Making Everything Easier!™

VMware Special Edition

Micro-segmentation

FOR DUMMIES®

A Wiley Brand

Learn to:

- Develop an inherently secure data center
- Prevent lateral spread of a data center attack
- Deploy a platform for advanced security solutions

Brought to you by

vmware®

Lawrence Miller, CISSP

Joshua Soto

About VMware

VMware is a leader in cloud infrastructure and business mobility. Built on VMware's industry-leading virtualization technology, our solutions deliver a brave new model of IT that is fluid, instant, and more secure. Customers can innovate faster by rapidly developing, automatically delivering, and more safely consuming any application. The company is headquartered in Silicon Valley with offices throughout the world and can be found online at **www.vmware.com.**

Micro-segmentation

FOR DUMMIES®
A Wiley Brand

VMware Special Edition

by Lawrence Miller, CISSP,
and Joshua Soto

FOR DUMMIES®
A Wiley Brand

Micro-segmentation For Dummies®, VMware Special Edition

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

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.

ISBN 978-1-119-17734-0 (pbk); ISBN 978-1-119-17735-7 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

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.

Publisher's Acknowledgments

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

Development Editor: Elizabeth Kuball

Copy Editor: Elizabeth Kuball

Acquisitions Editor: Katie Mohr

Editorial Manager: Rev Mengle

Business Development Representative: Karen Hattan

Production Editor: Kinson Raja

Special Help: Geoff Huang, Catherine Fan, and Kausum Kumar

Table of Contents

Introduction

Traditional approaches to securing data centers have focused on strong perimeter defenses to keep threats on the outside of the network — not unlike castle defenses during medieval times! Towering castle walls were fortified with battlements and bastions, and access was controlled with a firewall — uh, drawbridge. For an attacking force, breaching the perimeter and gaining entry to the castle was the key to victory. Once inside the castle, defenses were practically nonexistent, and the attackers were free to burn and pillage!

However, this model is ineffective for handling today's new and evolving threats — including advanced persistent threats (APTs) and coordinated attacks. What's needed is a more modern, sophisticated approach to data center security: one that assumes threats can be anywhere — and are probably everywhere — and then acts accordingly. Micro-segmentation not only adopts such an approach, but also delivers the operational agility of network virtualization that is foundational to a modern software-defined data center.

Cyber threats today are coordinated attacks that often include months of reconnaissance, vulnerability exploits, and "sleeper" malware agents that can lie dormant until activated by remote control. Despite increasing types of protection at the edge of data center networks — including firewalls, intrusion prevention systems, and network-based malware detection — attacks are succeeding in penetrating the perimeter, and breaches continue to occur.

The primary issue is that once an attack gets past the data center perimeter, there are few lateral controls to prevent threats from traversing inside the network. The best way to solve this is to adopt a stricter, micro-granular security model with the ability to tie security to individual workloads and the agility to provision policies automatically. Forrester Research calls this the “Zero Trust” model, and micro-segmentation embodies this approach.

With micro-segmentation, fine-grained network controls enable unit-level trust, and flexible security policies can be applied all the way down to a network interface. In a physical network, this would require deploying a physical firewall for every workload in the data center, so up until now, micro-segmentation has been cost-prohibitive and operationally unfeasible. However, with network virtualization technology, micro-segmentation is now a reality.

About This Book

This book provides a broad overview of micro-segmentation in the data center. After reading this book, you’ll have a good basic understanding of micro-segmentation — like you’d get from a college-level 101 class, but far more interesting than Microbiology 101 or Microeconomics 101 (and not as difficult either)!

Foolish Assumptions

It’s been said that most assumptions have outlived their uselessness, but we assume a few things nonetheless:

- You have a strong working knowledge of networking and security fundamentals, concepts, and technologies, and a good understanding of virtualization.
- You work in a large organization or enterprise that operates one or more data centers in a public, private, or hybrid cloud environment to support your critical business functions.
- You’re a security executive, such as a chief information security officer (CISO) or chief security officer (CSO), evaluating data center security strategies and solutions for your organization. If that’s the case, then this is the book for you!

Icons Used in This Book

Throughout this book, we occasionally use special icons to call attention to important information. Here's what you can expect:

REMEMBER

This icon points out information that may well be worth committing to your nonvolatile memory, your gray matter, or your noggin — along with anniversaries and birthdays!

TECHNICAL STUFF

You won't find a map of the human genome here, but if you seek to attain the seventh level of NERD-vana, perk up! This icon explains the jargon beneath the jargon!

TIP

Thank you for reading, hope you enjoy the book, please take care of your writers! Seriously, this icon points out helpful suggestions and useful nuggets of information.

WARNING!

Proceed at your own risk . . . well, okay — it's actually nothing *that* hazardous. These helpful alerts offer practical advice to help you avoid making potentially costly mistakes.

Beyond the Book

Although this book is chock-full of information, there's only so much we can cover in 72 short pages! So, if you find yourself at the end of this book, thinking, "Gosh, this was an amazing book — where can I learn more about micro-segmentation?," simply go to `www.vmware.com`.

Where to Go from Here

With apologies to Lewis Carroll, Alice, and the Cheshire Cat:

"Would you tell me, please, which way I ought to go from here?"

"That depends a good deal on where you want to get to," said the Cat — er, the Dummies Man.

"I don't much care where . . . ," said Alice.

"Then it doesn't matter which way you go!"

That's certainly true of *Micro-segmentation For Dummies,* which, like *Alice in Wonderland,* is also destined to become a timeless classic!

If you don't know where you're going, any chapter will get you there — but Chapter 1 might be a good place to start!

However, if you see a particular topic that piques your interest, feel free to jump ahead to that chapter. Each chapter is individually wrapped (but not packaged for individual sale) and written to stand on its own, so you can start reading anywhere and skip around to your heart's content! Read this book in any order that suits you (though we don't recommend upside down or backward).

We promise you won't get lost falling down the rabbit hole!

Chapter 1

Defending the Data Center on a Broken Foundation

In This Chapter

- Recognizing the impact of data center breaches
- Understanding how attacks exploit the unguarded inside of the data center
- Identifying what's wrong with data center security

Data centers have become the virtual bank vaults of the 21st century. Sensitive corporate, financial, and personal information stored on data center systems is potentially worth hundreds of millions of dollars for today's cybercriminals. Although dependence on these systems has grown dramatically over the past few decades, the underlying foundation for delivering advanced security for these systems remains relatively unchanged: a strong focus on external perimeter security with little to no attention focused on stopping threats inside the data center.

In this chapter, you explore data center breaches — how they happen and why traditional data center security approaches that leave the inside of the data center relatively defenseless are ineffective.

Data Breaches Continue to Occur

Despite a heightened focus on security in the enterprise as evidenced by increasingly stringent data protection laws, heavy investments in security technology, and ever growing and ever capable security teams, data center breaches continue to occur at an alarming rate — and each new breach dwarfs the last in terms of millions of records and dollars stolen.

While recent attacks on Anthem, Home Depot, Sony, Target, and others have been different, they all have one characteristic in common: Once the perimeter was breached, the attacks were able to propagate laterally from server to server within the data center with essentially no security controls in place to stop them from spreading. Sensitive data was then collected, exfiltrated, and exploited. These cases highlight a major weakness of modern data centers: Tremendous effort and technology is applied to securing the perimeter of the data center, but the same level of security does not exist inside the data center. To effectively address this weakness, any security technologies and controls that are applied to the perimeter of the data center need to be considered and implemented (if appropriate) *inside* the data center as well, in order to stop or isolate attacks once the perimeter is breached.

According to Verizon's *2015 Data Breach Investigation Report,* in 2014 there were 79,790 confirmed security incidents worldwide and 2,122 confirmed cases where sensitive data was compromised. In its *2014 Cost of Data Breach Study,* the Ponemon Institute calculated the total average cost of a data breach incident to U.S. companies at $5.85 million.

The costs of a data breach can, of course, be significantly higher. Sony Pictures Entertainment, for example, has had two data breaches in recent years. The June 2011 breach cost Sony Pictures $171 million. Based on what is known about the November 2014 breach, analysts believe the costs will likely reach $100 million. The 2014 attack on Sony Pictures forced a shutdown of its entire network for days — also becoming an enormously costly disaster recovery event.

The Lifecycle of a Data Center Attack

Today's sophisticated cyberattacks exploit a foundational vulnerability that exists in modern data center designs: the existence of little or no security controls inside the perimeter of the data center. Popular security models, such as the Lockheed Martin Cyber Kill Chain (see Figure 1-1), provide a simple framework for understanding the systematic process used by cybercriminals to breach a data center perimeter. Once inside the data center, an attacker relies heavily on the ability to move laterally throughout the data center in order to expand the attack surface and achieve the attack objectives.

Reconnaissance	Harvesting specific target information to breach the data center perimeter
Weaponization	Coupling exploit and backdoor into a deliverable payload
Delivery	Delivering weaponized bundle to breach data center perimeter defenses
Exploitation	Moving laterally to take advantage of vulnerabilities on data center systems
Installation	Installing malware on systems throughout the data center
Command & Control (C2)	Establishing remote communication and control channels inside the data center
Actions on Objective	Pursuing attack intentions with complete "hands on keyboard"access

Figure 1-1: The Lockheed Martin Cyber Kill Chain.

Unfortunately, these models reflect a grim reality: A tremendous — and disproportionate — amount of effort and resources has been applied to preventing a breach in the first place, by protecting the data center perimeter (corresponding to the first three steps in Figure 1-1). But breaches inevitably still happen far too often. Once inside the data center, an attacker can exploit vulnerabilities, install malware, establish a command and control (C2) infrastructure, and move laterally across systems throughout the data center with relative ease (see Figure 1-2).

C2 communication is critical to a successful attack and must, therefore, be stealthy in order to avoid detection. C2 traffic is often Secure Sockets Layer (SSL)–encrypted and uses proxies or tunneling within legitimate applications or protocols.

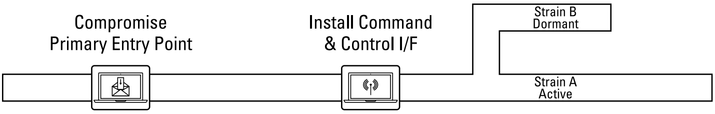


Figure 1-2: C2 enables further reconnaissance in the data center.

Next, an attacker installs additional C2 infrastructure on other devices and systems, covers any traces of the attack, and escalates system privileges in a multipronged attack that takes advantage of relatively weak or nonexistent security inside the data center (see Figure 1-3).

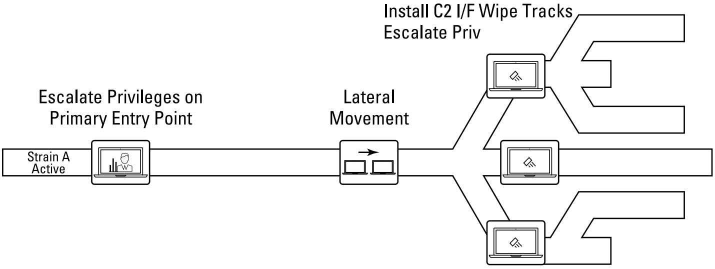


Figure 1-3: Additional C2 infrastructure is installed to ensure persistence as the attacker moves laterally through the data center.

Modern cyberattacks take advantage of relatively weak or nonexistent security within the data center to move freely between different systems and steal information. Chapter 2 explains how micro-segmentation blocks an attacker's lateral movement and helps prevent successful installation of a C2 infrastructure in the data center.

Modern, advanced attacks are designed to be persistent and resilient. Thus, if an active threat is discovered, the attacker can simply "wake up" a dormant malware strain on another infected system in the data center and continue the attack (see Figure 1-4). The lack of adequate segmentation and security controls, and the explosion of east–west traffic inside the data center, make it difficult — if not impossible — for incident response teams to effectively isolate an attack.

The attacker can then carry out any desired action against the target (see Figure 1-5).

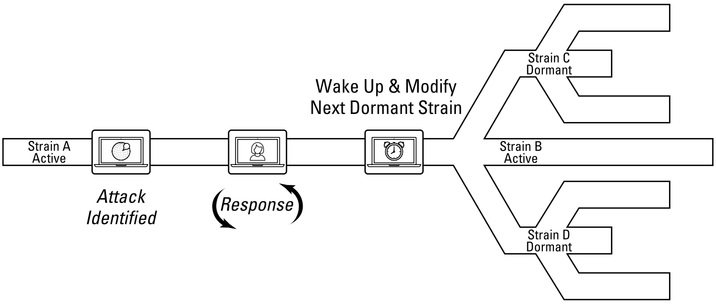


Figure 1-4: If an attack is discovered, the attacker simply makes a dormant strain active and continues the attack.

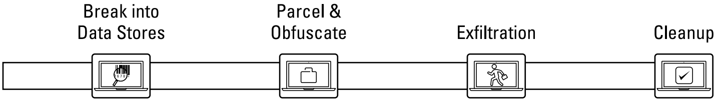


Figure 1-5: The attacker is then free to perform any desired actions on the data center objective.

If the intent is to steal sensitive information, the attacker parcels the data into small, encrypted payloads to avoid detection during exfiltration from the target network.

TIP

In Chapter 2, you learn how micro-segmentation prevents a successful attack by blocking an attacker's lateral movement in the data center, and with capabilities such as advanced security service insertion that enables deep packet inspection (DPI) — and blocking (if appropriate) — of encrypted outbound payloads from the data center, for example.

REMEMBER

Using a patient, resilient, multipronged, and stealthy attack strategy, an attacker inside your data center can move between systems relatively unencumbered, and steal sensitive data for months or even years before being detected.

Throwing Stones at the (Data Center) Perimeter Walls

Segmentation is a fundamental information security principle that has been applied to data center design for decades. At its most basic level, segmentation occurs between two or more networks, such as an internal network (the data center) and an external network (the Internet) with a firewall deployed at the perimeter between the different networks (see Figure 1-6).

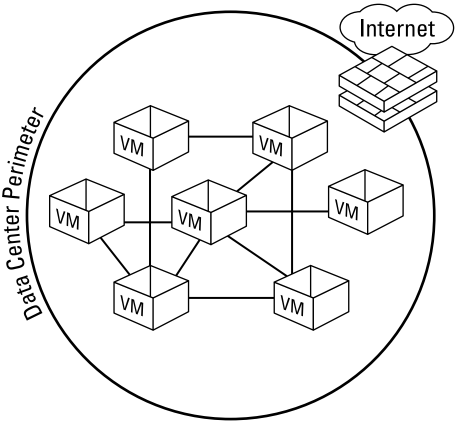


Figure 1-6: Perimeter-based security is insufficient in a data center where security is needed everywhere.

Although segmentation *does* exist in data centers today, the network segments are much too large to be effective and are typically created to restrict north–south traffic between the Internet and the data center or between client workstations and the data center. For example, a network may be segmented into multiple trust levels using additional firewalls to create a DMZ or separate department networks (such as finance, human resources, and R&D). To be completely effective, segmentation (and firewalling) needs to be possible down to the level of the individual workload. But a typical data center may have thousands of workloads, each with unique security conditions. And again, the primary focus is on controlling north–south traffic *in and out* of the data center, rather than the east–west traffic *within* the data center upon which modern attacks are predicated.

To effectively protect data centers from modern attacks, micro-segmentation down to the individual workload is needed. But deploying hundreds (or even thousands) of appliance-based firewalls inside the data center to protect each individual workload is financially and operationally infeasible. And virtual firewalls, while somewhat less expensive than hardware firewalls, still do not address the need to segment the data center network down to the individual workload. The bottom line: Maintaining unique and effective security policies for thousands of individual workloads as part of a comprehensive — and cohesive — enterprise security strategy using existing data center security technologies, controls, and processes has been impractical . . . until now. Network virtualization (explained in Chapter 3) makes micro-segmentation a reality in the data center and takes advantage of new or existing infrastructure.

TIP

You learn how to deploy micro-segmentation while leveraging and improving the performance of your existing security technologies and data center infrastructure in Chapter 5!

Many organizations logically partition their data center networks into different security segments, which then need to be translated to networking constructs, such as subnets and virtual LANs (VLANs). These techniques provide only rudimentary access control and result in security constructs that are too rigid and too complex, because security policies are largely defined by where a workload is physically deployed in the network topology (see Figure 1-7). Segmenting the data center with such large zones creates a significant attack surface and enables threats to move throughout large portions of the data center unrestricted, once an attacker has overcome the data center's perimeter defenses. These segmentation techniques also result in significant delays when deploying new workloads or changing existing workloads, because they must be manually configured to reflect a rigid and static network topology.

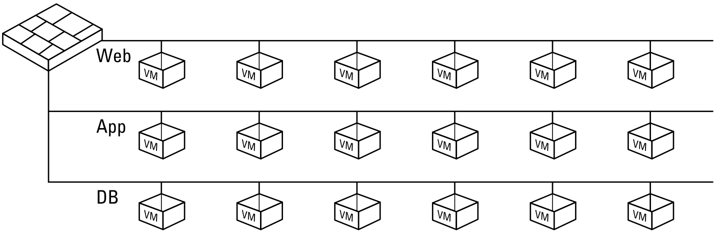


Figure 1-7: Today, security is tied to a rigid and complex network topology that is further complicated by a consolidated, multitier application infrastructure.

Subnets and VLAN changes can also be a frequent source of network outages, security compromises, configuration errors, and application deployment delays. Also, it's not always possible to thoroughly test proposed changes in a test environment that accurately replicates the production data center.

Different segments should be created inside the perimeter to limit the lateral spread of threats within the data center. Ideally, segmentation and policy enforcement should be available down to the level of the individual workload.

In addition to inadequate logical segmentation, another unfortunate consequence of traditional data center design that adds complexity and degrades network performance is hairpinning east–west server traffic — communications between servers that would not otherwise traverse a firewall boundary — through a firewall (see Figure 1-8).

Hairpinning is incredibly inefficient and greatly increases complexity in the data center by

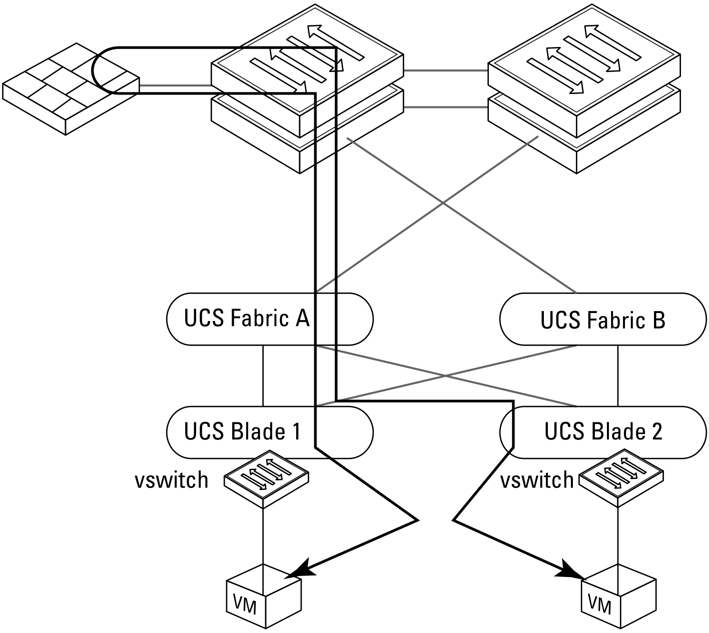


Figure 1-8: Host-to-host east–west firewalling or hairpinning traffic.

- Creating unnecessary performance choke points in the network and potential points of failure
- Backhauling as much as 60 percent of all network traffic across firewalls, adding congestion and latency on the network
- Contributing to firewall rule sprawl and performance bottlenecks as security administrators are increasingly reluctant to modify or remove complex rulesets when workloads are decommissioned, fearful of causing an outage or security breach

Hairpinning is particularly frustrating for east–west traffic between virtualized workloads that could otherwise communicate at close to wire speed — and traffic between workloads on the same host shouldn't even have to hit the wire!

Finally, many advanced security solutions have been deployed at the perimeter, including next-generation firewalls, anti-malware, intrusion prevention systems (IPS), distributed denial-of-service (DDoS) prevention, unified threat management (UTM), spam filtering, and many other technologies. Although these solutions bolster perimeter defenses, they are often designed to address specific threats with limited context and correlation between different security technologies, and the fundamental problem with data center security remains: When an attacker gets past the perimeter and is inside your data center, security controls are relatively weak or nonexistent and the attacker can roam freely (so to speak). To stop threats anywhere and everywhere that they occur, these solutions need to be deployed both at the perimeter and *inside* the data center, on a common platform that provides context and coordination across individual workloads and disparate technologies.

Outdated data center security methodologies are not aligned with modern business requirements for anywhere, anytime access to data center applications and data from any device, and are insufficient to address today's sophisticated attacks. These methodologies and challenges include the following:

- **Perimeter-centric foundation:** A strong perimeter is important, but security controls *within* the data center are weak or nonexistent. Layering on advanced security solutions, such as next-generation firewalls, IPS, DDoS prevention, and other technologies strengthens the perimeter, but it's insufficient to address threats *inside* the data center.

- **Lack of internal controls:** Attackers take advantage of weak or nonexistent security controls inside the data center to move laterally between workloads and quickly expand the attack surface.
- **Inability to scale:** Deploying hundreds — possibly even thousands — of firewalls to protect every workload in the data center is infeasible and impractical.
- **Security mapped to network topology:** Security policies determined by the physical location of a server workload in the data center do not address business and compliance requirements and are too rigid and too complex. This legacy approach to security often leads to significant delays in application deployment.
- **The inefficiency of hairpinning:** Forcing east–west traffic through firewalls creates choke points, unnecessarily backhauls server traffic, and contributes to sprawling firewall rulesets and complexity.
- **Large security zones:** Using firewall choke points inside the data center, in an attempt to segment it, creates coarse security zones that still allow threats to move relatively unencumbered throughout these large segments.

While perimeter-based security is an important element of security, it shouldn't be the foundation — just as the walls of a building form a critical structural boundary, but aren't the foundation. Instead, the foundation provides a platform upon which the building is constructed.

In Chapter 2, you learn what micro-segmentation is all about and how it prevents data center attacks from succeeding.

Chapter 2

Micro-segmentation Explained

In This Chapter

- Bolstering data center defense inside the perimeter
- Using the software-defined data center as a weapon against attacks
- Making zero-trust trust a reality in the data center
- Gaining persistence, ubiquity, and extensibility with micro-segmentation
- Recognizing what micro-segmentation isn't

Micro-segmentation enables organizations to logically divide the data center into distinct security segments down to the individual workload level, and then define security controls and deliver services for each unique segment. This restricts an attacker's ability to move laterally in the data center, even after the perimeter has been breached — much like safe deposit boxes in a bank vault protect the valuables of individual bank customers, even if the safe has been cracked. In this chapter, you learn what micro-segmentation is — and what it isn't.

Limiting Lateral Movement within the Data Center

Modern attacks exploit inherent weaknesses in traditional perimeter-centric network security strategies (discussed in Chapter 1) to infiltrate enterprise data centers. After successfully evading the data center's perimeter defenses, an attack can move laterally within the data center from workload to workload with little or no controls to block its propagation.

Micro-segmentation of the data center network restricts unauthorized lateral movement but, until now, hasn't been operationally feasible in data center networks.

Traditional packet-filtering and advanced next-generation firewalls implement controls as physical or virtual "choke points" on the network. As application workload traffic passes through these control points, network packets are either blocked or allowed to traverse the firewall based on the firewall rules that are configured at that control point.

There are two key operational barriers to micro-segmentation using traditional firewalls: throughput capacity and security management.

Limitations on throughput capacity can be overcome, but at a significant cost. It's possible to buy enough physical or virtual firewalls to deliver the capacity required to achieve micro-segmentation, but in most (if not all) organizations, purchasing the number of firewalls necessary for effective micro-segmentation isn't financially feasible.

The burden of security management increases exponentially with the number of workloads and the increasingly dynamic nature of today's data centers. If firewall rules need to be manually added, deleted, and/or modified every time a new VM is added, moved, or decommissioned, the rate of change quickly overwhelms IT operations. It's this barrier that has been the demise of most security teams' best-laid plans to realize a comprehensive micro-segmentation or least privilege, unit-level trust strategy (discussed later in this chapter) in the data center.

The software-defined data center (SDDC) leverages a network virtualization platform to offer several significant advantages over traditional network security approaches — automated provisioning, automated move/add/change for workloads, distributed enforcement at every virtual interface and in-kernel, scale-out firewalling performance, distributed to every hypervisor and baked into the platform.

Growth of east–west traffic within the data center

Over the past decade, applications have increasingly been deployed on multitier server infrastructures and east–west server–server communications now account for significantly more data center traffic than north–south client–server and Internet communications. In fact, traffic *inside* the data center now accounts for as much as 80 percent of all network traffic. These multitier application infrastructures are typically designed with little or no security controls restricting communications between the tiered systems — and increasingly leveraging 10 Gbps Ethernet connectivity for optimum throughput and performance.

Attackers have modified their attack strategy to take advantage of this paradigm shift in data center traffic, as well as the fact that prevailing perimeter-centric defense strategies offer little or no controls for network communications within the data center. Security teams must likewise extend their defense strategy *inside* the data center — where the vast majority of network traffic actually exists and is unprotected — instead of focusing almost exclusively on perimeter defenses.

Visibility and context

The growth of east–west traffic within the data center and the rise of server virtualization are two trends that have contributed to an alarming lack of visibility and context in the data center.

For the most part, east–west server communications in the data center do not pass through a firewall and are, therefore, not inspected. For all intents and purposes, this traffic is invisible to network security teams. When east–west traffic *is* forced through a firewall — using techniques such as hairpinning to backhaul the traffic through a firewall choke point — the result is a complex and inefficient communication path that negatively affects network performance throughout the data center.

Innovation in server virtualization has far outpaced the underlying network and security constructs in traditional data centers, which contributes to the problem of limited visibility and context in the data center. Deploying multiple virtual workloads on a single physical host configured with multiple network interface cards (NICs) is common in virtual server environments. Without virtual switches, the traffic going to and from individual VMs cannot be easily identified. This can cause significant issues for network teams attempting to identify and troubleshoot problems, and is fertile ground for an attacker.

The network hypervisor in a virtualized network is uniquely positioned to see *all* traffic in the data center (see Figure 2-1), down to the level of individual VM workloads. This level of visibility and context enables micro-segmentation based on attributes that are unique to each workload, such as the operating system, patch level, running services, and many other properties. This capability, in turn, enables more intelligent network and security policy decisions that can be defined with an understanding of the specific purpose of each individual workload in the data center. For example, unique policies can be specifically defined for the web tier of an order-taking application, or for an enterprise human resources management system, based on the needs of the individual workload rather than being constrained by the underlying network topology.

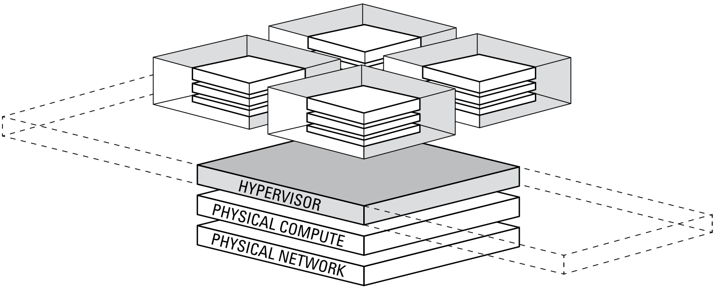


Figure 2-1: The network hypervisor is uniquely positioned to see all traffic in the data center.

Isolation

Isolation is an important principle in network security, whether for compliance, containment, or simply keeping development, test, and production environments separated.

Manually configured and maintained routing, access control lists (ACLs), and/or firewall rules on physical devices have traditionally been used to establish and enforce isolation in data center networks.

TECHNICAL STUFF

Forrester Research outlines its “Zero Trust model” of information security and isolation, in which perimeter security controls are extended throughout the entire data center. It requires organizations to protect external *and* internal data resources and enforce strict access controls. Zero Trust incorporates the principle of “Least Privilege” — a cornerstone of information security that limits access and permissions to the minimum required to perform an authorized function. Finally, “Trust, but verify” is so 1980s (with respect and apologies to President Reagan) — “Never trust, always verify” is the new paradigm for a safe and secure world.

Virtual networks are inherently isolated from other virtual networks and from the underlying physical network by design. This concept is distinctly different from the legacy approach of assuming some default level of trust within the data center. Isolation is inherent to network virtualization — no physical subnets, virtual LANs (VLANs), ACLs, or firewall rules are required to enable this isolation. Virtual networks are created in isolation and remain isolated unless deliberately and explicitly connected together.

Any isolated virtual network can be made up of workloads distributed anywhere in the data center and workloads in the same virtual network can reside on the same or separate hypervisors. Additionally, workloads in several isolated virtual networks can reside on the same hypervisor.

Isolation between virtual networks also allows for overlapping IP addresses. So, it’s possible, for example, to have isolated development, test, and production virtual networks, each with a different application version, but with the same IP addresses, all operating at the same time on the same underlying physical infrastructure.

Finally, virtual networks are also isolated from the underlying physical infrastructure. Because traffic between hypervisors is encapsulated, physical network devices operate in a completely different address space than the workloads connected to the virtual networks. For example, a virtual network could support IPv6 application workloads on top of an IPv4 physical network. This isolation protects the underlying physical infrastructure from any possible attack initiated by workloads in any virtual network. Again, all this is independent from any VLANs, ACLs, or firewall rules that would traditionally be required to create this isolation.

Segmentation

Related to isolation, but applied within a multitier virtual network, is segmentation. Traditionally, network segmentation is achieved with a physical firewall or router that allows or denies traffic between network segments or tiers — for example, segmenting traffic between a web tier, application tier, and database tier. Segmentation is an important principle in security design because it allows organizations to define different trust levels for different network segments, and reduces the attack surface should an attacker breach the perimeter defenses. Unfortunately, data center network segments are often far too large to be effective and traditional processes for defining and configuring segmentation are time consuming and prone to human error, often resulting in security breaches.

Network segmentation, like isolation, is a core capability of a network virtualization platform. A virtual network can support a multitier network environment — multiple layer 2 segments with layer 3 segmentation (or micro-segmentation) on a single layer 2 segment, using distributed firewalling defined by workload security policies. These could represent a web tier, application tier, and database tier, for example.

In a virtual network, network and security services — such as layer 2, layer 3, ACLs, firewall, quality of service (QoS), and others — that are provisioned with a workload are programmatically created and distributed to the hypervisor virtual switch and enforced at the virtual interface. Communication within a virtual network never leaves the virtual environment, removing the requirement for network segmentation to be configured and maintained in the physical network or firewall.

Automation

Automated provisioning enables the correct firewalling policies to be provisioned when a workload is programmatically created, and those policies follow the workload as it's moved anywhere in the data center or between data centers.

Equally important, if the application is deleted, its security policies are automatically removed from the system. This capability eliminates another significant pain point — firewall rule sprawl — which potentially leaves thousands of stale and outdated firewall rules in place, often resulting in performance degradation and security issues.

Enterprises can also apply a combination of different partner capabilities by chaining advanced security services together and enforcing different services based on different security situations. This enables organizations to integrate their existing security technologies to build a more comprehensive and correlated security capability inside the data center. Existing security technologies actually function better with micro-segmentation than otherwise possible, because they have greater visibility and context of individual workload VM traffic inside the data center, and security actions can be customized for individual VM workloads as part of a complete security solution. For example, a workload may be provisioned with standard firewalling policies, which allow or restrict its access to other types of workloads. The same policy may also define that if a vulnerability is detected on the workload during the course of normal vulnerability scanning, a more restrictive firewalling policy would apply, restricting the workload to be accessed by only those tools used to remediate the vulnerabilities (see the sidebar "Segmentation with advanced security service insertion, chaining, and traffic steering").

TIP

Security vendors can take advantage of the network virtualization platform to trigger advanced security service responses from a completely different security vendor's technology solution — an innovation that's simply not possible without network virtualization!

Segmentation with advanced security service insertion, chaining, and traffic steering

The VMware NSX network virtualization platform provides stateful inspection firewalling features to deliver segmentation within virtual networks. In some environments, there is a requirement for more advanced network security capabilities. In these instances, organizations can leverage the SDDC platform to distribute, enable, and enforce advanced network security services in a virtualized network environment. The NSX platform distributes network services into the vSwitch to form a logical pipeline of services applied to virtual network traffic. Third-party network services can be inserted into this logical pipeline, allowing physical or virtual services to be consumed in the logical pipeline.

Every security team uses a unique combination of network security products to meet the needs of their environment. The NSX platform is being leveraged by VMware's entire ecosystem of security solution providers. Network security teams are often challenged to coordinate network security services from multiple vendors in relationship to each other. Another powerful benefit of the NSX approach is its ability to build policies that leverage NSX service insertion, chaining, and steering to drive service execution in the logical services pipeline, based on the result of other services, making it possible to coordinate otherwise completely unrelated network security services from multiple vendors.

For example, VMware's integration with Palo Alto Networks leverages the NSX platform to distribute the Palo Alto Networks VM-Series next-generation firewall, making the advanced features locally available on each hypervisor. Network security policies, defined for application workloads provisioned or moved to that hypervisor, are inserted into the virtual network's logical pipeline. At runtime, the service insertion leverages the locally available Palo Alto Networks next-generation firewall feature set to deliver and enforce application-, user-, and content-based controls and policies at the workload's virtual interface.

Another example might use Trend Micro for malware detection. If malware is detected being uploaded onto a VM, Trend Micro blocks the malware and triggers an alert that adds the VM to a "Security Violations" group policy. A snapshot of the VM is immediately created for forensic purposes and all traffic to and from the VM is redirected through an IPS and simultaneously mirrored to a remote SPAN (RSPAN) session for collection and forensic analysis.

Today, the VMware NSX platform has significant integration with partners, including Palo Alto Networks, Rapid7, Trend Micro, Symantec, CheckPoint, Intel Security, and more, allowing for unprecedented advanced security.

Essential Elements of Micro-segmentation

As discussed later in this chapter, the network hypervisor is uniquely positioned to provide both context and isolation throughout the SDDC — not too close to the workload where it can be disabled by an attack, and not so far removed that it doesn't have context into the workload. Thus, the network hypervisor is ideally suited to implement three key elements of micro-segmentation: persistence, ubiquity, and extensibility.

Persistence

Security administrators need to know that when they provision security for a workload, enforcement of that security persists despite changes in the environment. This is essential, as data center topologies are constantly changing: Networks are renumbered, server pools are expanded, workloads are moved, and so on. The one constant in the face of all this change is the workload itself, along with its need for security. But in a changing environment, the security policy configured when the workload was first deployed is likely no longer enforceable, especially if the definition of this policy relied on loose associations with the workload like IP address, port, and protocol. The difficulty of maintaining this persistent security is exacerbated by workloads that move from one data center to another or even to the hybrid cloud (for example, a live migration or for disaster recovery purposes).

Micro-segmentation gives administrators more useful ways to describe the workload. Instead of relying merely on IP addresses, administrators can describe the inherent characteristics of the workload, tying this information back to the security policy:

- What type of workload is this (for example, web, application, or database)?
- What will this workload be used for (for example, development, staging, or production)?
- What kinds of data will this workload be handling (for example, low-sensitivity, financial, or personally identifiable information)?

Micro-segmentation even allows administrators to combine these characteristics to define inherited policy attributes. For example, a workload handling financial data gets a certain level of security, but a production workload handling financial data gets an even higher level of security.

Ubiquity

Traditional data center architectures prioritize security for important workloads, too often neglecting lower-priority systems. Traditional network security is expensive to deploy and manage, and because of this cost, data center administrators are forced into a situation where they have to ration security. Attackers take advantage of this fact, targeting lower-priority systems with lower levels of protection as their infiltration point into a data center.

In order to provide an adequate level of defense, security administrators need to depend on a high level of security being available to every system in the data center. Micro-segmentation makes this possible by embedding security functions into the data center infrastructure itself. By taking advantage of this widespread compute infrastructure, administrators can rely on the availability of security functions for the broadest spectrum of workloads in the data center.

Extensibility

Aside from persistence and ubiquity, security administrators also rely on micro-segmentation to adapt to new and unfolding situations. In the same way that data center topologies are constantly changing, so are the threat topologies inside data centers changing: New threats or vulnerabilities are exposed, old ones become inconsequential, and user behavior is the inexorable variable that constantly surprises security administrators.

In the face of emerging security scenarios, micro-segmentation enables administrators to extend capabilities by integrating additional security functions into their defense portfolio. For instance, administrators might begin with stateful firewalling distributed throughout the data center, but add next-generation firewalling and intrusion prevention for deep packet inspection (DPI), or agentless anti-malware for better server security. But beyond merely adding more security functions, administrators need these functions to cooperate in order to provide more effective security than if they were deployed in silos. Micro-segmentation addresses this need by enabling the sharing of intelligence between security functions. This makes it possible for the security infrastructure to act concertedly to tailor responses to unique situations.

As an example, based on the detection of malware, an antivirus system coordinates with the network to mirror traffic to an intrusion prevention system (IPS), which, in turn, scans for anomalous traffic. The extensibility of micro-segmentation enables this dynamic capability. Without it, the security administrator would have to preconfigure a different static chain of services upfront, each one corresponding to a different possible security scenario. This would require a preconception of every possible security scenario during the initial deployment!

Balancing Context and Isolation

Many IT security professionals instinctively view new innovations, such as the SDDC, as a potential new target. But the reality with the SDDC is that the positive impact to IT security is far greater than any changes to what needs to be secured. In other words, for IT security professionals, the SDDC is a weapon, not a target. The SDDC approach delivers a platform that inherently addresses some fundamental architectural limitations in data center design, which have restricted security professionals for decades.

Consider the trade-off that is often made between context and isolation in traditional security approaches. Often, in order to gain context, we place controls in the host operating system. This approach allows us to see what applications and data are being accessed and what users are using the system, resulting in good context. However, because the control sits in the attack domain, the first thing an attacker will do is disable the control. This is bad isolation. This approach is tantamount to putting the on/off switch for a home alarm system on the outside of the house.

An alternative approach, which trades context for isolation, places the control in the physical infrastructure. This approach isolates the control from the resource it's securing, but has poor context because IP addresses, ports, and protocols are very bad proxies for user, application, or transaction context. Furthermore, there has never been a ubiquitous enforcement layer built into the infrastructure — until now.

The data center virtualization layer used by the SDDC offers the ideal location to achieve both context and isolation, combined with ubiquitous enforcement. Controls operating in the data center virtualization layer leverage secure host introspection, the ability to provide agentless, high-definition host context, while remaining isolated in the hypervisor, safe from the attack that is being attempted.

The ideal position of the data center virtualization layer between the application and the physical infrastructure, combined with automated provisioning and management of network and security policies, kernel-embedded performance, distributed enforcement, and scale-out capacity is on the verge of completely transforming data center security and allowing data center security professionals to achieve levels of security that were operationally infeasible in the past.

Implementing Least Privilege and Unit-Level Trust with Micro-segmentation

Assuming that threats can be lurking anywhere in the data center is a prudent security approach that requires a least privilege and unit-level trust model. Together, least privilege and unit-level trust achieve a positive control model that implements a "never trust, always verify" security strategy at a very granular level, down to the individual workload.

TECHNICAL STUFF

A *positive control model* explicitly defines what is permitted on the network and implicitly blocks everything else. A *negative control model* explicitly defines what is not allowed on the network and implicitly permits everything else.

Least privilege begins with no default trust level for any entity or object in the data center — including network segments, server workloads, applications, and users. Unit-level trust requires enterprises to establish trust boundaries that effectively compartmentalize different segments of the data center environment at a very granular level, and move security controls as close as possible to the resources that require protection.

Least privilege and unit-level trust require continuous monitoring and inspection of all data center traffic for threats and unauthorized activity. This includes both north–south (client–server and Internet) and east–west (server–server) traffic. Micro-segmentation enables the implementation of an effective least privilege and unit-level trust security strategy by establishing multiple trust boundaries at an extremely fine level of granularity and applying appropriate policies and controls to individual workloads in the data center.

REMEMBER

Micro-segmentation allows organizations to adopt a least-privilege, unit-level trust strategy that effectively restricts an attacker's ability to move laterally within the data center and exfiltrate sensitive data.

What Micro-segmentation Is Not

The concept of micro-segmentation is nothing new. The reality of achieving micro-segmentation is new, made possible for the first time with network virtualization and security distributed to the hypervisor. Unfortunately, as with any new technological innovation, there is often a great deal of confusion about the capabilities and limits of micro-segmentation. So, it's time to debunk a few myths about micro-segmentation.

First, micro-segmentation — and, more specifically, the network and security services of a network virtualization platform — is not a replacement for hardware firewalls deployed at the data center perimeter. The performance capacity of hardware firewall platforms is designed to control traffic flowing to and from hundreds or thousands of simultaneous data center workload sessions.

Notwithstanding the performance of hardware firewalls deployed at the data center perimeter, the firewalling performance and capacity of VMware's NSX platform, for example, for east–west traffic within the data center is impressive. The NSX platform delivers 20 Gbps of firewalling throughput and supports over 80,000 connections per second, per host. This performance is applied to the VMs on its hypervisor and every time another host is added into the SDDC platform, another 20 Gbps of throughput capacity is added.

Next, micro-segmentation isn't possible with existing tools and technologies and can't be effectively implemented on an underlay network because it lacks context. Effective micro-segmentation requires intelligent grouping of individual workloads so that network and security policies can be dynamically applied at an extremely fine level of granularity completely independent of the underlying physical network topology (see Figure 2-2).

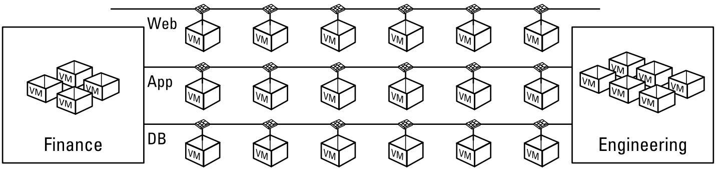


Figure 2-2: Intelligent grouping of data center workloads in a multitier environment that is completely independent of the underlying physical network is only possible with micro-segmentation.

Micro-segmentation enables organizations to define groups in creative ways that have never before been possible. Because of the ubiquitous nature of the network hypervisor and its unique visibility and understanding of individual workloads in the data center, intelligent network and security policy groups are based on characteristics (see Figure 2-3) such as

- ✓ Operating system
- ✓ Machine name
- ✓ Services
- ✓ Multitier
- ✓ Regulatory requirements
- ✓ Active Directory containers
- ✓ Unique tags

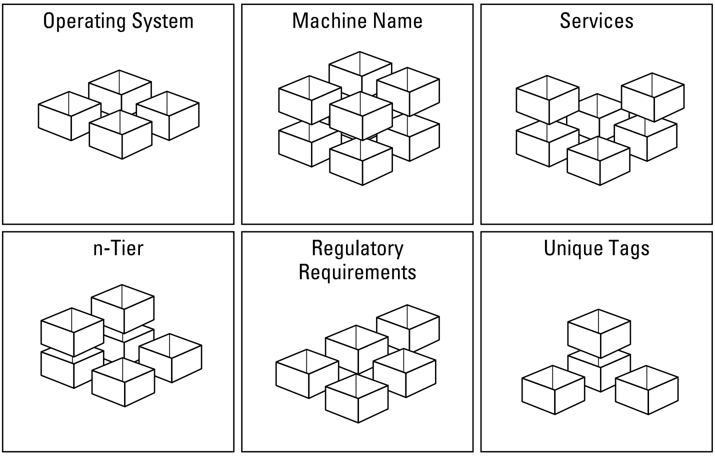


Figure 2-3: Intelligent grouping defined by customized criteria.

This level of granularity would potentially require deploying thousands of firewalls (physical, virtual, or both) inside the data center — which isn't feasible, financially or operationally, for most organizations. Although micro-segmentation enables the deployment and centralized management, automation, and orchestration of hundreds of thousands of individual firewalls at the individual workload level, micro-segmentation is neither a hardware-defined solution nor a virtual appliance. And firewall rule management and automation are both important enabling capabilities of micro-segmentation, but they don't define micro-segmentation in and of themselves.

Finally, micro-segmentation cannot be achieved with agent-based security installed on individual workloads. Micro-segmentation uses a group-based approach to apply network and security services to individual workloads. These policies move and change dynamically as business and technical requirements for an individual workload change. For example, a database that doesn't process, store, or transmit cardholder information would not be considered part of the cardholder data environment (CDE) and, therefore, it isn't subject to Payment Card Industry (PCI) Data Security Standards (DSS) compliance requirements. However, if cardholder information is erroneously stored in the database at some point, the security policy would need to be updated. Micro-segmentation enables this policy change to happen automatically, by moving the database to the appropriate group within the SDDC. An agent-based approach would require manual updating of the security policy for that specific database. Additionally, agent-based security lacks the appropriate level of isolation because it is often the first component an attacker will disable, effectively turning off the alarm.

In Chapter 3, you learn how the network virtualization platform enables micro-segmentation in the data center.

Chapter 3

Moving the Data Center to Software

In This Chapter

- Recognizing today's networking and security challenges
- Extending virtualization to the network
- Understanding how network virtualization works
- Putting together the building blocks of network virtualization

Enterprise data centers today are realizing the tremendous benefits of compute and storage virtualization solutions to consolidate and repurpose infrastructure resources, reduce operational complexity, and dynamically align and scale their application infrastructure in response to business priorities. However, data center networking and security have not kept pace and remain rigid, complex, proprietary, and closed to innovation — a barrier to realizing the full potential of virtualization and moving the data center to software.

In this chapter, you discover how network virtualization transforms the modern data center and builds the foundation for micro-segmentation.

Key Forces Driving the Need for Data Center Transformation

Compute and storage virtualization solutions have dramatically transformed the data center by delivering significant operational savings through automation, capital savings through consolidation and hardware independence, and greater agility through on-demand and self-service approaches to provisioning. However, networking and security teams, under continuous pressure to innovate and move faster, are constrained by a hardware-based technology model and manual operational processes. Working with this outdated foundation, networking and security teams constantly struggle to adapt quickly enough to meet rapidly evolving business requirements and keep pace with an ever-changing threat landscape.

TIP

Moving the data center to software creates a common networking and security platform and enables the integration of multiple vendor solutions for more effective and efficient operation throughout the data center.

The current operational model has resulted in slow, manual, error-prone provisioning of networking and security services to support application deployment. Network operators are dependent on terminal, keyboard, scripting, and command-line interfaces (CLIs) to manipulate a multitude of virtual LANs (VLANs), firewall rules, load balancers and access-control lists (ACLs), quality of service (QoS), virtual routing and forwarding (VRF), and media access control/Internet Protocol (MAC/IP) tables. Complexity and risk are further compounded by the need to ensure that changes to the network for one application do not adversely impact other applications.

Given the complexity of this situation, it's no surprise that several recent studies point to manual configuration errors as the cause for more than 60 percent of network downtime and/or security breaches. The result is that, in addition to the frequent, inevitable configuration missteps, IT response time to new business requirements is too slow, as rapidly repurposed compute and storage infrastructure must still wait for networking and security to catch up.

Gone are the days of a single monolithic application stack that could be provisioned in isolation on a physical server with a static security policy tied to it. Today's business applications demand workload mobility, speed to market, and accessibility in a way that simply did not exist before the data center began defining compute in software. But the current device-centric approach to networking and security confines workload mobility to individual physical subnets and availability zones. In order to reach available compute resources in the data center, network and security administrators are forced to perform manual box-by-box configuration of VLANs, ACLs, firewall rules, and so forth. This process is not only slow and complex, but also one that is fundamentally limited (for example, 4096 for total VLANs). Organizations often resort to expensive overprovisioning of compute capacity for each application/networking pod, resulting in stranded resources and suboptimal resource utilization.

This imbalance between the pace of business change and the stagnation of networking and security models has created a significant bottleneck — often weeks or months — in the provisioning of workloads to deploy new applications and services. And practically from the moment a workload is provisioned, changes inevitably occur that force the business to repeat the cycle and go back through the bottleneck, further straining the relationship between the business and IT (see Figure 3-1).

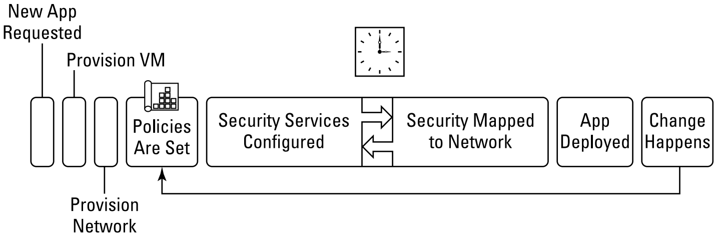


Figure 3-1: Rigid security policies mapped to network topology can't keep up with the pace of business changes.

At the same time, the demand for better security in the data center remains a constant and ever important business requirement. High-profile data breaches, prolific and sophisticated threats, and increasingly complex and stringent regulatory compliance requirements are top of mind for every chief information security officer (CISO) and security professional today. Maintaining a proactive and effective security posture in the data center while enabling an agile, dynamically changing business environment has become an insurmountable challenge.

Transforming Your Data Center with Network Virtualization

The software-defined data center (SDDC) extends core virtualization technologies to the entire data center. The SDDC transforms data center economics and business agility through automation and nondisruptive deployment leveraging existing physical compute, network, and storage infrastructure investments.

In much the same way that server virtualization programmatically creates, snapshots, deletes, and restores software-based virtual machines (VMs), network virtualization programmatically creates, snapshots, deletes, and restores software-based virtual networks. The result is a completely transformative approach to networking that not only enables data center managers to achieve orders of magnitude better agility and economics, but also allows for a vastly simplified operational model for the underlying physical network.

Network virtualization is a completely nondisruptive solution that can be deployed on any IP network, including traditional networking models and next-generation fabric architectures.

With compute virtualization, a software abstraction layer (server hypervisor) reproduces the physical attributes of an x86 physical server — processors, memory, storage, and network interfaces — in software, allowing them to be programmatically assembled in any arbitrary combination to produce a unique virtual machine, or VM (see Figure 3-2).

With network virtualization, the functional equivalent of a "network hypervisor" reproduces the complete set of layer 2 to layer 7 networking services — such as switching, routing, firewalling, and load balancing — in software. Thus, you can create any number of arbitrary network topologies, as needed. This also enables advanced security technologies to be chained together, in any combination or order, to deliver the most effective security controls possible (see Chapter 2 to learn more about this capability).

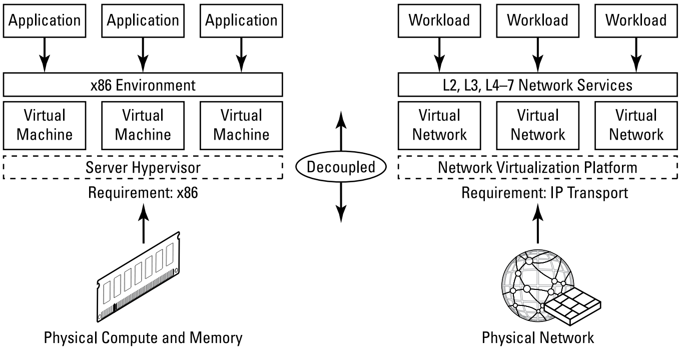


Figure 3-2: Compute and network virtualization.

Additionally, network virtualization enables you to distribute these same layer 2 through layer 7 networking and security services across all the workloads in the data center environment. This distributed approach makes advanced networking and security services available everywhere and minimizes the impact of any single failure. This approach also allows you to scale your capacity with the addition of each new workload. For example, each new workload added to the data center also adds additional firewalling capacity.

Not surprisingly, similar benefits between compute and network virtualization are also derived. For example, just as VMs are independent of the underlying x86 platform and allow IT to treat physical hosts as a pool of compute capacity, virtual networks are independent of the underlying IP network hardware and allow IT to treat the physical network as a pool of transport capacity that can be consumed and repurposed on demand.

Unlike legacy network architectures, virtual networks can be provisioned, changed, stored, deleted, and restored programmatically without reconfiguring the underlying physical hardware or topology. By matching the capabilities and benefits derived from familiar compute and storage virtualization solutions, this innovative approach to networking unleashes the full potential of the SDDC.

How Network Virtualization Works

Network virtualization programmatically creates, provisions, and manages virtual networks, utilizing the underlying physical network as a simple packet forwarding backplane. Network and security services in software are distributed to hypervisors and "attached" to individual VMs in accordance with networking and security policies defined for each connected application. When a VM is moved to another host, its networking and security services move with it. And when new VMs are created to scale an application, the necessary policies are dynamically applied to those VMs as well.

Similar to how a virtual machine is a software container that presents logical compute services to an application, a virtual network is a software container that presents logical network services — logical switching, logical routing, logical firewalling, logical load balancing, logical VPNs, and more — to connected workloads. These network and security services are delivered in software and require only IP packet forwarding from the underlying physical network.

Network virtualization coordinates the virtual switches in server hypervisors and the network services pushed to them for connected VMs, to effectively deliver a platform — or "network hypervisor" — for the creation of virtual networks (see Figure 3-3).

One way that virtual networks can be provisioned is by using a cloud management platform (CMP) to request the virtual network and security services for the corresponding workloads (see Step 1 of Figure 3-4). The controller then distributes the necessary services to the corresponding virtual switches and logically attaches them to the corresponding workloads (see Step 2 of Figure 3-4).

This approach not only allows different virtual networks to be associated with different workloads on the same hypervisor, but also enables the creation of everything from basic virtual networks involving as few as two nodes, to very advanced constructs that match the complex, multi-segment network topologies used to deliver multitier applications.

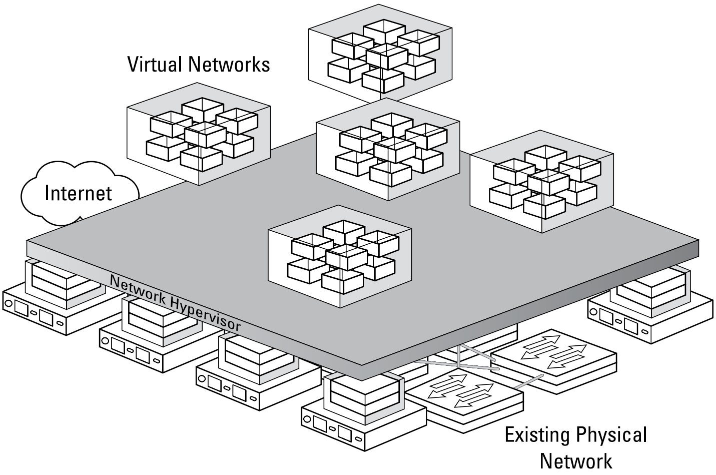


Figure 3-3: The "network hypervisor."

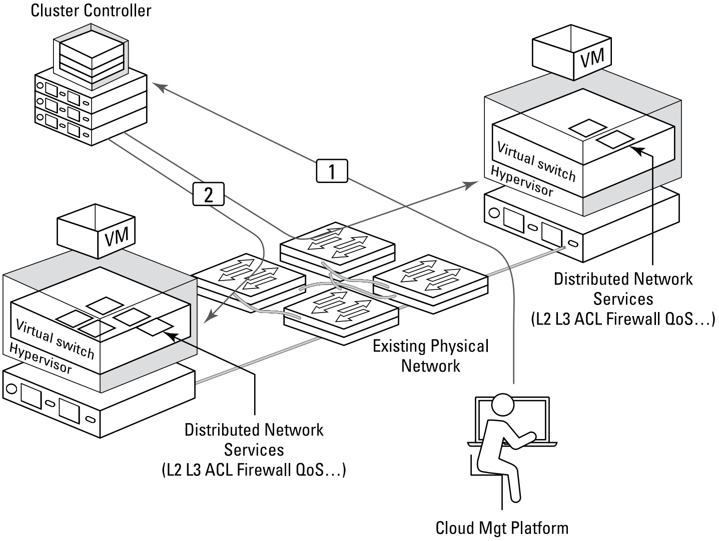


Figure 3-4: Virtual network provisioning.

To connected workloads, a virtual network looks and operates like a traditional physical network (see Figure 3-5). Workloads "see" the same layer 2, layer 3, and layer 4 through 7 network services that they would in a traditional physical configuration. It's just that these network services are now logical instances of distributed software modules running in the hypervisor on the local host and applied at the virtual interface of the virtual switch.

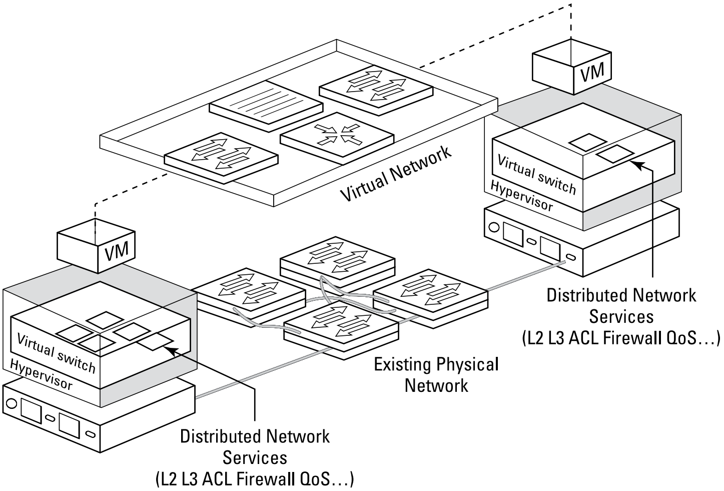

Virtual Network
VM
Virtual switch
Hypervisor
VM
Virtual switch
Hypervisor
Distributed Network Services (L2 L3 ACL Firewall QoS…)
Existing Physical Network
Distributed Network Services (L2 L3 ACL Firewall QoS…)

Figure 3-5: The virtual network, from the workload's perspective (logical).

To the physical network, a virtual network looks and operates like a traditional physical network (see Figure 3-6). The physical network "sees" the same layer 2 network frames that it would in a traditional physical network. The VM sends a standard layer 2 network frame that is encapsulated at the source hypervisor with additional IP, user datagram protocol (UDP), and virtual extensible LAN (VXLAN) headers. The physical network forwards the frame as a standard layer 2 network frame, and the destination hypervisor de-encapsulates the headers and delivers the original layer 2 frame to the destination VM.

TIP

The ability to apply and enforce security services at the virtual interface of the virtual switch also eliminates hairpinning (see Chapter 1) in situations where east–west traffic between two VMs on the same hypervisor, but in different subnets, is required to traverse the network to reach essential services such as routing and firewalling.

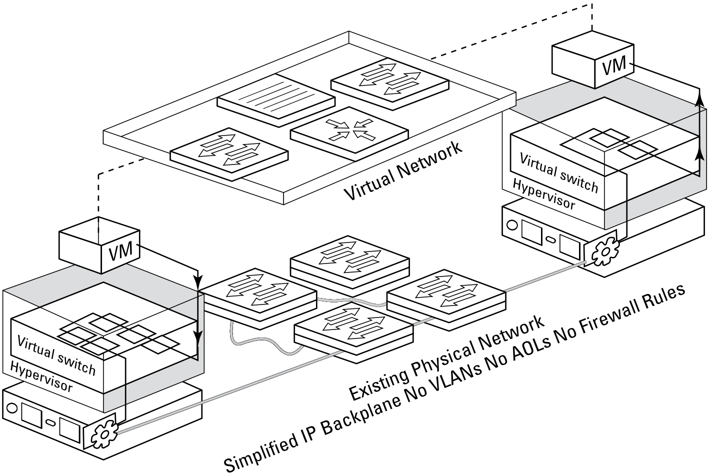

VM
Virtual Network
Virtual switch
Hypervisor
VM
Virtual switch
Hypervisor
Existing Physical Network
Simplified IP Backplane No VLANs No AOLs No Firewall Rules

Figure 3-6: The virtual network, from the network's perspective (physical).

Essential Elements for Network Virtualization

Network virtualization distributes layer 2 through layer 7 networking and security services to every workload in the environment, including switching, routing, load balancing, and firewalling. This principle of distribution is the panacea — the ultima Thule, if you will — of information security: centralized control with granular enforcement. Distribution with network virtualization also diffuses any data center failure points across the entire environment so that no single point of failure causes a more catastrophic problem. A network virtualization platform consists of a control plane, management plane, and data plane using an encapsulation protocol to abstract underlying physical and/or virtual networking components that provide an IP backplane.

TIP

Unlike software-defined networking (SDN) in which hardware remains the driving force, network virtualization technology is entirely hardware agnostic and truly decouples network resources from underlying hardware. Virtualization principles are applied to physical network infrastructure, abstracting network services to create a flexible pool of transport capacity that can be allocated, utilized, and repurposed on demand.

Just planes — no trains or automobiles

The *control* plane in a virtualized network runs in a controller and is responsible for routing traffic over the network. The control plane also performs encapsulation of network traffic in network virtualization technologies that use the VXLAN and NVGRE protocols (explained later in this chapter). Multiple controller nodes may be deployed in a network for high availability and scalability.

The *management* plane provides centralized configuration and administration of the virtualized network and may be integrated with a cloud management platform for deploying applications and provisioning functional network and security services, including the following:

- **Switching:** Enables extension of a layer 2 segment or IP subnet anywhere in the fabric irrespective of the physical network design.
- **Routing:** Routing between IP subnets can be done in the logical space without traffic going out to the physical router. This routing is performed in the hypervisor kernel and provides an optimal data path for routing traffic within the virtual infrastructure (east–west communication) and to the external network (north–south communication).
- **Distributed firewalling:** Micro-segmentation enables security enforcement at the kernel- and virtual network interface–level. This enables firewall rule enforcement in a highly scalable manner at near wire speed, without creating bottlenecks in physical appliances.
- **Logical load balancing:** Support for layer 4 through layer 7 load balancing with Secure Sockets Layer (SSL) termination capability.
- **Virtual private network (VPN):** Layer 2 and layer 3 SSL VPN services.
- **Connectivity to physical networks:** Layer 2 and layer 3 gateway functions provide communication between workloads deployed in logical and physical spaces.

The *data plane* transports the network traffic. In a virtualized network, data plane functions are typically implemented in a virtual switch. A virtual switch abstracts the physical network and provides access-level switching in the hypervisor. It's central to network virtualization because it enables logical networks that are independent of physical constructs such as VLANs. Here are some of the benefits of a virtual switch:

- **Support for overlay networking and centralized network configuration:** Overlay networking enables the following capabilities:
    - Creation of a flexible logical layer 2 overlay over existing IP networks on existing physical infrastructure without the need to re-architect any of the data center networks
    - Agile provision of communication (east–west and north–south) while maintaining isolation between tenants
    - Application workloads and virtual machines that are agnostic of the overlay network and operate as if they were connected to a physical layer 2 network
- **Facilitation of massive scale of hypervisors**
- **A comprehensive toolkit for traffic management, monitoring, and troubleshooting within a virtual network:** Multiple features accomplish this, such as port mirroring, NetFlow and IP Flow Information Export (IPFIX), configuration backup and restore, network health check, quality of service (QoS), and Link Aggregation Control Protocol (LACP).

Additionally, the data plane consists of gateway devices that can provide communication from the logical networking space to the physical network. This functionality can happen at layer 2 (bridging) or layer 3 (routing).

Encapsulation

Encapsulation (or "overlay") protocols decouple connectivity in the logical space from the physical network infrastructure. Devices connected to logical networks can leverage network functions — including switching, routing, firewalling, load balancing, VPN, and physical connectivity — independently from how the underlying physical infrastructure is configured. The physical network effectively becomes a backplane used to transport overlay traffic.

This decoupling solves many of the challenges traditional data center deployments are facing, such as the following:

- **Agile and rapid application deployment:** Traditional networking design represents a bottleneck slowing down the rollout of new applications at the pace that businesses are demanding. The time required to provision the network infrastructure in support of a new application often is counted in days if not weeks.
- **Workload mobility:** Compute virtualization enables mobility of virtual workloads across different physical servers connected to the data center network. In traditional data center designs, this requires extending layer 2 domains (VLANs) across the entire data center network infrastructure, affecting the overall scalability and potentially jeopardizing the overall resiliency of the design.
- **Large-scale multitenancy:** The use of VLANs as a means of creating isolated networks is limited to a maximum 4094. This number, while it may seem large for typical enterprise deployments, is becoming a serious bottleneck for most cloud providers.

Encapsulation enables logically separate network functions to be abstracted from their underlying protocols by wrapping packets with protocol information from the layer immediately above. For network virtualization, there are currently four encapsulation protocols available:

- **Virtual Extensible LAN (VXLAN)** encapsulates layer 2 frames within layer 4 UDP packets. Vendors supporting development of VXLAN include Arista Networks, Broadcom, Cisco, Dell, Juniper Networks, OpenBSD, Red Hat, and VMware, among others.
- **Network Virtualization using Generic Routing Encapsulation (NVGRE)** uses Generic Routing Encapsulation (GRE) to tunnel layer 2 packets over layer 3 networks. Vendors supporting development of NVGRE include Arista Networks, Broadcom, Dell, Emulex, F5 Networks, Intel, Hewlett-Packard, and Microsoft, among others.

- **Generic Network Virtualization Encapsulation (Geneve)** specifies a dataplane schema and decouples encapsulation from the control plane to provide flexibility across different deployment scenarios. Vendors supporting development of Geneve include Intel, Microsoft, Red Hat, and VMware.
- **Stateless Transport Tunneling (STT)** uses the TCP Segmentation Offload (TSO) capability on network interface cards (NICs) for hardware acceleration and decouples encapsulation from the control plane. Vendors supporting development of STT include Broadcom, eBay, Intel, Rackspace, and Yahoo!, among others.

TIP

Not sure which encapsulation protocol to use? Don't worry, they're all compatible and can coexist on the same network. Use any encapsulation protocol your particular network virtualization technology supports for your unique deployment scenario!

A VXLAN primer

Virtual Extensible LAN (VXLAN) has become the de facto standard overlay (or encapsulation) protocol with broad industry support. VXLAN is key to building logical networks that provide layer 2 adjacency between workloads, without the issue and scalability concerns found with traditional layer 2 technologies.

VXLAN is an overlay technology encapsulating the original Ethernet frames generated by workloads (virtual or physical) connected to the same logical layer 2 segment, usually named a logical switch (LS).

VXLAN is a layer 2 over layer 3 (L2oL3) encapsulation technology. The original Ethernet frame generated by a workload is encapsulated with external VXLAN, UDP, IP, and Ethernet headers to ensure that it can be transported across the network infrastructure interconnecting the VXLAN endpoints (virtual machines).

Scaling beyond the 4094 VLAN limitation on traditional switches has been solved by leveraging a 24-bit identifier, named VXLAN Network Identifier (VNI), which is associated with each layer 2 segment created in the logical space. This value is carried inside the VXLAN header and is normally associated with an IP subnet, similar to what traditionally happens with VLANs. Intra-IP subnet communication happens between devices connected to the same virtual network (logical switch).

(continued)

(continued)

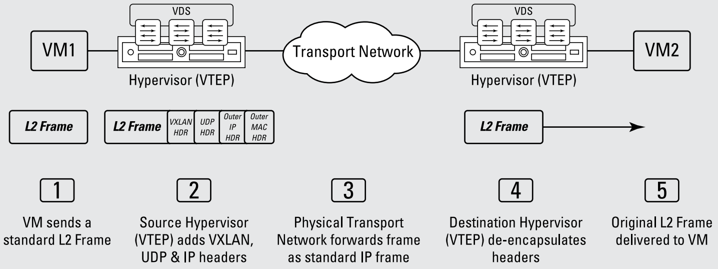
Hashing of the layer 2, layer 3, and layer 4 headers present in the original Ethernet frame is performed to derive the source port value for the external UDP header. This is important to ensure load balancing of VXLAN traffic across equal cost paths potentially available inside the transport network infrastructure.

The source and destination IP addresses used in the external IP header uniquely identify the hosts originating and terminating the VXLAN encapsulation of frames. This hypervisor-based logical functionality is usually referred to as a VXLAN Tunnel EndPoint (VTEP).

Encapsulating the original Ethernet frame into a UDP packet increases the size of the IP packet. Increasing the overall maximum transmission unit (MTU) size to a minimum of 1600 bytes for all the interfaces in the physical infrastructure that will carry the frame is recommended. The MTU for the virtual switch uplinks of the VTEPs performing VXLAN encapsulation is automatically increased when preparing the VTEP for VXLAN.

The following figure describes (at a high level) the steps required to establish layer 2 communication between VMs leveraging VXLAN overlay functionality:

1. VM1 originates a frame destined to the VM2 part of the same layer 2 logical segment (IP subnet).
2. The source VTEP identifies the destination VTEP where VM2 is connected and encapsulates the frame before sending it to the transport network.
3. The transport network is only required to enable IP communication between the source and destination VTEPs.
4. The destination VTEP receives the VLXLAN frame, de-encapsulates it, and identifies the layer 2 segment to which it belongs.
5. The frame is delivered to VM2.


VM1
VDS
Hypervisor (VTEP)
Transport Network
VDS
Hypervisor (VTEP)
VM2
L2 Frame
L2 Frame
VXLAN HDR
UDP HDR
Outer IP HDR
Outer MAC HDR
L2 Frame
1
VM sends a standard L2 Frame
2
Source Hypervisor (VTEP) adds VXLAN, UDP & IP headers
3
Physical Transport Network forwards frame as standard IP frame
4
Destination Hypervisor (VTEP) de-encapsulates headers
5
Original L2 Frame delivered to VM

Chapter 4

Automating Security Workflows

In This Chapter

- Identifying different security policy approaches
- Creating network and security policies for new application workloads
- Keeping up with continuous change
- Automating threat response in the data center
- Simplifying firewall management

Micro-segmentation enabled by network virtualization and the software-defined data center (SDDC) allows security workflows such as provisioning, moves/adds/changes, threat response, and security policy management to be automated for improved accuracy and better overall security in the data center. This chapter shows you how.

Creating Security Policies for the Software-Defined Data Center

Network virtualization provides the capability to micro-segment the SDDC to implement an effective security posture by intelligently grouping workloads based on their attributes and applying appropriate security policies.

Security policy rules can be created in various ways with network virtualization, as shown in Figure 4-1.

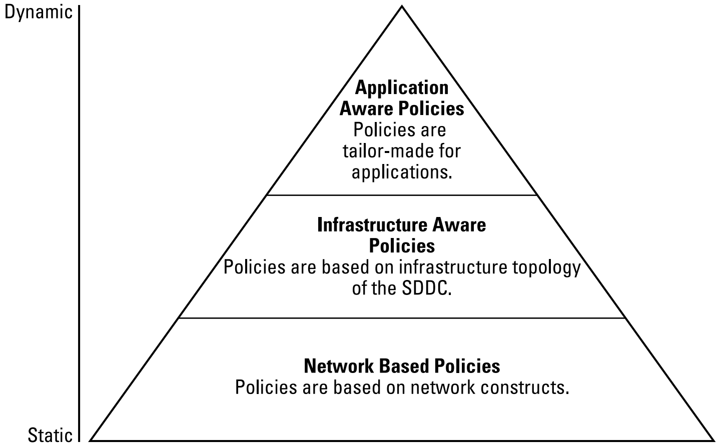

Dynamic
Static
Application Aware Policies
Policies are tailor-made for applications.
Infrastructure Aware Policies
Policies are based on infrastructure topology of the SDDC.
Network Based Policies
Policies are based on network constructs.

Figure 4-1: Network-, infrastructure-, and application-based policies. Network-based policies are intended for static environments. In more dynamic environments, policies need to evolve with the dynamic nature of the applications.

Network-based policies

Network-based security policies group elements based on Layer 2 or Layer 3 elements, such as media access control (MAC) or Internet Protocol (IP) addresses.

The security team needs to be aware of the networking infrastructure to deploy network-based policies. There is a high probability of security rule sprawl as grouping based on dynamic attributes is not used. This method of grouping works great if you're only migrating existing rules from different vendor firewalls.

In dynamic environments, such as self-service IT provisioning and cloud-automated deployments, where you're adding/deleting virtual machines (VMs) and application topologies change at a rapid rate, MAC address-based grouping approaches may not be suitable because there can be significant delay between provisioning a VM and adding the MAC addresses to the group. In data center environments with high workload mobility (for example, migrating VMs and high-availability), Layer 3 IP-based grouping approaches may also be inadequate.

Infrastructure-based policies

Infrastructure-based policies group data center infrastructure elements such as clusters, logical switches, and distributed port groups, among others. An example of this infrastructure-based policy would be to group a Payment Card Industry (PCI) cardholder data environment (CDE) in a single virtual LAN (VLAN) with appropriate security rules applied based on the VLAN name. Another example might use logical switches in your data center to group all VMs associated with a particular application onto a single logical switch. Effective infrastructure-based policies require close coordination between security and application teams to understand logical and physical boundaries in the data center.

If there are no physical or logical boundaries in your data center, then an infrastructure-based policy approach isn't feasible. You also need to be cognizant of where applications can be deployed in this scenario. For example, if you need the flexibility to deploy a PCI workload to any cluster that has adequate compute resources available, the security posture cannot be tied to a specific cluster. Instead, the security policy should move with the application.

Application-based policies

Application-based policies group data center elements based on a wide variety of customizable mechanisms, such as the application type (for example, VMs tagged as "Web_Servers"), application environment (for example, all resources tagged as "Production_Zone"), and application security posture. The advantage of this approach is that the security posture of the application is not tied to either network constructs or the data center infrastructure. Security policies can move with the application irrespective of network or infrastructure boundaries and policy templates can be created and reused across similar application types and workload instances.

To implement application-based policies, the security team only needs to be aware of the application that it is trying to secure based on the policies. The security policies follow the application life cycle from policy creation (when the application is deployed) to destruction (when the application is decommissioned). The application-based policy approach enables a self-service IT model. Concise and reusable security rules and templates can be created without knowledge of the underlying topology.

Infrastructure- and application-based policies provide the best security in virtualized networks using micro-segmentation.

Provisioning

Network virtualization provides the operational model of a VM for networks, streamlining provisioning of network and security services from weeks to seconds. Network virtualization significantly reduces the manual effort and cycle times associated with procuring, installing, and configuring traditional network hardware (see Chapter 6).

The powerful orchestration capabilities in a network virtualization platform programmatically distribute network services in lock step with VMs. Enterprises use these capabilities to standardize and maintain predefined templates that consist of the network and security topologies and services.

For example, a network engineer can create a template for a multitier application for development purposes. The environment can then be provisioned to an application developer in a matter of seconds via a self-service portal. The same can be done for quality assurance (QA), staging, and production environments — across multiple applications and services — with consistent configuration and security. These automation capabilities reduce operational expense, accelerate time to market, and speed IT service delivery.

Network virtualization also streamlines operations by consolidating configuration state and instrumentation data for all network connections — both virtual and physical. Administrators have complete operational visibility into what's occurring across the entire network infrastructure. This simplifies traffic management, monitoring, troubleshooting, and remediation.

Adapting to Change

Change is constant everywhere in today's world — except in the modern data center, which is a relatively static, "set it and forget it" utopian environment. Uh, no?!

Okay, data centers are constantly changing and IT organizations are struggling to keep up with ever demanding and increasingly dynamic business requirements. This capability and service delivery gap has become abundantly clear and been further exacerbated by current trends, such as server virtualization and cloud computing — trends that enable greater business agility and correspondingly more change (see Chapter 2 for more about these trends and challenges). Networking and security teams, in particular, have felt the strain as they're forced to use tools and solutions that haven't kept pace with the demands of the business and other functional IT areas.

Network virtualization untethers applications and services from physical network infrastructure, making networks as portable and agile as VMs. With network virtualization, networks are virtualized in the same software switch that is attached to the VM. When a workload is moved, its network and security services automatically move with it.

Enterprises use network virtualization to seamlessly migrate applications from one host to another, or from one data center to another. Real-world use cases include moving an application to leverage capacity at another location, complying with data residency laws, migrating to a new data center, or performing maintenance/refresh of the physical infrastructure.

For example, physical network topologies typically require IT to change IP addresses when applications are moved. In some cases, IP addresses are hardcoded into applications, which may require code changes and regression testing. With network virtualization, enterprises have the freedom to rapidly move applications without readdressing them. These conveniences significantly reduce operational cost and improve IT agility and responsiveness.

Responding to Threats

Modern attacks against the data center are sophisticated events that are rapidly evolving, and likewise require a rapid and adaptable response. Such a response can only be achieved effectively with automated security workflows. Today's adversary has the tools and resources necessary to automatically modify a threat or attack in order to sidestep static security controls and reactive countermeasures in the data center. Micro-segmentation provides the ability to match this capability and thwart attacks with equally sophisticated, granular security controls, and workflows applied to individual workloads in the data center.

For example, micro-segmentation enables security teams to enforce a security policy that provides both speed and security in a particular multi-tiered application architecture. Under normal operating conditions, this policy might perform basic security access control and malware scanning with minimal impact on application performance. However, if a malware threat is detected, the security policy can immediately isolate the application and its affected components from the rest of the network in order to prevent further exploitation of the application or any other applications in the data center. The newly applied policy might then require deep-packet inspection (DPI) by an integrated next-generation firewall in order to identify any other threats that may be using tactics, such as Secure Sockets Layer (SSL) hiding or port hopping, to evade detection and exfiltrate sensitive information.

Figures 4-2 and 4-3 illustrate an example of physical and logical views, respectively, of micro-segmentation in a multi-tiered application.

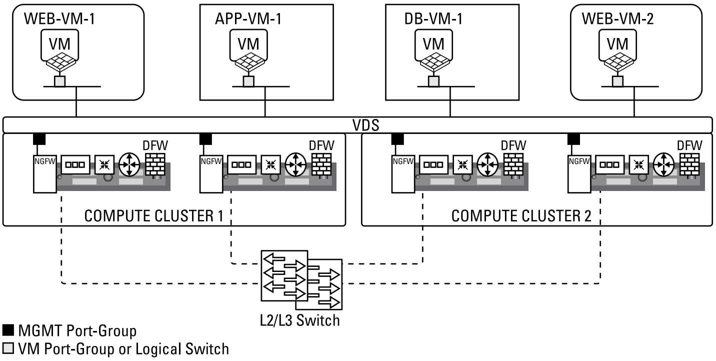


Figure 4-2: Physical view of micro-segmentation in a multi-tiered application.

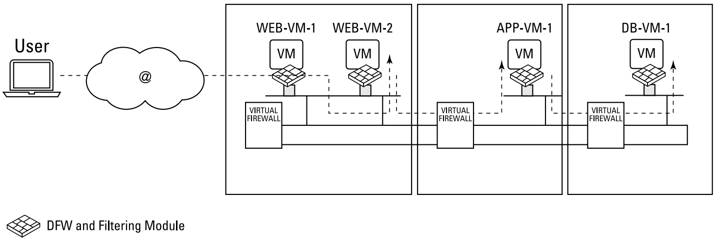


Figure 4-3: Logical view of micro-segmentation in a multi-tiered application.

TIP

You can find out more about the advanced security service insertion, chaining, and traffic steering capabilities that are possible with micro-segmentation in Chapter 2.

Firewalling Tens of Thousands of Workloads with a Single Logical Firewall

Finally, a network virtualization platform makes it possible to manage literally thousands of firewalls in the network fabric of the SDDC from a single "pane of glass" as a single firewall. Security administrators can automate workflows, policies, and rulesets within the firewall and configure other advanced firewalling capabilities — then propagate these configuration changes to thousands of firewalls — to protect the data center *inside* the perimeter and everywhere. Put another way, network virtualization enables distributed security policy enforcement with centralized management.

Chapter 5

Getting Started with Micro-segmentation

In This Chapter

▶ Deploying micro-segmentation in your data center

▶ Exploring micro-segmentation security use cases

In this chapter, you learn how to implement micro-segmentation in your data center and examine a few common micro-segmentation security use cases.

Achieving Micro-segmentation

Whether designing a software-defined data center (SDDC) built on a virtualized network with micro-segmentation or adding SDDC capabilities to a data center with existing infrastructure, a network virtualization platform enables enterprise architects to build a data center that is truly optimized for security and performance.

Core SDDC design principles with micro-segmentation (see Figure 5-1) include

- ✓ Isolation and segmentation
- ✓ Unit-level trust/least privilege
- ✓ Ubiquity and centralized control

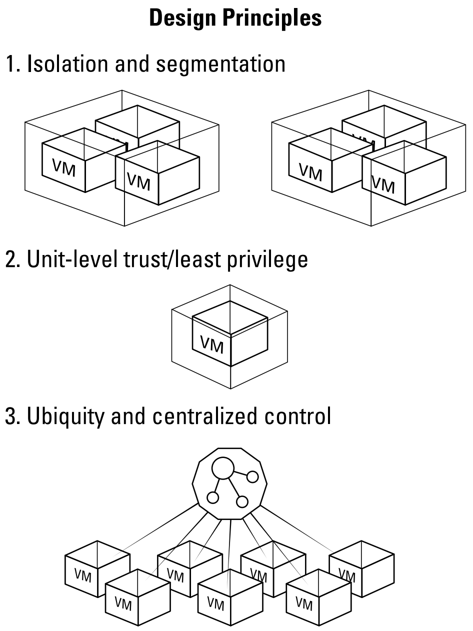

Design Principles
1. Isolation and segmentation
VM
VM
VM
VM
2. Unit-level trust/least privilege
VM
3. Ubiquity and centralized control
VM
VM
VM
VM
VM
VM
VM

Figure 5-1: Micro-segmentation, a new model for the software-defined data center.

See Chapter 2 for more information on these core micro-segmentation principles.

Traditional hardware-defined data center designs build a network around limited hardware resources and direct traffic through predetermined communication paths and security choke points. The SDDC can be built on commodity hardware that simply provides physical connectivity to the data center infrastructure — freeing data center architects to build innovative new security constructs that take advantage of the most efficient and simplified traffic flows throughout the data center.

To achieve a unit-level, zero-trust model (see Chapter 2) with micro-segmentation, start by understanding traffic flows within the data center. Then analyze relationships between the workloads. Finally, create a policy model that is aligned with the individual security needs of each workload.

Determine network flows

Understanding how network traffic flows into, out of, and within the data center is an important first step that often uncovers inefficiencies in traffic flows or security vulnerabilities that can potentially be exploited, which may have lain dormant for years.

Begin your traffic analysis by reviewing existing firewall rules on your perimeter firewalls and segregating north–south and east–west traffic. Various flow monitoring tools, such as IPFIX (NetFlow) or SYSLOG, can help you collect and analyze these traffic flows and will enable correlation with your existing firewall.

Flow patterns that are hairpinned (see Chapter 1) are typically indicative of east–west traffic flows. Analyzing existing firewall rules helps build an understanding of how to replace hairpinned traffic with logical switches and routing using virtualized network overlays.

Identify patterns and relationships

The rules from existing perimeter firewalls, when correlated with the flow patterns collected from flow monitoring tools, provide the initial set of security policies for the micro-segmentation model.

Flow patterns provide insight into the relationships that exist within your data center. For example, you can see how each of your workloads interacts with shared IT services, other applications or users, and across different environments (such as production versus development/test). Understanding these relationships will help you define appropriate micro-segments, and the rules that will govern the interaction between them. For example, you can create a micro-segment for each application, and then control communication to other micro-segments, such as shared IT services like Active Directory (AD), Domain Name Service (DNS), Network Time Protocol (NTP), and others.

TIP

Common examples of micro-segment definitions include by business department or tenant, environment, application, user access, and data classification or regulatory compliance.

Create and apply the policy model

To enable a unit-level, zero-trust model of micro-segmentation, start with a "default block" policy model, where no communication between the various workload relationships in the data center is allowed. In other words, start with every door and drawer in the bank locked. Based on your analysis of the traffic flow patterns and relationships, define security policies that incrementally open up specific communication channels between workloads, as needed. This is the best-practice method for protecting the data center with micro-segmentation.

Not all traffic flows and relationships in the data center may be fully understood. In these limited cases and with great discretion, use a "default allow" policy — essentially leaving the locks open — to prevent a service interruption of the application. Then block any inappropriate communication channels that are subsequently identified, to eliminate traffic between those micro-segments.

As workload and application/user/data contexts change over time, adjust your security policy model to align with the changing security needs of the workload, in order to constantly provide current and relevant security controls.

Security Use Cases

Enterprises are using network virtualization to deliver a multitude of new security use cases and high-value IT outcomes not previously possible with traditional networking infrastructure. IT is also performing existing operations faster and at a lower cost than ever before. Enterprises can often justify the cost of network virtualization through a single use case. At the same time, they establish a strategic platform that automates IT and drives additional use cases over time.

Popular use cases include disaster recovery, self-service research and development (R&D) clouds, cloud application portability and data center migration, IT automation and orchestration, and infrastructure optimization and refresh. The following sections highlight three additional use cases: server-to-server traffic, multi-tenancy, and virtual desktop infrastructure (VDI).

Network security inside the data center

Micro-segmentation brings security inside the data center with automated, fine-grained policies tied to individual workloads. Micro-segmentation effectively eliminates the lateral movement of threats inside the data center and greatly reduces the total attack surface.

East–west server communications have become more prolific within the data center as multitier application infrastructures built on virtualized server platforms are increasingly deployed.

This network traffic is typically unencumbered by traditional security controls and, instead, optimized for maximum performance and throughput because of a flawed security design, which assumes that threats are stopped at the perimeter firewall and anything inside the data center is trusted. As the fallacy of this security design is exposed with each high-profile data security breach reported in popular news media, organizations scramble to remedy the situation with largely inefficient practices such as hairpinning east–west traffic through firewall choke points in their existing data centers. Because micro-segmentation assumes a unit-level, zero-trust model of security that blocks all communication channels by default and requires lateral traffic to be explicitly allowed, these underlying foundational security concerns are eliminated.

TIP

See Chapter 1 for a complete discussion of hairpinning and Chapter 2 to learn about zero trust.

A network virtualization platform enables IT organizations to eliminate practices such as hairpinning by implementing extremely granular policies for individual workloads in the data center, using micro-segmentation.

DMZ anywhere

Today's fast-paced, global economy is driving businesses to demand data center access for their users from anywhere, at any time, and on any device. To securely enable the business and deliver access everywhere, IT needs a DMZ anywhere!

Micro-segmentation enables security controls to be assigned to the individual VM workload, rather than the underlying network topology. This capability makes it possible to apply advanced security services and restrict or grant Internet access to any system within the data center, irrespective of its location within the network. Thus, the network perimeter and DMZ are no longer defined by a physical firewall between the Internet and the data center. Instead, it's defined for each unique workload *inside* the data center.

Secure user environments

Many enterprises have deployed a virtual desktop infrastructure (VDI) to leverage virtualization technologies *beyond* the data center (see Figure 5-2). Micro-segmentation enables these organizations to extend many of the security advantages of the SDDC to the desktop — and even to mobile environments — including the following:

- Integrating key network and security capabilities into VDI management
- Eliminating complex policy sets and topologies for different VDI users
- Setting firewall and traffic filtering, and assigning policies for logical groupings
- Decoupling security policies from the network topology to simplify administration

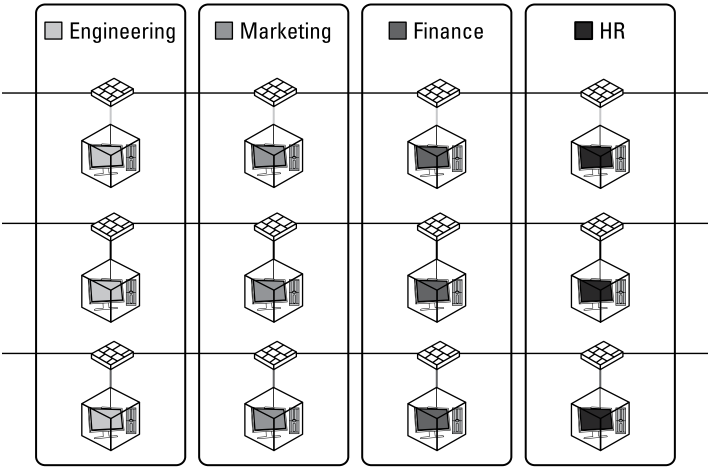


Figure 5-2: Micro-segmentation in a VDI environment.

These security and micro-segmentation use cases are just a few examples that demonstrate the many benefits of network virtualization. Other use cases include automating IT processes to keep pace with business requirements, securing a multi-tenant infrastructure, facilitating disaster recovery, enabling application continuity, and more.

In Chapter 6, you learn more about the business and security benefits of micro-segmentation in the data center.

Chapter 6

Ten (Or So) Key Benefits of Micro-segmentation

Micro-segmentation dramatically transforms network security inside the data center. In this chapter, you get the skinny on the business and functional benefits of micro-segmentation.

Minimize Risk and Impact of Data Center Security Breaches

If a threat infiltrates the data center, micro-segmentation contains and blocks its lateral movement to other servers, which dramatically reduces the attack surface and risk to the business. Micro-segmentation isolates each workload with its own security policy, preventing attackers from exploiting other systems and stealing valuable data.

By reducing the attack surface, micro-segmentation helps organizations avoid or minimize the cost and impact when a data breach occurs, including

- Direct legal costs such as actual and punitive damages, fines, and attorneys' fees based on the size and scale of the data breach
- Loss of customers from turnover or diminished acquisition rates
- Forensic analysis and investigations

- Lost productivity
- Miscellaneous costs such as free credit reports, identity monitoring subscriptions, customer communications, and outsourcing hotline support

As noted in Chapter 1, the average cost of a data breach for U.S. companies has been estimated at almost $6 million, and several high-profile breaches in recent years have far exceeded $100 million.

Automate IT Service Delivery and Speed Time to Market

Just as server virtualization transformed the operational model of computing, networking has been transformed by network virtualization which enables micro-segmentation and, in turn, has transformed security in the data center. Enterprises are using micro-segmentation to provision security services with the same agility, speed, and control as virtual machines (VMs) for computing.

With micro-segmentation, enterprises can provision security services for cloud-native or traditional applications in a matter of seconds. Application teams can have access to full self-service provisioning — no more waiting days or weeks for hardware to be procured, networking to be set up, and security to be configured. Plus, automation and orchestration capabilities eliminate the risk of manual configuration errors that can result in performance issues or, worse yet, security holes.

See Chapters 2 and 3 for a complete discussion on how server virtualization creates new business challenges for networking and security teams, and how network virtualization and micro-segmentation address these challenges.

Finally, micro-segmentation significantly shortens the time it takes to safely and securely bring new revenue-generating applications and services to market. This new level of speed and agility fuels rapid innovation and competitive advantage.

Simplify Network Traffic Flows

The volume of server-to-server (east–west) traffic generated by modern applications inside the data center continues to grow exponentially, which consumes network bandwidth, increases latency, adds complexity, and increases oversubscription on the data center network core.

Network virtualization and micro-segmentation enable direct east–west communication between server workloads through a virtual switch or aggregation fabric which

- Significantly reduces east–west traffic hops for better application performance (near wire speed for VMs on the same physical host)
- Eliminates inefficient *hairpinning* (forcing east–west traffic through physical firewalls), which creates choke points, backhauls excessive server traffic, increases complexity, and contributes to firewall rule sprawl
- Enables workload mobility by allowing individual workloads to be deployed anywhere in the data center with their own security policies, instead of being tied to the physical network topology

TIP

See Chapter 1 to learn about east–west traffic, hairpinning, and workload mobility in the data center.

Enable Advanced Security Service Insertion, Chaining, and Traffic Steering

Environments that require advanced, application-level network security capabilities can leverage micro-segmentation to distribute, enable, and enforce advanced network security services in a virtualized network context. Network virtualization distributes network services into the virtual network interface (vNIC) of individual VM workloads. This forms a logical pipeline of network and security services that can be applied to virtual network traffic for safe enablement of applications and complete threat protection at a granular level with micro-segmentation.

Another powerful benefit of micro-segmentation is the ability to build unit-level policies for individual VM workloads, which leverage service insertion, chaining, and steering to drive service execution in the logical services pipeline based on the result of other services. This capability makes it possible to coordinate and correlate otherwise completely unrelated network security services from multiple vendors.

Leverage Existing Infrastructure

Micro-segmentation is not an all-or-nothing proposition. Because virtual networks require no configuration changes to the underlying physical network (beyond allowing network virtualization encapsulated packets through existing firewalls), they can transparently coexist on the physical network — with as much or as little micro-segmentation of existing application workloads as needed.

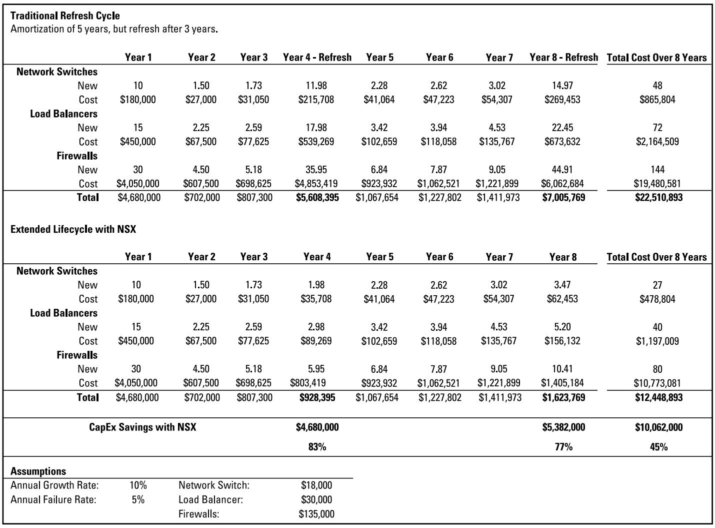
IT departments have the flexibility to virtualize and segment portions of the network simply by adding hypervisor nodes to the virtualization platform. In addition, gateways — available as software or top-of-rack switch hardware — deliver the ability to seamlessly interconnect virtual and physical networks. These can be used, for example, to support Internet access by workloads connected to virtual networks, or to directly connect legacy virtual LANs (VLANs) and bare metal workloads to virtual networks.

Enterprises are using micro-segmentation and network virtualization to bridge and simplify data centers without disruption. Micro-segmentation works with traditional multitier tree-type architectures and flatter next-generation fabric architectures. The result is a common platform with the same logical networking, security, and management model. Enterprises are also using micro-segmentation and network virtualization for a number of optimization and consolidation scenarios. For example, integrating and securing information systems following mergers and acquisitions, maximizing hardware sharing across tenants in multitenant clouds, and accessing islands of unused compute capacity.

All of this means that organizations can deploy micro-segmentation in their data centers at a pace that suits *their* unique business needs — whether in a proof-of-concept pilot project, a high-value multitiered application, or a full-scale greenfield software-defined data center (SDDC) build-out.

Chapter 5 explains how to get started with micro-segmentation in your data center.

With micro-segmentation, organizations can leverage their existing physical network and security equipment and, in many cases, significantly extend the useful life of their existing infrastructure. For example, you may be able to avoid the expense of expanding core capacity with more hardware by eliminating excessive traffic through network firewalls due to convoluted data center traffic patterns, such as hairpinning and backhauling east–west server traffic. Figure 6-1 provides an example of the savings an enterprise might typically expect for extending the useful life of existing network and security infrastructure.



Traditional Refresh Cycle
Amortization of 5 years, but refresh after 3 years.

	Year 1	Year 2	Year 3	Year 4 - Refresh	Year 5	Year 6	Year 7	Year 8 - Refresh	Total Cost Over 8 Years
Network Switches									
New	10	1.50	1.73	11.98	2.28	2.62	3.02	14.97	48
Cost	$180,000	$27,000	$31,050	$215,708	$41,064	$47,223	$54,307	$269,453	$865,804
Load Balancers									
New	15	2.25	2.59	17.98	3.42	3.94	4.53	22.45	72
Cost	$450,000	$67,500	$77,625	$539,269	$102,659	$118,058	$135,767	$673,632	$2,164,509
Firewalls									
New	30	4.50	5.18	35.95	6.84	7.87	9.05	44.91	144
Cost	$4,050,000	$607,500	$698,625	$4,853,419	$923,932	$1,062,521	$1,221,899	$6,062,684	$19,480,581
Total	$4,680,000	$702,000	$807,300	**$5,608,395**	$1,067,654	$1,227,802	$1,411,973	**$7,005,769**	**$22,510,893**

Extended Lifecycle with NSX

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Total Cost Over 8 Years
Network Switches									
New	10	1.50	1.73	1.98	2.28	2.62	3.02	3.47	27
Cost	$180,000	$27,000	$31,050	$35,708	$41,064	$47,223	$54,307	$62,453	$478,804
Load Balancers									
New	15	2.25	2.59	2.98	3.42	3.94	4.53	5.20	40
Cost	$450,000	$67,500	$77,625	$89,269	$102,659	$118,058	$135,767	$156,132	$1,197,009
Firewalls									
New	30	4.50	5.18	5.95	6.84	7.87	9.05	10.41	80
Cost	$4,050,000	$607,500	$698,625	$803,419	$923,932	$1,062,521	$1,221,899	$1,405,184	$10,773,081
Total	$4,680,000	$702,000	$807,300	**$928,395**	$1,067,654	$1,227,802	$1,411,973	**$1,623,769**	**$12,448,893**
CapEx Savings with NSX				**$4,680,000**				**$5,382,000**	**$10,062,000**
				83%				**77%**	**45%**

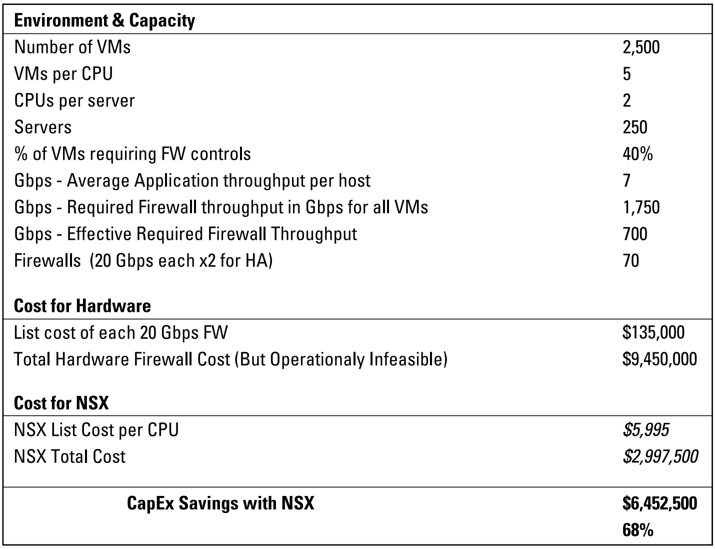
Assumptions

Annual Growth Rate:	10%	Network Switch:	$18,000
Annual Failure Rate:	5%	Load Balancer:	$30,000
		Firewalls:	$135,000

Figure 6-1: Hardware lifecycle capital expenditure savings.

Reduce Capital Expenditures

Deploying additional physical firewalls to control increasing volumes of east–west traffic inside the data center is cost prohibitive for most enterprises. Additionally, the sheer number of devices needed and the effort required to set up and manage a complex matrix of firewall rules make such an approach operationally infeasible. Micro-segmentation enables complete control of individual workloads in the data center without purchasing additional physical firewalls for each workload, resulting in significant savings in enterprise data centers. Figure 6-2 illustrates this use case for a typical enterprise data center.



Environment & Capacity	
Number of VMs	2,500
VMs per CPU	5
CPUs per server	2
Servers	250
% of VMs requiring FW controls	40%
Gbps - Average Application throughput per host	7
Gbps - Required Firewall throughput in Gbps for all VMs	1,750
Gbps - Effective Required Firewall Throughput	700
Firewalls (20 Gbps each x2 for HA)	70
Cost for Hardware	
List cost of each 20 Gbps FW	$135,000
Total Hardware Firewall Cost (But Operationaly Infeasible)	$9,450,000
Cost for NSX	
NSX List Cost per CPU	*$5,995*
NSX Total Cost	*$2,997,500*
CapEx Savings with NSX	**$6,452,500** **68%**

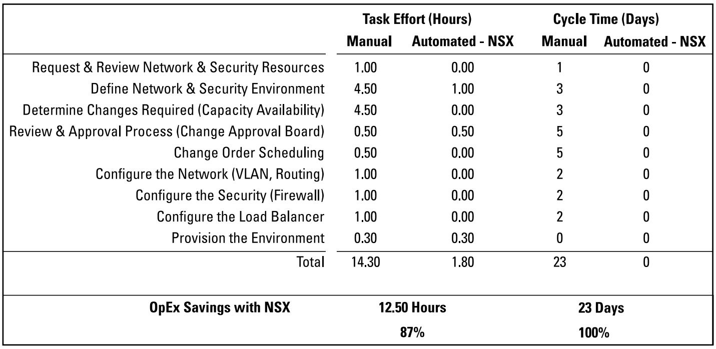
Figure 6-2: Micro-segmentation eliminates the need for additional physical firewalls.

Lower Operating Expenses

Micro-segmentation dramatically reduces the manual effort and cycle time for security tasks, including provisioning, change/adaptation, scaling, and troubleshooting/remediation.

Generally, micro-segmentation reduces the effort from hours to minutes and the cycle times from days to minutes. If you consider all the manual tasks required to provision and manage security for a physical network — across development, testing, staging, and production environments — and the fact that micro-segmentation automates these tasks, you begin to see all the opportunities for reducing operational costs.

As the analysis in Figure 6-3 shows, micro-segmentation dramatically speeds the initial provisioning of security services into production. With traditional hardware, the associated cycle time to provision security services for a new application forces enterprises to wait 23 days. Network virtualization reduces that to minutes — nearly a 100 percent reduction and a massive time-to-market win. Likewise, provisioning security services for a new application takes 14 person hours or close to two days of person effort. Micro-segmentation reduces that to less than 2 person hours — a substantial 87 percent reduction.



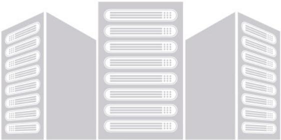
	Task Effort (Hours)		Cycle Time (Days)	
	Manual	Automated - NSX	Manual	Automated - NSX
Request & Review Network & Security Resources	1.00	0.00	1	0
Define Network & Security Environment	4.50	1.00	3	0
Determine Changes Required (Capacity Availability)	4.50	0.00	3	0
Review & Approval Process (Change Approval Board)	0.50	0.50	5	0
Change Order Scheduling	0.50	0.00	5	0
Configure the Network (VLAN, Routing)	1.00	0.00	2	0
Configure the Security (Firewall)	1.00	0.00	2	0
Configure the Load Balancer	1.00	0.00	2	0
Provision the Environment	0.30	0.30	0	0
Total	14.30	1.80	23	0
OpEx Savings with NSX	12.50 Hours		23 Days	
	87%		100%	

Figure 6-3: IT automation operational expenditure reductions.

Securely Enable Business Agility

The benefits of micro-segmentation through network virtualization are immense. Businesses have historically been forced to choose between speed and security as IT security teams are often unfairly perceived to inhibit business agility instead of safely enabling the business. This conflict strains the relationship between IT security teams and business units, often leading to counterproductive cat-and-mouse games between users attempting to circumvent controls so they can perform their job functions, and IT security administrators trying to protect those same users from themselves by enforcing unwieldy, maligned security policies — or worse, responding to resulting security incidents.

Network virtualization makes micro-segmentation in the software-defined data center a reality and enables businesses to rapidly — and *securely* — innovate to achieve competitive advantage, while maintaining ubiquitous and persistent security in the data center. Businesses everywhere are enjoying the many security and performance benefits of micro-segmentation in the data center, and will continue to discover innovative uses and applications for this truly disruptive technology.

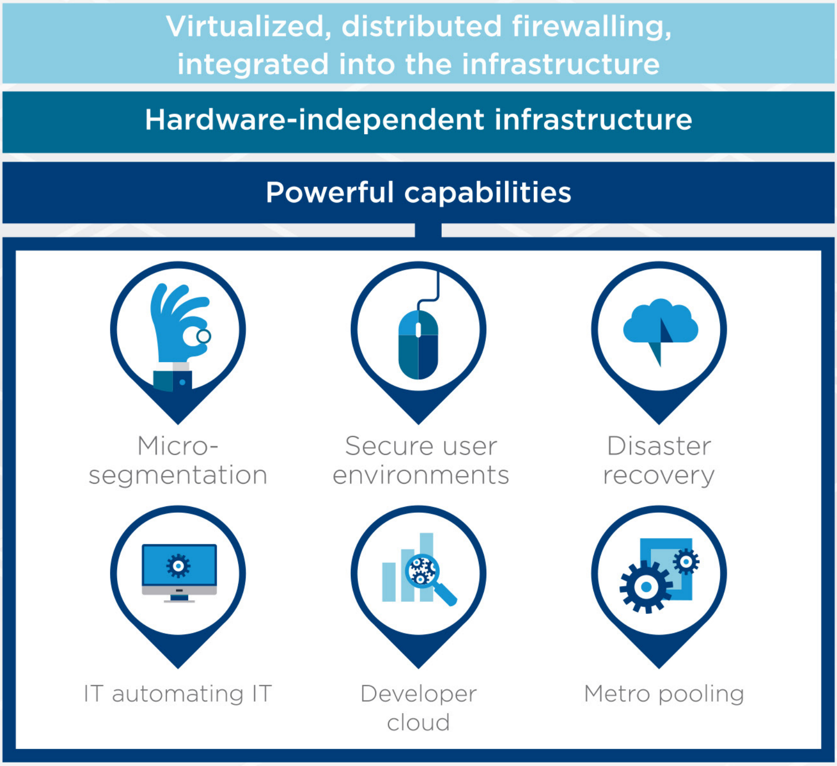


vmware

NETWORK VIRTUALIZATION AND SECURITY

The Software-Defined Data Center - is critical to modern IT as a focal point for innovation.

Network virtualization is a core component of the SDDC, delivering security controls that are native to the infrastructure, giving you greater agility, protection of assets, and improved security.



Virtualized, distributed firewalling, integrated into the infrastructure

Hardware-independent infrastructure

Powerful capabilities

Micro-segmentation

Secure user environments

Disaster recovery

IT automating IT

Developer cloud

Metro pooling

GET STARTED NOW

Find out more about how to virtualize your data center network

www.vmware.com/products/nsx

One **CLOUD.** Any **APPLICATION.** Any **DEVICE.**™

vmware

Micro-segmentation is the new foundation for security to block threats *inside* the data center

Micro-segmentation dramatically transforms network security inside the data center perimeter by containing and blocking the lateral spread of threats to other servers. This book explains how to implement micro-segmentation with your existing data center infrastructure to dramatically reduce the data center attack surface and risks to your business.

- ***Learn what's wrong with the data center security foundation* — and how to fix it with micro-segmentation**
- ***Make Zero Trust in the data center a reality* — with unit-level trust and security policies defined down to the individual workload level**
- ***Automate security workflows* — and combine different security technologies by chaining advanced security services to improve security response and performance**
- ***Understand the security benefits of micro-segmentation* — and how it transforms data center security with innovative use cases**

Open the book and find:

- **Why micro-segmentation in the data center hasn't been feasible — until now**
- **How attackers move laterally in the data center — and how they exploit the explosion of east–west data center traffic to expand their attacks**
- **How to leverage micro-segmentation with your existing infrastructure and improve data center security and performance**

Go to Dummies.com® for more!

Lawrence Miller has worked in information security in various industries for more than 25 years. **Joshua Soto** is a product marketing manager for the VMware NSX platform.

FOR DUMMIES®
A Wiley Brand

Also available as an e-book

ISBN: 978-1-119-17734-0
Not for resale