THE GORILLA GUIDE TO...™

Hyperconverged Infrastructure for Data Center Consolidation

Scott D. Lowe
Consultant & Industry Veteran

HELPING YOU NAVIGATE THE TECHNOLOGY JUNGLE

Hyperconverged Infrastructure for Data Center Consolidation

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

ActualTech Media

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.
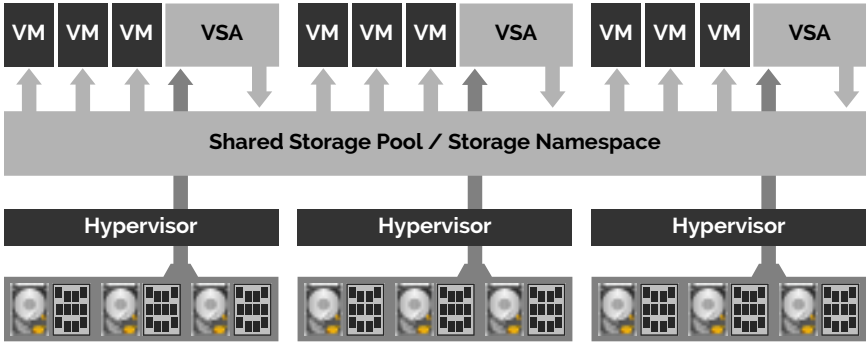

Figure 1-1: An overview of a Virtual Storage Appliance

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

Data Center Consolidation

Mergers and acquisitions. Cost cutting. New business initiatives. There are all kinds of reasons why companies make the decision to consolidate data centers. Maybe your company undergoes explosive, barely controlled growth or they may decide to stop, take a pause, and reconsider how IT does business in their organization. Or maybe your company decides to buy out another company, and you suddenly inherit a whole series of data centers that you're not prepared to handle.

Want to know a secret? It will be up to you to figure it out. Furthermore, you'll probably be asked to do it with the same budget you already have.

Here's the thing, though. Data center consolidation isn't always just about reducing the number of data centers from a big number to a smaller number. Sometimes, it's about reducing the amount of stuff strewn about the data centers that you already have.

Today's IT organizations generally buy and integrate numerous point solutions, from a plethora of vendors, each with its own training courses, licensing, refresh cycles, and mode of operation. These point products are the result of years of planning and investments to support business applications.

We hear the same story time and time again. Does this sound familiar? You virtualized about 5 to 8 years ago and naturally a data protection strategy project came directly after that. The SharePoint implementation project for your marketing organization took eons to complete and required purchasing a new SAN. It feels like you just bought that SAN yesterday, but you blinked, three years flew by, and it's time to refresh . . . again. The decision to buy all of these products made sense at the time, but today data center complexity can feel overwhelming and discourage innovation.

All of this has really cooked up several challenges for IT organizations, including:

- Time overhead spent on operational tasks

- Mobility and management of virtual machines (VMs)

- Budget constraints

- Breaching service level agreements

- Operational efficiency

- Application performance

The list goes on and on.

Ask yourself these questions to determine where you might have pain points:

- As data growth explodes, can you continue to operate the same way you always have?

- Are your legacy technologies designed for virtualized environments?

- How much time are you spending on maintenance, upgrades, deployments, provisioning and management tasks instead of building more valuable innovation for the business?

- Do you have the necessary expertise to manage each of these products separately?

- Are you struggling to meet your Service Level Agreements (SLAs) with the business?

- Are you missing data protection objectives like Recovery Point Objectives (RPOs) and Recovery Time Objectives (RTOs)?

If you answered "No" to one or more of these questions, you might be a candidate for hyperconverged infrastructure.

Consolidation with Hyperconvergence

In every chapter of this book so far, you've learned about how hyperconverged infrastructure solutions can reduce the variety of hardware and software you have to manage in the data center. Every time you eliminate a class of hardware or software in your data centers, you are on your way to answering "Yes" to all of the questions outlined in the previous section.

That's the ideal scenario.

At the most basic level, hyperconverged infrastructure consolidates storage and compute, enabling you to eliminate the monolithic SAN environment. From there, some hyperconverged infrastructure

vendors make things pretty interesting. For example, Nutanix and Hewlett Packard Enterprise both provide something in the way of data reduction via deduplication and compression. HPE SimpliVity, however, takes this to the extreme through the use of their accelerator card, which forms the basis for what they call their HPE OmniStack Data Virtualization Platform.

By enabling global inline deduplication and compression with a solution like HPE SimpliVity, you suddenly need less overall capacity, which means you need less overall hardware. With constant data reduction, you no longer need:

- WAN accelerators to reduce data over the wide area network because data stays reduced

- Separate backup software to protect the data in your environment

- Separate deduplication appliances

- SSD arrays

Instead, you can massively reduce the amount of hardware and software that you're operating, maintaining, getting trained on, and, maybe even struggling with. With less stuff to manage and worry about, you can better focus on the business and on improving SLAs, RTOs, and RPOs. You get to focus on the business rather than on the technology.

If you're in a situation in which you need to reduce the number of data centers you're managing, hyperconverged infrastructure can help

you there, too. How? For the same reasons that we just discussed. Rather than just taking all of the hardware from the various sites and combining it all into one supersite, you can rethink the whole model. In addition to cutting down physical locations, you can also minimize complexity.

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.
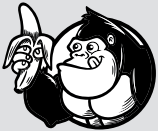
Visit *www.hpe.com*

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.