THE GORILLA GUIDE® TO...

Hybrid Cloud Storage

Scott D. Lowe

Partner, ActualTech Media

Tintri

HELPING YOU NAVIGATE THE TECHNOLOGY JUNGLE

The Gorilla Guide To...

Hybrid Cloud Storage

Written by

Scott D. Lowe

Partner, ActualTech Media

The Gorilla Guide to Hybrid Cloud Storage

Author: Scott D. Lowe

Editors: Rachel Burdick, WordNerd Services
Hilary Kirchner, DreamWrite Creative, LLC

Printed in the United States of America

First Printing, 2016

ISBN ebook: 978-1-943952-16-8
ISBN print book: 978-1-943952-17-5

ActualTech Media
Okatie Village Ste 103-157
Bluffton, SC 29909

www.actualtechmedia.com

About the Author

Scott D. Lowe, vExpert

Scott Lowe is a vExpert and a partner in and co-founder of ActualTech Media. Scott has been in the IT field for over twenty years and spent ten of those years 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 Tintri

Tintri offers an enterprise cloud infrastructure built on a public-cloud like web services architecture and REST APIs. Organizations use Tintri all-flash storage with scale-out and automation as a foundation for their own clouds—to build agile development environments for cloud native applications and to run mission critical enterprise applications. Only Tintri guarantees the performance of their applications, automates common IT tasks to reduce operating expenses, troubleshoots across their infrastructure, and predicts an organization's needs to scale—the underpinnings of a modern data center. That's why thousands of cloud service providers and leading enterprises including Comcast, Chevron, NASA, Toyota, United Healthcare and 20% of the Fortune 100 trust Tintri with enterprise cloud.

Gorilla Guide Features

Executive Corner

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 expand 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.

Table of Contents

Chapter 1:
The Storage Journey 1

The Path from Virtualization to Cloud 2

Key Milestones Along the Way 2

The Cloud Conundrum 5

A Tale of Three Clouds 6

Public Cloud 6

Private Cloud 9

Hybrid Cloud 10

Chapter 2:
Current Trends Impacting Hybrid Cloud 12

Operations Trends 12

Increasing Virtualization 12

Analytics-Driven Monitoring and Planning 14

Autonomous Operation 14

Industry Trends 15

DevOps 15

Web-services 16

Containers 17

OpenStack 18

Serverless Computing 19

Summary 20

Chapter 3:
Hybrid Cloud Requirements....................................21

Storage Platform and the Public Cloud.......................................21

Scalability...22

Automation and Self-Service..25

Summary...26

Chapter 4:
Hybrid Cloud Storage Foundation.........................27

Challenges with Legacy Block Storage and Hyperconverged Infrastructure...27

Physical-era Architecture..27

Storage is Not a Web Service...30

APIs are Not Clean...30

Infrastructure Components Must be Welded Together....31

Hyperconverged Infrastructure Linear Scale and Lack of Choice..31

Native Hybrid Cloud Storage Requirements..........................32

Virtualization-specific Architecture...............................32

Automation, Orchestration, and Clean APIs...................33

Unified Model Across Different Components...............35

Summary...38

Chapter 5:
Hybrid Cloud Storage Vision & Potential.......... 39

Cloud as a Journey, Not a Destination........39

Manage a Cloud........40

DevOps........40

Self-Service........41

Chargeback / Showback........42

Autonomic Operation........43

Summary........44

1

The Storage Journey

"Storage sucks."

That's what so many people have said over the years. It's expensive; it's complex; it's inflexible; it's slow; it's too small; it takes up too much space in the data center.

For far too long, your business has been held captive by bad storage… storage that just barely met your needs, didn't integrate with anything, and was a nightmare to troubleshoot when it came to performance issues.

For quite some time, the storage industry was just… there. Not a whole lot of excitement was in the offing.

And then… things changed. Storage started to become, well, fun again. All kinds of new and exciting storage opportunities flooded the market. Then the public cloud hit the scene and there was suddenly a lot of chatter about how to make public cloud work with the local data center.

But it all had to be able to work together. Suddenly, storage wasn't sucky anymore, as long as you had the right storage system to meet your cloud needs.

If you're deploying or planning for cloud, storage needs to be part of your equation—after all, it can account for more than one-third of your cloud investment.

That's the purpose of this Guide, and this chapter, to explain how we got to this point; from a storage industry focused solely on the local data center to one that is now a source of much innovation and central to cloud.

The Path from Virtualization to Cloud

Since the 1990s, the storage industry has meandered along a path that has been impacted by the advent of ancillary technologies, such as virtualization, flash, and the cloud. While meandering down this path, storage vendors of all shapes and colors have emerged with solutions intended to meet the critical business needs of the day, but over the years serious challenges have remained.

Back before VMware propelled virtualization into our daily lives, storage was a simpler place, although that term is somewhat relative. Before the advent of the SAN, we relied on storage directly attached to servers, which proved to be monumentally inefficient. The monolithic SAN emerged as the answer to all that was wrong in that server-centric storage world and enabled companies to centralize all of those islands of storage and manage it as a cohesive whole.

And then virtualization came along and ruined everything.

Key Milestones Along the Way

Of course, you already know the story of virtualization. It turned out to be a boon for IT, for the business, and for a thriving ecosystem of product vendors out there. It enabled IT to make far more efficient use of IT resources, a particularly important outcome as organizations across the globe seek to continually rein in operational expenses, which includes cutting back on technology spend. Beyond the expense benefits wrought by virtualization, however, there are some really important operational benefits that have emerged, including far improved business agility and the materialization of critical capabilities such as streamlined disaster recovery.

But all was not well…

For quite some time after the serious adoption of virtualization, the storage market became an active hindrance in terms of supporting the burgeoning technology. Thanks to the shifting I/O patterns induced by virtual hosts, some storage systems simply crumbled and could no longer keep up with performance requirements. This fact became a painful reality for far too many early adopters of virtual desktop infrastructure (VDI) systems leading to the failure of many nascent efforts in this space. Making matters worse was the fact that it was incredibly difficult to actually figure out where storage-induced performance challenges were taking place. But it wasn't just VDI that felt the pernicious performance pain. While VDI's sheer I/O intensity brought those problems immediately to the forefront, other I/O-intensive applications, such as databases and analytics systems and even enterprise applications such as large Exchange environments eventually began to be impacted.

Storage systems had become a stubborn challenge with regard to growing the environment as capacity and performance needs dictated. The lack of insight into storage performance challenges forced companies to undergo truly maddening rituals, which generally involved *throwing hardware at the problem*. In other words, rather than being able to solve for a specific performance issue, companies would simply add more disks to their existing environment with the sole desire to increase performance.

Flash Forward

This is the part where you're saying, "Hey dummy! You know that we have this thing now called flash storage, right?" And you're absolutely correct in that flash has helped us think differently about how to solve these tough performance challenges. Even better, as the cost of flash media continues to plummet, it's become more accessible to more organizations. Even all-flash arrays are becoming more commonplace and can handily support lots of random I/O.

But flash doesn't really solve the underlying cause of storage pain sprung by virtualization. It simply masks it and it simply another way to throw hardware at a problem rather than trying to actually fix it. Virtualized workloads demand a storage environment that is tailored to virtualization and that is VM-Aware and understands the unique needs of these systems.

Realizing that the world was moving inextricably toward 100% virtualization, new storage vendors emerged with innovative arsenals of tools to help businesses solve their most serious virtualization-centric storage performance challenges. This new class of storage was built with intelligence, which imbued it with the sentience that it is part of something beyond itself; it is aware of the fact that it is operating in a virtualized environment and is highly tuned to support the unique needs therein.

VM-Aware Storage (VAS) was born (cue musical fanfare here)!

The combination of VAS and flash was a truly potent combination, unleashing theretofore unseen levels of intelligence and performance. The results of this combination have created an entirely new storage market laser focused on virtualization and capable of helping datacenter administrators quickly understand and address even their most-vexing performance challenges.

And then, a funny thing happened on the way to the 100% virtualized data center. A shadow fell across the on-premises data center landscape.

The cloud had arrived, and it wasn't going to take "No" for an answer.

The Cloud Conundrum

You know the stories. The cloud is going to take all our jobs and is going to relegate traditional IT to an afterthought as a part of the new world order. You've all heard the mantra that enterprise IT workloads are marching slowly (or quickly, in some cases) but surely to the cloud. In this context, the assumption is that workloads are exiting the local data center and making their way to public cloud providers such as Amazon and Microsoft. While this is certainly true for some applications, it's absolutely false for others.

With that in mind, and before we get too much deeper, let's consider a few items. First of all, consider *why* organizations are evaluating the public cloud for some of their workloads. For smaller businesses, the decision ultimately boils down to cold, hard cash. A public cloud provider is able to amortize their fixed costs across dozens, hundreds, or thousands of clients. They are able to enjoy significant economies of scale that individual organizations simply cannot achieve on their own. Additionally, in most cases, deploying a new application in the cloud requires very little or no initial capital expenditure. If you think about how we generally buy data center infrastructure, it makes a lot of sense for SMBs to avoid hefty upfront CapEx in favor of public cloud. Rather than tying up a bunch of capital on hardware for which a return may be months or years away – or is really nebulous – why not move to a solely OpEx-based computing model in which you basically rent someone else's infrastructure? And then, as you need more computing power, you simply request it from the provider.

For Enterprises, while CapEx reduction is certainly a motivation, the move to the cloud goes well beyond financial outcomes. In many cases, there are potential operational benefits to be had. For example, agility—standing up a new workload becomes a really easy task. Deploying a new virtual machine in Azure is very simple. In fact, it's so simple that your end users can – and probably are – doing it. That's not a statement intended to denigrate the intelligence of end users. Rather, the intent of that statement is to demonstrate that those with deep technical skills in an organization are not necessarily the technology gatekeepers, as was the case in the past. For these kinds of super-simple processes to work, however, cloud providers – which include public cloud providers and even software-as-a-service-(SaaS)-based tools, of which there are thousands – have spent considerable time automating backend processes and orchestrating the activities that take place when a user makes a request.

A Tale of Three Clouds

"The cloud" is a—forgive the punnyness—nebulous term at best and a dangerous one at worst since it can mean so many different things and the context in this case has a whole lot of meaning.

Public Cloud

Think Amazon, Azure, SoftLayer and the like. These are those cloud providers that do it all. You can stand up entire services in them and never even build out a single server in your local data center. This infrastructure-as-a-service (IaaS) are the stuff of a server hugger's nightmares.

You can also consider some software-as-a-service (SaaS) applications to be public cloud as well. These include applications such as Microsoft Office 365 and Salesforce.

A few years ago, invoking the phrase "the cloud" was sometimes considered an indictment of local IT's ability to carry out their infrastructure and workload responsibilities. If things didn't improve, the thinking went, the business would just move to the cloud and fire everyone. These were real concerns when public cloud first came on the scene.

There are four key characteristics that organizations are seeking from their public cloud:

- Economies of scale
- OpEx-based economics
- Easy and granular scale
- Agile operation

Each of these is described in the sections below. Note that the following sections focus on outcomes derived from the *public cloud*. We'll chat about the differences between public, private, and hybrid clouds shortly.

Economies of Scale

Non-IT companies really don't want to hire hordes of technical IT specialists to manage key business systems. People are expensive and can be difficult to manage. However, with traditional deployment models, a lot of companies find themselves hiring network administrators, storage admins, end user specialists, virtualization administrators, database administrators, technical business analysts, and a whole lot more. And then, for redundancy, there is cross-training or hiring additional people.

Since it's the business they're in, cloud providers are well-suited to staffing many of the infrastructure-related functions identified above. And, as stated before, because they can amortize these costs over their full customer base, they can provide such services to individual customers at a lower rate. Best of all, the customer doesn’t necessarily need to have full-time people in those local roles.

Of course, not all of these jobs can simply transfer to a cloud provider. Even if your entire data center resides in AWS, you’re still going to need local network admins to maintain connectivity, and your virtualization administrators will probably become your cloud admins. As for the others, you may be able to redirect those staffing lines to more business-facing technology roles, or potentially eliminate them. Of course, most companies aren’t operating 100% in the cloud, so simply eliminating critical roles is not an option.

OpEx-based Economics

With someone else handling the infrastructure, you’re basically renting it in exchange for a monthly fee. No upfront CapEx spend means everything becomes an OpEx play. For CFOs, this can be a dream come true since ROI calculations become far easier; you just pay for what you use and nothing more.

Here’s why this has become a key driver for cloud: as budgets constrict, CFOs want to wring as much value as humanly possible out of everything they buy and they don’t want to wait years for that ROI. Consider the data center hardware lifecycle. Most of your equipment is probably on a three- to five-year cycle and, at the beginning of a new cycle, you procure equipment that will get you through that replacement cycle. You buy what you need for today and enough overhead to grow into over a period of years.

You end up with a significant amount of infrastructure for which you don't have any ROI for a long time and you might never see ROI on some of it if you don't grow into it.

Easy and Granular Scale

This one is easy and is associated with the previous point. With cloud, as you need to expand, you just do it and pay the delta cost between what you have now and what you're adding. There is no need to wait weeks for new expansions to arrive and be integrated into what you already have. Just as importantly, there is little to no effort involved when you need to scale.

Agile Operation

Many cloud services are really easy to use. End users can quickly and easily deploy new services. Here, though, we're focused on services such as AWS and Azure. These services, are really easy to work with for IT pros, especially considering that there is no hardware and virtualization deployment to deal with. Self-service is a reality—new workloads can be spun up as needed (by IT generalists, not storage specialists), and then torn down without penalty.

As has been mentioned previously, a lot of the operational efficiency has come about because cloud providers need all of the economies of scale they can get to maintain their pricing level and to bring them down over time as competition heats up.

Private Cloud

The term *private cloud* is chock full of misconceptions. Many believe that a heavily virtualized data center makes it eligible for the private cloud label.

Not so much.

In reality, you need to consider the full breadth of benefits that are generally realized by cloud before you can bestow this label upon your data center. There are some key characteristics that help describe a real private cloud:

- High levels of automation. Individual tasks are generally automated, requiring little administrator intervention.
- High levels of orchestration. Having the ability for many services to automatically interoperate is a key characteristic of the private cloud.
- Some level of self-service. Do you see cloud provider staff manually carrying out your every whim? Probably not. You rely on the self-service tools that have been deployed. Likewise, in order for your data center to be cloud-like, you, too, need some level of self-service capability.

Hybrid Cloud

If you're 100% in the public cloud, your organization has just a public cloud environment. If you are running everything 100% in your local data center and you've implemented some level of orchestration, automation, and user self-service, you're a private cloud. If, however, you're doing both—that is, consuming some public cloud services while also retaining your private cloud environment—you're operating a hybrid cloud. As the name implies, hybrid cloud is a combination of public and private cloud services.

Organizations need the benefits of both the public cloud as well as the private cloud and will want the option to deploy different workloads into different environments. Where it makes sense, companies will deploy workloads into their local private cloud, but in other cases, will need to deploy workloads into public cloud environments. In addition, companies are procuring services from the thousands of software-as-a-service providers, such as Salesforce.com and Microsoft Office 365, and will continue to do so. With that in mind, it's clear that many organizations are seeking a hybrid scenario under which they can deploy workloads into different environments depending on their business needs at the time and the needs of the new application.

This is not just a random thought, either. In fact, IDC estimates that, by 2020, 48% of IT spend will be on cloud infrastructure, with public cloud spending leading private cloud infrastructure spending.

Simply put, hybrid cloud is the future of IT. It's highly unlikely that we'll see everyone simple move everything to public cloud providers, but it's just as unlikely that we'll see pure private cloud environments carry the day. With that in mind, the hybrid cloud is the focus of our next chapter.

2

Current Trends Impacting Hybrid Cloud

You learned previously that the hybrid cloud—a compelling combination of hand-picked public cloud services combined with strategic retention of a local private cloud—is the future of IT. Although there is a lot of focus on the word *cloud* these days, cloud is not really an endgame. In reality, the journey toward cloud principles is the key driver for organizational improvement. In this chapter, you will learn about how current trends—part of the journey—are impacting people's hybrid cloud plans.

Operations Trends

The data center is a trendy place right now (it feels kind of weird and awesome to write that). There are a number of local data center trends that deserve to be called out for their role in helping organizations bring cloud-like characteristics to the local data center.

Increasing Virtualization

Virtualization does not equal cloud. Yes, we said this earlier in this book as well, and it's still true. However, virtualization is perhaps the most critical element in getting to an environment that can be considered cloud-like. As you look at cloud providers out there, one fact becomes abundantly clear:

All of them run workloads that are virtualized in some way. Amazon's workloads run on Xen; Microsoft Azure's workloads run on a combination of Azure VMs and Hyper-V.

Why is this? There are a whole lot of reasons, and here are the most important:

- **Hardware efficiency**. Virtualization has made it possible to push hardware to its very limits. This is a good thing and it reduces overall hardware cost since you don't need to buy individual servers for every workload.
- **Abstraction**. From an operation perspective, the ability to abstract operating systems and applications from hardware has made it possible to consider underlying hardware as almost an afterthought. As long as it has the capability to support its workloads, it doesn't matter on which system workloads actually operate. Here's the key item to remember: abstraction turns hardware-based servers and components into software. Software is far more easily managed and manipulated than hardware, which brings us to…
- **Workload mobility**. Enabled by abstraction, mobility is the ultimate outcome from virtualizing workloads and has imbued supported workloads with the ability to be shifted to new platforms and data centers for availability and copied to various places for disaster recovery purposes.

Continuously increasing the level of virtualization for new workloads enables more and more automation and efficiency in data center operations.

Analytics-Driven Monitoring and Planning

You can't make good decisions without good data! Executive teams have known this forever and they're always looking for good organizational metrics. In more recent years, however, this kind of thinking has come to IT infrastructure as well. Today, IT decision makers consider what's actually happening in their infrastructure environments in order to plan for the future and to react to operational performance issues. The ability to access key performance indicators at *every* level of the infrastructure and application stack is absolutely critical to maintaining a high level of performance and avoiding downtime.

Autonomous Operation

The traditional data center requires a ton of attention from a variety of people, each with vastly different, and often expensive, skill sets. The ultimate destination for the local data center, however, is one that manages itself without the need for a lot of people.

This is not to say that there won't be an IT staff that has to take some part in managing the data center, but it does mean that the skills that these IT pros bring to bear may look very different than the ones we see today.

Autonomous data center operations is actually viable, too… when you fill your data center chock full of virtualization and then sprinkle on a dash of analytics. With the right foundation—virtualized—combined with the right level of analytics, it is possible to have a data center that is able to handle its own routine operations. For example, is your CRM application being overloaded by too many new connections from potential customers? Allow your autonomous data center to spin up a series of new virtual web servers to help handle the load. Is your key customer-facing application beginning to run out of storage capacity? Allow your autonomous data center to provision temporary capacity in the public cloud or move low-value data from the local data center to a cloud provider for archiving to free up local capacity.

Industry Trends

Operations trends are one thing, but there are several overall industry trends that are important to understand as organizations make their way to the hybrid cloud.

DevOps

A portmanteau of "development" and "operations", DevOps is a movement that joins together development processes and infrastructure with eventual product outcomes. Under a DevOps-driven methodology, the infrastructure becomes a part of the development process, which streamlines the lifecycle and leads to more frequent deployments with higher levels of success. The ultimate goal of DevOps is to increase an organization's overall business performance.

According to Puppet Lab's 2016 State of DevOps report, highly performing organizations get a number of benefits, including:

- 200 times more frequent code deployments
- 2,555 times faster lead times
- 24 times faster mean time to recover
- 3 times lower change failure rate

All of this leads to lower costs, increased efficiency, much faster time-to-value, and increased revenue.

In a DevOps world, you may hear the phrase *programmable infrastructure*. In essence, programmable infrastructure is generally software-defined and enables developers to treat various infrastructure components just like they would treat a software object. They can treat infrastructure as code, thus enabling their software creations to fully manipulate the very foundation on which those creations operate. Remember the operations trend of autonomous operations?

Web-services

Because there were not enough confusing terms in the IT landscape, someone decided to throw another one into the mix. We're used to big, monolithic architectures, but we also know that such environments are far from flexible and agile. As we make changes to such systems, we risk unintended consequences. Here's a little limerick for you:

99 little bugs in the code.

Take one down, patch it around,

127 little bugs in the code…

Unfortunately, that's a reality for a lot of developers of big systems. To try and combat this and to make it easier to develop systems, web-services have entered the equation.

Loosely coupled with one another, the intent behind web-services is to enable modularity and separation of core functionality in order to improve security and to vastly accelerate development efforts. Web-services depend heavily on APIs that enable deep interfacing between components. They don't need to all be built and managed by a single company, either. Many of today's applications and services have rich APIs that enable inter-product communication, so many web-services from many sources can actually interoperate quite nicely. Of course, a single company can choose to leverage web-services for all its development, too. The beauty of a web-services framework is that it enables discrete development on small components. This is the part that makes it possible to very quickly iterate and accelerate the introduction of new functionality into a development project. There is less testing to do and less that can go wrong.

Containers

If you like virtualization because it allows you to cram more workloads onto a single server than was possible in the glory days of the physical server, then you're gonna love containers! And, if you're hooked on web-services, you'll quickly see why containers have become so popular at DevOps parties.

Imagine, if you will, a world in which you were forced to deploy a new virtual machine for every web-service component. After all, just like you do with monolithic applications, you might want to keep your individual web-services running individually as well, to improve security and performance. With the need to create a software-based server (a virtual machine) and install an operating system, patch it, and then deploy the web-service, you're looking at a *lot* of overhead, as you can see in **Figure 2-1** later in this chapter.

With small web-services, the overhead gets pretty intense. You need a more granular way to deploy these services. That's where containers come into the picture. Rather than a separate operating system deployment for every virtual machine, containers all run atop the same operating system instance. On top of that operating system runs a container engine—Docker is the most popular and well-known such engine—and each individual application or web-service gets its own access to the binaries and libraries that support it. **Figure 2-1** gives you a look at how virtual machines and containers compare.

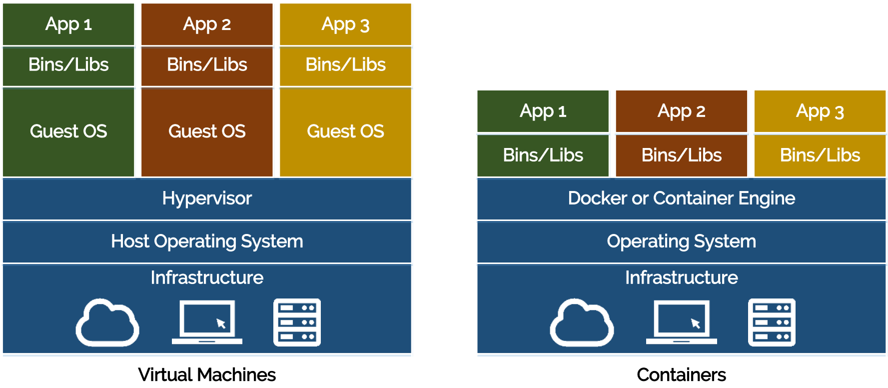

Figure 2-1. Virtual machines vs. containers

OpenStack

Everything you've heard so far is great, but for one thing. You need a place to run all this stuff. Further, you want that place to be something that gives you cloud-like capability and resources, including compute, storage, and networking, among other components and services. Well, back in 2010, NASA and Rackspace walked into a lab one day and walked out the next day having created the platform known as OpenStack.

OpenStack provides organizations with an infrastructure-as-a-service offering that is fully manageable via a command line, a GUI, and via powerful APIs. Every six months, the OpenStack development community releases new updates to correct problems and add new functionality. OpenStack is a highly componentized system, enabling fast development on individual modules without the risk of those development efforts impacting other modules.

Since this is a book about storage, let's focus there for a minute. OpenStack's Cinder service provides shared block storage functionality to the platform and also provides the capability for third party storage providers to provide plugin drivers. These drivers enable OpenStack to leverage that company's storage assets as if they were a core part of OpenStack itself.

Why is this important? In an open world, one size does not fit all. The people behind OpenStack understand that other solutions can provide far more functionality than the platform itself can ever do. For example, through a Cinder driver, OpenStack adopters can deploy a VM-Aware storage solution, which operates directly at the virtual machine level, eliminating the need for complex mapping to LUNs and volumes. This dramatically simplify and improves how you manage storage in an OpenStack cloud deployment.

Serverless Computing

Finally, let's talk about the trend known as *serverless computing*. Let's get one thing out of the way, though. Servers are still involved, but apparently some marketing person won the battle of the moniker. In reality though, serverless computing is much more about abstraction than it is about actually eliminating the use of servers. The goal is to enable developers to focus on their code rather than on infrastructure, which includes both physical and virtual servers. Today, developers often have to worry about these elements, but this worry shouldn't be necessary.

Let's look at a real-world serverless computing service—AWS Lambda. With Lambda, you're able to upload code to the service and execute this code without having to provision the underlying resources, such as virtual machines. The service itself silently and automatically provisions whatever resources are necessary to carry out your will. And then, once the code is done executing, all the virtual infrastructure that was provisioned is destroyed. Under such a service, you're only paying for the resources that were in use while your code was running.

Why is this important? Think about the model of simply shifting your virtual machines to a cloud provider. If those virtual machines are running 24/7, you're paying 24/7, even if you only use some of them once a week.

Under a serverless computing model, you can move ever close to a true "pay as you go" economic model for workload operations.

Summary

We talked a lot about trends in this chapter. All of this is great in theory, but we need action! In Chapter 3, you'll learn about what makes the hybrid cloud actually tick.

3

Hybrid Cloud Requirements

We've talked about the journey and how we got to this point. We've chatted a bit about the trends that are assisting with a move to the hybrid cloud. Now, let's talk a bit about what needs to happen for the hybrid cloud to actually come to fruition.

Storage Platform and the Public Cloud

In any environment, data is king. An organization's data is its lifeblood. As a result, storage is the only non-ephemeral resource in the data center. Compute and even networking frameworks can be built and destroyed at will, but storage requires much more care since it often needs to live forever (or, at least, for a very long time).

Regardless of environment, storage must be well-supported and there must be permanent mechanisms in place for its safety and retention. Moreover, because storage persists for a long time, it also needs to be protected from prying eyes. The financial and reputational backlash in failing to do so can be devastating.

Beyond ensuring ongoing security, storage in the hybrid cloud must be flexible when it comes to where it lives. After all, the very nature of the hybrid cloud is that it straddles both public cloud providers and private cloud environments. With that in mind, a basic assumption behind hybrid cloud storage is that it has at least some hook into the public cloud. For some vendors, this may mean running software-based versions of their operating system in a public cloud provider to enable a full feature set both on-premises as well as in the cloud. For others, it may mean having the ability to use the public cloud as a replication or data protection target. And of course, it may be a combination of these kinds of features or evolve into something else entirely.

Scalability

One of the key characteristics of cloud—whether it's public, private, or hybrid—is the ability to scale in whatever dimension makes sense for your workloads' needs. The traditional direction for storage scaling was up—adding more shelves of capacity as needed. But now that capacity generally comes with a lot of all-flash IO, and so the controller becomes the bottleneck. And piling up shelves of capacity only creates a growing failure domain—lose the controllers and lose the data on ALL the shelves to which they are attached.

Inside the Data Center: Scale Up vs. Scale Out

For years, a silent war has been waged inside your data center… one with the hallmarks of some of the greatest historical matchups: Coke vs. Pepsi, Mac vs. PC, Kirk vs. Picard. In this war, however, unsuspecting organizations are the losers.

Scale up vs. Scale out: To the untrained eye, both are just ways to add more capacity to an existing storage environment.

Scale up architectures allow you to add more disks to an existing array or add expansion shelves of disks atop a processing unit, which contains the processing and memory resources for the storage environment.

In a scale out system, every time you add capacity, you also add the underlying resources upon which that capacity depends. Add a shelf of disks and you also add more processors, RAM, and network/storage fabric ports.

This is really important and here’s why: Predictability. As you add storage capacity to the data center, you shouldn’t have to give up performance, but that’s exactly what can happen in a scale up environment. Eventually, you overtax the shared processors and fabric connections and you begin to suffer from storage performance problems. When that happens, you must start being careful about where you’re placing workloads and virtual machines so that you avoid the hot spots.

In the proper world of scale out storage, you don't run into such issues since you’re adding *all* of the resources your storage needs to function. You can scale workloads and maintain predictable levels of performance. Even better as you add new workloads and virtual machines, the right system can optimize placement of these items without worry that necessary resources won’t be available.

Happy users, happy business.

And so more organizations have shifted the direction of their scale plans to out. This entails adding more nodes—interconnecting arrays to join their controllers together, and adding both capacity and performance. This adds redundancy of controllers and increases your ability to put all that performance to effective use.

The challenge here is to keep storage scale-out as simple as say… scaling compute. If you need more compute resources for your cloud workloads, you simply install a new virtualized server, add it to the resource pool and automatic live migration optimizes the VMs across the pool.

Fortunately, the previously mentioned category of VM-aware storage has this visibility across your pool and capability to optimize the placement of every VM. Simply add another node of VM-aware storage, and intelligent algorithms re-distribute VMs to best balance capacity and performance requirements. Best of all, it all happens automatically without admin intervention.

Now remember that scale also means spinning workloads up and down. When we're talking about scale in the context of cloud, bear in mind that it should be possible to expand the size of a resource pool as well as shrink resource pools when workload demands begin to diminish.

But, there is a real difference in achieving scale in a public cloud as opposed to a private cloud, which means that the two sides of your hybrid cloud will have somewhat different capabilities here. On the public cloud front, you can temporarily provision a resource and then, when you're done with it, deprovision that resource on the fly without incurring any capital expenditure costs. On the private side of your hybrid cloud, if you run out of a resource, you'll need to buy more and deploy it before you can provision it for use. If you have a temporary workload, even if you deprovision the resource when you're done with it, the resource still physically exists in your data center and you've paid for it. In a private data center, by deprovisioning resources, you're making them available for use by other workloads.

Put concisely, on the public side, you get the full cloud experience, which includes both economics and operational efficiency. On the private side of your hybrid cloud, the focus is on making things easier to expand and contract and to make more efficient use of your resources.

Automation and Self-Service

In cloud, automation of routine tasks is a defining characteristic. In any hybrid cloud scenario, you should be able to automate the deployment of new resources, such as virtual machines, networks, and storage resources. Even if you're not a DevOps-drinking developer, you can still adopt some of the infrastructure-as-code methodologies that are bringing new possibilities to the application workload environment. For example, you can create self-service systems that enable individual business units to provision their own resources. Or, consider this scenario: suppose you're working at a college and want to make sure every new student gets their own virtual desktop and dedicated storage space. You can write a routine that creates these resources automatically each time someone in the Admissions department accepts a new student and changes their record in the database.

If you haven't enabled deep automation in your data center, you don't have a hybrid cloud. Remember, self-service is one of the virtues of the cloud; without automation, you can't introduce self-service into the environment. This doesn't necessarily mean that every user in your company needs to be able to provision resources, but it does mean that the ability to use resources should not be limited to just the operations staff. In a DevOps world, this capability should extend to, at a minimum the development staff.

Summary

These are just some of key requirements for a hybrid cloud in general. In the next chapter, we'll go deep in what's necessary for the storage resource.

4

Hybrid Cloud Storage Foundation

Up to this point, we've talked about cloud and storage in relatively general terms. But here's where we get into the details and focus on storage challenges today and the foundation of hybrid cloud storage.

Challenges with Legacy Block Storage and Hyperconverged Infrastructure

There is no shortage of storage options at your disposal in today's marketplace. These choices include legacy block storage systems and more modern hyperconverged infrastructure solutions. While these solutions have enjoyed a lot of positive feedback, they do come with some associated challenges that may not make them suitable for cloud requirements.

Physical-era Architecture

Although there are some people in the world who truly love managing storage and all the elements that come with it, the reality is that the constructs we've gotten used to are there due to inherent limitations that were present in legacy storage systems. We accept these constructs because *that's the way we've always done it*—as poor an excuse as any.

There are two common constructs in use in legacy storage systems:

- Logical Unit Numbers (LUNs). In the days of physical storage, LUNs represented physical devices that existed in the SCSI bus. Since then, however, LUNs have also become a bit virtualized themselves and can represent specific portions of carved up RAID groups as just one example. They have evolved to represent logical groupings of storage elements to hosts. LUNs are generally SAN-centric and are presented to hosts for the creation of volumes.

- Volumes. Volumes are areas of storage capacity carved out of one or more LUNs. Generally, volumes are created on host systems, such as vSphere hosts.

These terms are sometimes used a bit differently by different storage vendors, but these are the basic ideas. The end result is the same, though—they suffer from underlying fundamental architectural challenges.

In the modern data center, there are several needs that simply can't be easily met with physical-era architectures. These include:

- Simple scalability of both capacity and performance (addressed in the previous chapter)

- Predictable performance as the environment grows—Physical-era infrastructure in a scale up environment or stuffing too many virtual machines inside too few LUNs, making them difficult to isolate, thus subjecting them all to unpredictable performance

- Application of granular quality of service policies, for example, at the individual VM level to isolate VMs and eliminate conflict over resources

- Quick and easy identification of performance problems, including root cause determination of the exact point in the stack (compute, network and/or storage) that is causing the poor performance

As storage systems grow to span multiple data centers and even multiple clouds, these problems become even more profound. While technologies such as VMware's Virtual Volumes (VVols) have been developed to help address some of these challenges, VVols exhibits its own problems:

- Extreme dependence on storage array vendors to implement the complete VVols specification, which is carried out with varying levels of success and attention
- VVols doesn't fundamentally change the underlying architecture; it just masks architectural issues to help vSphere associate virtual machines with disks, which *does* help to solve some problems; VVols can make it possible for your storage to provide more granular per-VM policy management instead of per-LUN management. For example, per-VM controls on performance and redundancy.
- VVols only works in a complete way if you're using Sphere 6 or higher; if you're using an older version of vSphere or another hypervisor altogether, you're out of luck

That last point is particularly important to understand. While VMware's vSphere today enjoys a commanding market presence, ActualTech Media research shows a surge of interest in alternatives, most notably Microsoft Hyper-V. With the risk of VMware beginning to lose share, there are potential challenges for storage vendors that depend solely on VVols.

Storage is Not a Web Service

Storage is really important. In fact, of the resources in the data center, it's the *most* important. Why? It holds the keys to the business kingdom. As a business, if a server completely fails or if your network crashes, you can recover, even if you haven't done a good job building contingency plans. If your storage suffers catastrophic failure and you've not put into place contingency plans—e.g. backup systems—you're essentially out of business.

As we look at some of the emerging storage systems on the market, the components responsible for managing storage are not always as robust as we might like. For example, although there are hyperconverged infrastructure systems on the market that are quite good, there are others that have bolted storage on to the platform. In a data-centric world, companies can't afford to rely on systems that treat storage as "just another service." Moreover, as the market continues to extend support for container-based workloads, new challenges emerge, such as the need to maintain storage persistence inside these ephemeral constructs. More robust storage layers that can provide a persistent storage layer for non-persistent containers will be an increasing need as container adoption accelerates.

APIs are Not Clean

Traditional storage arrays lack any programmatic management. They are managed by a proprietary interface and, other than storage protocols, interface with no other external services than perhaps logging and statistical gathering.

Policy-based management and automation are a critical piece of any cloud construct. As companies move to leverage enterprise cloud, they require storage systems that support the automation necessary in a cloud model.

By using modern storage that provides a RESTful API you'll be able to integrate your storage with automation/orchestration tools, provisioning, chargeback, and more.

Infrastructure Components Must be Welded Together

When you look at a legacy data center environment, it quickly becomes clear just how much welding you must do to make things work in a reasonable way. Even when things are designed to interoperate, you still need to create all kinds of constructs—such as the aforementioned LUNs and volumes—to make it all work.

There are obvious seams around these welding points, too. These are the seams that create security concerns, but that is just a small part of the problem. Instead, each of these welding points results in additional friction in the data center. Each is a point that needs to be touched all the time.

Data center friction causes latency in achieving the goals of the business and drives business units into the waiting arms of outside providers that can move faster. The hybrid cloud storage environment requires a storage architecture that is a smoother and more flexible than traditional physical-era environments allow.

Hyperconverged Infrastructure Linear Scale and Lack of Choice

Hyperconverged infrastructure has emerged as a popular way for organizations to quickly implement data center environments. While hyperconvergence is a good solution for many, for others, it carries with it some inherent challenges.

First, the need to scale hyperconverged infrastructure node resources and software licenses in a linear way is a detriment. You're adding more compute power than you probably need as you add more storage capacity. This scale methodology may sound like scale-out storage, but there's a key difference: hyperconverged infrastructure nodes often have less overall capacity than storage arrays, so you're adding *way* more compute than you probably need. In addition, every time you add a new node, you also have to pay for another processor license for your favorite hypervisor. Those costs can add up really fast.

Moreover, you don't get choice in your storage. You have to use what you're provided and may not have the flexibility to choose an optimal combination of compute and storage.

Native Hybrid Cloud Storage Requirements

Although physical-era storage environments can coexist with and support hybrid cloud storage initiatives, in terms of efficiency and flexibility, a lot is left on the table. In the world of the hybrid cloud, virtualization is a core component and, as such, the use of a filesystem purpose-built for virtualization is important.

The sections below expand on this thinking.

Virtualization-specific Architecture

As the saying goes, "you can put lipstick on a pig" but, at the end of the day, you just have a pretty pig, not an awesome cheetah. You're still hampered by the limitations inherent in the underlying animal. In the world of storage, you can add all the extensions you like to a legacy architecture, but you will never be able to get the same level of benefits that you do from a modern storage system.

Rather, it's important to customers to look at storage systems that don't depend on uncontrollable and inconsistently delivered API implementations—VVols is an API implementation—to deliver value. VM-Aware storage platforms are far better positioned for hybrid cloud storage environments since they eschew complexity in favor of simple implementations that increase performance and manageability. As organizations seek to implement hybrid cloud architectures, improved performance potential and easier administration are significant parts of their journey.

Cloud is a Journey

Consider the Cloud. Often, when thinking about the Cloud, Amazon and Azure immediately come to mind. However, for many organizations, there is hesitancy around simply throwing workloads over the fence to these public cloud providers. Rather, organizations considering the Cloud are really looking for the kinds of *outcomes* that are achievable with public cloud services, such as improved economics and frictionless operations. For them, the Cloud is not about a destination, but about a journey of improvement. You'll learn more about this in the next chapter.

Automation, Orchestration, and Clean APIs

Legacy storage architecture is, well… not friendly. Much has been written about the complexity of legacy storage systems and the negative impacts that this complexity has on IT and the business as a whole. Remember, your storage is the lifeblood of your business. If you have difficulty with that resource, it will, thanks to the hooks it has into everything you do, create challenges for you.

Modern storage with a virtualization-specific architecture eschews this complexity, which also makes it easier to automate. Comprehensive automation is a key tenet of the Cloud and, without it, you can't implement higher order features, such as self-service capabilities, inter-system orchestration, and DevOps-friendly capabilities. In **Figure 4-1**, note that automation drives orchestration and self-service.

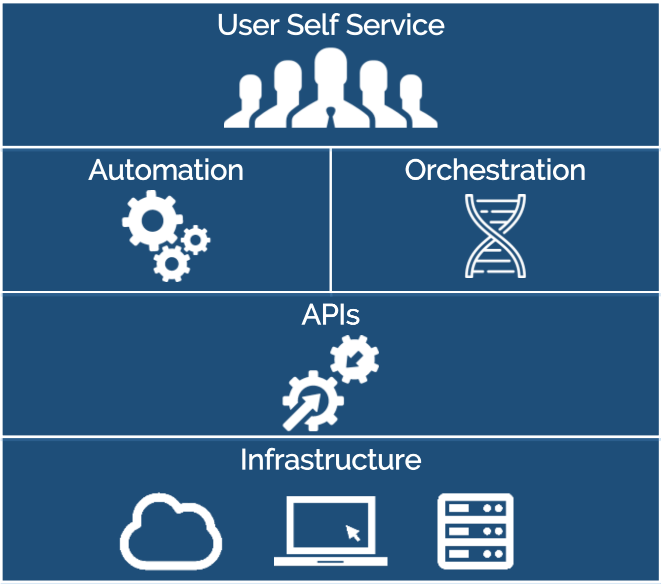

Figure 4-1. APIs, automation, and orchestration are key characteristics of hybrid cloud storage services

Clean REST-based APIs are an additional key underpinning for enabling automation, orchestration, and user self-service in the world of hybrid cloud storage. A complete set of APIs enables a storage system to be fully managed without a person clicking a mouse.

Let's for a moment consider a web-based, web-scale application that straddles both a local data center and the cloud. Imagine, if you will, this common scenario: The company that developed this application experiences occasional surges in its use. Perhaps the application is one that supports an ecommerce site and holidays drive additional traffic. During such surges, there is a need to continually adjust resources assigned to support the application. Without automation and orchestration driven by a clean API, this work will need to be handled manually.

In other words, someone will need to keep constant watch over the application's performance levels and storage capacity usage and then make reactive adjustments to resource levels. This reactive, manual method basically ensures that some customers will experience very poor results while they wait for IT to assign new resources.

Now, let's assume that the organization deployed their application atop an API-laden system and one that embraces a modern virtualization-specific architecture. As a part of the application development process, the developer can integrate deeply with the storage environment. They can write routines that keep an eye on storage latency and capacity and, as latency increases due to load or as storage capacity dwindles, enable the application to proactively deploy new virtual machines and storage services without an IT staff person having to be involved.

That's the potential power behind the software-centric nature of today's virtualization-aware storage products.

Unified Model Across Different Components

Virtual machines are just the tip of the iceberg in the modern data center. Today, containers are making a big splash in the data center pond, and with good reason: They are far more efficient than virtual machines.

If you consider the original reasons that people made the switch from physical servers to virtual ones, it was all about resource utilization. With physical servers, you had a ton of hardware to deploy, and you only ended up using between 5% and 15% of that hardware's potential, on average. There was just a ton of wasted capacity. With virtual machines, we used the same hardware, but we could cram more workloads onto it because we abstracted that hardware and turned each server into a software construct, eliminating a lot of hardware overhead and massively increasing the utilization of the hardware.

Virtualization brought relatively high levels of efficiency and flexibility to IT, particularly when compared to the world of the physical workload. With virtualization, companies can run operating systems from many vendors and different versions of each of those operating systems on the same hardware. Those workloads can be easily shifted between hosts, between data centers, and even to and from cloud providers.

But, for many, virtualization is inefficient. Every single virtual machine gets an operating system. Think about that for a second. If you install 100 identical virtual machines to support an application, that's 100 individual operating systems running in your environment. Every operating system imposes steep overhead by consuming CPU, memory, and storage resources.

To combat this inefficiency, *containers* have hit the market. Containers have an implicit assumption that the operating system for many environments is, in fact, the same, and can be safely abstracted, thus eliminating the overhead of individual operating systems. This has the immediate effect of increased density in the environment since you can cram even more workloads onto the hardware.

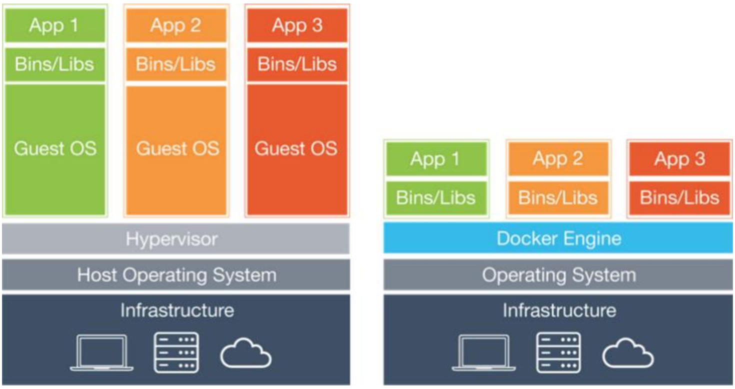

Figure 4-2: Containers vs. Virtualization

In terms of hybrid cloud storage, it's important for this resource to be able to equally support virtual machines as well as containers while continuing to enable deep understanding of performance bottlenecks. Containers will continue to increase in usage, particularly for web-scale applications and for tools that leverage web services, which we discussed earlier in this book). They provide a more efficient operational framework than is offered by virtual machine paradigms.

Containers in the Real World

Containers are still in their infancy and will undergo the same interest, scrutiny, and deployment challenges that initially faced virtualization, but with some twists. Whereas virtualization generally enabled companies to simply drag and drop applications from the physical environment to the virtual one, with containers, this is not the case. Applications must be purpose built for containers, at least for now. Over time, expect to see more tools put into place to enable more seamless support for migrating workloads to containers. For now, however, understand that the market is still digesting exactly how containers will ultimately fit into the IT landscape of the future.

Summary

Understanding where the storage market used to be and where it's heading is critically important to understand the eventual vision and potential of the environment. This is the topic of the next chapter.

5

Hybrid Cloud Storage Vision & Potential

Cloud as a Journey, Not a Destination

We touched on this in a Chapter 4 sidebar and even covered some back in the beginning of this book, but it's so important that it's worth a little repetition:

Companies don't want "cloud." They want the benefits that come with cloud. If that means they must adopt cloud services, they will, but it doesn't mean that the public cloud is the only option.

The Cloud as a public service inherently provides a number of benefits that have been discussed in this book. From turning traditional data center economics upside down to transforming how new services are deployed, the public cloud has changed how people think about the infrastructure upon which IT operates.

However, the outcomes are a journey, not necessarily a destination. If you're able to provide your company with improved IT economics and can feasibly implement the kinds of architectures that enable simple workload deployment, you will have done your job. With hybrid cloud storage systems offering robust APIs and a virtualization-first methodology, you will be well on your way to cloud success.

Let's dig a little deeper into the potential vision and outcomes that you can expect with the architecture described above.

Manage a Cloud

Once you have decided that a hybrid cloud is for you and that you need cloud-like characteristics embedded in your IT function, there are a number of achievements that you can unlock for your enterprise.

DevOps

We talked a bit about DevOps earlier on in this book, but it's emerging as a key consideration for how IT moves forward with data center architecture and infrastructure, or lack thereof (cloud). As you consider undertaking a hybrid cloud journey, consider the potential good that is imbued into your DevOps environment and processes. A fungible infrastructure solution, such as one based on cloud principles, enables an agile development environment for developers, who can now create and destroy virtual infrastructure on a whim in support of their development and testing efforts.

The goal here is to enable an agile environment to help push the business to meet and exceed its goals. This can happen with local infrastructure, but the environment can also extend to the public cloud with the development and support of cloud-native applications that seamlessly integrate with local applications.

Self-Service

Let's use a development example again. Developers are constantly in need of fresh infrastructure as they seek to develop bug free software that operates with high levels of performance. As they sully good virtual machines with test versions of their code, they need to continuously refresh the environment. It's doubtful that even small development shops want to constantly involve IT operations for these needs, so they need a way that they can create and refresh their own environments.

Likewise, individual business unit IT specialists may need ways to create centralized services that reside in the corporate data center. Although they may not be a part of the central IT staff, they want the ability to meet their own regular needs without always involving IT.

Such on-demand, self-service activities are increasingly important. On the storage front, enabling self-service capabilities requires the creation of templates that specify, among other things, quality of service (QoS) and data protection policies. As IT, you create a series of templates ahead of time and then allow users to simply choose from them. Behind the scenes, your API-driven automation and orchestration services do the heavy lifting.

In a multitenant environment, you want control of Quality of Service (QoS). But legacy storage can only set QoS at the LUN level, and that leaves all the resident VMs still fighting over assigned resources. Contrast that with QoS per-VM, enabled only by VM-aware storage. QoS at the VM-level allows you to guarantee resources for an individual VM; you can set a maximum ceiling on a rogue VM or a minimum floor for a critical VM. Cloud Service Providers can use QoS per-VM to establish performance tiers and premium services for their customers.

Chargeback / Showback

IT resources aren't free in any economic model, whether that's an OpEx-heavy cloud model or a CapEx-intensive traditional model. At some point, those services need to be paid for. For a variety of reasons, organizations often prefer to make sure that individual departments are charged for their data center activity. For example, if the sales group requires ten virtual machines in the data center environment, they cover the costs of those resources. This paints a much clearer picture from an organization-wide budgeting perspective than does the model where all data center purchasing is a part of the IT group's budget.

Chargeback vs. Showback

Chargeback is the process of actively billing business units for their resource utilization and is a common central IT cost recovery model. Showback is "chargeback light" in that it shows a business unit how much they're using in the way of resources, but doesn't actively bill them.

In a public cloud environment, it's easy to decide who pays for what since cloud provider environments are, by their nature, multitenant-oriented with comprehensive billing capabilities. In the virtualized world of a local data center, it can be a bit more challenging sometimes, particularly with legacy storage systems that aren't capable of neatly delineating who is using which resources. In a modern storage environment that leverages VM-aware storage, it can be much easier to identify in-use resources and map them to their consumers.

Autonomic Operation

Raw computing power continues to increase as Intel releases processors jam-packed with more and more cores. With all this compute power, there's plenty to spare for a core or two to keep a watchful eye on the data center environment. And, with an environment that has a lot of great APIs and automation and orchestration capabilities, we can do some really interesting things.

For example, if you're a typical company, you probably have people that work from 8 AM to 6 PM and that's your heavy period. Or, you may be an ecommerce company that experiences a surge at certain times of the day, but you're relatively quiet at other times of the day.

What if you could create an environment that simply managed itself based on the workload? This goes beyond the application-centric DevOps scenario that was discussed earlier. In this scenario, the infrastructure senses that there's no need for all thirty vSphere hosts to be operational since the current workload requires only eight. As such, the management layer migrates all of those workloads to a few hosts and shuts down the rest to save electrical and cooling costs. As demand dictates, hosts are brought back online and the virtual machine workload distributed accordingly. The same kind of process can be used with the storage layer, as long as that storage layer has some intelligence.

You can even create an environment in which test and development data is automatically refreshed from production at certain intervals so that your developers are always working with current information. This is sometimes referred to as copy data management, but many providers have even more sophisticated interpretations that reduce operational overhead and storage consumption. Copy data management refers to a process in which the storage system manages copies of data by creating virtual copies of data as they are needed. New data copies aren't actually physically created, thereby reducing overall capacity needs. These capabilities are useful for development needs as well as for times when you need to create a copy of a virtual machine to, for example, perform point-in-time data analysis.

Summary

With a virtualization-centric architecture, VM-aware storage can help companies achieve these goals. By leveraging a public cloud-like web services architecture and clean, complete APIs, VM-aware storage enables greater automation, orchestration, scale and self-service than is possible with legacy storage solutions. Companies that are investing hybrid cloud are investing in these kinds of technologies to realize the full potential of their cloud vision.

The hybrid cloud isn't just a financial play, although the right technology *can* help to rein in capital and operational expenses. Rather, it's about the future of the business and the ways that you can best support your customers, employees, and partners. It's about possibilities and potential—the potential to propel your business ahead of the competition; the potential to eliminate any silos that may exist; the potential to transform the environment into an unstoppable force that can adapt as quickly as the business environment demands.