Brought to you by:

NUTANIX

Enterprise Cloud

for dummies®

A Wiley Brand

Achieve public cloud agility and elasticity

Accelerate innovation and customer service

Retain on-premises control and security

Nutanix 2nd Special Edition

Scott D. Lowe

About Nutanix

Nutanix is a global leader in cloud software and hyperconverged infrastructure solutions, making infrastructure invisible so that IT can focus on the applications and services that power the business. Companies around the world use Nutanix Enterprise Cloud OS software to bring one-click application management and mobility across public, private, and distributed edge clouds so they can run any application at any scale with a dramatically lower total cost of ownership. The result is organizations that can rapidly deliver a high-performance IT environment on demand, giving application owners a true cloud-like experience. Learn more at **www.nutanix.com** or follow **@nutanix** on Twitter.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Enterprise Cloud

Nutanix 2nd Special Edition

by Scott D. Lowe

for dummies®
A Wiley Brand

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Enterprise Cloud For Dummies®, Nutanix 2nd Special Edition

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

Copyright © 2019 by John Wiley & Sons, Inc., Hoboken, New Jersey

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Trademarks: Wiley, For Dummies, the Dummies Man logo, Dummies.com, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries, and may not be used without written permission. Nutanix and the Nutanix logo are trademarks or registered trademarks of Nutanix. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.

For general information on our other products and services, or how to create a custom *For Dummies* book for your business or organization, please contact our Business Development Department in the U.S. at 877-409-4177, contact info@dummies.biz, or visit www.wiley.com/go/custompub. For information about licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

ISBN 978-1-119-57270-1 (pbk); ISBN 978-1-119-57268-8 (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
Associate Publisher: Katie Mohr
Editorial Manager: Rev Mengle
Business Development Representative: Karen Hattan
Production Editor: Siddique Shaik

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Table of Contents

INTRODUCTION 1
About This Book 1
Foolish Assumptions 2
Icons Used in This Book 2
Beyond the Book 2

CHAPTER 1: **Surveying the State of IT within the Enterprise** 3
Trends Shaping IT Infrastructure Today 3
Flash storage 4
Software-defined functionality 5
Hardware commoditization 6
Hypervisor commoditization and the emergence of containers 6
The (hyper)converged revolution of compute and storage 7
Modern application architectures 7
Security evolution 8
Adoption of multi-cloud environments 9
The State of Public Cloud 10
The growth of all things cloud 10
Increasing viability of public cloud 11
Understanding security and trust in the cloud 12
Embracing any cloud 13

CHAPTER 2: **Why Enterprise Cloud?** 15
Frictionless IT: Why Public Cloud 16
IT as a business enabler 16
Understanding fractional consumption 16
Near-instant deployment 17
Seamless infrastructure refresh 18
Security and trust in the cloud 18
Control over IT: Why Private Cloud 18
Considering public cloud limitations 18
Deciding: Owning versus renting infrastructure 19
Controlling and optimizing public cloud costs 21
Understanding data proximity and locality 21
Linking custom-tailored SLAs and performance characteristics 22
Security and trust in on-premises 22

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Achieving Flexibility, Agility, and Choice with Enterprise Cloud....23
The Enterprise Cloud....23

CHAPTER 3: **What Is an Enterprise Cloud?**....25
Defining the Enterprise Cloud....25
Integrating the ethereal: On-premises + multi-cloud....26
Understanding the need for a single control plane....26
Understanding Historical Private Cloud Roadblocks....26
Defining Enterprise Cloud....28
Unified governance....28
Full-stack infrastructure and platform services....30
Zero-click operations....32
Instant elastic consumption....35
Integrated security and control....37
Application mobility....39

CHAPTER 4: **Building an Enterprise Cloud**....41
Adjusting Your Perspective....41
Why shadow IT should be understood and then managed....42
Why the 80/20 rule no longer rules....43
Why you need to address bimodal IT....44
Why automation and APIs are the way of the future....45
Preparing Your People....45
Leveling up IT organizational skills....45
Empowering end-users....46
Adapting Your Processes and Infrastructure....46
Rethinking infrastructure economics....46
Understanding disaggregation....48
Consolidating storage services....49
Ensuring consistent performance....49
Choosing the right public cloud....50

CHAPTER 5: **Exploring an Enterprise Cloud Deployment**....51
Deploying Infrastructure....51
Deploying Workloads and Applications....53
Administering the Enterprise Cloud Environment....54
Enabling Self-Service....55
Managing Enterprise Cloud Costs....56
Scaling the Environment....57

CHAPTER 6: **Ten Reasons Why Enterprise Cloud Is the Future of IT**....59

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Introduction

Since I wrote the first edition of this book, a lot has changed. The public cloud continues making inroads into the enterprise. At the same time, organizations have begun to understand the limits imposed by the public cloud. Some find that public cloud is not the best fit for their workloads and have migrated back to on-premises operations, while many others are building robust hybrid clouds to get the benefits of both.

Meanwhile, Nutanix has also evolved. The company today is bigger and broader than it was when the first edition came out.

However, some things are still the same. 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 that 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 64-page journey helps you understand enterprise cloud and how it fits into your datacenter paradigm. By the end of this book, you will see how enterprise cloud can help you propel your business into the 22nd century.

In the new chapter in this edition (Chapter 5), you go through life as John, a sysadmin who faces common challenges. This chapter shows how you can solve these key challenges in an enterprise cloud and, specifically, how Nutanix brings tools to bear to solve them.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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.

REMEMBER

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.

TIP

You should 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.

TECHNICAL STUFF

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.

WARNING

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`.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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 while creating new challenges

Chapter 1
Surveying the State of IT within the Enterprise

Today's IT departments face challenges that are both familiar and unknown. Even in a world in which the phrase *digital transformation* is bandied about, IT faces the same kinds of resource constraints it's always had, but new and different solutions now provide ways to address them. Deploying and managing infrastructure is still a challenge, but ongoing innovations offer a path that mitigate these 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 were driven by demands that businesses placed on IT and set the stage for even greater transformation over the next few years.

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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Flash storage

It's not possible to overstate the impact that flash storage has had on the modern IT environment. 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 often threw 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.

REMEMBER

In recent years, the cost of NAND-based flash storage has continued to plummet by double-digit percentages while capacity continues to increase. 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 30TB behemoth costs far more than the same spinning disk capacity, but it also means you can achieve all-flash capacity density that is far better than that of disk. In a world where physical size matters, that's no small thing.

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 data must traverse the network, 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.

Some 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 nonissue 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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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.

REMEMBER

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

And there's even more! Understanding that flash needed to be a first-class storage citizen, the storage market responded by introducing an interface and protocol designed with flash in mind. NVMe-based flash devices have hit the market with gusto. Prior to the advent of NVMe, flash devices were hobbled by being supported by controllers and protocols that were developed for spinning disk. Binding flash to spinning disk constructs made it impossible to push flash to its fullest potential.

No more. NVMe allows access to hundreds of areas of the storage simultaneously and massively simplifies the technical storage underpinnings, which yields even more performance improvements.

Software-defined functionality

At the same time that flash and NVMe storage have become common in the datacenter, Intel has continued to release processors with massive numbers of cores just begging to be set free. The abundance 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.

REMEMBER

Today, the industry is 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Hardware commoditization

Intel is also driving 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 silos — 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 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.

REMEMBER

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.

However, turn your attention to the public cloud for a moment. VMware dominates the private datacenter, but that is far from being the case with public providers, none of which use VMware vSphere as the cornerstone for their services. In fact, two of the biggest public cloud providers on the planet — Amazon and Google — use a KVM core as the heart of their entire solution.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Today, it's less critical to ask, "What hypervisor is being used?" and more critical to ask, "Can my workloads run well?" Is there a management and automation overlay that treats your non-VMware workloads like first-class citizens?

If you consider Google and Amazon, KVM has emerged as a leading hypervisor. Of course, Nutanix AHV leverages KVM at its core, too, bringing significant benefits to on-premises workloads, such as improved financials and more fine-grained control of the environment.

Meanwhile, containers have emerged as an alternate abstraction technology allowing applications to be developed, tested, and deployed quickly and easily. It's important for the infrastructure platform of today 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 simplified datacenter administration, decreased costs, and increased end-user and customer satisfaction.

REMEMBER

Hyperconvergence is just one aspect of a bigger picture. It solves the question of "Where do I run my on-premises workloads?" but the right solution goes much further. It encompasses all those on-premises workloads by bringing cloud-like constructs to bear.

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.

Mode 1 IT

First, organizations have a reasonable expectation that IT will continue to support what might be considered "legacy applications."

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

This is what Gartner refers to as "Mode 1" IT. 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 require a rock-solid foundation, high availability, and a light touch when doing updates — and these updates 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.

Mode 2 IT

On the flip side of the equation, a new breed of applications is popping up, and they 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. These efforts are exploratory, with developers undertaking experimentation to address business needs. This more experimental approach is what Gartner refers to as "Mode 2" IT.

REMEMBER

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.

Security evolution

Although security has never been an afterthought — at least not in a very long time — today, it's a front-and-center issue with boardroom appeal. As attackers become increasingly savvy, malicious, and bold, organizations must put in place multi-layer defense strategies. Such systems include technology deployments that do more than ever before to protect themselves from attack. In addition to these systems, the human element is critically important. It is beyond the scope of this book, though it deserves your careful attention.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

One way that modern architectures can be built to limit the potential for widespread intrusion is to build in network micro-segmentation. I discuss this more later in this book, but for now, understand that *microsegmentation* enables granular control and management of all traffic into and out of individual or groups of virtual machines. This granularity helps implement boundaries that prevent the spread of malicious activity across the network.

REMEMBER

Security is a challenge everywhere, including your favorite public cloud provider! Humans remain the most significant security risk in every organization. In recent years, several high-profile breaches were not due to vulnerabilities but to user misconfigurations of public cloud resources.

Adoption of multi-cloud environments

Cloud has been around for a long time and will continue to grow. In the past, the question was, "Are you going to adopt cloud?" Today, the question is, "Which clouds are you going to adopt?"

Welcome to the new normal of the multi-cloud world. This is a place where organizations weave application and data fabrics that span all possible universes, from the private cloud, to AWS, to Azure, to Google, and everything in between.

As you might guess, getting from zero to a well-crafted multi-cloud environment is more easily said than done. You have innumerable configuration options, which means you have to wrangle it all. You need to tighten up super-exciting things like your data governance strategy and compliance plan. You also need to rein in what can be out-of-sight costs in the public cloud.

Why are companies doing this? Simply put, businesses seek to adopt multi-cloud architecture to address multiple use cases, avoid lock-in, and optimize between cost and SLAs, while fostering greater agility. However, the complexity of adopting a variety of services across multiple clouds leads to lack of cloud interoperability, limited visibility, and uncontrolled spending.

TIP

Cloud teams require solutions that turn cross-cloud management into one-click operations and offer centralized governance. This governance must continuously optimize consumption and enforce compliance across all cloud environments — public and private.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

The State of Public Cloud

Take a trip with me to yesteryear.

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 creeping up, 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.

It didn't happen quite that way.

REMEMBER

The post-IT future hasn't — and won't — come to pass. In fact, it's become clear that, although a "cloudy" world is both the present and the future, IT pros are continuing to thrive. According to research performed by Nutanix, applications are returning to on-premises environments at a significant rate. In this research, 80 percent of respondents indicated that they're moving 50 percent of their applications on-premises within two years. Moreover, 35 percent of respondents indicated that they exceeded their budgets on cloud spending.

But that doesn't mean that the public cloud is a failure.

Even though organizations have increased use of the public cloud as they've discovered applications and use cases where the public cloud makes sense, many continue to look for creative ways to get cloud-like benefits from their on-premises environments.

However, 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

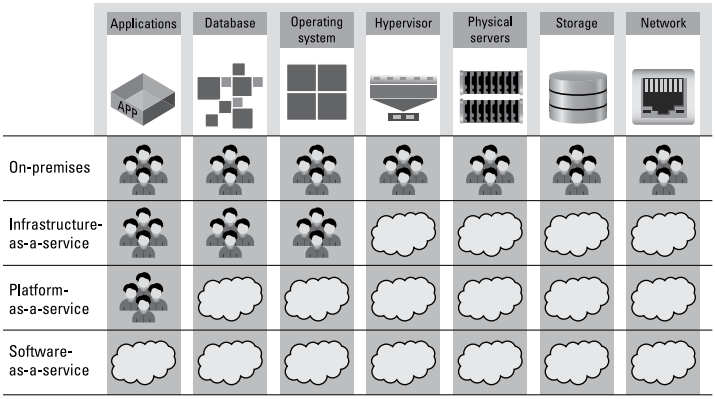


FIGURE 1-1: Comparing public cloud service types.

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). 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. For more stable or predictable workloads, the cloud is less 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 somewhere for a while.

In some cases, cloud adoption acts like a slingshot — a business builds and deploys an application on a public cloud service, but when the application reaches a certain scale, 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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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.

REMEMBER

People have finally realized 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. In fact, some of the big players have developed versions of their services that are specially hardened to meet industry-specific security and compliance requirements.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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.

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 of Amazon, although that is changing. While Amazon has been the public cloud leader and is certainly dominant, it is far from the only viable public cloud option. Over the past few years, Microsoft and Google have become formidable players as well.

All kinds of "as-a-service" cloud options that go far beyond Amazon are available to you. With all these services, enterprises must 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.

TIP

The right multi-cloud environment can help. By not putting all your public cloud eggs in one basket and by building an interconnected web of clouds, you make it easier to protect yourself if you need to make a quick jump from one of your providers.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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

Chapter 2
Why Enterprise Cloud?

Enterprise IT teams desire true cloud flexibility and elasticity, though traditional third-party cloud environments don't truly integrate with on-premises infrastructure. Organizations have thrown together massive amounts of IT services and third-party products to fuse disparate enterprise and public cloud technologies. However, the end result is yet another IT silo — with distinct operational and expertise requirements as well as higher overall expenses.

What is needed is a native and integrated approach that builds on a common cloud operating system to deliver a true extension of the datacenter environment. This is much like what Apple has done for IOS-supported devices with iCloud. It's a consistent experience across both stacks.

This cloud OS subsumes compute, storage, virtualization, and automation, eliminating fragmented IT. Cloud-based services should be invoked and implemented on demand, via one click from a single management pane. This pane should be able to provision and deploy applications that span multiple clouds, as well as migrate data seamlessly, to eliminate boundaries.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Instead of looking at this as 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 existing business-critical 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.

REMEMBER

Because of the need to ensure ongoing availability and performance levels for existing and next-gen applications, spending an inordinate amount of time tending to infrastructure is often frowned upon by the business, which may not 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

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.

However, 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 likely 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 a portion of their lifecycle.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

This purchasing practice is far from efficient. Most organizations overbuy and overprovision resources so they don't run out mid-cycle. This approach provides insurance, but it means money is being left on the table.

REMEMBER

The public cloud offers enterprises a *fractional consumption* model for IT resources. This model enables pay-as-you-grow economics, which allows you to buy resources on demand. This was one of the original 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 the 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 upfront. 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. Deployment time is one of the biggest ways by which latency increases in new business projects. 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, and configure the new hardware to integrate with your existing environment.

REMEMBER

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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Seamless infrastructure refresh

Every so often, the replacement cycle rears its ugly head. It's time to replace your SAN. Nothing strikes fear into a CIO's heart like having to deal with a refresh of the storage environment.

Refresh cycles can be expensive and risky, and they often require downtime. Replacing monolithic structures such as a SAN requires big capital expenditures, and you must pay staff off-hours to perform the migrations.

REMEMBER

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

People have far more trust and faith in the security of the cloud than they did in the past, although, in something of a surprise, a lot of CIOs would rather keep things closer. That said, 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 the good, the public cloud is not a panacea.

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

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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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.

In fact, managing predictable workloads is where IT shines. IT teams have been doing that for decades and do it 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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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.

REMEMBER

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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

TIP

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.

Controlling and optimizing public cloud costs

The opportunity to make good decisions around workload placement is one of the key reasons organizations have adopted and continue to adopt cloud. More often than not, these decisions revolve around cost, although performance, deployment time, and other technical factors may come into play, too.

As you move more into a multi-cloud scenario, though, keeping track of spending across these locations becomes increasingly challenging. One of the most common complaints among public cloud adopters is the surprise monthly invoice that shows a crazy leap in costs. Often, this is due to poor oversight of various cloud platforms.

You need a one-stop shop that allows you to see all your cloud spending in one place, enforce departmental budget rules, and more easily manage cloud costs.

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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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.

REMEMBER

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.

Security and trust in on-premises

For many, concerns about security also make on-premises deployments more attractive — the ability to control your own destiny trumps relinquishing such control to public cloud providers. The idea is that organizations can fully tailor their environments to

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

include just what they need . . . and nothing more. A more narrow deployment presents a smaller attack surface. Additionally, a growing ecosystem of tools is available to provide deep levels of security for on-premises environments.

Moreover, keeping full control of the infrastructure means that there is less opportunity for an unknown element to place that environment at risk.

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 anymore.

REMEMBER

These days, businesses want to use the public cloud where appropriate — for example, for backup, disaster recovery, and applications with highly unpredictable resource needs — 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)

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

- No local infrastructure operational overhead (although there is still operational overhead and the need for specialized skills and training)
- 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.

With enterprise cloud, enterprise IT can

- Drive simplicity for multi-cloud governance with an open approach that values flexibility and optionality.
- Empower end-users with self-service of on-premises and public resources.
- Automate deployment and management of applications across multiple environments without compromising on governance and control.
- Unify governance across all cloud environments and teams for optimal resource utilization and compliance.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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

Chapter 3
What Is an Enterprise Cloud?

Chapter 2 tells you why an enterprise cloud is important. In this chapter, I define what makes an enterprise cloud. I introduce the five critical characteristics that define an enterprise cloud and show 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Integrating the ethereal: On-premises + multi-cloud

This book uses the word *multi-cloud* a bunch of times, and with good reason: Multi-cloud is here. In the multi-cloud world, enterprises can procure many cloud services and stitch those services together to meet their needs. For example, they may use Azure for Microsoft applications, Google for analytics, and so on. One drawback, however, is that a multi-cloud approach risks creating multiple operations silos.

Enterprise cloud integrates all these ethereal elements under one umbrella. With an understanding that, for most companies, on-premises environments simply aren't going away, the enterprise cloud makes the private cloud component a first-class citizen in the architecture hierarchy.

Understanding the need for a single control plane

Perhaps one of the most critical components in an enterprise cloud environment is a single, end-to-end control plane. Without one, administrators are forced to rely on and correlate across multiple consoles and teams as they seek to accomplish individual tasks.

A single control plane makes it possible for administrators to manage a vast array of infrastructure and services, regardless of whether the infrastructure or service resides in the datacenter down the hall or in a public cloud vendor's datacenter in Singapore.

More importantly, a single control plane provides a consistent view of the entire distributed environment. A cloud control plane acts as the broker among the discrete services that comprise an enterprise cloud. Such a control plane arms administrators, and sometimes developers, with data they need about where workloads will run and how they are managed.

Understanding Historical Private Cloud Roadblocks

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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Well, others have tried and 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. Although OpenStack had good beginnings, it proved to be complex and just added another domain silo.

That said, the goal to deliver self-service provisioning, monitoring, billing, and chargeback persists. However, the underlying infrastructure is still based on storage accessed over a storage network that is deployed and scaled in big chunks. What's needed is a replatforming 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. 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 complexity and unpredictability, scenarios 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 not only combines the ability to leverage scale-out storage while maintaining data locality, it does so without introducing complexity and additional management overhead.

REMEMBER

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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Defining Enterprise Cloud

Here's how to define *enterprise cloud:* It delivers the agility, simplicity, and fractional consumption of public cloud services while providing control over performance, location of data and services, and choice of platforms.

Six key components comprise the enterprise cloud:

- **Unified governance** models are the policy-centric instantiations of the single control plane discussed earlier in this chapter.
- **Full-stack infrastructure and platform services** 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** deliver operational simplicity through automation and insights.
- **Rapid elastic consumption** allows businesses to buy and use only the IT resources they need, and nondisruptively scale when demands grow. It helps embrace an OpEx model within and outside of the datacenter.
- **Integrated security and governance** covers the entire infrastructure stack across private and public clouds, leveraging automation to maintain a security baseline.
- **Application-centric mobility** lets businesses place and move applications anywhere, with no infrastructure lock-in.

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

Unified governance

One of the critical missing pieces for many organizations used to be comprehensive governance that spanned the organization. Sure, organizations would *say* they had super-duper governance structures in place, but pockets of shadow IT lurked throughout the company.

Bad CIOs reacted by trying to stomp out shadow IT through policy. Good CIOs listened and learned why business units were avoiding IT and doing their own thing. These CIOs learned that existing IT was slowing down business units, standing in the way of projects.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Business units wanted more autonomy and speed, and shadow IT was the easy answer.

These good CIOs then created updated governance models in their organizations that made technology-based decisions more inclusive. IT would still have to provide input into solutions to ensure security and compatibility and to minimize service overlap, but business units would have some ability to control their own destiny. In such models, IT becomes a broker of services with a governance model that recognizes this shift. Of course, IT still provides services via on-premises infrastructure, but using an approved service catalog, business units are free to buy what they need from where they need it.

USING SHADOW IT TO YOUR ADVANTAGE

Shadow IT is a set of symptoms that can help you figure out exactly where IT is not meeting organizational goals or where this is governance weakness. Every shadow system in use is an example of one of two things:

- A need that could not be satisfied by IT
- A department that worked around established governance processes and, if it hadn't been for that meddling CIO, would have gotten away with it

I see shadow IT as a menu for CIOs. It's a gift. It's a list of places that need attention in some way. Of course, this doesn't mean CIOs should wander around and find out who's broken the rules. Rather, it's a chance for a CIO to establish support mechanisms that bring former shadow systems into the light and eliminate those shadow systems that should not exist.

That's right . . . not every shadow system should exist. Some such systems are security risks, and others may place the organization at some other kind of risk. For example, I've seen colleges in which individual departments were creating their own databases, resulting in figures that didn't match with other departments. The risk is paralysis. Which data set is the truth? With the right set of tools, CIOs can help their organizations get beyond shadow IT and shore up their governance models at the same time.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

The governance model provides support for the unified world of enterprise cloud and encompasses technical, financial, security, and compliance requirements.

Full-stack infrastructure and platform services

Regardless of where you decide to run your critical applications, you need 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 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.

TIP

Rather than buying a super-expensive monolithic SAN, buy infrastructure that you can leverage to meet the needs of your current and emerging 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.

However, virtualization today, particularly in the public cloud, should be treated as assumed and available, not a separate product to procure.

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 *web-scale* engineering characteristics:

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

- Software-defined
- Distributed *everything* to avoid single points of failure
- 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. Moreover, the infrastructure should be full-featured with offerings that support a myriad of workload needs. On storage, for example, the infrastructure you provide should be capable of supporting file and block storage alongside virtual machine storage. Different applications have different storage requirements, after all.

The platform needn't be limited to the private cloud or your local datacenter. You should be able to support hybrid delivery of applications — that is, to provide choice between on-premises infrastructure and public cloud services for your business-facing applications. In this context, "platform" also means the administrative experience.

Your enterprise cloud environment must also 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 managing infrastructure challenging because you have to manage each silo separately. The enterprise cloud must support bare metal, virtualized, and containerized environments for any application.

Finally, you simply can't forget about security. In a full-stack environment, security comes in many layers. Your platform of choice must have security baked in so that you don't have to try to bolt something on later.

REMEMBER

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. However, unless infrastructure is built to scale out, includes comprehensive APIs, is less complex, provides native services, and doesn't include single points of failure, you will not end up with a cloud-like environment.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Figure 3-1 provides a look at the web-scale world.

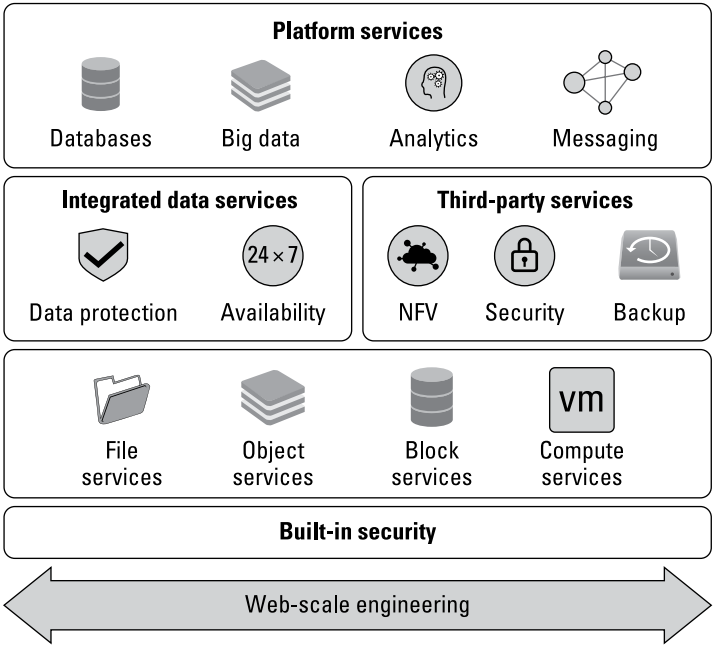


FIGURE 3-1: The web-scale world.

The similarities between Figure 3-1 and modern hyperconverged infrastructure platforms shouldn't come as a surprise. Hyperconvergence has emerged as a powerful enabler for organizations, providing comprehensive infrastructure service and forming the foundation for enterprise cloud efforts.

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.

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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

REMEMBER

As you deploy an enterprise cloud and a combination of public and private cloud, the ability to reduce administrative overhead via simpler infrastructure management, automation, and orchestration becomes critical. The resulting 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 1980s 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. 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, humans 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

REMEMBER

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.

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 return on investment (ROI) by keeping the IT administrative paradigm simple.

REMEMBER

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 — while not compromising on performance or capabilities to run the most demanding applications. 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

How do you make automation happen? One way is through the use of application programming interfaces (APIs). Modern infrastructure products often include comprehensive and powerful APIs that control every aspect of the environment.

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.

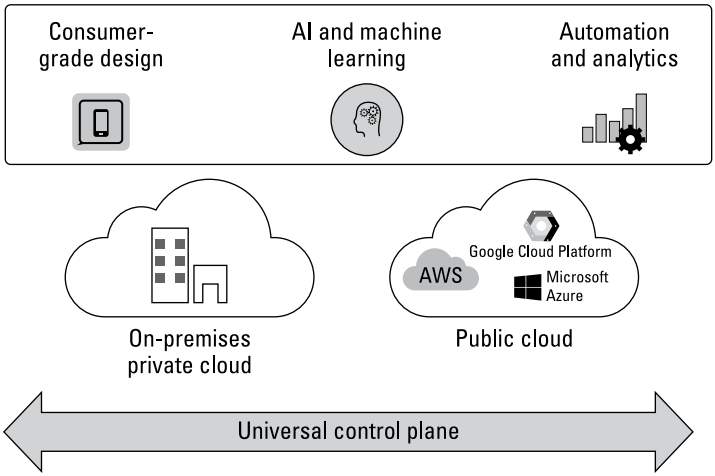


FIGURE 3-2: How the universal control plane supports enterprise cloud.

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 upfront, but the discounts might be less generous when your back is against the wall and your capacity is dwindling.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

- **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.

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

With 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.

REMEMBER

With 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 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.

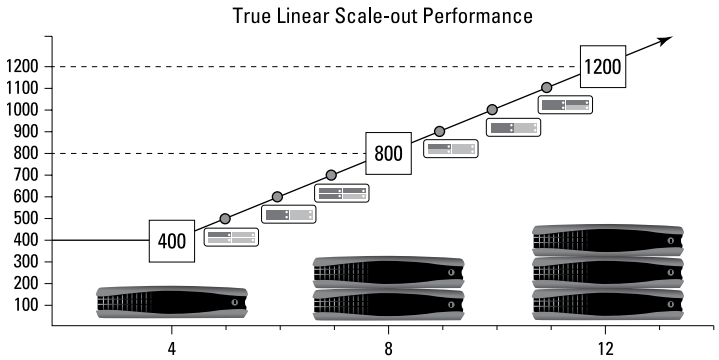


FIGURE 3-3: Linear scale-out infrastructure supports the needs of the enterprise cloud.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Integrated security and control

Organizations are working hard to secure their environments against attacks from within and from the outside. As the potential fallout from security issues increases — bad public relations, fines, lost business — companies must 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 that 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. You're messing around with security when you should be working on business-facing activities.

Life is 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.
- **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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

WHY MICROSEGMENTATION IS SO CRITICAL

Microsegmentation is an important concept that's worth a closer look:

- The network is the foundation for everything that happens in the connected enterprise. Networks extend from the desktop to the cloud to the airwaves. The modern network is a ubiquitous beast.
- With increasing application complexity, driven by distributed services and rapid growth in SaaS and cloud-based offerings, the need to easily visualize, analyze, and govern network communications is critical.

The connected enterprise requires an integrated software-defined virtual networking solution that doesn't compromise and that doesn't add cost and complexity. That capability starts with visualization. Visualization allows administrators to better understand traffic patterns between virtual machines.

From there, microsegmentation provides granular control and governance of all traffic into and out of individual virtual machines or groups of VMs. It ensures that only permitted traffic between application tiers is allowed and protects against advanced threats propagating within the virtual environment.

After all, this is where real security challenges exist. Companies have spent tons of money on protecting the network perimeter, but the bad guys get in anyway. Microsegmentation helps you limit the exposure of a breach by monitoring, managing, and restricting inter-VM traffic.

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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

- **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.

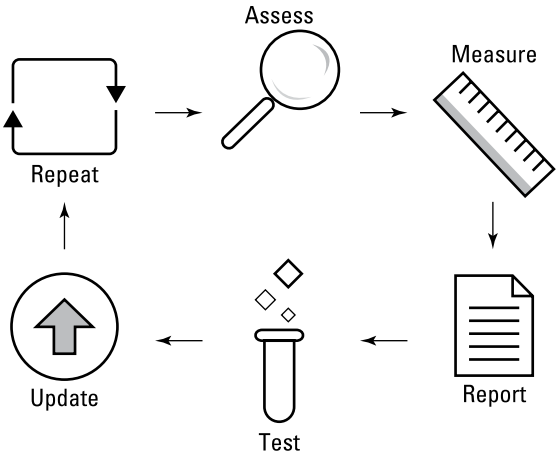


FIGURE 3-4: A regular security posture review of your infrastructure and application environment is essential.

Application mobility

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 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 higher levels of performance than others.

On the economic 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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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

REMEMBER

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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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

Chapter 4
Building an Enterprise Cloud

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 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.

REMEMBER

As with so many things, you must adapt or become irrelevant. Look at how many mainframe operators didn't survive the wave of decentralization that defined IT in the 1980s and 1990s.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Change is inevitable. It's time for you to adjust your perspective on several fronts — beginning right now — so that you stay relevant into the 2020s and beyond.

Why shadow IT should be understood and then 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. Many end-users rival and exceed IT staff technical knowledge *and* they are the subject matter experts in their areas.

TIP

Users don't want to be slowed down by IT anymore, either. They want to do their jobs on their own terms.

This phenomenon is 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 and compliance:** Although end-users have become far more savvy about technology, they are often far less knowledgeable than IT about security and compliance laws. 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, that data may have inadequate security measures.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

- **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 has 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 extend its portfolio and begin encompassing services that were stood up by end-users.

REMEMBER

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. 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. The rule has evolved to mean 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

REMEMBER

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.

TIP

The 80/20 rule is a remnant of a different time. With business needs changing quickly, 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.

You may never fully escape the 80/20 rule, however. As you improve your practices, a new version of the basics emerges — but ideally they'll be far more productive and efficient than your earlier practices.

Why you need to address bimodal IT

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 a legacy environment and other 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 building 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Why automation and APIs are the way of the future

If you aren't using APIs today, you're missing out. APIs are present in just about everything and there are even web services that bring massive API exchanges to your fingertips.

The purpose of an API is to automate something. APIs provide you with the means to granularly control target hardware and applications with snippets of code. With the right APIs, you can do everything from spinning up new virtual machine clusters in your on-premises datacenter and in Amazon, to creating the networking fabric that will enable communication between these new virtual machines.

In an infrastructure context, some companies refer to this as *infrastructure as code,* intending to convey that APIs basically turn hardware devices into programmable objects that can be bent to your will. Ever-present APIs are making it much easier to build solutions that are, in fact, greater than the sum of their parts. You can cobble together an incredible solution from a bunch of little pieces and build something that appears as if it's a single, functional unit.

This is why APIs are the future and they're the most common way that automation will be handled as time goes on.

Preparing Your People

Your people include your IT staff and your end-users. Both groups need attention in an enterprise cloud.

Leveling up IT organizational skills

Compared to dealing with people, the technology is easy! However, you need to ensure that your people are prepared for the changes to your IT organization and 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

REMEMBER

Hyperconverged infrastructure-driven enterprise cloud systems require a different set of skills that allow a focus on higher-level outcomes, rather than tedious drudge work.

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 complex technology in the datacenter.

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.

Empowering end-users

The enterprise cloud, though, isn't about IT. It's about empowering your user base as your company jump-starts its digital transformation efforts. With a service catalog at the center of the user experience, the enterprise cloud can be configured to allow individual business units to deploy their own IT resources using rules that are defined by central IT. This way, rather than calling the help desk every time they need something, a business unit can deploy a new web front end to help support, for example, a marketing initiative.

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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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.

REMEMBER

Most IT departments buy what they think they'll need for the duration of the replacement cycle. These estimates are difficult because business needs change. What is correct today may not be a year from now. One thing is certain, though. For some period of time, you won't use 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. This is the *zero return on investment zone*. More than three years go by before the 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.

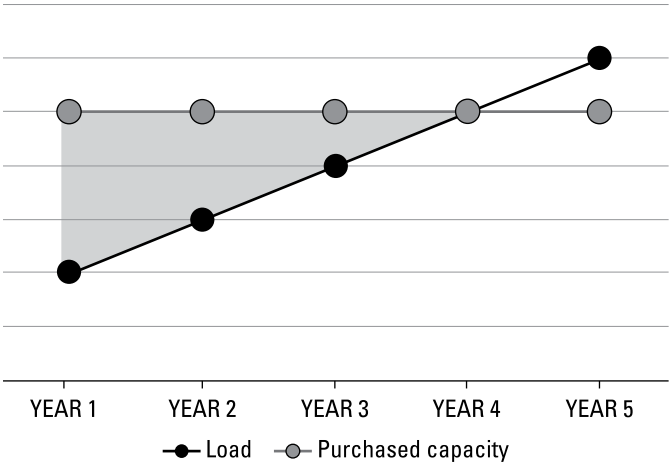


FIGURE 4-1: Traditional infrastructure procurement economics is not a viable solution.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

REMEMBER

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 cloud-like 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.

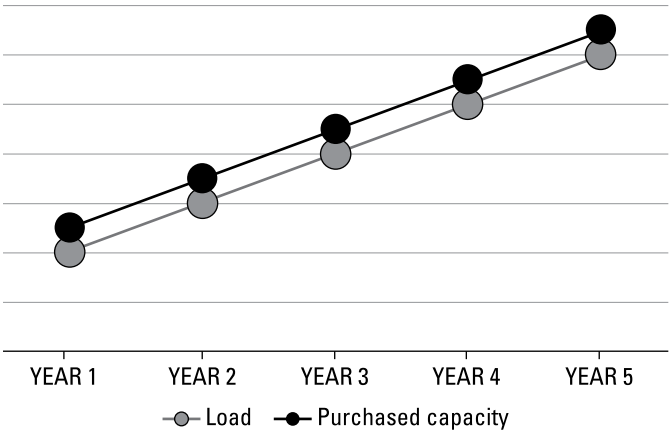


FIGURE 4-2: Hyperconvergence and enterprise cloud can help you reinvent the IT budget.

Understanding disaggregation

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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

running low on storage capacity, your new node can be storage heavy, although it will have CPU and RAM as well.

REMEMBER

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.

Consolidating storage services

Speaking of disaggregation, let's talk about storage. A full-featured enterprise cloud platform allows you to deploy all the storage-centric workloads that you need, regardless of type and without having to deploy a duplicate hardware environment. The storage layer can handily support virtual machines, block storage services, and object storage services.

With a broad storage foundation, the on-premises portion of the enterprise cloud platform can be configured to support the internal requirements imposed by workloads running directly inside that platform or can be used to support the needs of other standalone systems.

Ensuring consistent performance

Making sure a multi-cloud environment is cost effective is great, but if the applications that you're operating in a multi-cloud environment run about as fast as a turtle hightailing it through a field of molasses, then you have a problem.

You can buy any number of third-party monitoring tools to keep tabs on your applications, but once you've purchased them, you must go through a deployment process and align resources so you're watching the right stuff.

Alternatively, you can get a solution that's already baked right into the enterprise cloud stack. By doing so, you leverage that solution's inherent understanding of your company's infrastructure environment, no matter how wide it is. A good solution also includes a default dashboard for common applications via a marketplace and provides equal support for infrastructure systems as well as running applications.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Choosing the right public cloud

Deciding which private cloud services you want to adopt can be difficult without the right tools in place. After all, you don't want to have an aneurysm after getting your first cloud bill. To that end, you need your enterprise cloud platform to provide tools that constantly watch cloud pricing and keep you abreast of what's the most cost-effective. This doesn't mean that you'll bounce your workloads around between cheap providers in what will surely be a race to the bottom. It means that, upfront, you want to find the right public cloud so that you can build a fabric that extends from that public cloud all the way down to your individual virtual machines.

Keep two factors in mind: pricing and technical fit. All public clouds are not the same. There are reasons to choose Google over Azure, or Azure over Google, and so on. By abstracting the fabric — that is, looking at it and operating from a level above the cloud provider — you enable flexibility of choice so that you can choose the right cloud for the right application at the right time. You no longer need to choose one cloud and shoehorn everything in. Now, you match individual workload needs against cloud provider technology and pricing and then choose the provider that makes the most sense.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

IN THIS CHAPTER

- Discovering how John the sysadmin goes from zero to enterprise cloud
- Learning how John deploys self-service capabilities for his users
- Studying the tools that make the enterprise cloud a reality today

Chapter 5
Exploring an Enterprise Cloud Deployment

Chapters 2 through 4 prepare you to think more deeply about enterprise cloud, and this chapter helps you put your new knowledge into practice! In a series of "take your admin to work days," I offer scenarios to show how you can go from zero to hero for your company's enterprise cloud efforts. I also talk about how Nutanix technologies address deployment across the stack — infrastructure to applications — and enable self-service for end-users.

Deploying Infrastructure

No matter how much you talk about software, hardware can't be overlooked in the enterprise cloud. In this case, that hardware requirement is satisfied by any one of Nutanix's hyperconverged infrastructure platforms or, if you prefer to run on any number of supported platforms, the Nutanix software can run on non-Nutanix provided hardware, too.

Put yourself in the shoes of John, a standard admin, charged with deploying infrastructure for his company. John's current infrastructure environment is a cobbled-together conglomeration of

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

hardware and software from a bunch of different vendors. He knows that the company needs to simplify this morass and he's recommended considering hyperconverged infrastructure as the foundation for a new enterprise cloud architecture.

John takes a peek at Nutanix's AOS, shown in Figure 5-1, which combines storage, compute, networking, and virtualization into a single solution. The platform also provides support for containers, which are a key part of John's company's future roadmap. It comes with data protection and encryption capabilities, and ensures that data is collocated with the workloads that need it. Moreover, built into the platform is the ability to move the company's expensive and unreliable disaster recovery (DR) service out in favor of one that leverages the cloud as a DR target.

REMEMBER

On the virtualization front, John decides to use Nutanix's AHV across the board. AHV is based on the popular open-source KVM hypervisor, but has been significantly extended to bring first-class features to the platform. For example, AHV includes a dynamic scheduling feature, which provides intelligent workload placement and resource contention avoidance. The best part? It's included by default in the platform. There's no need to pore over licensing matrices to figure out what's included in the product.

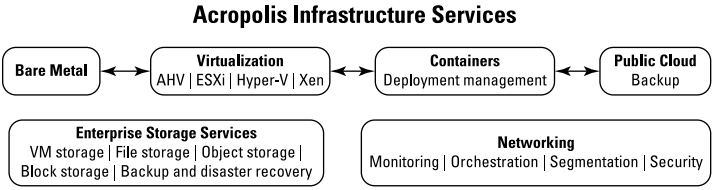


FIGURE 5-1: An overview of Nutanix Acropolis.

John knows that he can't turn off his old stuff overnight. Transitions of this kind always require some cutover period. That's why he's pleased that his shiny new platform provides Volumes, Files, and Buckets. Volumes is a native block storage service for physical applications. Files is a file storage solution for unstructured files, such as user home directories and the like. Buckets is object storage that addresses long term retention, backup, and developer needs.

With these services, John can easily integrate his new platform into his existing environment and migrate at a pace that makes sense for his organization.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Deploying Workloads and Applications

As often happens in IT, new work comes along and John has new tasks, which now include deploying new workloads and applications into the environment. Fortunately, an enterprise cloud environment simplifies these tasks, which is good because he still has his original job to do as well.

The company has a lot of complex applications to support, so John decides to use Nutanix Calm to help maintain sanity. Calm provides a number of features that the company wants to leverage. These features will help it get critical apps deployed today, and, more importantly, make it easier to implement some advanced capabilities, such as user self-service, down the line.

REMEMBER

Calm's most critical feature revolves around application lifecycle management. With Calm, John can fully automate the provisioning, scaling, and deletion of traditional multitiered applications, as well as modern distributed services. Moreover, he can create application blueprints for these applications, making redeployment much easier later on. Blueprints include all elements of each app, such as relevant VMs, configurations, and related binaries. Figure 5-2 shows an application configuration in Calm.

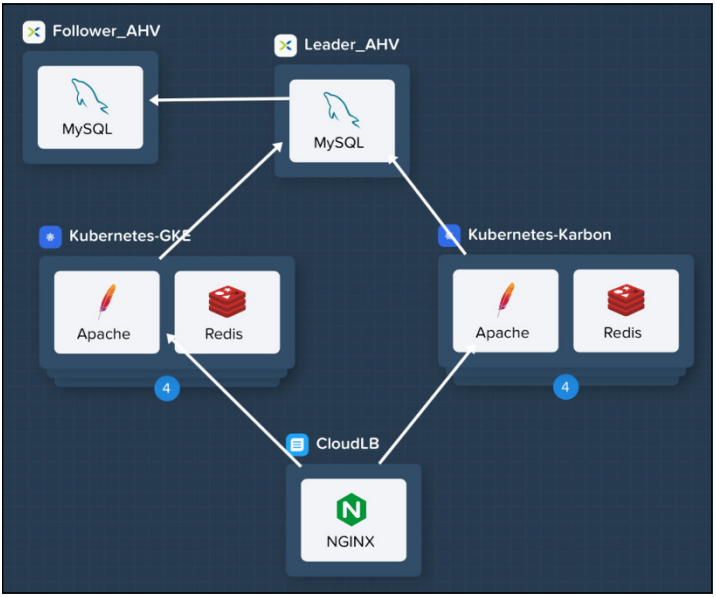


FIGURE 5-2: Configuring a distributed application in Nutanix Calm.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Best of all, once a blueprint is built, administrators can publish it to the Marketplace, an enterprise app store. This enables application owners and developers to request IT services that can then be instantly provisioned.

John also needs a better way to manage the company's burgeoning database challenges. Developers are keeping dozens of copies of databases strewn about the environment, wasting resources and making it difficult to remember which database is which. John decides to use Nutanix Era to help solve the sticky challenge of database copy data management.

REMEMBER

Era allows developers to create space- and time-efficient database snapshots and to clone databases to any point in time. This is the marquee feature of Era. With a core construct called Time-machine, Era captures all database states for any given service-level agreement (SLA) and allows DBAs to not only create fully functional database copies, but also refresh existing copies.

Administering the Enterprise Cloud Environment

The company asks John to make sure that the administration of the enterprise cloud environment is handled in a way that adheres to the requirements specified in this book — namely, that there is a single interface for administration that spans the entire environment.

John has all along been using Nutanix Prism, which is a complete enterprise cloud management solution. Enterprise clouds require machine intelligence and automation to simplify complex operations and get as close to zero-touch management as possible. Prism offers a consumer-grade management experience for virtualized datacenter environments by combining several aspects of administration and reporting.

Behind all of this is advanced machine learning technology. Prism mines large volumes of system data to automate common tasks to optimize the administration experience and minimize the need to touch infrastructure. Figure 5-3 provides a look at Prism.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

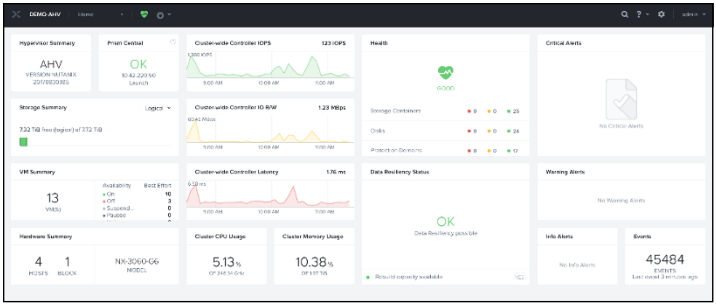


FIGURE 5-3: Nutanix Prism in action.

REMEMBER

The Prism console enables John to manage entire clusters of infrastructure, manage virtual networks, manage storage, and get an overall visualization of the environment. From the same console, John receives detailed capacity planning advice and alerts when he's getting low on infrastructure or when some other anomaly occurs.

Enabling Self-Service

Under a company initiative to increase efficiency, John has been tasked with figuring out how to deploy self-service capabilities to his organization. He discovers that Nutanix Calm exists entirely within Prism and investigates its capabilities for the assist.

To ensure good self-service, organizations must provide two things:

- Users need an easy way to procure workloads.
- Good governance rules and policies must be in place to make sure that only authorized users are allowed to deploy certain workloads and that any unique per-department policy is followed. For example, there's likely no reason for the accounting assistant to be allowed to create a new DHCP server, though engineering doesn't require the same data protection policy as our friend the accounting assistant.

Calm provides the marketplace and governance capabilities so that John's users can get the applications they need without slogging through the traditional IT procurement cycle.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

REMEMBER

On the governance front, Calm maintains control with role-based governance that limits user operations based on permissions. Additionally, all activities and changes are centrally logged for end-to-end traceability, aiding security teams with key compliance initiatives — you know who did what and when they did it.

Managing Enterprise Cloud Costs

John's next task is to help figure out which enterprise cloud components do the best job at ensuring that costs remain in check.

REMEMBER

A couple of components fit the bill. The first is Nutanix AHV, a component of the Acropolis platform. AHV is a license-free hypervisor that can replace an existing hypervisor in your private cloud environment. Replacing an expensive hypervisor with one that carries no separate licensing fee can be a boon to John's stretched IT budget. Moreover, it becomes possible to remove the complexity of what are likely multiple components that were required to manage the other platform, such as a separate operations tool, automation tool, site recovery tool, and more.

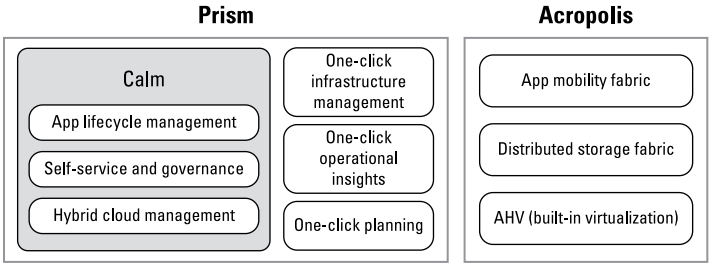


FIGURE 5-4: The Nutanix Enterprise Cloud platform.

AHV includes all the critical features John needs, as well as some capabilities that make his migration to enterprise cloud far more seamless. AHV includes live migration of VMs between hosts, and, more importantly for John's migration needs, it includes the ability to migrate VMs between hypervisors, meaning he can easily move away from his old environment.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

AHV also includes management of VM images, VM high-availability capabilities, network virtualization, and much more. It's all managed via Prism, as you'd want with any enterprise cloud platform.

Scaling the Environment

This is where the real beauty of what's been built shines through. Consider, for a moment, the bad old days. John would have needed to spend weeks — maybe months — deciding what he needs. He'd have to spend time on every resource, poring over usage reports and trying to figure out how quickly his company's individual resource needs are changing.

With the Nutanix enterprise cloud platform spanning John's on-premises environment and the public cloud, the guesswork is gone. The powerful data gathering capabilities of Prism Central tells him exactly when his on-premises environment is going to run out of resources.

Best of all, when John needs to add more, all he has to do is plug in an HCI node that contains the right mix of CPU, RAM, and storage for those additional needs. That's it. No muss and no fuss.

The same happens on the public cloud side of the equation. As John's public cloud environment gets close to capacity, he receives a warning and has the opportunity to spin up more resources. Alternatively, using the platform's APIs, he can configure these resources to be added automatically as needed.

REMEMBER

As you peel back the layers of what's available in the Nutanix enterprise cloud platform, it becomes clear that all the pieces are there and that they're well integrated and ready to help organizations achieve their enterprise cloud vision.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

IN THIS CHAPTER

- Understanding the enterprise cloud model
- Looking to the future

Chapter 6
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:

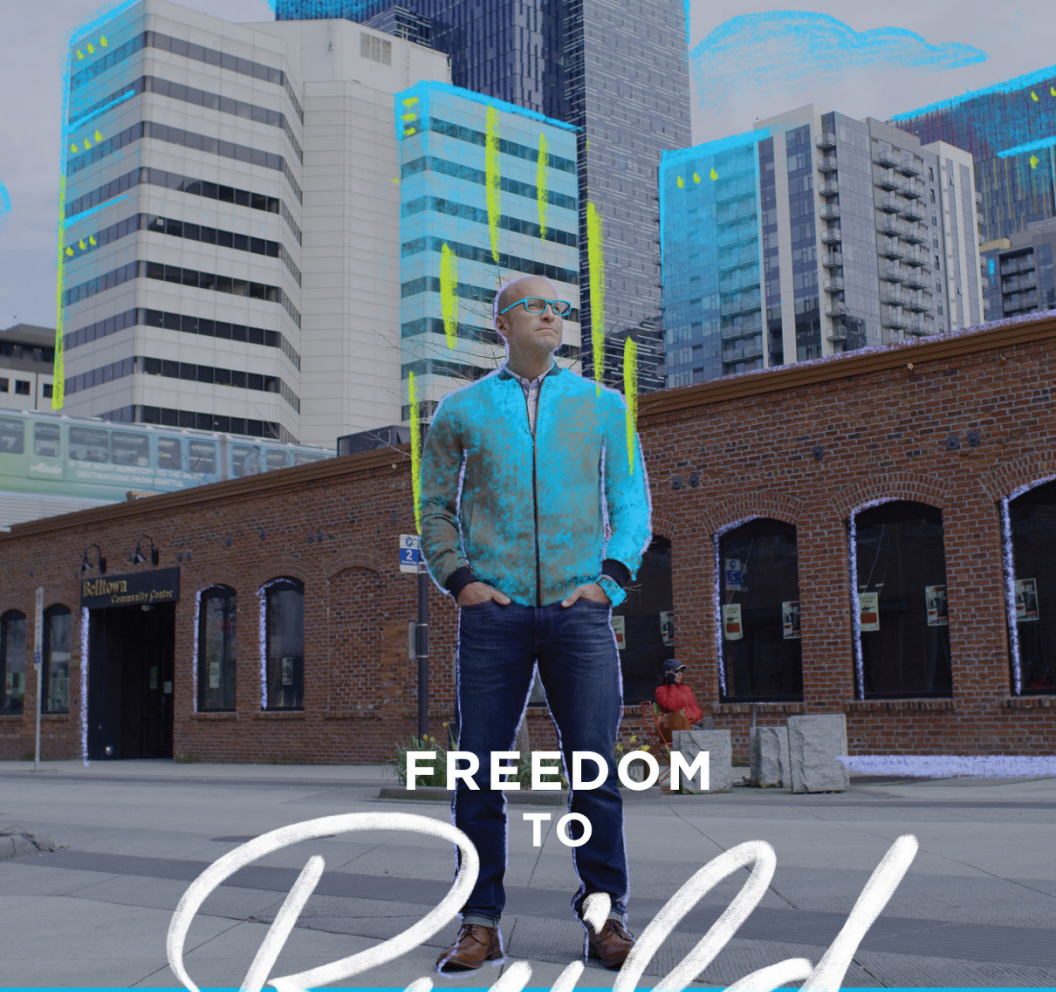
- **A brand new economic model:** Your old legacy-based IT economic model is no longer sufficient. With enterprise cloud, you can adopt the pay-as-you-go characteristics of the public cloud while providing a common foundation upon which to run both legacy and new-style or cloud-native apps.
- **A focus 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

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

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 and extend those capabilities to your public cloud-centric workloads.
- **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 is here to stay. With enterprise cloud, you get the outcomes promised by bimodal IT without worrying about the inefficiencies that this bifurcated 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.

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.



FREEDOM TO *Build*

Stop building data centers. Start building clouds. Transform your enterprise with the leader in hyperconverged infrastructure.

nutanix.com/freedom

NUTANIX™

YOUR ENTERPRISE CLOUD

These materials are © 2019 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Get the best of cloud and on-premises

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 delivers public cloud benefits with on-prem security and control. This book helps you understand what the enterprise cloud is and how it can help your organization compete and innovate in the era of digital transformation.

Inside...

- Embrace the digital era
- Leave three-tier limits behind
- Eliminate operational complexity
- Understand multi-cloud technology
- Explore an enterprise cloud deployment
- Focus on the applications and services that power your business

NUTANIX™

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.

Go to Dummies.com®
for videos, step-by-step photos, how-to articles, or to shop!

ISBN: 978-1-119-57270-1
Not For Resale

for dummies®
A Wiley Brand

Also available as an e-book

WILEY END USER LICENSE AGREEMENT

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