

The Versatility of Hyperconvergence In Any Data Center

James Green, vExpert

ActualTech Media

January 2016



Table of Contents

Table of Contents	2
Introduction	3
The Biggest Misconception.....	3
Hyperconvergence Defined.....	4
What Is It?	4
Core Definition.....	4
Hyperconvergence Misconceptions	4
What Applications Are Suitable for Hyperconvergence?	5
Hyperconvergence Changes the Game.....	5
Hyperconvergence Is Efficient	5
Is Hyperconvergence Software or Hardware?	6
Hardware vs. Software	6
Business Tradeoffs	6
Hyperconvergence Concerns Addressed	7
The Link Between Compute and Storage	7
Hyperconvergence is Flexible	7
How Should It Scale?	7
The Potential Scale of Hyperconverged Infrastructure	8
Now Very Small and Very Large.....	8
Conclusion.....	8



ActualTech Media © 2016. All rights reserved.

Under no circumstances is this document be sold, copied, or reproduced in any way except with written permission.

The information contained with the document is given in good faith and is believed to be accurate, appropriate, and reliable at the time it is given, but is provided without any warranty of accuracy, appropriateness or reliability.

The author does not accept any liability or responsibility for any loss suffered from the reader's use of the advice, recommendation, information, assistance or service, to the extent available by law.

All trademarks are the property of their respective owners.

Introduction

In many organizations today, hyperconvergence is known as little more than an industry buzzword. However, many organizations have found wild success by leveraging this data center architecture.

One of the major challenges every new paradigm must face, regardless of industry, is the challenge of overcoming general unawareness and misunderstanding in the market.

There are those who have historically approached new ideas with hostility, but those individuals are in the minority. *Most* individuals (and the organizations that those individuals comprise) are open to new ideas - especially when there's potential for business upside. The problem is that they lack the information and education about the new idea to make an informed decision.

In this paper, you'll learn valuable information about the basics of hyperconvergence and how it fits in a modern data center. You'll also learn the truth about some of the most common misconceptions in the market related to hyperconverged infrastructure.

Armed with this information, you'll be able to develop an informed view of the hyperconverged infrastructure paradigm and begin to see where it might fit in your organization.

[The Biggest Misconception](#)

Perhaps the most damaging and misinformed notion that pockets of the data center and IT market hold about hyperconvergence is that it's only a fit for niche use cases, or organizations of a certain size.

This paper will posit that as the technology to support the architecture has grown to maturity, hyperconverged infrastructure is a fit for all but the most specialized and unique data centers. For most use cases, hyperconvergence in it's current state of maturity is a viable option for data center architecture.

Hyperconvergence Defined

WHAT IS IT?

Hyperconvergence as a technology is in the ‘buzzword’ stage, much the way “cloud” was half a decade ago. People want to know about it, and vendors ask whether you’re thinking about it. In the same way the (still) rather ambiguous term “cloud” needs to be defined, the question that is always hovering over HCI (hyperconverged infrastructure) conversations is: “What does ‘hyperconvergence’ mean, exactly?”

Core Definition

At its core, hyperconvergence is a quest for simplicity and efficiency. Every vendor with a hyperconverged platform approaches this slightly differently, but the end goal is always the same: combine resources and platforms that are currently disparate, wrap a management layer around the resulting system, and make it *simple*. So, what are the disparate resources being combined, and how do HCI vendors make it simple?

Consider an enterprise data center from the past few years, or maybe just poke your head in your own data center. It contains clusters of “compute” (typically x86 server platforms) and one or more monolithic storage arrays. There’s a deduplicating backup appliance, a backup software, a replication platform, and a few WAN accelerators. Hyperconvergence, on the other hand, leverages a return to direct-attached storage (DAS). It uses software to combine varying platforms (compute, storage, backup, multi-site federation, and so on) into one platform.

Hyperconvergence Misconceptions

A common misconception is that ‘hyperconvergence’ means servers and storage in the same box. Pooling storage is a good example of SDS (Software-defined Storage) which is *a part* of hyperconvergence, but it is not the whole picture. HCI aims to bring as many platforms as possible under one umbrella, and storage is just one of them.

Once these once disparate systems are owned by the hyperconvergence platform, they’re accessible from the same management interface as all the other systems. Different vendors do this piece differently; it may be a dedicated user interface or it may be a plugin to an existing interface. Whatever the foundation, the end goal is to manage everything (compute, storage, backups, replication, snapshots) from one management interface. This provides the simplicity that administrators are in desperate need of in the modern datacenter.

WHAT APPLICATIONS ARE SUITABLE FOR HYPERCONVERGENCE?

In the world of enterprise IT, “high performance” has traditionally been synonymous with “complex.” It has been necessary to tolerate certain levels of complexity to gain the highest performing solution possible. Because of the assumption that high-performing solutions must be complex, there’s an inherent implication that a simple solution must be low-performing.

Hyperconvergence Changes the Game

Hyperconvergence defies the industry’s notions about complexity and performance by being both simple *and* complex, yet still high performing. The software that makes a hyperconverged solution work under the hood is actually quite complex. There’s substantial work and intellectual property that goes into making the solution as high performing as possible. However, there’s an equal or higher proportion of dedication to abstracting this complexity from the consumer of the product. What this means, in effect, is that while it looks simple on the surface and is easy to manage, it can keep up with the complex solutions in terms of performance.

There are certainly still cases where intricate solutions can be meticulously built and outperform HCI. But for most workloads, performance is not going to be a concern.

Thanks to the current state of the technology, and given the correct configurations, HCI is 100% capable of running business-critical applications like enterprise database, messaging, and application servers. The only workloads which tend to not be a good fit for hyperconvergence are those that require special hardware.

Hyperconvergence Is Efficient

Due to the ability to run disparate business-critical systems on one platform, there can also be substantial savings in the way of efficiency. Often, a separate segment of infrastructure is provisioned to serve a single service (e.g. messaging). By deploying many services on the same segment, high data efficiency ratios can be attained and higher density/consolidation is possible.

Almost any application can be run successfully on a hyperconverged system.

Is HYPERCONVERGENCE SOFTWARE OR HARDWARE?

In reality, it's both. A hyperconverged platform will always have a software component and may sometimes have a hardware component as well. Some vendors will support their choice to include or not include propriety hardware as the best and only correct option; but the truth is that the “correct” choice comes down to what you prefer.

Hardware vs. Software

The VSA (virtual storage appliance) that most hyperconverged offerings use to pool storage, or the kernel module in the case of something like VMware VSAN, are examples of the software that make a hyperconverged platform.

On the other hand, some solutions like SimpliVity's OmniStack do include custom hardware like the OmniStack Accelerator Card. The purpose of custom hardware is typically to increase performance and/or efficiency. In the specific case of the OmniStack Accelerator Card, the purpose is to offload data efficiency and integrity processes so that the system's main CPUs don't lose valuable cycles performing those operations and can continue serving virtual machine workloads.

Business Tradeoffs

There's not a right or wrong to whether or not hardware is included in an HCI platform. If special hardware is included, it limits your choice in regard to what equipment can be used to run the platform. But it likely increases stability, performance, and capacity on a node (all else being equal).

The opposite view is that leveraging a VSA and no custom hardware opens up the solution to a wide variety of hardware possibilities. While flexible, the downside of this approach is that it consumes resources from the hypervisor which would have served VM workloads in a traditional design. This can potentially add up to a considerable amount of overhead.

Which direction ends up being the best choice is dependent on a myriad of different variables and is unique to each environment.

Hyperconvergence Concerns Addressed

THE LINK BETWEEN COMPUTE AND STORAGE

The relationship between compute and storage is commonly misunderstood by prospective hyperconvergence adopters. And it's certainly a valid misunderstanding. One notices that compute resources and storage are in the same building block and surmises that the two are now inextricably linked. So, scaling one must mean scaling the other.

Hyperconvergence is Flexible

The conclusion that storage and compute must be scaled in equal proportions to each other is a reasonable assumption to make based on the way things appear. But it's a false assumption across almost all of the hyperconverged industry. Most vendors know this would be a major challenge and provide a method to cope with it. There are two possible ways of handling this issue:

- Allow the addition of “compute only” nodes which do not contain storage, but are still managed by the platform and access the storage shared amongst other hosts
- Offer broad node configurations so that although storage and compute always scale together, the amount of each can be tuned such that the desired capacity for both resources is attainable. (For example, nodes with powerful compute but minimal storage, or vice versa)

How Should It Scale?

Second, there's the question of whether scaling the resources independently is actually desirable. In many cases it may be the best decision to scale compute and storage in tandem, even if it's not necessary for capacity's sake. There are two reasons for this: increased IO performance potential due to extra storage controllers and spindles, and narrowed failure domains due to expanding the breadth of the deployment. In the latter benefit, data is spread across more nodes, resulting in the failure of a given node having less impact.

In the end, the answer to the question “what if I only need more [storage or compute]?” is: you can do just that. You are not bound to scaling both resources at once in most solutions.

THE POTENTIAL SCALE OF HYPERCONVERGED INFRASTRUCTURE

At inception, hyperconverged infrastructure had some limited use cases, simply due to immaturity. This is expected with any new technology; when x86 virtualization first became mainstream, it had a limited use case until it could be fine-tuned. But as the architecture matures and vendors perfect little pieces of the puzzle, HCI begins to be applicable to most use cases and most organizations.

A common question a potential HCI candidate in the SMB space will ask is “Can this be small enough to be affordable to me?” On the opposite end of the spectrum, enterprise customers might ask, “Can this scale high enough to meet my needs?”

Developing a new technology - and launching a business in general - is all about developing a product that will appeal to the largest segment of your target market first. Once that is developed and stabilized, it may make sense to extend outwards to the ends of the spectrum. In this regard, hyperconvergence is no exception. The first iterations of the architecture were the best fit for the small to medium business that needed anywhere from 3 to 8 nodes in a cluster. That covers a *substantial* segment of the addressable market.

Now Very Small and Very Large

More recently, solutions have been announced like VSAN 6.1, offering a two-node deployment (with witness appliance for quorum), or SimpliVity's OmniStack 3.0, doubling the number of sites that can be managed in a given federation. HCI has now reached the maturity that offerings are starting to reach the edges of the spectrum, where the smallest and largest customers are both served well.

The question of scale doesn't only apply to technological limits, however. “Does it work?” is an important question; equally important is, “How easy is it to use at [insert scale]?” This is the beauty of the HCI paradigm – no matter the scale, great attention will be given to making sure that simplicity doesn't begin to fade. Except for in the smallest of small environments and the largest of large environments, hyperconverged infrastructure can likely serve the needs of any environment.

CONCLUSION

Hyperconvergence is easily one of the most revolutionary paradigm shifts in the data center industry in recent years. Find out how SimpliVity can help your organization with the transformation by visiting <https://www.simplivity.com/>.