ACTUAL/TECH MEDIA

WWW.ACTUALTECHMEDIA.COM

1000

2014 State of Web-scale Storage Report

Scott D. Lowe, vExpert, MVP Hyper-V

Developed in Cooperation with



Λ

Table of Contents

Executive Summary	2
Introduction	4
Part 1: Respondent Demographics	6
Part 2: Web-scale Storage Defined	18
Part 3: Modernizing The Data Center Storage Environment	29
Summary	31
Appendix: Additional Webscale Storage Perceptions	33

Executive Summary

By 2017, the IT industry will have hit the 50% mark in implementing web-scale approaches to IT and storage. This, according to Gartner, will provide organizations with new ways to provide business-focused technology services that expand flexibility and resiliency. Web-scale approaches to storage will enable organizations to maintain expected levels of service while preventing widespread outages and enabling quicker recovery times when outages do occur.

<u>ActualTech Media</u>, in cooperation with <u>Coho Data</u>, performed a in-depth survey of over 200 enterprise executives and administrators in order to identify market perceptions,

readiness, and concerns around the concept of web-scale storage. The results show a market in flux, both in readiness around the underpinnings of web-scale storage as well as a broad lack of understanding and agreement on a number of the foundational tenets that comprise the web-scale model. The results also provide insight into the current state of the hypervisor wars, a

WEB-SCALE STORAGE HELPS MAINTAIN EXPECTED LEVELS OF SERVICE WHILE PREVENTING WIDESPREAD OUTAGES AND ENABLING QUICKER RECOVERY FROM FAILURE WHEN OUTAGES DO OCCUR.

battle that Microsoft is pushing hard to win and even showing signs of success, with more than 40% of respondents running Hyper-V somewhere in their organization.

On the web-scale storage-readiness front, customers in organizations both large and small have undertaken some of the steps necessary to succeed, as shown by the increasing 10 GbE adoption rates among respondents, now registering above 50%. Moreover, when it comes to balancing storage capacity and storage performance, organizations now fall squarely in the middle. Whereas managing capacity used to be the key driver in making storage decisions, it's obvious that emerging I/O-intensive application workloads have dramatically shifted the conversation to one that places equal value on both of these metrics. In fact, for many respondents storage performance is a far more critical metric than storage capacity and these organizations measure success on an IOPS level rather than at a terabyte level.

All that said, it's still no surprise that overall storage cost remains a key factor for many companies. This only makes sense. For many organizations, the bottom line is the business' key performance metric, and maintaining a reasonable cost structure with storage fits well with this goal.

The level of interest in the burgeoning web-scale space makes it clear that this is a trend that deserves attention and respondents to the survey agreed. Most respondents are either actively undertaking, or are planning to evaluate, the potential merits of web-scale storage for their own organizations. However, when it comes to the actual term "webscale", the survey reflects some market confusion on outcomes. In short, the vendor community needs to continue its outreach efforts to help people understand that webscale storage is a data center storage design that provides:

- A 'shared-nothing' architecture. With no single points of failure, a web scale architecture allows for higher levels of availability than can be achieved with traditional services.
- Horizontally-scaled. Traditional vertically-scaled environments eventually begin to suffer performance issues as more and more load is placed on shared components. A horizontally scaled, or scale-out, architecture enables linear resource expansion. As such, these systems can achieve expected levels of service even as they continue to grow. Furthermore, the loss of one node in a scale out cluster does not generally have a deleterious impact on partner nodes.
- Greater simplicity (or less complexity). Very few CIOs ask their staff members to find ways to make it more difficult to manage IT. Rather, CIOs and other business unit leaders are pushing IT to be more nimble, more flexible, and less costly. This is translating into a need for IT to find ways to make the job simpler. Web-scale storage architectures fit perfectly into this vision as they can help IT eliminate resource islands and can provide an easier administrative experience, which equates to lower ongoing costs.

The data center is undergoing fundamental changes that are intended to help IT take a more business-centric approach to operations. However, the full 2014 "State of Web-scale Storage Report" shows that enterprises continue to struggle to extract performance and scalability gains from existing infrastructure.

Introduction

While every aspect of the data center resource spectrum — networking, compute, storage — is critical when it comes to completing the full infrastructure picture, storage has emerged in recent years as the current focus for many vendors — and many new trends — that have hit the IT landscape. One of the most prevalent trends is what Gartner has coined *web-scale IT* and, with that, the rise of what's become known as *web-scale storage*.

In Gartner's estimation, by 2017, web-scale IT will be found in 50% of enterprises. Cameron Haight, research vice president for Gartner, has this to say about web-scale IT:

"Web-scale IT looks to change the IT value chain in a systemic fashion. Data centers are designed with an industrial engineering perspective that looks for every opportunity to reduce cost and waste. This goes beyond redesigning facilities to be more energy efficient to also include in-house design of key hardware components such as servers,

"WEB-ORIENTED ARCHITECTURES ALLOW DEVELOPERS TO BUILD VERY FLEXIBLE AND RESILIENT SYSTEMS THAT RECOVER FROM FAILURE MORE QUICKLY." -GARTNER GROUP

storage and networks. Web-oriented architectures allow developers to build very flexible and resilient systems that recover from failure more quickly."

The primary challenge with these kinds of prognostications and the associated trends is the market confusion that ultimately results as different vendors lay claim to the trend and work to redefine it to meet the features of their particular product. Ultimately, end users are left with varying descriptions of what the much-hyped trend is supposed to accomplish, diluting the potential positive outcomes that can result in rethinking IT architectures. There is a struggle to gain clarity as to action items and which potential resource partners can help to achieve constantly evolving technology-related organizational goals. In order to gain a better view into potential web-scale IT adopters' understanding of the web-scale storage trend, Coho Data partnered with ActualTech Media — the owner of <u>VirtualizationSoftware.com</u> and <u>EnterpriseStorageGuide.com</u> — to conduct a market survey which asked respondents to provide feedback on both their existing data center environments as well as on overall views of the web-scale trend. This report provides results and analysis around the feedback provided by respondents with a focus on web-scale storage.

Part 1: Respondent Demographics

In order to provide overall context for the analysis performed for this report, it's important to understand the organizations that provided data and feedback to the initial survey, which consists of 206 respondents.

Overall Organizational Size

Organizational size will often dictate how an organization leverages technology in pursuit of improving the bottom line. Of course, there are always exceptions.

As shown in the chart to the right, respondents to the web-scale storage survey work in organizations that are sized across the spectrum from small companies with less than 100 employees up to enterprise-level firms with 5000+ employees.

Organizational Size



IT Staff Size

The IT function is of growing importance in businesses large and small. Organizations with small numbers of IT staff relative to the size of their overall business are generally under even more pressure than more IT-focused firms to "do more with less" and must find ways to make routine — or even eliminate —some of the ongoing administrative functions that are often in place. The average size f respondents' IT organizations is in the 26 to 50 range, but the end points of the IT

Number of IT Staff



department size spectrum were certainly favored, with 25% of respondents working in organizations with 100+ IT staff and 37% working with an IT staff of ten or less.

Industry Verticals

Some industries rely more heavily on technology than others. For example, in the world of finance, technology downtime translates into real dollar lost. In education, even though technology is critically important, traditional classes can still be taught when systems are down.

Of course, as time goes on, even verticals that could once operate without technology are becoming increasingly reliant on its availability. This makes it even more important to look for ways to implement more efficient data center operations. The top three verticals represented in the survey were high tech/telecommunications, government, and finance.



Respondents' Industry Verticals

Respondents' Storage Environment

A critical element of the demographics section of the survey asked respondents to provide important details about their storage environments. This information provides the context by which decision-makers in these organizations must consider how they should move forward in terms of their storage environments. Should they maintain legacy status quo systems or should they adopt a web-scale approach to solve what might be critical challenges?

Total Storage Environment Size

One-half of respondents manage storage environments that are sized under 50 TB with the other half reporting total storage above 50 TB. It comes as no surprise that total storage grows in pretty much a straight line as organization size grows, as shown in the

chart to the left. Of course, there are some outliers in each category, but on average, storage grows along with an organization. As is the case with many answers to survey questions, this linear growth perfectly demonstrates the need to implement storage systems that can seamlessly grow as the company grows and as the company's storage needs grow. As needs

Organization Size	Average Total Storage
0-100	< 20 TB
101-500	< 50 TB
501-1000	< 100 TB
1001-5000	< 100 TB
5000+	< 500 TB

grow, organizations don't simply want to add support resources in the same linear fashion, so it becomes increasingly important to implement storage that can scale in a way that maintains cost-efficiency.

Total Storage Capacity Dedicated to VDI

Gaining an understanding as to how storage is used is a significant first step in understanding how the environment can be improved. Perhaps one of the most challenging aspects of legacy storage environments is the fact that many environments have been broken down into many parts in order to address different needs that are imposed by different services. Nowhere is this fact more prevalent than virtual desktop infrastructure (VDI) environments. VDI environments have long been considered challenging implementations when it comes to storage. VDI environments carry with them very different I/O patterns and capacity needs than traditional services.

In the chart below, it's easy to see that most respondents maintain storage environments that are in the 10 TB range. However, there is a caveat: During the initial data collection phase for this report, the question that asks "How much storage respondents have dedicated to VDI?" did not have a none option, so it is possible that many of the 128 people who responded that they have less than 10 TB of VDI-dedicated capacity actually have none. That said, it remains apparent that there are a great many organizations dedicating specific storage resources to their virtual desktop environments.



Storage Dedicated to a Specific Uses

Given VDI's vastly different I/O challenges when compared to server virtualization, this does not come as a surprise. In fact, due to storage-related performance challenges with VDI, many organizations discovered that their existing server-focused storage could not keep up with VDI's demands. It's unfortunate, too. Being required to deploy multiple islands of storage just to meet the demands of a specific application adds to an environment significant costs, complexity, and administrative overhead — the very items that most organizations want to reduce. Having storage that can handle both VDI and general workloads lowers costs and increases utilization.

10 Gb Ethernet Penetration

Web-scale storage requires a robust communications fabric in order to deliver on its promised goals. Over the years, organizations have adopted a wide range of storage communications technologies, including Fibre Channel, iSCSI, Infiniband and more. At the same time, venerable — but ubiquitous — Ethernet technology has risen as a prime choice across many aspects of the data center. The popularity of this medium has risen as improvements corrected what were considered deficiencies and as vendors released ever-faster versions of the technology.

While gigabit Ethernet has been an incredibly successful option in the data center, as organizations seek to adopt architectures based on horizontal and linear scale, 10 gigabit Ethernet (10 GbE) is becoming all but a requirement. Fortunately, as has been the case with previous renditions of the technology, 10 GbE costs are coming down as time goes on and as vendors continue to explore even faster speeds.



Do you currently have 10 GbE deployed?

A full one-half of survey respondents have deployed 10 GbE technology with about 13% more planning to deploy technology within the next year.

When it comes to 10 GbE deployment, larger organizations certainly have the edge as the vast majority have already deployed the technology. As such, large enterprises are already primed to leverage emerging web-scale storage opportunities. However, even in the 0-100 employee space, there is reasonably good uptake for 10 GbE and with the cost for 10 GbE coming down all the time, this fast version of Ethernet is becoming more accessible even at the lower end of the market, so it should not be an inhibiting factor when it comes to adopting web-scale storage.

Managing Storage Capacity Growth Remains a Critical Function

For well over a decade — and maybe longer for some — storage administrators have been charged with managing the growth of what is an increasingly important resource. At the same time, users and the business itself have often demanded expansion that can force IT departments into what might be short-term solutions that end up staying in place for the long term.



How important is managing storage growth for your business?

Some estimates place enterprise storage growth at a whopping 25% year over year. End users contribute significantly to this statistics. Most end users don't want to have to constantly focus on keeping their storage use in check and simply find ways to fill up whatever capacity they are provided. Survey respondents overwhelmingly agree that keeping careful tabs on this resource is of primary importance. Over 75% of respondents indicated that managing storage growth is a major concern in the data center.

This fact supports the need to consider storage systems that don't require constant attention to growth and that can easily and significantly scale. Web-scale storage provides a foundation for enabling linear horizontal scale without unreasonable limitations.

Managing Storage Performance Growth

Close to 78% of survey respondents consider managing storage performance to be a critical activity. Yes... *performance*.



How important is managing storage performance for your business?

In recent years, storage performance has become as critical a metric as storage capacity, particularly as new kinds of workloads have drastically changed traditional I/O patterns and made it close to impossible to design storage systems around specific patterns, as was easily accomplished in the days of the physical server. One of the primary culprits behind these changing I/O patterns is virtualization of both servers and desktop systems. Before virtualization, storage was designed around individual workloads, but the shared nature of virtualization makes this kind of configuration extremely difficult, if not impossible.

As IT is pushed to focus more on the business and as businesses seek to expand their virtual environments to encompass more and bigger applications — including missioncritical tier 1 applications — paying attention to normalized scalability in the storage environment becomes increasingly important. Rather than building islands of storage to meet individual application needs — which adds unnecessary administrative overhead — businesses need to implement storage systems that can keep pace with performance growth and that don't break the budget. Monitoring capacity and performance as separate metrics is fine but managing them separately shouldn't be necessary. As customers manage storage performance growth they shouldn't have to constantly reconfigure their storage. It should adapt to their work loads without requiring constant management as web-scale does.

Capacity and Performance Are Now Equally Important Metrics

Surprisingly, in looking at the survey outcomes, respondents actually gave an edge to performance when it came to comparing the relative importance of scaling for capacity versus scaling for performance. This result substantiates the observation that storage performance has become a critical data center metric. With a data center that is as focused on managing performance as on managing capacity, administrators gain new challenges as there are sometimes some mutually exclusive decisions that must be made. Fortunately, as web-scale storage technologies become more ubiquitous in end user environments, web-scale's ability to manage both of these factors without major administrative overhead makes the technology a compelling choice.



What is more important to your organization as it pertains to storage?

Leading web-scale storage vendors have taken a hybrid storage approach in their products. Hybrid storage hardware includes a combination of both solid state storage and spinning disk that are brought together to provide customers with a balanced approach to ongoing storage needs. The flash portion of these solutions can often be used as a performance tier and store information for applications that demand raw IOPS. This performance layer is also used as a caching mechanism sitting in front of the spinning disk layer, or capacity tier. By providing a storage tier for the spinning disk, all operations that depend on spinning disk can be accelerated.

The vast majority of respondents place equal importance on scaling for capacity and scaling for performance. In this instance, the market is exactly matching consumer needs as these hybrid solutions continue to take hold. Web-scale storage vendors such as Coho Data perfectly match these emerging needs.

Controlling Storage Cost is Critical

Cash is king and 80% of survey respondents agree, which is really no surprise. Ensuring that storage costs don't skyrocket out of control is of critical importance to these respondents. What is surprising is that 20% of the people polled didn't cite cost as critically important. Perhaps these organizations measure their storage environment value differently or perhaps they just consider this a part of the cost of doing business.



How important is managing storage *cost* for your business?

The fact that cost is so important to so many would seem to lend support to implementing a storage environment that carries with it a relatively predictable cost of acquisition, support, administration, and expansion. Web-scale storage's linear scalability nature would seem to be a good fit.

Many Storage Protocols = Additional Workload

Today's data center environment seems to be chock full of a variety of storage protocols. Although this fact may enable organizations to leverage any storage at any time, the more protocols that are in production, the more complexity and administrative overhead that the system requires. Each protocol carries with it very different monitoring needs and skill sets. One of the basic tenets of web-scale storage is the use of a single protocol for all operations — one generally based on ubiquitous Ethernet technology. Only 40% of survey respondents indicated that they've standardized on a single storage fabric, with 9% indicating that they operate four or more storage protocols.

When considering the future of the data center, it appears as if Ethernet technologies have already overtaken other communications fabrics, including Fibre Channel. In fact, a full 87% of respondents are already running some kind of Ethernet-based storage protocol — iSCSI, NFS, or SMB — in their data centers.



Which protocols are you using to connect to your storage?

Whereas this technology was shunned by storage aficionados not that many years ago,

Ethernet now has a massive foothold in an environment that was once dominated by Fibre Channel, which less than one-half of respondents indicate they continue to operate in their data centers. It's no surprise that Fibre Channel continues to be popular in larger organizations, though, where there is opportunity for a more diverse IT skill set and where there may have been workload reasons for running this robust fabric. For smaller organizations, though, the simplicity and relatively low cost of an Ethernetbased storage fabric seems to have been a compelling story, and dominates in those organizations' section of the market.



Part 2: Web-scale Storage Defined

Web-Generation companies like Amazon, Google and Facebook have consciously avoided traditional monolithic arrays in favor of a new type of storage architecture that relies on commodity hardware and intelligent software that can handle quick, efficient and cost-effective scaling. Along with introducing this emerging space — Web-Scale Storage — Gartner suggests that this model is not just for web companies and in fact will be present in over half of enterprises by 2017 due to benefits in both lower acquisition costs and automation and operational efficiency.

In the decades since Information Technology systems first became prevalent in the enterprise, the field has undergone continuous change, but this time span has also been marked by a few major architectural upheavals. These upheavals — for example, the moves from mainframes to distributed networks and the introduction of virtualization — have led organizations to new ways of managing and leveraging IT. In the realm of storage, web-scale is the next progression in the cycle of change.

Characteristics Web-scale Storage Systems

While thoughts about how to position the trend may vary — thoughts which are discussed later in this report — there are a number of traits that characterize web-scale storage, each of which is described in the following section. These traits include horizontal scaling, "shared nothing" architecture, simpler storage tiers, and a high degree of standardization.

Horizontal Scaling

The days of vertical scaling — sometimes referred to as scale up — of storage systems are quickly fading as organizations seek ways to ensure that all storage metrics — capacity, networking/communications throughput, data processing speed, and performance/IOPS — grow as the organization expands its use of a storage device. In legacy scale up storage systems, customers had the ability to easily scale the capacity and IOPS portions of their storage systems, but while doing so, more and more load was placed on the single processing units and network connections that these devices leveraged.

This approach eventually leads to network- and CPU-based performance challenges while, at the same time, the environment is exposed to a single point of failure in what is basically a shared architectural component: the processing unit, which connects to the storage communications fabric.



Scale Up Architecture – Increasing Burden on Shared Components and Single Points of Failure

Web-scale storage relies on a different architectural paradigm known as horizontal scaling, which is often referred to as a scale out architecture. In horizontally scaled environments, new resources — generally in the form of new clustered nodes — are added in a linear fashion and, as added, each new node adds all necessary resources — storage, IOPS, networking, and CPU — necessary to stand alone. These individual nodes, although they include everything necessary to function on their own, participate as members of the clustered environment and contribute their individual components to the whole.

Disk Shelf

Storage Head/Processo

Disk Shelf

Storage Head/Processor

Disk Shelf

Storage Head/Processo

Disk Shelf

Storage Head/Processor

Disk Shelf

Storage Head/Processo

The cluster management software itself controls where data is placed on these individual

nodes using globally-defined granular policies that enable high levels of data protection and availability. If a node experiences a fault or other availability issue, another node in the cluster can transparently assume the

THE FACT THAT THE FAILOVER PROCESS IS TRANSPARENT IS A KEY CHARACTERISTIC OF WEB-SCALE STORAGE SCENARIOS.

responsibilities for the failed node. The fact that this failover process is transparent is a key characteristic of web-scale storage scenarios.

As new nodes are added to these kinds of clusters, the cluster management layer takes control of the new resources. This process includes adding the individual node resources to the global cluster and, if necessary, rebalancing other cluster nodes to leverage any performance or availability opportunities that may be afforded by the new resources. Again, these processes are generally transparent to both administrators and end users.

"Shared-Nothing" Architecture

In the data center, great effort has been taken over the years to reduce reliance on shared components. The reason is simple: When a shared component fails, any second order services that depend on that component will also fail, resulting a wider impact. Vertically scaled environments are particularly susceptible to failures that result from shared component failures. Moreover, such environments are susceptible to performance issues for the reasons described in the previous section — at some point, shared resources will become overwhelmed.

A shared nothing architecture implies that there are no single points of failure in the data center. A shared nothing architecture that relies on commodity hardware controlled by a comprehensive software layer carries with it the understanding that hardware components can fail and they will fail. This should not be construed as a weakness in the system. In fact, it should be considered a strength as the software layer carries with it necessary intelligence in order to transparently handle hardware failures in a way that maintains the integrity of the stored data as well as expected performance levels.



Simpler Storage Tiers

Fact: Administrators should not need to manually control storage tiering. For those unfamiliar with the term storage tiering, this is a process by which storage is carved up into smaller chunks. Each of these chunks carries with it individual parameters and characteristics. For example, an administrator may create one storage tier for general purpose file storage. This tier would be configured with high capacity and relatively low performance specifications. Another tier might be created for database applications and may have only modest capacity needs but require significant IOPS to keep up with with business needs. Further, storage tiers may be created based on how data is accessed. Manual moving data around a storage system's tiers is a tedious process that is operationally expensive and that limits IT's short-term flexibility. With simpler and automated storage tiering, if data has been accessed recently, it may reside on a hot tier, while stale data may be moved to a cold tier.

When it comes to opportunities to scale the storage environment, breaking apart storage in these kinds of separate constructs begins to add complexity to the environment. Some storage systems do a good job at handling this complexity behind the scenes so that the administrator doesn't need to constantly be involved in tiering decisions. However, these auto-tiering solutions aren't available in every storage environment, particular in environments in which the storage architecture has been kludged together over the years. Moreover, even automated tiering systems can be more reactionary in nature than proactive. With a web-scale storage environment, the storage system tracks historical data usage patterns in order to predict when particular data will be used This enables such systems to take a much more proactive approach to storage management.

High Degree of Standardization

The phrase economies of scale plays a major role in a web-scale storage environment. These environments are highly standardized, which has positive impact on a customer's

total cost of ownership. Standardization also improves the overall administrative paradigm and improves outcomes provided by vendor NORMALIZED — AN support personnel. The more standard — or normalized — an environment, the easier it is to administer. Consider for a moment the

THE MORE STANDARD — OR ENVIRONMENT, THE EASIER IT IS TO ADMINISTER.

plight of the legacy storage administrator. In many cases, as an environment has had to grow, organizations may have opted to move to vertical scale solutions simply to address immediate capacity or IOPS needs. Worse yet, they may have found it necessary to adopt storage systems that fell into one of the categories below:

- Were from different vendors.
- Were from the same vendor, but fell into different product lines due to different storage needs or vendor product availability.
- Were not intended to scale horizontally and were managed with per unit or per node administrative tools.

However, with web-scale storage, the software and processes layer is not the only element that carries with it a high degree of standardization. The hardware itself is also highly standardized and is often based on commodity off-the-shelf hardware. Besides being less expensive than systems that require high levels of custom hardware engineering, this kind of hardware is readily available from original equipment

manufacturers, making creation of web-scale systems easier. Customers do not have to wait inordinate amounts of time for new nodes to be delivered, meaning that customer scaling needs can be accomplished relatively quickly.

People often believe that the word commodity relates directly to the word cheap, as in low quality. While it may be true that commodity hardware is in use in many modern data center architectures, including those that take a web-scale approach to storage, what is not true is that these components are necessarily low quality. A commodity hardware-based approach to web-scale storage enables storage vendors to focus more on building the comprehensive software layer that is going to become more important as businesses continue to push IT to greater limits. This approach untethers storage vendors from specific hardware and enables what would have been hardware engineers to focus on building and improving the software layer. As time goes on, IT will live in an increasingly software-defined world, a trend in and of itself. That has the potential to reap significant benefits for both IT and the business.

"Commodity" Doesn't Have a Consistent Definition

As is discussed later in this report, web-scale storage and overall web-scale IT systems generally ride atop general purpose, commodity hardware. With primary goals of being scalable and affordable, the ability to leverage commodity hardware is a key differentiator of emerging architecture solutions — including web-scale storage — when compared to proprietary sibling technology. End users carry with them mixed messages and opinions when it comes to how they feel about commodity hardware and survey responses were all over the map. Here are just a few opinions, pulled directly from survey responses:

- Cheap, generic, utilitarian servers and storage. Little competitive differentiation.
- Easy to obtain cheap hardware.
- COTS Common off the Shelf x64/x86, NAND based SSD, Industry standards based hardware, SAS, SATA.
- SuperMicro, inexpensive, support which isn't as "cohesive" as for one of the name brands.
- "Commodity" and "Enterprise" may have different performance and reliability qualities. Choose carefully.

- "The end of [vendor 1], [vendor 2], [vendor 3] and others of that blood sucking proprietary pack."
- "386 servers from K-mart."

5%^{3%} 24% Wery Comfortable Mostly Comfortable Neutral Doubtful Not Comfortable

How comfortable are you with reliance on commodity hardware managed by modern, complete software?

Overall, most respondents — 59 % — are comfortable with the potential use of commodity hardware in the data center storage environment. 33% are neutral on the subject, but 8% have grave concerns about whether or not commodity hardware-based solutions can really address enterprise needs in the way that proprietary solutions have done for decades.

A COMMODITY HARDWARE-BASED APPROACH TO WEB-SCALE STORAGE ENABLES STORAGE VENDORS TO FOCUS MORE ON BUILDING THE COMPREHENSIVE SOFTWARE LAYER THAT IS GOING TO BECOME MORE IMPORTANT AS BUSINESSES CONTINUE TO PUSH IT TO GREATER LIMITS.

The reasons for discomfort with commodity hardware are varied, and many responses focused their discomfort with commodity storage around their poor experience with other technologies. Some respondents indicated that their discomfort stems from their lack of experience around the use of commodity storage. Others still understand that the

experience around commodity storage depends completely on the managing software, with one respondent indicating concern around "Redundancy and data loss potential. It all depends upon how reliable the managing software is."

Designed For Virtualization

Believe it or not, there remain organizations that have yet to adopt virtualization. Some still prefer to maintain a fully physical environment as there is either a higher degree of comfort with these systems or there is an application dependency that requires doing so. However, for the vast majority of organizations, virtualization has become the de facto default layer upon which new applications are deployed. In fact, many organizations have made the leap to a multi-hypervisor environment, leveraging different hypervisors to meet different business needs.

VMware Isn't the Only Hypervisor in Town Anymore

Survey respondents clearly remain in favor of vSphere as their primary hypervisor. Almost 82% of survey respondents indicated that vSphere is one of their implemented hypervisors. The next closest hypervisor — Hyper-V — garnered a 41% positive response from survey respondents.



Perhaps the more interesting survey statistic is the number of respondents that are running multiple hypervisors in their organizations. As shown in the in the chart above, a

full 40% of respondents are running at least two hypervisors. 33% of respondents are running two hypervisors, 5% are running three hypervisors, and 2% of respondents are running either four or five hypervisors. This statistic is an interesting one and supports observations that multi-hypervisor use is growing. However, whether the multi-hypervisor trend is a result of organizations deciding to use different hypervisors for different purposes or because they are migrating between hypervisors is an unknown.

Virtualization Continues to Grow as a Platform of Choice

If virtualization is an assumed underpinning of the web-scale storage operation, survey

respondents are certainly well prepared. In fact, many emerging data center architectures carry with them the basic assumption that, moving forward, most — if not all — of an organization's applications will operate in a virtual environment. Web-scale storage is no exception. One of the reasons that web-scale storage companies are growing quickly is due to the storagerelated challenges that come along with virtualization's benefits. As the virtual environment grows and organizations



need to expand their storage, administrators often have to migrate existing data to new storage systems, a process that carries with it significant amounts of time as well as significant amounts of money. This is particularly important given the fact that the storage component in a virtual environment usually bears the majority of the infrastructure cost.

As organizations seek to expand their virtualization efforts, having the ability to seamlessly and linearly expand the storage component of these environments becomes critically important. Web-scale storage imbues a virtualized environment with this seamless expansion capability.

Many Factors Contribute to Stalled Virtualization Initiatives

As already mentioned in this report, there are myriad storage challenges that organizations have had to overcome when it comes to virtualization. However, while storage is certainly a culprit in these challenges, not all virtualization challenges revolve around the storage resource. Respondents indicate normal reasons, including the fact that some applications simply aren't suited for a virtualization environment. Others indicate the need for specialized hardware in support of their applications as a key factor. Another respondent cited anomalies with Microsoft SQL Server clustering as a reason. Finally, others cite lack of executive buy-in on the benefits of virtualization as a key issue holding them back from the 100% virtualized panacea.

There Is No Clear Agreement On The Meaning Of "Web-scale"

As is the case with many of the technology industries emerging trends, web-scale is a trend that is not always consistently defined by industry partners. Further, thanks to the so-called hype cycle and the lack of consistent definition, end users are rightly wary — and even dismissive —of what may appear to be nothing but talk.

Survey respondents were all over the map when it came to providing a free form answer to the question "When you think of "web-scale IT, what is the first thing that jumps to mind?" Responses to this question ran the gamut from not having heard the term prior to receiving the email notification regarding the survey to a deep understanding of the generally accepted definitions of the trend and its opportunities.

It may prove to be an unfortunate name for a trend, however, as many respondents equated web-scale literally with "the Web" while others consider the trend to be directly related to cloud. Depending on how these thoughts are framed, they may fit into the definition of web-scale. That said, the many blank responses and responses with respondents indicating lack of awareness of the trend means that companies providing web-scale IT and web-scale storage products and services must work harder to improve their outreach and education attempts. It's imperative that vendors take care in their messaging to prevent splintering of the definition and to ensure that there is broad understanding of potential outcomes from web-scale storage.

Part 3: Modernizing The Data Center Storage Environment

With a clear understanding for the definition and benefits of web-scale storage, what are people doing with regard to this emerging architecture and what might be keeping people from adopting it and modernizing their data centers? Answers to those very questions were asked as a part of the web-scale survey.

"Web-scale" is Really a Thing

Most of the respondents to the survey indicate that they are either evaluating or intend to evaluate web-scale storage technologies. That said, there are a number of respondents that also indicated that they plan to stick with the traditional scale-up storage stack for their environments. The most cited reason for this is that the respondents are not currently experiencing storage issues that are requiring them to reconsider their approach. it's feasible to believe that, as replacement cycles allow and as new storage issues arise, some of these companies will consider these systems.

However, other respondents have actively considered web-scale approaches and made the decision to stick with their existing systems. Fortunately for web-scale aficionados, there were very few respondents that cited negative reasons for not considering webscale storage. Of those that did, less than fifteen individual respondents indicated concerns around the following:

- Overall maturity of web-scale storage approaches. It's true that web-scale storage is a relatively new infrastructure model and some respondents are more comfortable taking a more legacy-based approach while the new technology proves itself.
- Lack of in-house skills. Any time a new technology is introduced, there is concern around skill set needs. In this case, however, it seems as if the web-scale experience would simplify the environment and potentially reduce the need for some of the skills that are currently needed in the data center.
- Lack of enterprise-class support. Enterprises rely heavily on technology partners to provide deep support in the event of a problem. With new architectures, there is some perception that there is lack of support from the partner ecosystem. However, there are very competent web-scale partners including Coho Data that can solve this issue.

Summary

Many new trends in the IT space end up with lots of hype and fizzle out as promises are broken and the technology ultimately fails to live up to hopes and expectations. Webscale storage, however, has real underpinnings that make it a compelling option for many organizations, whether or not they currently suffer from storage-related challenges. Further adopting web-scale storage can help companies provide streamlined IT services that are more scalable and easier to manager while, at the same time, improving the overall performance of those services. As evidenced by the responses in this State of Web-scale Storage survey, there is ample opportunity for web-scale storage vendors to make their case to customers. Customer environments — in many cases — are ready to make the jump to the new architecture.

About the Author

Scott is Co-Founder of Actual/Tech Media and serves as Senior Content Editor and Strategist. Scott is an enterprise IT veteran with close to twenty years experience in senior and CIO roles across multiple large organizations. A 2014 VMware vExpert Award winner, Scott is also a microanalyst for Wikibon and an InformationWeek Analytics contributor. Scott also develops training resources,



technical white papers and e-books for a variety of Fortune 100 technology firms.

About Coho Data

Led by a team of XenSource/Citrix virtualization and storage industry veterans, Coho Data is enabling businesses of all sizes to build their own high performance web-scale storage for their private cloud. Inspired by the highly scalable,



commodity-hardware based approaches of public clouds, the company is developing the first flash-tuned scale-out storage architecture designed for private clouds that delivers unparalleled performance and simplified management at public cloud capacity pricing.

Founded in 2011, Coho Data is funded by Andreessen Horowitz, Ignition Partners and has offices in Vancouver, British Columbia, Cambridge UK or and Sunnyvale, California.

Appendix: Additional Webscale Storage Perceptions

Is your organization interested in looking at web-scale storage alternatives to traditional arrays that leverage commodity hardware, scale-out software, and automated storage management processes?



If you're looking at a web-scale storage approach, which of the following paths is likely for your organization:



If you're not looking at a web-scale storage approach yet, what is preventing you?

