HPE PRESENTS

The Gorilla Guide to...®

FOURTH EDITION

Hyperconverged Infrastructure Economics and Impact on the Budget

Scott D. Lowe

INSIDE THE GUIDE:

- Discover the basics of hyperconverged infrastructure as a technology
- Identify the budget areas that require analysis after deployment of a hyperconverged infrastructure solution
- Understand how hyperconvergence enables a renewed focus on the business rather than technology

Helping You Navigate The Technology Jungle!

ActualTech Media
www.actualtechmedia.com

In Partnership With

Hewlett Packard Enterprise

THE GORILLA GUIDE TO...

Hyperconverged Infrastructure Economics and Impact on the Budget

By Scott D. Lowe

ACTUALTECH MEDIA

6650 Rivers Ave Ste 105 #22489

North Charleston, SC 29406-4829

www.actualtechmedia.com

PUBLISHER'S ACKNOWLEDGEMENTS

EDITOR
Keith Ward, ActualTech Media

PROJECT MANAGER
Wendy Hernandez, ActualTech Media

EXECUTIVE EDITOR
James Green, ActualTech Media

LAYOUT AND DESIGN
Olivia Thomson, ActualTech Media

WITH SPECIAL CONTRIBUTIONS FROM
Charles Wood

ENTERING THE JUNGLE

CALLOUTS USED IN THIS BOOK

The Gorilla is the professorial sort that enjoys helping people learn. In the School House callout, you'll gain insight into topics that may be outside the main subject but are still important.

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

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

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

Discusses items of strategic interest to business leaders.

ICONS USED IN THIS BOOK

DEFINITION

Defines a word, phrase, or concept.

KNOWLEDGE CHECK

Tests your knowledge of what you've read.

PAY ATTENTION

We want to make sure you see this!

GPS

We'll help you navigate your knowledge to the right place.

WATCH OUT!

Make sure you read this so you don't make a critical error!

CHAPTER 1

Introduction to Hyperconverged Infrastructure

The world of technology is changing at both a broader and faster pace than ever before. In years past, IT practitioners and decision makers might have had to deal with a few new hardware options and some software upgrades. Today, changes come every day as companies break the data center walls in favor of hybrid cloud, as cloud providers continue to grow, as software updates deploy silently behind the scenes, and as entire new architectures come and go.

As employees devour technologies such as smartphones, tablets, wearables, and other devices, and as they become more comfortable with solutions such as Slack, Teams, Dropbox, and Zoom, their demands on enterprise IT intensify. On top of this, management and other decision makers are also increasing their demands on enterprise IT to provide application environments that have higher levels of availability and easier growth capability, but with the kinds of pay-as-you-grow economics that they see from the cloud. Unfortunately, enterprise IT organizations often don't see much, if any, associated increases in funding to accommodate these demands, particularly as spending increases on other business priorities, such as improved security, analytics, remote work, and digital transformation efforts.

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 its 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.

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 quickly 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. Moreover, organizations find themselves increasingly turning to technology-centric solutions to solve real-world crises, placing further strain on IT resources.

Even as the public cloud continues to gain favor for certain workloads, the real center of workload action remains the on-premises data center. Unfortunately, traditional data center architectures actively work against modern goals, because with increasing complexity comes increased costs—and things have definitely become more complex. Virtualization was a fantastic opportunity for companies, but with virtualization came new challenges, including major issues with storage. With virtualization, enterprise IT 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.

Further, legacy data center infrastructure was often not very intelligent. Once appropriately cobbled together, the IT department was able to provide a generally suitable workload operating environment, but it was sometimes operated via the *brute force* administrative method of bashing on keyboards and dragging mice to perform a multitude of

manual operations. These legacy environments didn't have the kinds of software-centric, data-driven intelligence that now enables autonomous operation.

Here are some other challenges legacy data center admins 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's increasingly complex, decreasingly flexible, 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 extremely 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.
- **A focus on security.** Security has always been important, but never more than it is today. Increasingly a boardroom issue, security spending is accelerating, sometimes at the expense of other areas of IT, a situation that data center architects need to accommodate.
- **Taking efficiency to the next level.** Businesses today want an IT function focused on business workloads, not one that has to

constantly wrangle storage into submission with esoteric scripts. Next-level efficiency includes imbuing infrastructure with an AI-led, data-driven operational framework that bludgeons the complexity out of the environment to enable faster results to help accelerate business initiatives.

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

Hyperconverged Infrastructure from 30,000 Feet

A number of years ago, a new data center architectural option, dubbed *hyperconverged infrastructure*, came on the scene and exploded faster than anyone could have imagined. Hyperconvergence is a 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 with modern iterations of the technology. This modern technology includes intelligence features that streamline overall management and bring in aspects of AIOps.

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, while providing an infrastructure that's self-managing.

Why are storage and compute at the core of hyperconverged infrastructure? Simply put, storage had become an incredible challenge for many companies. It's traditionally been one of—if not the—most expensive resources in the data center, often requiring a highly skilled person or team to keep it running.

Moreover, for many companies, storage became a single point of failure. When it failed, swaths of services were negatively impacted.

Finally, with the increase in the growth of data volume, legacy storage architectures were beginning to crumble. This data explosion

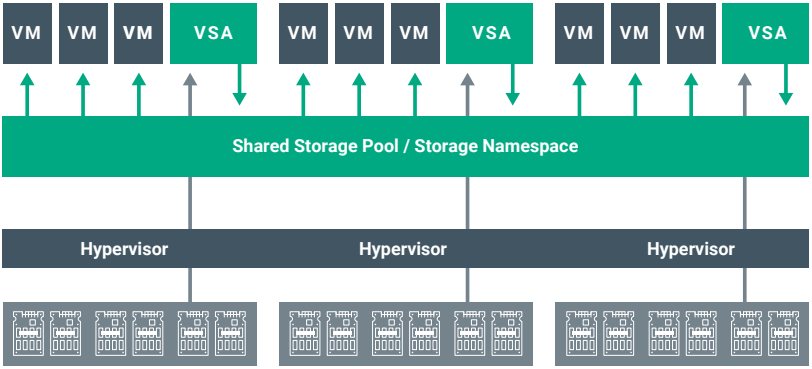

Figure 1-1: An overview of a Virtual Storage Appliance

overwhelmed systems that simply weren't designed to handle the flood of information coming their way.

Combining storage with compute is in many ways a return to the past, but this time serious brains have been wrapped around it. Before virtualization and before SANs, many companies ran physical servers with direct-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 without a centralized coordinating mechanism. Virtualization solved this resource sharing problem, but introduced its own problems, which were discussed earlier.

Hyperconverged infrastructure distributes the storage resource among the various nodes that comprise a cluster. Built using standard 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 communicates with all of the other VSAs in the cluster over an Ethernet link, thus forming a distributed file system across which virtual machines (VMs) are run.

The fact that hyperconverged systems leverage standard off-the-shelf hardware is critical. The power behind hyperconverged infrastructure

lies in its ability to corral resources—RAM, compute, and data storage—from hardware that doesn't have to be custom-engineered. This is the basis for hyperconverged infrastructure's ability to scale granularly, work more efficiently and save money.

Resources to Consolidate

The basic combination of storage, servers, and a hypervisor is the undergirding of all hyperconverged infrastructure. And the more hardware devices and software systems that can be collapsed into a hyperconverged solution, the easier it becomes to manage and less expensive to operate.

Here are some data center elements that can be integrated in a hyperconverged infrastructure. Often, these simply become invisible features of the solution and don't even require separate management.

Deduplication Appliances

In order to achieve the most storage capacity, deduplication technologies are common in today's data center. Standalone storage solutions and dedicated appliances are available that handle complex and CPU-intensive deduplication tasks, ultimately reducing the amount of data that has to be housed on primary storage.

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.hpe.com/HCI/fordummies

But even in these cases, deduplication isn't always 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. In essence, you may end up in a scenario in which data is being constantly reduced and rehydrated as it meanders around the environment. It's a terribly inefficient use of resources that could be better targeted to support more workloads.

SSD Caches, Hybrid Storage, and All-Flash Arrays

To address storage performance issues, companies increasingly deploy either solid-state disk (SSD)-based caching systems, hybrid storage arrays, or full SSD/flash-based storage arrays. However, all of these 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's 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 service provided by 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, "The Ins and Outs of Backup and Recovery"—are both shrinking metrics upon which IT needs to improve.

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, some baseline solutions include comprehensive

backup and recovery capability that can enable extremely short RTO windows, enabling very small RPO metrics—both *very* good characteristics!

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 replication 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 in traditional scenarios, companies implement a variety of other data center services. For example, to minimize replication impact on bandwidth, companies deploy WAN-acceleration 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 operation costs 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.

The Ins and Outs of Backup and 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 **The Gorilla Guide To...® Hyperconverged Infrastructure Implementation Strategies**.

Up Next

With an understanding of hyperconverged infrastructure and knowledge about many of the resources that can be consolidated into such solutions, let's move on to discuss the ramifications that the technology may have on your organization's IT staff and budget.

CHAPTER 2

Hyperconvergence Economics: The Impact on the IT Budget

Hyperconverged infrastructure has the potential to transform more than just the data center. By unlocking staff time and other resources, hyperconvergence can help your organization transform IT from a "keeping the lights on" cost center to a top-line revenue driver.

Focus on the Business, Not the Tech

When the technology becomes too complex for IT to fully manage, or constantly requires the addition of new IT staff and skills, the focus shifts from the business to the infrastructure. In other words, people focus on building the infrastructure itself instead of enhancing what the infrastructure can do for the business. That can lead to a long-term problem and is one of the primary reasons that you'll start seeing misalignment.

The goal for most organizations must be to reduce the amount of "technical overhead." Just like other areas of the business, you reduce overhead to lower costs. You can do the same thing in IT with its own version of overhead. Reduction activities include simplifying administration, improving utilization of existing assets, and limiting the staff time spent on the care of feeding the existing infrastructure.

The result is often reduced budgetary needs, and it can also mean that the company is better able to seize new business opportunities as they

arise. That alone can have a dramatic positive impact on an organization's finances and the perception of IT's value to the business.

Where Do the Savings Emerge?

This next question revolves around the actual source of savings. As you evaluate hyperconverged infrastructure solutions for yourself, these are some of the areas in which you'll need to focus to properly calculate cost and expense differential between your traditional environment and the hyperconverged environment.

- **No need to separate workloads.** By having the ability to consolidate workload silos into fewer or even just one silo, you can more easily increase overall utilization of what you're running. This harkens back to the early days of server virtualization, which carried increased utilization as a key driver for adoption.
- **Coexistence with existing systems.** Although this isn't possible for all hyperconverged infrastructure solutions, the savings can be significant for those that can, in fact, coexist with existing infrastructure. In this scenario, you can more easily continue to use some existing systems (generally, existing vSphere hosts) in conjunction with the new hyperconverged environment.
- **Reduction in electrical and cooling.** If you eliminate a bunch of servers, a slew of hard drives, and the accompanying infrastructure (such as WAN accelerators and SSD caches) from your data center, you will massively reduce your electrical and cooling costs. Less equipment translates to less power. Fewer moving parts means less generated heat, which leads to lower cooling costs. Whether or not you're a fan of the Green New Deal, it's easy to agree that reducing electrical costs, and therefore emissions, is a laudable goal.
- **Don't discount these potential savings.** They can be significant and are direct reductions on the operational side of the budget.

- **Reduced maintenance contracts.** You probably have annual maintenance contracts on everything in your data center. What would be the impact if you had fewer such contracts because you no longer required those hardware or software components in your data center? The likely outcome is that you'd be saving money in the operational budget. As has been stated throughout this book, hyperconverged infrastructure has the potential to transform IT operations, including the operational budget.

- **Get those digital transformation efforts off the ground.** Less "tech" to manage means that the same or even fewer staff can manage it. There are fewer discrete skill sets needed in a hyperconverged infrastructure, meaning you can begin to redirect efforts toward goals that are more directly business-facing.

- **Training.** Top-down management (workload) vs. bottom-up (component) means less training will be required. Consider an iPhone. The iPhone handily replaces several distinct devices. But with the iPhone, there's only one interface to learn vs. the component-level management of a GPS, music player, camera, video recorder, timepiece, and so on.

- **Operational and capital expense improvements.** You've seen a number of ways that operational efficiency can be improved, which reduces operational expense, but hyperconverged infrastructure can also have a dramatic impact on capital budgets. A lower cost of acquisition is just the beginning. As was mentioned previously in this book, you can implement a "rolling upgrade" paradigm. This can even out many of the spikes and valleys inherent in many capital budgets and enable easier scaling, which doesn't rely on predetermined financial schedules.

Initial Investment Analysis

The initial investment in a solution is often one of the only financial elements considered by many organizations. That's because it's really important. They need to know how much a solution is going to cost right now. However, there are a ton of flaws inherent in this myopic model. First, you don't get a good feel for what the TCO will be for the solution. Forrester recently revealed that hyperconverged infrastructure can deliver a 192% ROI and a payback period as short as 7 months.[1]

Furthermore, many traditional procurement models aren't even looking at what the business needs today. Instead, because of the way that most organizations have established budgets, they're buying the solutions that they need three or even five years from today. In other words, you're buying resources you'll have to grow into, not what you actually need immediately.

Let's examine this in more detail. In the **Figure 2-1**, you'll see two lines. The flat line is the purchased resource capacity at the inception of the

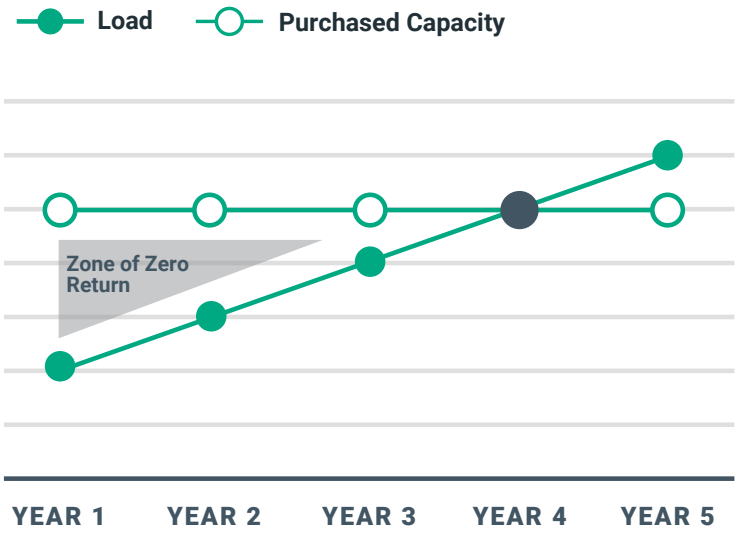

Figure 2-1: Load vs. purchased capacity in a traditional data center infrastructure

[1] https://www.hpe.com/us/en/resources/integrated-systems/simplivity-economic-impact.html

CapEx vs. OpEx Financial Models

There are two kinds of expenses that you need to keep in mind when you consider data center economics:

Capital expenditures (CapEx)—These align pretty closely with the initial cost of a solution and are often considered the one-time costs associated with a purchase.

Operational expenditures (OpEx)—An operational expenditure (OpEx) financial model is where you consume IT services on your own terms, and only pay for what you use.

HPE GreenLake for HPE SimpliVity is a service that delivers IT outcomes with hardware, software, and expertise on-premises in a pay-per-use model. Start with the capacity you need today and grow with business demand, leveraging active capacity management, variable monthly payments based on metered usage, and help eliminate overprovisioning. Included services help to remove the heavy lifting in IT, freeing up IT staff to focus on delivering business value.

current replacement cycle. The sloping line is the actual resource need in the data center. The shaded area is a zero ROI zone.

For the first few years of this solution, there's massive waste in resource and budget. IT has to buy this way today because many data center solutions don't easily lend themselves to incremental scale. Instead, IT staff have to attempt to project the data center's needs for the next three to five years, and they won't always get it right. This is not a knock on IT pros; after all, stock market gurus often fail to predict markets, too. It's just the way things are based on the tools we have today.

1T = One-time
OG = Ongoing

	Traditional		Hyperconverged		Difference	
	1T	OG	1T	OG	1T	OG
Server						
Hypervisor						
Storage/SAN						
Backup/Recovery Tools						
Disaster Recovery Tools						
WAN Accelerator						
SSD Cache						
Deduplication Appliance						
Power & Cooling						
Dedicated Staffing						
Other						
Totals				Savings		

Figure 2-2: Traditional vs. hyperconverged infrastructure worksheet

Hyperconverged infrastructure solutions can help you break out of this cycle and more closely match data center resources with current business needs. By enabling granular and simple scaling, you can buy what you need today and, as the business grows, just add more nodes.

Besides being technically simple, this also enables you to rethink the budgeting process. By not having to buy for needs that might be in place three to five years from now, you can focus on improving what the business needs today.

Your Financial Evaluation Criteria

As you consider implementation of hyperconverged infrastructure, **Figure 2-2** shows a simple worksheet you can use to help determine which solution makes the most financial sense—a traditional approach or a hyperconverged infrastructure.

That's a Wrap!

You're now versed in the organizational and financial considerations that accompany hyperconverged infrastructure. You've made it through the jungle and can move on to more harmonious locales. We hope your journey here has helped you better understand this important topic.

To learn more about hyperconverged infrastructure and key use cases, look for The Gorilla Guide To...® Hyperconverged Infrastructure Implementation Strategies.

ABOUT HPE

Hewlett Packard Enterprise

Hewlett Packard Enterprise is an industry-leading technology company that enables customers to go further, faster. With the industry's most comprehensive portfolio, spanning the cloud to the data center to workplace applications, our technology and services help customers around the world make IT more efficient, more productive, and more secure.

For more information, visit www.hpe.com

ABOUT ACTUALTECH MEDIA

ActualTech Media is a B2B tech marketing company that connects enterprise IT vendors with IT buyers through innovative lead generation programs and compelling custom content services.

ActualTech Media's team speaks to the enterprise IT audience because we've been the enterprise IT audience.

Our leadership team is stacked with former CIOs, IT managers, architects, subject matter experts and marketing professionals that help our clients spend less time explaining what their technology does and more time creating strategies that drive results.

For more information, visit www.actualtechmedia.com