Making Everything Easier!™

Nutanix Special Edition

Enterprise Cloud

FOR DUMMIES®

A Wiley Brand

Learn to:

- Prepare your organization for the new era of IT infrastructure
- Get the frictionless benefits of public clouds without sacrificing the control of on-premises infrastructure
- Leverage new best practices for enterprise storage, virtualization, and cloud technologies

Brought to you by

NUTANIX™

Scott D. Lowe

About Nutanix

Nutanix makes infrastructure invisible, elevating IT to focus on the applications and services that power the business. The Nutanix enterprise cloud platform leverages web-scale engineering and consumer-grade design to natively converge compute, virtualization, and storage into a resilient, software-defined solution with rich machine intelligence. The result is predictable performance, cloud-like infrastructure consumption, robust security, and seamless application mobility for a broad range of enterprise applications. Learn more at **www.nutanix.com** or follow @nutanix on Twitter.

Enterprise Cloud

FOR DUMMIES®

A Wiley Brand

Nutanix Special Edition

by Scott D. Lowe

FOR DUMMIES®

A Wiley Brand

Enterprise Cloud For Dummies®, Nutanix Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

ISBN 978-1-119-31763-0 (pbk); 978-1-119-31699-2 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Publisher's Acknowledgments

We're proud of this book and of the people who worked on it. For details on how to create a custom *For Dummies* book for your business or organization, contact info@dummies.biz or visit www.wiley.com/go/custompub. For details on licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

Some of the people who helped bring this book to market include the following:

Project Editor: Martin V. Minner
Acquisitions Editor: Katie Mohr
Editorial Manager: Rev Mengle
Business Development Representative: Karen Hattan
Production Editor: Kumar Chellappan

Introduction

If you aren't willing to adapt, you'll die. Evolution isn't only for species anymore. It applies to whole industries, companies, departments, and even individual careers. Legacy disaggregated infrastructure is becoming less sustainable, and IT is under assault from all directions — from the business that demands more, from providers that are often doing IT better, and even from within IT itself.

"The cloud" has become one way end-users are seeking to empower themselves and unshackle their fortunes from IT. Unfortunately, the public cloud is fraught with its own challenges and may not always be suitable (although for some applications, it represents a perfect platform). That's why more and more organizations are turning to enterprise cloud — which carries public cloud support and characteristics — to meet critical business needs.

About This Book

There is more to the cloud than meets the eye. This 48-page journey helps you understand enterprise cloud and how it fits into your datacenter paradigm. By the end of this book, you see just how enterprise cloud fits into the landscape and how it can help you propel your business into the 22nd century.

Foolish Assumptions

For this book, I assume you have at least a basic understanding of virtualization, storage, cloud, and datacenter computing. The general audience for this book is anyone in IT who wants to learn more about how enterprise cloud can help address evolving business needs. The audience is intended to be technical staff as well as managerial and executive staff.

Icons Used in This Book

Throughout this book you find a number of icons intended to help you better understand and remember key concepts.

I use this icon when you need to stop for a second and make sure you recall a key concept before forging ahead in a chapter.

You'll want to keep certain details in mind as you analyze your own datacenter environment. When you see the Tip icon, put that information in your back pocket to save for later.

Although I don't go super-deep into technical stuff in this book, I provide some technical elements for you in various places. You can find these marked with the Technical Stuff icon.

Sometimes you need a little extra nudge to watch out for certain things that can become problems for you. Throughout this book, I point out places where you might need to take some extra care.

Beyond the Book

There's only so much I can cover here. To learn even more about enterprise cloud, keep an eye on `www.nutanix.com`.

Where to Go From Here

Even if you choose to ignore everything else you've read in this book, never forget that, no matter how much money you save, if the users can't get their work done in a reasonable way, IT will be seen as a failure. It's up to forward thinking technologists to figure out how to transform the IT function and make it a leader of the business. Enterprise cloud is a powerful way to help end-users gain the benefits of public cloud without assuming the risks and downsides of that medium.

Chapter 1

Surveying the State of IT for the Enterprise

In This Chapter

- Discovering the key trends that affect the way IT does business
- Learning why flash storage and hyperconverged infrastructure have revolutionized the datacenter
- Finding out how the public cloud can enable IT but creates new challenges to overcome

The challenges facing today's IT function are both familiar and strange. IT faces the same kinds of resource constraints it's always had, but new and different solutions now provide ways to address them. Infrastructure is still a challenge, but recent innovations offer a path beyond the rough spots. This chapter explores the current state of enterprise IT.

Trends Shaping IT Infrastructure Today

In the past decade, IT infrastructure has undergone a revolution stemming from a number of evolutions across various resource silos. These changes have led IT to where it is today and they set the stage for the fundamental transformation that the industry is poised to undergo over the next few years.

From storage to servers to software, no area of the datacenter has been spared.

Flash storage

Not so long ago in a datacenter not so far away, solving storage performance issues was about as likely as a Stormtrooper hitting a target. Storage administrators were often witnessed throwing hardware at a problem. They had to add spindles — more spinning disks — to imbue their storage environments with sufficient IOPS to meet workload demand.

And then a funny thing happened on the way to Tatooine. Flash storage started to become a viable option for the enterprise. As this solid state storage became more popular, vendors began to work in earnest on ways to address the two major issues with the technology: cost and longevity.

In recent years, the cost of NAND-based flash storage has plummeted by double-digit percentages while capacity has increased. Today, when considering a standard disk-based form factor, you can buy an SSD that has even more capacity than a disk. Of course, that 16TB behemoth costs far more than the same spinning disk capacity, but it also means you can achieve all-flash capacity density that is better than that of disk.

Just as important as the ability to leverage flash is the ability to get at data quickly. This is where data locality comes into play. The closer that data is to processors and RAM, the more quickly that data can be retrieved and consumed. This is one area in which even all-flash storage arrays can be challenged. Storage in such environments sits in a separate silo and must traverse the storage fabric, which adds latency to the computation. The farther away from your application the data lives, the greater the latency and the lower the throughput. As you consider flash or hybrid storage solutions for your datacenter, keep this point in mind. A solution that enables data storage right in the server chassis will enjoy far better overall performance than solutions that require data to traverse a slow network.

Many people today still worry about flash "wear" that can cause drives to fail in place. As flash has become a staple of the datacenter, however, the wear concern has become a non-issue for most organizations. Drive manufacturers and array vendors have begun to implement all manner of mechanisms

intended to help keep drives alive. From *wear leveling* — in which a flash controller prevents a drive from pounding the same cells over and over — to *active write avoidance* techniques — such as deduplication and compression, which reduce the need to write data in the first place, the issue of whether a flash disk will fail during its usable life has been practically solved.

The short version is this: Flash is here. It isn't going anywhere. It's fast; it's durable and dependable. And it's becoming more affordable every month.

Software-defined functionality

At the same time that flash storage has become common in the datacenter, Intel has continued to release processors with massive numbers of cores just begging to be set free. The plethora of computing performance is being wrangled into submission through the use of powerful software tools, which are steadily replacing functions that used to be handled solely in hardware.

Why is this change important? In most cases, customized hardware is expensive, particularly when the hardware is performing a task that can easily be solved by using a commodity CPU with software. ASICs and FPGAs require occasional respinning — or updating — to remain viable. Over time this solution becomes expensive, particularly when the functionality can easily be replaced with a pure software component.

Today, we're seeing the rise of what has become known as the software defined datacenter (SDDC), a phenomenon enabled by commoditization of hardware. SDDCs allow far greater flexibility in datacenter configuration while also helping to reduce overall costs.

Hardware commoditization

Remember when I mentioned that Intel processor in the previous section? Well, that company is at the core of another revolution in the datacenter: hardware commoditization.

These days, you find all sorts of storage arrays that look practically identical to servers, and there's a good reason for that: They *are* servers. Rather than build a bunch of custom hardware and spend all their time on hardware engineering, resource-specific siloes — storage and networking — are increasingly turning to off-the-shelf servers and components to power their solutions. In essence, many of today's fastest growing storage and networking companies are truly software companies. They buy existing hardware that makes sense for their solution and build their software around it. Because the existing hardware is standards-based, the storage or networking company can easily swap components out as necessary, which helps a great deal with reducing cost and complexity.

Hypervisor commoditization and the emergence of containers

Back in the early days of virtualization, there was one company — VMware — to rule it all. Today, although VMware is still the leader in the hypervisor space overall, other commercial and open source hypervisor offerings are eating away at VMware's leadership position.

On a feature-by-feature basis, modern hypervisors generally have all the features that organizations really need in order to succeed. Sure, some have some extras here and there, but the capabilities — such as workload migration and high availability mechanisms — that initially drove virtualization adoption are common across almost any hypervisor choice.

Feature-rich hypervisors have led to a scenario in which the hypervisor can be considered a commodity for many organizations. The necessary features are guaranteed to be there, so switching to different hypervisors — such as Hyper-V, KVM, or a variant — becomes feasible.

At the same time, containers are emerging as an alternate abstraction technology allowing applications to be developed, tested, and deployed quickly and easily. It's important for the infrastructure platform of the future to support containerized applications.

The (hyper)converged revolution of compute and storage

Thanks to the rise of flash storage and the commoditization of the compute and storage layers, recent years have seen the tremendous rise of hyperconverged infrastructure. In such an environment, storage and compute — servers — are collapsed into a single unit of infrastructure, effectively eliminating expensive and complex SAN environments.

Hyperconverged infrastructure enables organizations to easily manage and scale their datacenter environments. This architectural option has been a boon for many customers because it's enabled far easier administration of the datacenter and has led to decreased costs and increased end-user and customer satisfaction.

Modern application architectures

If you've never heard the phrase *bimodal IT,* here's a quick rundown for you: IT has dueling priorities these days. First, organizations have a reasonable expectation that IT will continue to support what might be considered "legacy applications." In reality, such applications likely will continue to be mainstays of the business foundation for the foreseeable future. These hardy survivors include client/server enterprise resource planning (ERP) systems, collaboration systems, and local database applications.

These applications traditionally have required a conservative approach to maintenance. As mission-critical applications, they need a rock solid foundation, high availability mechanisms, and a light touch when doing updates, which must be painstakingly planned. The goal is to reduce risk to the business by ensuring that crucial applications are always on. The need to minimize risk to these applications is one of the reasons some IT departments have reputations of being stodgy and unyielding. In fact, the IT group is simply trying to keep the business running. Change makes that a difficult charge.

On the flip side of the equation, a new breed of applications is popping up. These innovative application types might exist in

the cloud, locally, or even as apps. Where traditional applications require deliberate maintenance, the new apps require nimble, agile practices, which are often contrary to what has been considered best practice for application support.

Modern application architectures are driving enterprise IT needs. In the short term, a split has emerged in enterprise IT teams, depending on whether they cater to traditional IT applications or next-gen applications. Organizations must find ways to balance these conflicting goals.

The State of Public Cloud

In the early days of the public cloud, IT departments were quaking in their boots as they foresaw the potential to lose their jobs and their place in the organization. The era of the public cloud was upon us and the hype was real. Doomsayers predicted that IT pros would be out on the streets en masse offering their administration and programming skills to passersby. Businesses would thrive as they saved trillions of dollars in their budgets by eliminating all IT capital expenditures.

The post-IT future hasn't — and won't — come to pass. However, that doesn't mean organizations have eliminated the public cloud from usage. Instead, they've massively increased use of the public cloud as they've discovered applications and use cases where the public cloud makes sense.

But, the industry is far from the doomsday scenario that was envisioned early on.

The growth of all things cloud

Clouds come in all shapes and sizes, and the different options even have cute, little names. Figure 1-1 gives an overview of the various cloud types. It shows which entity — you or the cloud provider — handles each of the elements that comprise the infrastructure.

The industry is increasing adopting the three primary kinds of public cloud: infrastructure-as-a-service (IaaS), platform-as-a-service (PaaS), and software-as-a-service (SaaS). In fact, 451 Research indicates that the cloud computing

"as-a-service" marketplace is likely to triple in size through 2019 (source: `https://451research.com/report-short?entityId=87624&referrer=marketing`).

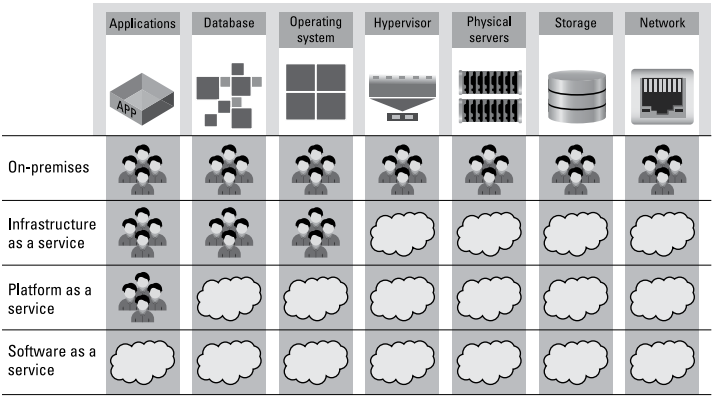

Figure 1-1: Comparing public cloud service types.

Large IaaS providers now deliver platform capabilities, such as databases and message queues, that allow applications to be built quickly using packaged building blocks.

Clouds are great for unpredictable or highly variable workloads because you pay only for what you use. But for more stable or predictable workloads, the cloud is not as economical. Renting is good for the short term or when you don't know what the future holds, but owning is more economical when you know you're going to stay in a place for a while.

In an interesting dynamic, cloud adoption appears to act like a slingshot. For a while, a business builds and deploys an application on a public cloud service. Then, when the application reaches a certain scale or becomes predictable, the business brings it back in-house.

Increasing viability of public cloud

Early on, even with the analyst hype about public cloud decimating IT departments and forcing CIOs out of their jobs,

public cloud providers had to contend with a number of daunting challenges:

- **Bandwidth:** There was — and in many cases, still is — concern around how certain areas of the world are served with Internet bandwidth. Many locales remain woefully underserved, making it difficult to deploy mission-critical services to an environment that relies on an Internet connection. Although this issue is being corrected, improvement is coming slowly. In addition, many places that have decent bandwidth still have only a single connection, which makes cloud somewhat unpalatable. That said, the situation today is *far* better than it was just ten years ago.
- **Loss of control:** At the beginning, the public cloud was an island. You had to manage it with completely separate tools, and a wall stood between it and your local datacenter environment. Today, a plethora of tools exist to help organizations seamlessly manage both local datacenters — private clouds in some cases — and public cloud environments. Control is no longer an issue.
- **Skills:** When any new technology hits the streets, building up adequate skills to support it takes time. Today, with years of experience under their belts, plenty of people with the necessary skills are available to maintain public cloud infrastructure and services.

Understanding security and trust in the cloud

Because security is so important, it gets its own section rather than being relegated to a bulleted list! With regard to security, the public cloud has made massive strides in the past decade.

Security in the cloud is orders of magnitude better than it was in the early days. In fact, many providers make available hardened environments so that they can properly secure sensitive workloads for their customers.

People's willingness to trust the cloud is evidenced by the massive growth of clouds of all types. Microsoft continues to report that growth of Office 365 — Software-as-a-Service — continues to explode, and Amazon is reporting record growth with Amazon Web Services.

The three-letter threat

I frequently do speaking engagements in the United States, the United Kingdom, and Canada. In the U.S., people's concerns about cloud security are quite different from those in my Canadian and U.K.-based audiences. In non-U.S. locales, *data locality* is a major concern. People there fear their data may end up being housed in the U.S. on U.S.-based servers, which could expose their business to spying by the U.S. intelligence community. With that in mind, many cloud providers have located datacenters all over the world. Even many SaaS-based services can be run from these global locations that are housed outside the U.S. As businesses, banks, and governments continue to look for ways to embrace the public cloud, where their data lives is a critical decision.

People are finally realizing that the public cloud is not a threat. It's simply another application delivery option that CIOs have at their disposal. The industry is realizing that, with the right provider, even sensitive workloads can be supported.

Beyond Amazon — Embracing any cloud

Just as Kleenex is associated with sneezing and Google is associated with web searching, when IT pros think about the word *cloud,* they often immediately think Amazon. While Amazon, as the public cloud leader, is certainty formidable, it is far from being the only option available for public cloud consumption.

All kinds of "as-a-service" cloud options that go far beyond Amazon are available to you. Enterprises must always have an exit strategy that enables them to switch providers quickly. If a provider goes out of business or increases pricing to unsustainable levels, you may need to move quickly. You should always have a way to support any cloud, any time.

Chapter 2

Why Enterprise Cloud?

In This Chapter

- Understanding why the *public* cloud is the absolute right choice for all your workloads
- Understanding why the *private* cloud is the absolute right choice for all your workloads
- Recognizing why neither one of these is the right answer and why you need an architecture based on *characteristics* rather than labels

"The public cloud rules all!"

"You'll move my apps out of my private cloud only after you pry them out of my cold, dead hands!"

It seems that people have chosen sides in the public cloud-versus-private cloud debate, and both sides have good arguments.

Instead of looking at this is a public-versus-private debate, it's useful to consider the individual outcomes that each side is trying to achieve and then architect around those.

Frictionless IT: Why Public Cloud

Perhaps one of the biggest challenges facing IT departments is *friction.* As Chapter 1 shows, today's IT departments face competing needs. IT must support legacy applications that require constant uptime and low risk. At the same time,

business leaders want IT capabilities that allow them to test new business models and roll out new services quickly.

Because of the need to ensure ongoing availability and performance levels for legacy and next-gen applications, fooling around with legacy local infrastructure is often frowned upon by the business, which often doesn't appreciate or understand the level of effort that it takes to keep systems going.

That's where the public cloud comes in.

IT as a business enabler

But you need to understand the *why* behind all of this first. IT has generally done a good job *supporting* the business by running tools for enterprise resource planning (ERP), collaboration, and so on.

But so much more is at stake.

With the right technology environment and the right mindset, IT can move from a behind-the-scenes supporting role intended to reduce expenses into a revenue-driving role. IT can become an active *enabler* of the business.

Moving from a support mindset to an enabler mindset takes lots of rethinking of IT infrastructure and services.

Understanding fractional consumption

Consider your legacy datacenter environment. You probably have a bunch of servers, a monolithic storage environment based on SAN technology, and various networks connecting it all together.

On the economics side, you likely prepay for all your resources — and you pay full price even if you use only one-half of the resources during their life cycle.

This purchasing practice is far from efficient. Most organizations overbuy and overprovision resources so they don't run

out mid-cycle. This approach provides good insurance, but it means money is being left on the table.

The public cloud offers enterprises a new consumption model for IT resources. *Fractional consumption* enables pay-as-you-grow economics, which allows you to buy resources on demand. This is one of the primary economic drivers behind cloud technology. CFOs and other financial decision makers love being able to buy only what is needed. However, fractional consumption in cloud also occurs as you *spin down.* That is, as your business needs change, you can reduce your resource utilization and, in turn, your monthly payment.

The new model shifts the cost paradigm from a CapEx-intensive activity to one based almost solely on OpEx. Cloud enables an OpEx focus because you have no need to buy a bunch of hardware up front. The upfront CapEx-intensive purchasing paradigm is one that plagues legacy IT.

Near-instant deployment

Perhaps one of the biggest downsides of legacy infrastructure is the time needed to deploy new resources, which drives up project latency. The result is frustration from business users who have become accustomed to on-demand services.

Public cloud services enable this kind of agile deployment. As you or your business users want more services, your imagination is the limit.

You don't have to

- Wait weeks for new hardware to arrive
- Rack and stack the new hardware
- Configure the new hardware to integrate with your existing environment

Instead, with cloud services, you can spin up infrastructure and platform resources on demand with a single click to build, test, and deploy applications. Building blocks — databases, message queues, and so on — are available to deploy new applications with zero wait time. This innovation drastically

reduces application development time and can massively increase time-to-value for new initiatives.

Seamless infrastructure refresh

Nothing strikes fear into a CIO's heart like having to deal with the replacement cycle of the IT infrastructure.

Refresh cycles can be expensive and risky, and they often require downtime. Replacing monolithic structures such as a SAN requires big capital expenditures and off-hours work.

With the cloud, you have no operational overhead in deploying, managing, and refreshing infrastructure. That's the provider's problem. As a user, you simply key in a credit card number and, on a management console, instantly provision resources.

Security and trust in the cloud

Although still not 100 percent, people have far more trust and faith in the security of the cloud than they did in the past. Businesses are even starting to trust public cloud services for security, governance, and risk management. Organizations can focus on innovation and services without diverting resources to maintain security of the on-premises infrastructure stack.

Control Over IT: Why Private Cloud

With all these great benefits of the public cloud, you might be wondering why you haven't walked into your datacenter and set it ablaze. Well, for all of the good, the public cloud is not a panacea. It still has challenges to overcome.

For many organizations, operating a private cloud makes far more sense.

Virtualization ≠ private cloud

The term *private cloud* is shockingly misused. Many believe that getting close to 100 percent virtualized means they've successfully deployed their private cloud and they can now enter the annals of cloud history for their accomplishment. Not so fast! The word *cloud* carries some implicit assumptions about architecture, workload manageability, automation, and user self-service. Only after you've successfully deployed an infrastructure that has the right architecture, at least some level of provisioning automation and user self-service can you start to consider it a private cloud. Without those features, you're nothing more than a highly virtualized datacenter. Virtualization is only one component of the private cloud.

Further, some businesses are trying to build private clouds using OpenStack/vCloud/Azure and a matching virtualization solution, but still relying on traditional scale-up storage and three-tier architecture. Although you get self-service provisioning and scale-out computing, you still deal with the complex infrastructure life cycle of infrastructure. You end up with silos for different applications (for example, all-flash arrays and bare metal for high performance, virtualization with hybrid storage for VDI, and so on). This setup may look like a cloud on the surface but doesn't deliver many of the benefits that public clouds like AWS enjoy. Bear that in mind as you read the rest of this book.

Considering public cloud limitations

The public cloud, for all its benefits, has limitations. While the public cloud is a viable, cost-effective option for elastic workloads where demand is highly variable or unpredictable, it is not as cost-effective as on-premises infrastructure for more predictable workloads.

In fact, managing predictable workloads is where IT shines. We've been doing that for decades and we do it really well. On the economic front, it's often less expensive to implement and maintain your own environment for predictable workloads than it is to pay monthly expenses for cloud infrastructure.

Here's why: Your predictable workloads often include such applications as ERPs, end-user productivity tools, and

business intelligence and analytics suites. These applications often require consistently high levels of performance and, particularly for applications that use a legacy client/server model, the network connectivity between the server and the connecting clients must be very low latency and very high bandwidth.

With public cloud providers, you pay far more, for example, for all-flash storage in a public cloud environment than for spinning disk. For any applications that require consistently high levels of CPU, you pay monthly for that peak usage. On the network front, you pay far more for a very high bandwidth, low latency connection to the public cloud provider than you would pay to implement such a network in your own environment.

Deciding: Owning versus renting infrastructure

Consider this scenario: Pretend for a minute that each of your travels for work and for pleasure equates to an enterprise workload use case. So, that trip you took to the Caribbean might represent a VDI deployment. The business trip you took to London might stand for a CRM deployment.

As you undertake each of these journeys, you need transportation, which is analogous to infrastructure. Now, as you arrive at the destination airport for each of these trips, do you make your way to an auto dealership and buy a brand new car to use while you're there?

Of course not! Economically, that would be ludicrous and wasteful. You'd also catch the attention of your finance department, who would laugh at your audacity as security escorts you out of the building.

Instead, when you're at home with your predictable travel needs, you likely own a car, or maybe you lease one so that you can replace it every three years. When you travel, or you have unpredictable travel needs, you typically rent a car for the time you need it.

In essence, you're making an ownership decision based on each individual use case. Likewise, businesses want to balance owning and renting infrastructure, choosing between private (owned) and public (rented) infrastructure depending on application workload characteristics. In some situations, renting makes sense. In other cases, owning is a better choice.

Public cloud providers benefit from economies of scale in terms of lower costs, operational efficiencies through automation, and appropriate resource sharing, which they pass along as cost savings to customers.

Private clouds are better suited to predictable, well-established workloads. For these workloads, you decide that owning the infrastructure is a better economic decision.

Before I move on, you should consider one more possibility. Suppose you travel *all the time* to the Caribbean. In this case you might want to own a car at that location, even though it isn't your primary residence. Owning may be less expensive than continually renting cars. Translating this example to the cloud, many organizations are discovering that, once they've moved a certain amount of workload to the cloud, the economics begin to break down. A time comes when, regardless of the kind of workload, pulling some of the workloads back to the private datacenter makes the most sense.

At a certain point, the economies of scale tip back to favor on-premises private cloud environments. Make sure you have a deep understanding of your organization's needs so you know which workloads to run in which location.

Understanding data proximity and locality

Chapter 1 discusses situations in which people in certain regions of the world want to avoid having their data reside in certain other regions of the world. Public cloud providers have begun to address this issue by deploying new datacenters in new regions, but the need to maintain high-level economies of scale can make doing so somewhat challenging.

In addition, economies of scale stop providers from offering a more differentiated experience for individual customers and applications. One of the biggest roadblocks to public cloud adoption is that customers want control over where the data sits and how it is accessed. This level of control is not always possible with the public cloud.

With public cloud, customers may not always know exactly where their data resides. Is it in their state or even in their country? With a patchwork of data security and privacy laws worldwide, not knowing where data resides can create compliance and security issues for customers. Some organizations desire all sensitive data to be under their direct control, effectively eliminating public cloud as a locale.

However, with the private cloud, data locality and proximity are 100 percent in your control. You get to decide exactly where data resides and how close it sits to end-users and applications.

Linking custom-tailored SLAs and performance characteristics

Although public cloud providers have become far more adept at offering granular service level agreements (SLAs), nothing compares to what you achieve with your own infrastructure.

When you're considering SLAs, be sure to consider these two points:

- **Availability:** Uptime is critical. Innumerable statistics show that the hourly cost of downtime can be insanely expensive when you consider lost business, lost employee productivity, overtime for IT to bring services back into operation, and public relations costs.
- **Performance:** Poor infrastructure performance can plague your financials just as much as downtime. As employees struggle to get their work done and as customers struggle to do business with you — and eventually give up — performance issues can be a huge drain on the coffers.

Today's businesses want carefully tailored performance and availability SLAs for their mission-critical applications, but they also want some choice — soft SLAs for less critical applications, and stringent SLAs for business critical applications. For example, for a particularly important application, the business may require that at least three copies of data be maintained separately — public cloud services cannot offer this level of granular control.

Achieving Flexibility, Agility, and Choice with Enterprise Cloud

Once upon a time, not very long ago, business users simply accepted whatever IT gave them, whether that service was fantastic or poor.

Not any more.

These days, businesses want to use the public cloud where appropriate — for example, for backup, disaster recovery, and applications with highly unpredictable IT requirements — and switch between private and public easily.

They want three things:

- **Flexibility:** The option to run workloads where it makes financial and operational sense
- **Agility:** The ability to quickly and easily stand up new applications and scale as business needs demand
- **Choice:** The capability to shift workloads between providers without worrying about downtime or business impact

The Enterprise Cloud

Here's a quick recap of what today's businesses demand.

Businesses want the public cloud for

- Fractional consumption and pay-as-you-grow economics
- Infrastructure and platform resources on demand (agility)
- Zero operational overhead
- Delegated infrastructure security, governance, and risk management

But they still want

- Balance between owning and renting, especially as workload characteristics change
- Proximity of data and services
- Tailored SLAs for specific applications
- Flexibility and choice of platform

Today, enterprise IT offers control, which you need for many applications, but when business users need frictionless agility and ease of use, they are going to the cloud. The two worlds are segmented, and bridging them is difficult.

The vision for the future is to have hybrid environments where the boundary between private and public disappears. You may have an application that has some parts/components in the cloud (for example, deep storage) and others on-premises. You can also have situations where the cloud is used as a backup/DR target while the production environment is on-premises. Finally, an application can be on the public cloud early on in its life when demand is unpredictable, but as demand becomes more stable, it may be migrated back to the on-premises environment.

To meet all these requirements, you need a new paradigm for the enterprise datacenter that can deliver on both sets of requirements and provide a seamless experience between on-premises infrastructure and public cloud services.

That's where the enterprise cloud comes in.

Chapter 3

What Is an Enterprise Cloud?

In This Chapter

- Discovering the components that define the enterprise cloud
- Understanding how public cloud characteristics are associated with your enterprise cloud environment
- Finding out how traditional infrastructure can fail to meet modern application needs

Now that you know a bit about why an enterprise cloud is important, I'd like to chat a bit about what an enterprise cloud actually is. In this chapter, you find out about the five critical characteristics that define an enterprise cloud and discover how each of these characteristics is vital to your enterprise cloud journey.

I also spend a bit of time discussing security, an increasingly important consideration for all organizations large and small.

Defining the Enterprise Cloud

Chapter 2 shows that the enterprise cloud is a collection of characteristics of the public and private cloud. The beauty of the enterprise cloud is that it infuses an organization with an infrastructure that is flexible and agile, and provides complete choice of where to run workloads.

The enterprise cloud is a model for IT infrastructure and platform services that delivers the advantages of public cloud services for enterprise applications without compromising on the value provided by private datacenter environments.

Think of it this way: You get the best of both worlds!

Understanding Historical Private Cloud Roadblocks

If private cloud — upon which enterprise cloud is based — carries with it such benefits, why hasn't it been done before?

Well, others have tried . . . and they have failed, or at least their efforts have not proven completely successful. In general, it's the fault of storage as a resource.

The vision of private and hybrid clouds is not new. Businesses have tried to deploy private clouds using cloud management platforms, such as OpenStack, that deliver self-service provisioning, monitoring, billing, and chargeback. However, the underlying infrastructure is still based on scale-up storage accessed over a storage network that is deployed and scaled in big chunks. What's needed is a re-platforming of the enterprise datacenter. You cannot build cloud capabilities on traditional three-tier infrastructure with scale-up storage.

Scale-up storage has hard limits. At a certain point, the shared components — controllers and the network fabric — get overwhelmed. It's inevitable. The question is not *if* this will happen, but *when.* As a result, many scale-up storage systems are bundled with spec sheets that tell customers that they can grow only so far before they have to add more shared components. Adding these components adds complexity to the system.

The end result is unpredictability, a scenario that cannot be tolerated in the modern datacenter. Businesses must be able to operate with the expectation that their workloads will operate continuously at predictable levels. In scale-up, as you add more burden to the shared resources, performance levels can be affected.

Even many of today's array-based scale-*out* storage methodologies begin to crumble under their own weight as they grow. Much of this has to with data locality, which I discuss in Chapter 1. The bigger these constructs grow, the more data has to traverse a storage networking fabric. Eventually, as data gets farther and farther from the CPU and RAM, performance problems ensue.

A reliable datacenter infrastructure combines the ability to leverage scale-out storage while maintaining data locality.

Storage continues to be the resource that holds back progress on the datacenter journey.

Defining Enterprise Cloud Key Ingredients

Here's a high level look at how to define the enterprise cloud: The enterprise cloud delivers the frictionless agility, simplicity, and fractional consumption of public cloud services while providing control over performance, location of data and services, and choice of platforms.

Five key components comprise the enterprise cloud:

- **Full-stack infrastructure and platform services** that deliver turnkey infrastructure for any app at any scale, anywhere, delivered through a combination of on-premises datacenters and public cloud services
- **Zero-click operations and machine intelligence** that deliver operational simplicity through automation
- **Instant elastic consumption** that allows businesses to buy and use only the IT resources they need, only when they need them, spinning resources up and down on demand, and eliminating overprovisioning and prediction risk
- **Integrated security and control** that covers the entire infrastructure stack, leverages automation, and simplifies maintenance of the security baseline using automation
- **Application-centric mobility** that lets businesses run applications anywhere with no infrastructure lock-in

The next section looks at each of these ingredients in a bit more depth.

Full-stack infrastructure and platform services

Regardless of where you decide to run your critical applications, you need a full set of infrastructure to do it. However, before you run out to buy a bunch of storage to connect to your servers, you should know a number of things.

In Chapter 1, I briefly discuss the concept of the software-defined datacenter (SDDC). Although a datacenter based on SDDC principles requires hardware, the hardware is not the focus.

Instead, with the SDDC, you transition to hardware components that are easily programmable. Organizations should consider infrastructure that is delivered as a set of software-defined services, including file, block, and object storage, with integrated data services such as protection and availability for applications.

Rather than buying a super-expensive monolithic SAN, buy infrastructure that you can compose to meet the needs of your individual workloads.

It goes without saying that virtualization is — and will remain — at the core of everything IT does. Virtualization should be a default and key component in any platform you use. Make sure you choose an environment in which server virtualization capabilities are built into the infrastructure stack. Virtualization should be treated as a *feature,* not a separate product.

Most businesses don't plan to stay stagnant. Most intend to grow as they onboard new customers and begin delivering new products. To maintain customer and product growth, you need to be able to easily grow the environment.

Your entire infrastructure stack should be built with these web-scale engineering characteristics:

- Software-defined
- Distributed *everything*

- Resilient and self-healing
- Extensive automation

In short, you need an infrastructure that allows you to scale without limits and without single points of failure.

The platform needn't be limited to the private cloud or your local datacenter either. You should have the ability to support hybrid delivery of applications — that is, you should be able to provide choice between on-premises infrastructure and public cloud services for your business-facing applications.

Your enterprise cloud environment also must offer powerful data protection and disaster recovery options, analytics to streamline operations, and other critical services.

With the growing diversity in infrastructure needs for applications, natural silos appear on the infrastructure side. For example, some demanding Oracle and SQL Server databases may be run in bare metal environments while others are virtualized. These silos make the process of managing infrastructure incredibly challenging, because you have to manage each silo separately. The enterprise cloud needs to deliver infrastructure capabilities that can support bare metal, virtualized, and containerized environments for any application.

Infrastructure is the fundamental building block for the enterprise cloud. Past attempts at building private clouds have focused on the software layer — such as on cloud management platforms — that sit on top of infrastructure and deliver self-service, monitoring, billing, and chargeback capabilities. But unless infrastructure is built to scale out and without single points of failure, you will not end up with a cloudlike environment.

Figure 3-1 provides a look what the web-scale world looks like.

Zero-click operations

Even if you're the master of the console in your virtualized datacenter, you likely still perform lots of clicks to get your work done. You may have deployed tools that help you achieve the beginnings of automation, but most organizations have yet to take these capabilities to their desired and natural conclusion: complete automation.

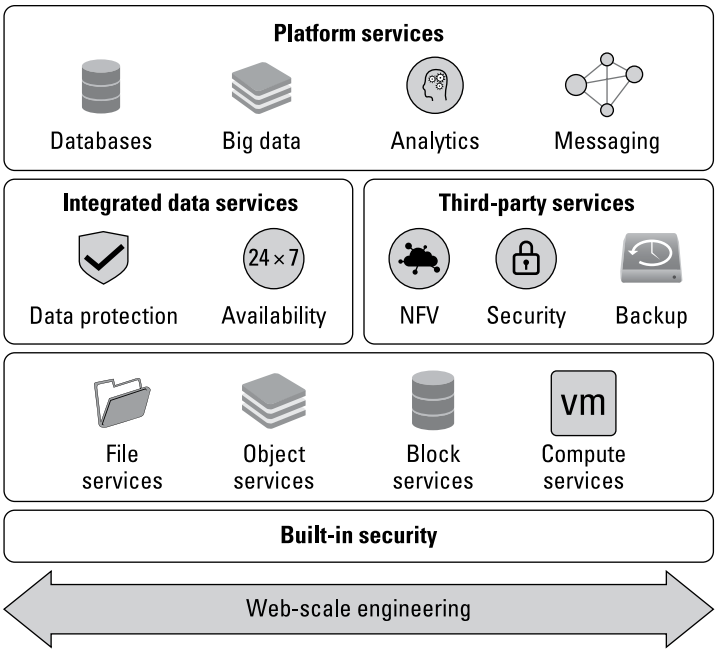

Figure 3-1: The web-scale world.

This is not to say that you'll be able to walk away from your datacenter and throw away the keys. You'll need manual steps here and there, but you shouldn't routinely get involved in ongoing operations.

As you begin to deploy an enterprise cloud and a combination of public and private cloud, the ability to reduce administrative overhead becomes critical. Smaller overhead is one of the ways organizations can reduce OpEx associated with datacenter management.

Universal control plane

Your enterprise cloud should have a universal control plane for all environments, including your public and private cloud. The control plane in an enterprise cloud is the management layer. By using a single universal control plane, you effectively eliminate the need to switch management silos as your business goes from one environment to another for applications. You can see for yourself if you're moving along the path toward a universal control plane. Do you have a separate

management infrastructure for single components, such as virtualization? If you do, you're building management silos and should reevaluate how you're moving forward.

Machine intelligence

Although the 1980's film *The Terminator* foretold the rise of the artificial intelligence SkyNet, humanity did not heed the warning and we're plowing ahead with efforts to turn decision-making over to robots and other technology-based constructs. That said, never fear! What's the worst that could happen?

Seriously, though, we've come a long way with machine learning tools. Humans now can program systems that actively learn about their environment and can help administrators automate many mundane, yet critical, datacenter operations.

With the tools at our disposal, we can implement machine intelligence and self-learning capabilities to drive end-to-end automation where the platform becomes smarter with decision making and recommendations over time.

Imagine a world in which you walk into the datacenter and find half of your nodes burned out, but you didn't even know because your management layer shifted those workloads to operational nodes. Or, imagine a scenario in which your management layer can sense that your web tier is hitting a capacity ceiling, and it automatically spins up an additional node to handle overload.

That kind of capability is here today.

In other words, you gain seamless infrastructure optimization and error remediation as part of a tight control system.

Consumer focus

Today's consumer electronics have plummeted in cost while growing in capabilities. Even better, they have become dead simple to use. Until recent years, enterprise hardware and software was just the opposite. You practically needed a PhD in storage to manage lots of arrays. You needed years upon years of background to even understand what you were clicking. Even worse, IT pros demanded increasing numbers of what have become known as "nerd knobs," an unfortunate term, but one somewhat grounded in reality.

My, how times have changed!

Today, hiding things seems to be the norm, and for a good reason. The right solution hides complexity from you. What you're provided on-screen is an outcomes-based paradigm, not a bunch of knobs where you manage inputs. Some companies have realized that they can achieve better ROI by keeping the IT administrative paradigm simple.

In the enterprise cloud, every aspect of the management experience must be built around the principle of consumer-grade design to enable ease of use. Minimize the ramp-up time needed to learn and become productive on the platform.

Automation and analytics

The goal is to remove operator involvement from everyday tasks. You need to provide true self-service capabilities so that users can request their own resources without constantly interrupting IT staff. Self-service requires high levels of automation so that results can happen without additional IT resources. For example, perhaps a developer can independently build a test/dev environment without working through an operations person.

At the same time, high levels of automation mean that you need comprehensive analytics. Why? In essence, the goal of enterprise cloud is to move IT into an exception-handling function. The routine things should *just happen,* whether that's through built-in machine intelligence mechanisms or user self-service. When an exception occurs, an IT operations person should be immediately notified to take appropriate action. Automation doesn't mean IT never touches infrastructure again; it simply means IT needn't touch infrastructure on a daily basis.

Figure 3-2 helps you envision how the pieces fit together. At the bottom is a universal control plane upon which the enterprise infrastructure — public and private — resides. Above the environments are the three principles of zero-click administration: consumer grade design, artificial intelligence, and automation/analytics.

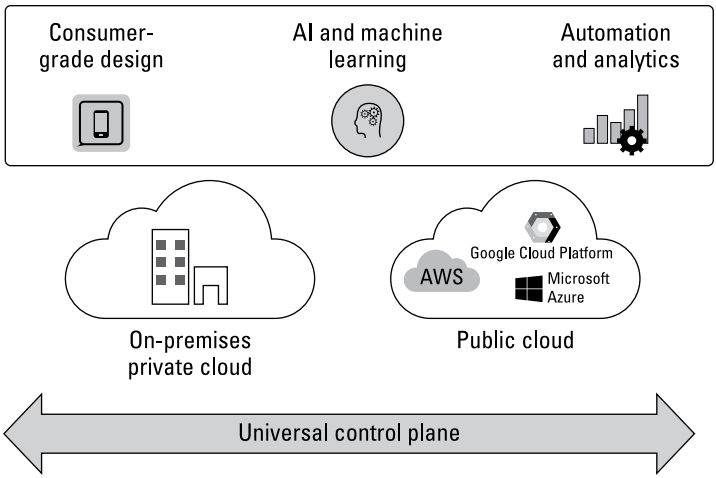

Figure 3-2: Understand how the universal control plane supports enterprise cloud.

Instant elastic consumption

The best part about implementing services on the public cloud is the ability to expand and contract usage on the fly. If you consider typical legacy enterprise environments, this isn't generally the case, for a variety of reasons:

- **Overbuying is rampant.** Because of the way replacement cycles operate and the services that have been available, many IT pros overbuy hardware, especially storage. You don't want to end up buying more storage mid-cycle. Further, many SAN vendors make their best offers up front, but the discounts might be a bit less generous when your back is against the wall and your capacity is dwindling.
- **Expansion can be difficult.** Adding capacity requires downtime and is sometimes fraught with risk because you must match firmware versions on controllers, disks, and other elements.
- **Resources can be tough to align.** Because you must scale resources individually in legacy infrastructure environments, growth can require lots of planning and effort.

Although you can use the public cloud to counter these issues, you know that the public cloud isn't always an option.

With a solution that enables enterprise cloud capabilities, you gain the ability to deploy workloads that can flex, much like in the public cloud. You get pay-as-you-grow scaling. If resources become low, you simply add a hyperconverged infrastructure appliance. You don't need to overprovision storage, for example. Further, you avoid infrastructure sitting idle.

With hyperconverged infrastructure-supported enterprise cloud services, you can adopt a "just in time" infrastructure mentality that is super-easy to scale. You simply call your vendor, ask for another node, and deploy it. The infrastructure should be all but invisible to the users. They shouldn't have to worry about the underlying infrastructure. They should only be focused on their workloads.

With the right enterprise cloud-centric infrastructure, you can scale up and down on demand. Figure 3-3 shows how easily you can add resources to grow an environment. As you add more nodes, you can scale resources in a linear manner.

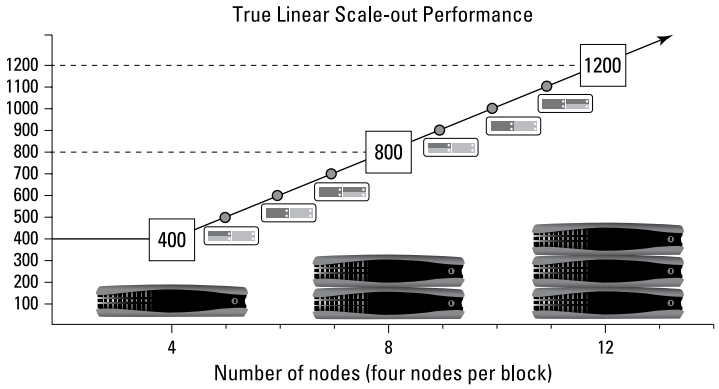

Figure 3-3: Linear scale-out infrastructure supports the needs of the enterprise cloud.

Integrated security and control

Organizations across the globe are working hard to secure their environments against attacks from within and from the outside. As the potential fallout from security issues

increases — bad PR, fines, lost business — companies need to ensure they can adequately secure their technology environments.

Security starts with the infrastructure. However, maintaining security with traditional infrastructure environments is challenging for a number of reasons. First among those reasons is the architecture of an infrastructure stack that is comprised of products from multiple vendors with a narrow and limited view of security.

Validating and maintaining a security baseline through software upgrades, for example, is time-consuming and often involves error-prone manual processes that take away from innovation and productivity. You're messing around with security when you should be able to work on business-facing activities.

Life is a bit different in the world of the enterprise cloud. In the cloud era, security must be an integral and invisible attribute of enterprise infrastructure.

Here are the ways the enterprise cloud helps organizations address security:

- **Security-first design:** Security specification and testing must be built into every step of product development rather than at the end.
- **Hardened infrastructure stack:** You must do away with a piecemeal approach to security and shift to comprehensive end-to-end infrastructure security. When you think of public cloud services from a security perspective, everything below the app is the cloud provider's responsibility. With the enterprise cloud, on-premises security must be equally seamless and invisible.
- **Hands-off:** The infrastructure must include automated security validation and self-healing capabilities to make security maintenance efficient.

Infrastructure and application security does not end at the boundaries of datacenters. The control fabric can take security policies defined and configured in one environment and port them over to the target environment, be it a private datacenter or the public cloud, automatically.

You should regularly assess the security posture of your infrastructure and application environment through a six-step, best practices-based process (shown in Figure 3-4):

- **Assess:** Gather the current baseline security posture of the environment.
- **Measure:** Determine where you are falling short.
- **Report:** Notify the appropriate people about the issues and ask them to fix the problems.
- **Test:** Check the outcome of the remediation efforts.
- **Update:** Update your baseline documentation with the new security posture.
- **Repeat:** Perform these steps as often as necessary, based on your organization's security policies.

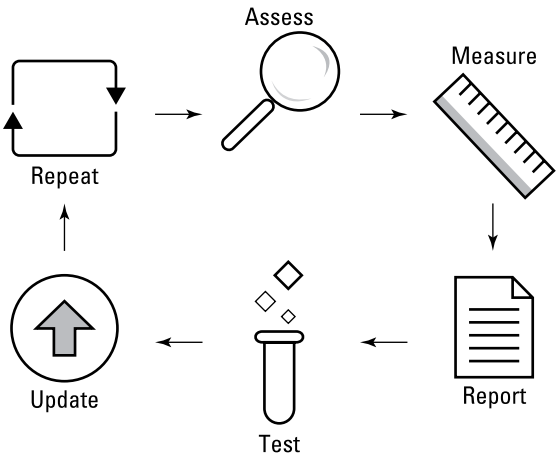

Figure 3-4: Linear scale-out infrastructure supports the needs of the enterprise cloud.

Application mobility

Perhaps one of the most important aspects of the enterprise cloud is application mobility. When applications are not bound by the constraints imposed by the infrastructure platform, enterprise IT can pick the best physical and virtual infrastructure platforms for its applications based on its needs today.

Every environment presents different performance characteristics. Those must be coupled with business and financial drivers. On the business front, IT must adhere to service level agreements (SLAs) that dictate the levels of performance availability that must be supported by the datacenter environment. Only with an environment that has predictable levels of performance can strict SLAs be adhered to. Further, different applications may require different SLAs. Some may require high levels of performance and availability while others may be able to get by with less.

On the economic and financial front, the datacenter environment must be affordable, both at inception and on an ongoing basis. If you buy an environment that provides high levels of performance and availability, but most of your applications require less, you're leaving money on the table. The same goes for available capacity in different areas of the environment. You must ensure that your environment can support different kinds of applications and has sufficient capacity in each area to support the workloads that will run there.

In other words, you run applications *in the right environment at the right time* — one of the core tenets of enterprise cloud. Applications must be able to move freely between hypervisors, to any public cloud service, and to container-based environments.

Any cloud, any time.

For this any-to-any freedom to be effective, it should

- Require no application changes
- Preserve application state, configuration, and environmental requirements to minimize risk
- Translate SLAs across different environments

Traditional infrastructure presents several barriers to application mobility:

- **Data gravity** (a phenomenon in which data must remain local to its application in order for that application to operate properly)
- **Tight coupling** between applications and runtime environments

- **The need for manual configuration and setup** when applications move from one environment to another
- **The need to learn new management tools, constructs, and paradigms** for each platform

Application mobility requires tackling each of these challenges. Your enterprise cloud environment must have the capability to overcome these barriers.

Virtualization has helped in many ways to get to this point. With enterprise cloud-based environment running on hyperconverged infrastructure, which requires workloads to be virtualized, you can quickly and easily decouple many elements of the datacenter. You can even decouple applications and their runtime environments. You effectively eliminate the data gravity issue and, because everything runs with a common management layer, you have no worries as you move applications between environments.

Further, you don't need to learn a bunch of new management tools. In your enterprise cloud environment, all aspects are handled on a common management layer.

Chapter 4

Building an Enterprise Cloud

In This Chapter

- Learning why you should embrace shadow IT
- Discovering why the 80/20 rule should no longer apply
- Learning how to prepare your people for the new paradigm
- Finding out how enterprise cloud affects economics and the replacement cycle

You've learned about the current state of IT and the "why" and "what" behind enterprise cloud. You're an expert on all the reasons enterprise cloud is a great path forward.

Just one problem remains: getting from here to there. That requires a number of activities on your part — changing the way you think about managing IT, ensuring that your staff is ready for the changes, and modifying processes and infrastructure.

So, this chapter starts at the beginning.

Adjusting Your Perspective: Change Is Coming

Change is hard. Everyone knows that. However, in an industry that often leads change in organizations, there is some irony in how difficult it can be for IT professionals to accept change themselves.

As with so many things, you must adapt or you will be relegated to the dustbin of irrelevancy. Look at how many mainframe operators who failed to stay current didn't survive the wave of decentralization that defined IT in the 1980s and 1990s.

History teaches that change is inevitable. It's time for you to adjust your perspective on a number of fronts — beginning right now — so that you stay relevant into the 2020s and beyond.

Why shadow IT should be embraced, but managed

"If IT isn't providing it, no one should be doing it."

Variations of this phrase have been around for a long time. My, how times have changed.

Today's business units can stand up services with nothing more than a credit card and, frankly, *many* do so. Services have become vastly more consumable than they were just a few years ago. You can thank the cloud for much of this. Thousands upon thousands of software services are available for your users to quickly and easily stand up and consume.

Further, end-users have become far more tech-savvy than they used to be. Luddites no more, many end-users rival and exceed IT staff technical knowledge in many ways *and* they are the subject matter experts in their areas.

Users don't want to be fully reliant on IT anymore, either. They want to do their jobs on their own terms.

This phenomenon has become known as *shadow IT* because it often happens in the dark corners of the company. End-user departments build systems that they need because IT either can't or won't build them. Or, users perceive IT as being so slow that it makes no sense to engage IT in the first place.

CIOs and other IT leaders see shadow IT as a threat to be abolished, and sometimes they have good reasons:

- **Security:** Although end-users have become far more savvy about technology, they are often far less knowledgeable than IT about security. Further, IT is charged

with maintaining infrastructure and application security, so it's challenging when IT doesn't have a full view into what's happening across the organization. As individual business units start sharing data with cloud providers, for example, that data may not be properly secured.

- **Consistency:** When you're working with business intelligence, maintaining a *single version of truth* is critical. Results shouldn't vary as different departments view data. All data elements should be consistent so that the organization can rely upon the decisions made with that data.
- **Cost:** When individual users start to procure their own IT services, economies of scale become far more difficult to achieve, which can increase overall costs.

So, it sounds like you should do everything in your power to stop shadow IT in its tracks, right? Well, not so much.

Shadow IT rises because the organization has a need that is not being met. Whether the need is real or perceived is generally irrelevant. Even if the need is only the *perception* of a failing, something happened in the organization to cause that perception.

It's time for CIOs and IT leaders to embrace shadow IT. Find out why the shadow systems were set up, and look for the underlying shortcomings in IT's services. That may even require IT to extends its portfolio and begin encompassing services that were stood up by end-users.

In general, IT governance processes should provide support for how shadow services can be brought under at least a semblance of IT management. The governance process should outline how departments stand up such services so that they comply with organizational security and data guidelines.

You can no longer ignore shadow IT, but you can't run in and put a stop to it, either. Instead, you need to implement constructs to help you discover the true needs of the business and ensure that your environment meets those needs within the confines of organizational policy.

Why the 80/20 rule no longer rules

If you've worked in IT for any length of time, you've probably heard of the 80/20 rule. Also known as the Pareto Principle, the rule states that 80 percent of the IT budget and IT's efforts go to keeping the lights on, while only 20 percent is dedicated to innovation and propelling the business forward. Okay, if you're a purist, the *original* Pareto Principle states that 80 percent of the effect comes from 20 percent of the causes, but the meaning has been extended to IT operations in the way I describe.

As CIOs, other IT leaders, and the executive team look for ways to better address key business problems, the 80 percent of the budget that goes to sunk costs can look enticing. After all, if you can improve efficiency just a little, you can change the ratio to 60/40 or 50/50. Improve efficiency enough, and you can create an IT organization that spends only 20 percent on the basics and 80 percent on value added.

This is where IT needs to head. The 80/20 rule is a remnant of a different time. With business needs changing at a faster pace than ever before, IT needs to reduce the 80 percent figure and focus on revenue generating activities.

By deploying enterprise-cloud-enabling hyperconverged infrastructure, IT can begin to shift some of that 80 percent toward other activities. An enterprise cloud infrastructure includes automation capabilities and user self-service, which helps users reduce their reliance on IT and frees up IT staff to focus on the business. Further, with a revamped economic model that enables just-in-time infrastructure and easy scaling, that 80 percent of the IT budget can get even lower.

Why bimodal IT is a short-term fix

A common school of thought says that IT departments need to fully embrace bimodal IT, discussed in Chapter 1. Under a bimodal support paradigm, you'd have people supporting legacy environment and a separate group of people supporting modern apps.

The problem is that this kind of support is expensive and inefficient. All you're doing is patching a symptom rather than addressing a root cause.

Rather than attempt to build structures around different application and infrastructure support models, a more sensible plan is to deploy infrastructure that can support both modes of support.

That's exactly what you get with an enterprise cloud deployment. You gain an infrastructure model that can support legacy applications as well as modern apps.

Preparing Your People

Compared to dealing with people, the technology is easy! However, you can't ignore the need to ensure that your people are prepared for the changes that you need to make to your IT organization and to your infrastructure.

For a long time, businesses have had to hire specialists for each area of the IT infrastructure. As organizations move into the brave new world, IT seems to need an ever-increasing number of specialists to keep the burgeoning set of resources operational.

Or not.

Hyperconverged infrastructure driven enterprise cloud systems don't require specialists to operate. You don't need to hire expensive storage administrators, virtualization administrators, and systems administrators to maintain your legacy environment. You don't need to hire cloud experts to deal with your public cloud systems.

With an enterprise cloud foundation, you need IT pros who have a breadth of knowledge, though they don't require massive depth. These IT generalists are the future of datacenter support. They'll be in the forefront as organizations seek to simplify a complex morass of technology in the datacenter.

At first, this change can be threatening to existing staff, particularly those who define themselves based on their subject matter expertise. Because every area is still represented in

the new paradigm, you can move existing people into new roles that are more general in nature. However, they can also have a more business-facing component that helps shift IT's focus from infrastructure to the bottom line.

Adapting Your Processes and Infrastructure

Beyond people, you also need to rethink how you handle some of your IT processes and your infrastructure.

Rethinking infrastructure economics

To find a starting point, consider the current IT replacement cycle. For this scenario, I assume that the organization has a five-year replacement cycle, a visual depiction of which is shown in Figure 4-1.

When you buy infrastructure, you probably overbuy, *even if you run out of capacity.* How can this be possible? In Figure 4-1, the horizontal line depicts the overall capacity that you've purchased. In this context, *capacity* doesn't refer only to storage; it also refers to the amount of processing (CPU) power and RAM that you have available.

Most IT departments buy what they think they'll need for the duration of the replacement cycle. Sometimes the estimate is correct and sometimes it isn't, but one fact is always true: For some period of time, you won't even come close to using all the capacity you've purchased.

In Figure 4-1, the diagonal line depicts the actual workload demand for the organization that purchased this infrastructure. The lines intersect in Year 4. The shaded area before Year 4 depicts the "waste" that this organization is suffering from. I refer to this as the *zero return on investment zone.* More than three years go by before the fictional company grows into what it purchased.

Additionally, the organization did not reach the end of its replacement cycle before running out of capacity. This means the company must make an out-of-cycle infrastructure purchase to add capacity.

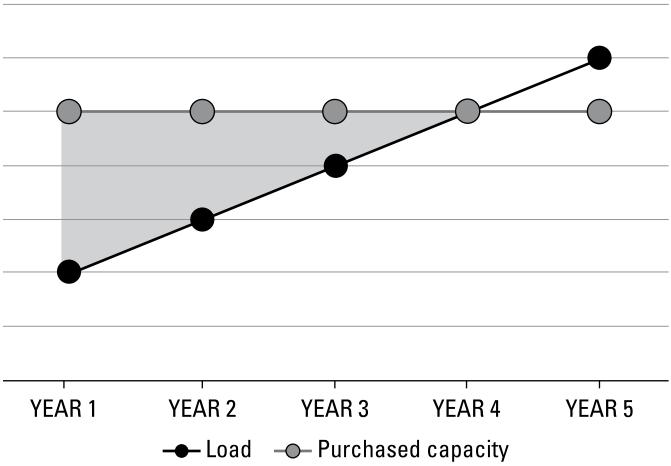

Figure 4-1: Traditional infrastructure procurement economics is not a viable solution.

With enterprise cloud and hyperconverged infrastructure, you can begin to adopt a just-in-time approach to datacenter resources. This method allows you to also adopt cloudlike pay-as-you-go economics. Figure 4-2 shows what such a scenario might look like. In Year 1, you buy what you need for that year, making sure to keep your purchased capacity just a little ahead of your workload needs.

Under this model, you have no zero ROI zone. You're effectively using what you've purchased. Your upfront economics are far better than they are with traditional infrastructure. In short, you aren't wasting your capacity.

Notice that your organization didn't run out of capacity in Year 4. Instead, your company simply added more nodes to its hyperconverged infrastructure-based enterprise cloud environment. You've successfully operationalized the changes you've made to the datacenter environment.

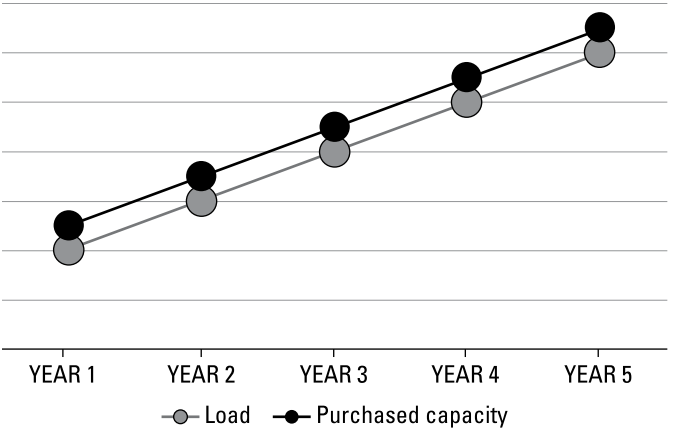

Figure 4-2: Hyperconvergence and enterprise cloud can help you to reinvent the IT budget.

Understanding disaggregation and the replacement cycle

With enterprise cloud based on hyperconverged infrastructure, you no longer have to worry about trying to manage resources separately. Instead, resources are aggregated and scaled together in a somewhat linear manner. Hyperconverged infrastructure vendors typically make it possible for end-users to focus on needed resources as they add new capacity. For example, if you're running low on storage capacity, your new node can be storage heavy, although it will have CPU and RAM as well.

A disaggregated replacement cycle is far more difficult to manage than one based on aggregation, as is the case with hyperconverged infrastructure. As you seek to scale your enterprise cloud environment, you don't need to focus on individual resources. You simply worry about the needs of your workloads, and add nodes as necessary.

Chapter 5

Ten Reasons Why Enterprise Cloud Is the Future of IT

Enterprise cloud has a bright future in IT, and for good reasons. Here's a look at ten reasons why enterprise cloud is the future of IT:

- **A brand new economic model:** Your old legacy-based IT economic model is no longer sufficient. With enterprise cloud, you can adopt the pay-as-you-go characteristics of the public cloud while providing a common foundation upon which to run both legacy and new-style apps.
- **A focus on the end-user:** Your users are demanding new services, and you may not even know it. Until you discover why your users are deploying shadow IT systems, you may not understand their needs. An enterprise cloud can help you better focus your efforts on addressing the deficiencies seen by end-users.
- **Faster response from IT:** IT has a perception of being slow, especially when compared with cloud providers, who can instantly deploy new infrastructure for you. With enterprise cloud, you gain the ability to provide instant infrastructure for your business users.
- **Refocusing IT on the business:** The 80/20 rule can become the 20/80 rule if you do things right. Enterprise cloud can help you to make this shift. You can tailor your IT department's services to activities that generate revenue rather than simply keeping the lights on.
- **Public cloud simply makes sense:** You get built-in economies of scale, instant deployment, and powerful management tools. Enterprise cloud can help you gain these abilities with your own infrastructure.

- **Private cloud simply makes sense:** You know that public cloud doesn't always address issues such as data locality, security, and compliance in a way that works well for your company. However, private cloud does so. By deploying enterprise cloud, you get the best parts of both public and private cloud, with the ability to seamlessly use both depending on application needs.
- **The trends are on your side:** Lots of trends came together to make enterprise cloud viable. You needed fast storage, which you get through flash. You needed the ability to deploy hardware in an economical way, which you get through commodity hardware. These trends have coalesced to enable enterprise cloud deployments.
- **Choice is key:** Any cloud any time. Your organization needs a choice of where to run workloads. You shouldn't be forced into a single public cloud provider. With the right enterprise cloud foundation, you choose the cloud and don't end up with workloads trapped somewhere.
- **You need to think beyond bimodal IT:** This book helps you understand why bimodal IT isn't necessarily the best path forward for the long term. With enterprise cloud, you get the outcomes promised by bimodal IT without worrying about the inefficiencies that this model can introduce.
- **Users are pretty smart:** Your users are far ahead of where they were a few years ago. Your infrastructure environment must reflect this fact by enabling user self-service and automation, both of which are supported in an enterprise cloud scenario.

Get the frictionless benefits of public clouds without sacrificing the control of on-premises infrastructure

Businesses want the agility, simplicity, and pay-as-you-grow economics of public clouds in their own datacenters. That's why more and more organizations are turning to enterprise cloud — which carries public cloud characteristics and support — to meet critical business needs. This book helps you understand what the enterprise cloud is and how it can help you propel your organization into the future.

- ***Adapt or perish* — your business is at a critical crossroads, and legacy infrastructure is not keeping pace**
- ***Adjust your perspective* — prepare your organization for change**
- ***Embrace the cloud era* — discover the power and potential of enterprise cloud**
- ***Elevate your IT* — focus on the applications and services that power the business**

Open the book and find:

- **A breakdown of critical trends that are shaping IT infrastructure today**
- **Expert perspectives on the state of IT for the enterprise**
- **Ten reasons why enterprise cloud is the future of IT**
- **Next steps to prepare your people and adapt your processes**

Go to Dummies.com® for more!

Scott D. Lowe is a former CIO and co-founder of ActualTech Media, a content creation and demand generation firm focused on creating content, conducting market research, and connecting technology companies with the right audience.

FOR DUMMIES® A Wiley Brand

Also available as an e-book

ISBN: 978-1-119-31763-0
Not for resale

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.