with an eye on hyperconverged infrastructure solutions that provide both. First, there are some downsides to just RAIN-based replication (Replication Factor 2 or RF2). There are solutions on the market that provide RF2. Systems based on RF2 will lose data if any two nodes or disks in a cluster fail, or if even just one node should fail while any other node is down for maintenance.

To make things a bit more resilient, you could bump up to RF3, but this replication factor then requires a minimum of five nodes at each site that uses RF3 and imposes an additional 50% penalty on capacity. With RF3, you can also start to think about using erasure coding, but this requires RF3 and carries with it a lot of CPU overhead due to the way that erasure coding works. This may not be suitable when trying to support high-performance applications.

How about combining RAID and RAIN into a single solution? Maybe you combine the use of local RAID 6 on individual nodes so that *any* node can tolerate double disk failures and keep virtual machines up and running. With each individual node very well protected, the likelihood of losing an entire node is reduced. From there, you apply RAIN so that, in the event that a complete node

is lost, you can tolerate that, too. The strategic combination of RAID and RAIN enables tolerance against a broad set of failure scenarios.

What is Erasure Coding?

Erasure coding is usually specified in an N+M format: 10+6, a common choice, means that data and erasure codes are spread over 16 (N+M) drives, and that any 10 of those can recover data. That means any six drives can fail. If the drives are on different appliances, the protection includes appliance failures, so six appliance boxes could go down without stopping operations.

Courtesy: www.networkcomputing.com/storage/raid-vs-erasure-coding/a/d-id/1297229

Backup and Recovery

Despite your best efforts, there will probably come a day when you need to recover data lost from production. Data losses can happen for a variety of reasons:

- **Human error** — People make mistakes. Users accidentally delete files. Administrators accidently delete virtual machines. IT pros can sometimes accidentally pull the wrong disk from a storage system or unplug the wrong server's power cord.

- **Hardware failure** — When hardware fails, sometimes it fails spectacularly. In fact, hardware failure may not even be the result of failed IT hardware. You may end up in a situation, for example, in which the data center cooling systems fail and server automatically shuts down as the temperature rises. This could be considered a server hardware failure because of the outcome (the server going down), when in fact the server is actually doing exactly what it's supposed to do in this case.

- **Disasters** — Hurricanes, tornados, floods, a new Terminator movie. Disasters come in all kinds of forms and can result in data loss.

The HPE SimpliVity story on protecting production data and availability

by Brian Knudston

Being a hyperconvergence platform, HPE SimpliVity first provides the compute and storage infrastructure for customer's production applications. As data is ingested from the hypervisor, we stage the VM (virtual machine) data into DRAM on the HPE OmniStack Accelerator Card across two of our nodes within a single data center. With data now protected across multiple nodes, in addition to supercapacitor and flash storage protecting the DRAM on each HPE OmniStack Accelerator Card, we acknowledge a successful write back to the VM and process the data for deduplication, compression and optimization to permanent storage on the Hard Disk Drives (HDDs) on both nodes. Once this process is complete, every VM in a HPE SimpliVity data center can survive the loss of at least two HDDs in every node, in a data center AND the loss of a full HPE SimpliVity node.

Disaster Recovery

Disaster recovery takes backup one step further than the basics. Whereas *backup* and *recovery* are terms that generally refer to backing up data and, when something happens, recovering that data, *disaster recovery* instead focuses on recovery beyond just the data.

Disaster recovery demands that you think about the eventual needs by starting at the end and working backward. For example, if your data center is hit by an errant meteor (and assuming that this meteor has not also caused the extinction of the human race) recovering

your data alone will be insufficient. You won't have anything onto which to recover your data if your data center is obliterated.

Before we get too fatalistic, let's understand what the word *disaster* really means in the context of the data center. It's actually kind of an unfortunate term since it immediately brings to mind extinction-level events, but this is not always the case for disaster recovery.

There are really two kinds of disasters on which you need to focus:

- **Micro-level disasters** — These are the kinds of events that are relatively common, such as losing a server or portion of a data center. In general, you can quickly recover in the same data center and keep on processing. Often, recovery from these kinds of disasters can be achieved through backup and recovery tools. With that said, these events will probably still result in downtime.

- **Macro-level disasters** — These are the kind of life-altering events that keep IT pros awake at night and include things like fires, acts of {insert deity here}, or rampaging hippos. Recovery from these disasters will mean much more than just restoring data.

Business Continuity

Since *disaster recovery* is kind of a loaded term, a lot of people prefer to think about the disaster recovery process as "business continuity" instead. However, that's not all that accurate. Business continuity is about all the aspects to a business continuing after a disaster. For example, where are the tellers going to report after the fire? How are the phone lines going to be routed? Disaster recovery is an IT plan that is a part of business continuity.

Thinking about the disaster recovery process with the end in mind requires that you think about what it would take to have everything back up and running — hardware, software, and data — before disaster strikes.

Yes, your existing backup and recovery tools probably play a big role in your disaster recovery plan, but that's only the very beginning of the process.

Disaster recovery plans also need to include, at a bare minimum:

- **Alternate physical locations** — If your primary site is gone, you need to have other locations at which your people can work.

- **Secondary data centers** — In these locations, or in the cloud, you need to have data centers that can handle the designated workloads from the original site. This includes a space for the hardware, the hardware itself, and all of the software necessary to run the workloads.

- **Ongoing replication** — In some way, the data from your primary site needs to make its way to your secondary site. This is a process that needs to happen as often as possible in order to achieve desirable RTOs and RPOs. In an ideal world, you would have systems in place that can replicate data in minutes after it has been handled in the primary data center. The right hyperconverged infrastructure solution can help you to achieve these time goals.

- **Post-disaster recovery processes** — Getting a virtual machine back up and running is just the very first step in a disaster recovery process. RTO is a measure of more than just the restoration of the VM. From there, processes need to kick off that include all the steps required to get the application

and data available to the end user. These processes include IP address changes, DNS updates, re-establishment of communication paths between parts of an n-tier application stack and other non-infrastructure items.

HPE SimpliVity's answer to full spectrum DR

by Brian Knudston

HPE SimpliVity alone makes it simple for you to achieve the first part of disaster recovery, which is making sure that virtual machines are always available, even if a data center is lost. Hewlett Packard Enterprise has made a focus to provide integration into other tools that can help automate and orchestrate all of the remaining steps of the disaster recovery process, including pre-built packages of HPE SimpliVity functionality within VMware's vRealize Automation and Cisco's UCS Director, and supporting partners in the development of tools on top of HPE SimpliVity APIs like VM2020's EZ-DR.

Data Reduction in the World of Data Protection

We're going to be talking a lot about data reduction – deduplication and compression – in this book. They're a huge part of the hyperconverged infrastructure value proposition and, when done right, can help IT to address problems far more comprehensively than when it's done piecemeal.

When it comes to data protection, data reduction can be really important, especially if that data reduction survives across different time periods – production, backup, and disaster recovery. If that data can stay reduced and deduplicated, some very cool possibilities emerge. The sidebar below highlights one such solution.

The Data Virtualization Platform and disaster recovery

by Brian Knudston

To protect data at specific instances of time, Hewlett Packard Enterprise designed backup and restoration operations directly into the DNA of the HPE OmniStack Data Virtualization Platform, enabled by our ability to dedupe, compress and optimize all the VM data. This results in backups and restores that can be taken in seconds, which can help reduce Recovery Point Objectives (RPOs) and Recovery Time Objectives (RTOs), while consuming almost no IOPS off the HDDs.

When protecting data across datacenters, HPE SimpliVity maintains awareness of data deduplication across the different sites. If a VM is configured to backup to a remote data center, the receiving data center determines which unique blocks need to be transported across the WAN and the sending data center only sends those unique blocks. This drastically reduces the WAN bandwidth necessary between sites, increasing the frequency of backups to remote sites and eliminate IOPS by reducing the amount of data that needs to be read from and written to the HDDs.

Fault Tolerance

Availability is very much a direct result of the kinds of fault tolerance features built into the infrastructure as a whole. Data center administrators have traditionally taken a lot of steps to achieve availability, with each step intended to reduce the risk of a fault in various areas of the infrastructure. These can include:

- **Using RAID** — As previously mentioned, RAID allows you to experience drive failures within a hyperconverged node and keep operating.

Simplified Storage Systems

Bear in mind that RAID, and storage in general, becomes far simpler to manage in a hyperconverged infrastructure scenario. There is no more SAN and, in most cases, RAID configuration is an "under the hood" element that you don't need to worry about. This is one less component that you have to worry about in your data center.

- **Redundant power supplies** — Extra power supplies are, indeed, a part of your availability strategy, because they allow you to experience a fault with your power system and still keep servers operating.

- **Multiple network adapters** — Even network devices can fail, and when they do, communications between servers and users and between servers and other servers can be lost. Unless you have deployed multiple switches into your environment and multiple network adapters into your servers, you can't survive a network fault. Network redundancy helps you make your environment resilient to network-related outages.

- **Virtualization layer** — The virtualization layer includes its own fault tolerance mechanisms, some of which are transparent and others require a quick reboot. For example, VMware's High Availability (HA) service continuously monitors all of your vSphere hosts. If one fails, workloads are automatically restarted on another node. There is some downtime, but it's minimal. In addition to HA, VMware makes available a Fault Tolerance (FT) feature. With FT, you actually run multiple virtual machines. One is the production system and the second is a live shadow VM that springs into action in the event that the production system becomes unavailable. However, with all that said, there are some limitations inherent in hypervisor-based fault tolerance technology, described in the sidebar

entitled *Fault Tolerance Improvements in vSphere 6*. This is why some hyperconverged infrastructure vendors eschew hypervisor-based fault tolerance mechanisms in favor of building their own, more robust solutions.

Fault Tolerance Improvements in vSphere 6

Frankly, Fault Tolerance (FT) in vSphere has been all but useless, except for the smallest virtual machines. Here's an excerpt from VMware's documentation explaining the limitations of FT: "Only virtual machines with a single vCPU are compatible with Fault Tolerance." This limitation is one of the many items that holds back FT from being truly usable across the board. vSphere 6 increases Fault Tolerance capabilities to virtual machines with up to 4 vCPUs. This is still a significant limitation when you consider than many VMs are deployed with 8 vCPUs or more, particularly for large workloads.

End Results: High Availability, Architectural Resiliency, Data Protection, and Disaster Recovery

No one wants downtime. It's expensive and stressful. Most organizations would be thrilled if IT could guarantee that there would be no more downtime ever again. Of course, there is no way to absolutely guarantee that availability will always be 100%, but organizations do strive to hit 99% to 99.999% availability as much as possible.

High availability is really the result of a combination of capabilities in an environment. In order to enable a highly available application environment, you need to have individual nodes that can continue to

work even if certain hardware components fail and you need to have a cluster that can continue to operate even if one of the member nodes bites it.

Hyperconverged infrastructure helps you to achieve your availability and data protection goals in a number of different ways. First, the linear scale-out nature of hyperconverged infrastructure (i.e., as you add nodes, you add all resources, including compute, storage, and RAM), means that you can withstand the loss of a node because data is replicated across multiple nodes with RAIN. Plus, for some hyperconverged solutions, internal use of RAID means that you can withstand the loss of a drive or two in a single node. With the combination of RAIN+RAID providing the most comprehensive disaster recovery capabilities, you can withstand the loss of an entire data center and keep on operating with little to no loss of data.

As you research hyperconverged infrastructure solutions, it's important to make sure that you ask a lot of questions about how vendors provide availability and data protection in their products. The answers to these questions will make or break your purchase.

About the Author

Scott D. Lowe, vExpert

Scott Lowe is a vExpert and partner and Co-Founder of ActualTech Media. Scott has been in the IT field for close to twenty years and spent ten of those years in filling the CIO role for various organizations. Scott has written thousands of articles and blog postings and regularly contributes to *www.EnterpriseStorageGuide.com* & www.*ActualTech.io*.

About the Editor

David M. Davis, vExpert

David Davis is a partner and co-founder of ActualTech Media. With over 20 years in enterprise technology, he has served as an IT Manager and has authored hundreds of papers, ebooks, and video training courses. He's a 6 x vExpert, VCP, VCAP, & CCIE# 9369. You'll find his vSphere video training at *www.Pluralsight.com* and he blogs at *www.VirtualizationSoftware.com* and *www.ActualTech.io*.

About ActualTech Media

ActualTech Media provides enterprise IT decision makers with the information they need to make informed, strategic decisions as they modernize and optimize their IT operations.

Leading 3rd party IT industry influencers Scott D. Lowe, David M. Davis and special technical partners cover hot topics from the software-defined data center to hyperconvergence and virtualization.

Cutting through the hype, noise and claims around new data center technologies isn't easy, but ActualTech Media helps find the signal in the noise. Analysis, authorship and events produced by ActualTech Media provide an essential piece of the technology evaluation puzzle.

More information available at **www.actualtechmedia.com**

Hewlett Packard
Enterprise

About Hewlett Packard Enterprise

Hewlett Packard Enterprise (HPE) is an industry leading technology company with the industry's most comprehensive portfolio, spanning the cloud to the data center to workplace applications. HPE technology and services help customers around the world make IT more efficient, more productive and more secure.

Early in 2017, the company acquired SimpliVity and now offers HPE SimpliVity hyperconverged systems, complete hardware-software solutions that are designed, built, and supported by HPE.

Visit *www.hpe.com*

Gorilla Guide Features

In the Book

These help point readers to other places in the book where a concept is explored in more depth.

The How-To Corner

These will help you master the specific tasks that it takes to be proficient in the technology jungle.

Food For Thought

In the these sections, readers are served tasty morsels of important information to help you expand your thinking .

School House

This is a special place where readers can learn a bit more about ancillary topics presented in the book.

Bright Idea

When we have a great thought, we express them through a series of grunts in the Bright Idea section.

Dive Deep

Takes readers into the deep, dark depths of a particular topic.

Executive Corner

Discusses items of strategic interest to business leaders.