

The Storage Journey: From Virtualization to Cloud

Author: Scott D. Lowe
Partner, ActualTech Media
November 2016

© 2016 ActualTech Media

Introduction

“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 as bad 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.

Contents

Introduction.....	1
Contents	1
The Path from Virtualization to Cloud.....	2
Key Milestones Along the Way	2
The Cloud Conundrum.....	3
A Tale of Three Clouds	4
Public Cloud.....	4
Economies of Scale.....	5
OpEx-based Economics	5
Easy and Granular Scale.....	6
Agile Operation	6
Private Cloud	6
Hybrid Cloud.....	7

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,

Flash Forward

Performance problems? 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.

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.

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 punniness—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 simply 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 paper.