Making Everything Easier!™

2nd VMware Special Edition

Micro-segmentation

FOR DUMMIES®

A Wiley Brand

Learn to:

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

Brought to you by vmware®

Matt De Vincentis

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

2nd VMware Special Edition

by Matt De Vincentis

FOR DUMMIES®

A Wiley Brand

Micro-segmentation For Dummies®, 2nd VMware Special Edition

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

ISBN 978-1-119-44854-9 (pbk); ISBN 978-1-119-45337-6 (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
Executive Editor: Katie Mohr
Editorial Manager: Rev Mengle
Business Development Representative: Karen Hattan
Production Editor: Magesh Elangovan
Special Help: Shinie Shaw, 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 security 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.

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

The primary issue is that once an attack gets into the network, there are few controls to prevent threats from moving laterally from system to system. The best way to solve this is to adopt a stricter, more 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" security model — in other words, the principle of least privilege applied to the network. 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 of an individual workload. In a physical network, this would require an enormous number of physical firewalls to be deployed, so up until now, micro-segmentation has been cost-prohibitive and operationally infeasible. However, with the mainstream adoption of 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 I 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 an 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 networking or security practitioner or decision maker, evaluating data center security strategies and solutions for your organization.

If these assumptions are true, then this is the book for you!

Icons Used in This Book

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

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!

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!

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.

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 I 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/go/nsx`.

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 I don't recommend upside down or backward).

I 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 the problem with traditional approaches to data center security

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

In this chapter, you explore data center breaches — how they happen and why traditional data center security approaches are ineffective, leaving the inside of the data center relatively defenseless in the event of an attack.

Data Breaches Continue to Occur

Despite a heightened focus on security, as evidenced by increasingly stringent compliance requirements, data protection laws, heavy investments in security technology, and

ever growing and ever capable security teams, data breaches continue to occur at an alarming rate. And each new breach seems to dwarf the last in terms of the number of records stolen and the cost of the breach to the business.

Verizon's *2017 Data Breach Investigation Report* studied 42,068 reported security incidents from 2016 alone, which resulted in 1,935 confirmed data breaches for the year. This does not, of course, represent any of the data breaches that went unreported. In its *2017 Cost of Data Breach Study,* the Ponemon Institute calculated the average total cost of a data breach globally was $3.62 million. Whichever way you slice the numbers, the frequency of data breaches and the associated costs to an organization are astounding.

The now famous attacks on organizations like Sony, Home Depot, Target, and Yahoo! all have one characteristic in common: Once the perimeter was breached, attackers were able to move laterally from server to server within the data center with essentially no security controls in place to stop this movement. Sensitive data was then collected and exfiltrated. These cases highlight a major weakness of traditional data center security strategies: 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, security technologies and controls that are applied to the perimeter of the data center need to be considered and implemented *inside* the data center as well, in order to stop or isolate attacks once the perimeter is breached.

The Life Cycle of a Data Center Attack

Today's sophisticated cyberattacks exploit a foundational vulnerability that exists in modern data center security design: 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 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).

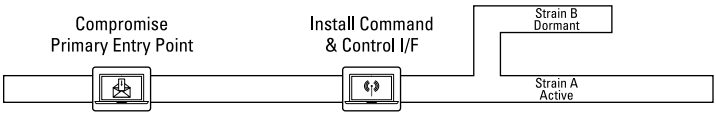


Figure 1-2: C2 enables further reconnaissance in the data center.

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.

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).

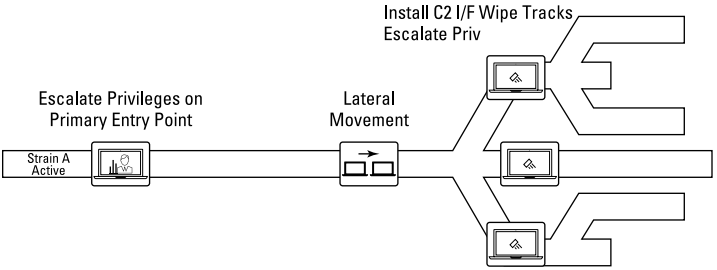


Figure 1-3: Additional C2 infrastructure is installed to ensure persistence as the attacker moves laterally through the data center.

Modern cyberattacks take advantage of relatively weak or non-existent security within the data center to move freely between different systems in order to 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 persistent and resilient. 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 (server-to-server) traffic inside the data center, make it difficult — if not impossible — for incident response teams to effectively see and isolate an attack.

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

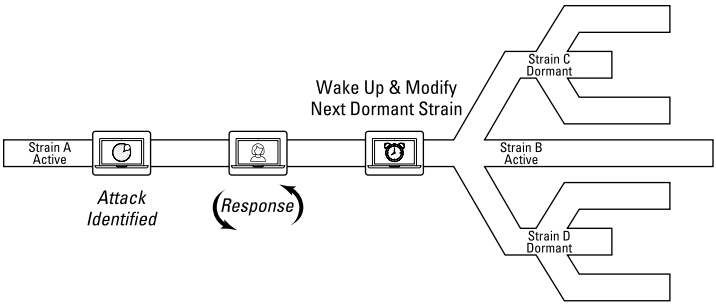


Figure 1-4: If an attack is discovered, the attacker simply makes a dormant strain active and continues the attack.

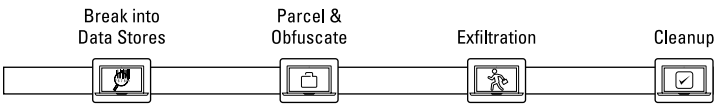


Figure 1-5: The attacker is then free to perform any desired actions on the data center objective.

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.

In Chapter 2, you learn how micro-segmentation prevents successful attacks by blocking an attacker's lateral movement in the data center, and with capabilities such as advanced security service insertion, service chaining, and traffic steering.

Once inside your data center, an attacker 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).

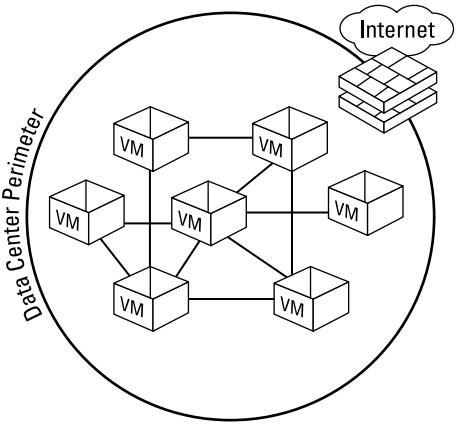


Figure 1-6: Perimeter-based security is insufficient in a data center where security is needed everywhere.

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, between client workstations and the data center, or between different security zones. For example, a network may be segmented into multiple trust levels using additional firewalls to create a DMZ or between different application tiers (such as web, app, and DB tiers). 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 needs. And again, the primary focus has been 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 sometimes less expensive than hardware firewalls, come with most of the same pitfalls as physical firewalls and 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 — security strategy

using existing technologies, controls, and processes has been impractical . . . until now. Network virtualization (explained in Chapter 3) makes micro-segmentation a reality in the data center, without needing to change any existing physical infrastructure.

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.

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

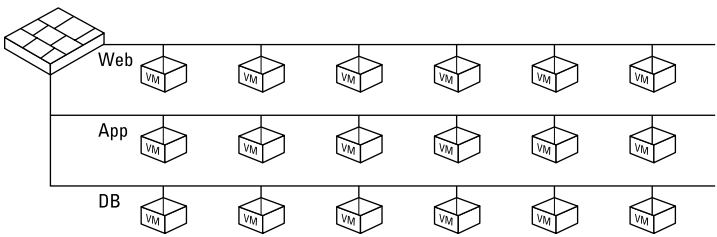


Figure 1-7: Today, security is tied to a rigid and complex network topology that is further complicated by a consolidated, multitier application infrastructure.

Different segments should be created inside the perimeter to limit the lateral spread of threats within the data center. To be most effective, segmentation should be defined and security policy should be enforced down to the individual workload level.

In addition to inadequate segmentation, another unfortunate consequence of traditional data center design that adds complexity and degrades network performance is hairpinning east–west traffic (see Figure 1-8).

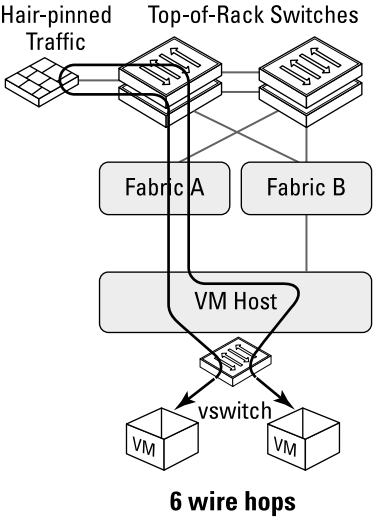


Figure 1-8: Firewalling east–west traffic causes hairpinning in a traditional design.

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

- 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 inefficient when it comes to virtualized workloads running on the same host that need to be firewalled from each other, that could have otherwise communicated securely without even having 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), and many other technologies. Although these technologies bolster perimeter defenses, they are often designed to address specific threats with limited context sharing and correlation between each other, 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 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 in stopping threats from spreading *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 financially and operationally infeasible.
- **Security mapped to network topology:** Security policies determined by the physical location of a server workload in the data center are too rigid and too complex. This legacy approach often a manual process, leading to significant delays in application deployment.
- **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.
- **The inefficiency of hairpinning:** Forcing east–west traffic through firewalls unnecessarily backhauls server traffic, creates choke points, and contributes to sprawling firewall rulesets and complexity.

While perimeter-based security is important, 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 and the shortcomings of agent-based security

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 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 to implement.

Traditional packet-filtering and next-generation firewalls place controls in the form of 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 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.

Network virtualization enables micro-segmentation to deliver several significant advantages over traditional network security models — automated provisioning, automated moves/adds/changes, 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 in a multitier server architecture 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 more than 80 percent of all network traffic in most data centers. These multitier application architectures are typically designed with little or no security controls restricting communications within each tier.

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 has contributed to an alarming lack of visibility and control of a significant amount of data center traffic.

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. This can cause significant issues for network and security teams attempting to identify problems, and is fertile ground for an attacker.

The hypervisor in a virtualized environment is uniquely positioned to see *all* traffic in the data center, as all packets must pass through it. This provides organizations with unprecedented visibility and context into data center traffic, enabling security policy to be defined in creative ways that have never before been possible. Because of the ubiquitous nature of the hypervisor and its unique visibility and context of individual workloads, micro-segmentation policy can be created based on workload characteristics (see Figure 2-1) rather than static constructs such as IP address, port, and protocol.

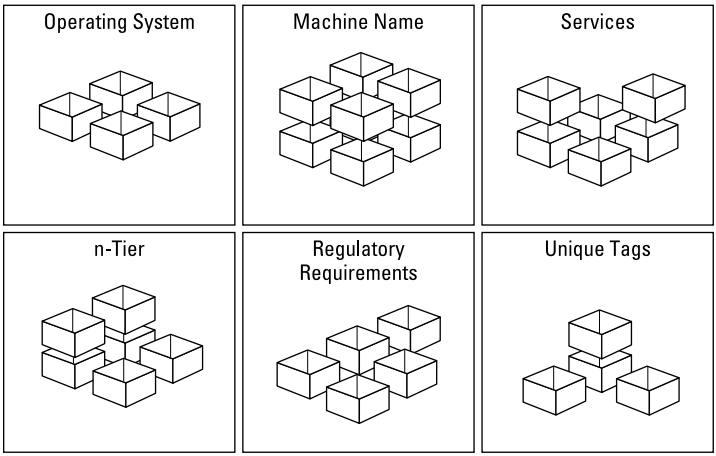


Figure 2-1: Intelligent grouping defined by customized criteria.

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.

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.

Forrester Research outlines its "Zero Trust model" of information security and isolation, in which 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 or virtual LANs (VLANs) 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, without causing conflicts. 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. This isolation protects the underlying physical infrastructure from possible attack initiated by workloads in a virtual network. Again, all of this is independent from any physical subnets or VLANs that would traditionally be required to create this isolation.

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, exposing the network to the possibility of a security breach.

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 physical 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 switching, routing, firewalling, and others — are provisioned with a workload, 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 in physical firewalls.

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, resulting in security holes.

Organizations can also apply a combination of different 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 into the virtual networks 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, service chaining, and traffic steering").

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, service chaining, and traffic steering

The VMware NSX network virtualization platform provides stateful firewalling to deliver segmentation within virtual networks. In some environments, particularly those under regulatory compliance, there may be a requirement for more advanced network security capabilities. In these instances, organizations can leverage the NSX platform to automatically distribute, enable, and enforce advanced network security services in a virtualized network environment. The NSX platform inserts these services to form a logical pipeline of capabilities that can be applied to virtual network traffic. This includes third-party services not offered by NSX natively, allowing physical or virtual services from NSX partners to be inserted at the right place and at the right time.

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, service 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 on a VM, Trend Micro could block the malware and trigger an event that automatically 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 Dell Technologies, Palo Alto Networks, Checkpoint, Fortinet, Trend Micro, Symantec, Intel Security, and more.

Essential Elements of Micro-segmentation

As discussed earlier in this chapter, the hypervisor is uniquely positioned to provide both context and isolation throughout the data center — 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 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 across clouds (for example, a live migration or for disaster recovery purposes).

Network virtualization 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 micro-segmentation 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. Network virtualization makes this possible by embedding security functions into the hypervisor itself. By taking advantage of this widespread nature of the hypervisor, administrators can rely on the availability of security functions for all the workloads within the data center, making micro-segmentation operationally feasible for the first time.

Extensibility

Aside from persistence and ubiquity, micro-segmentation must also be able 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, network virtualization enables administrators to extend capabilities by integrating additional security functions into their defense portfolio. For instance, administrators might begin with micro-segmentation with stateful firewalling distributed

throughout the data center, but add next-generation firewalling and intrusion prevention systems for deep packet inspection (DPI), or agentless anti-malware for better data center endpoint 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 innovations, such as the server virtualization, as a potential new target for attack. But the reality is that virtualization opens up new possibilities in the way you secure your applications. In other words, for IT security professionals, virtualization is a weapon, not a target. Virtualization 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, you place controls in the host operating system. This approach allows you 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 having a switch on a home alarm system that a burglar could just turn off!

An alternative approach, which trades context for isolation, places the control in the physical infrastructure. This approach isolates the control from the attack surface, 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 offers the ideal location to achieve both context and isolation, combined with ubiquitous enforcement. Controls operating in the data center virtualization layer leverage secure guest introspection, the ability to provide agentless, high-definition context of the guest workload, 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 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.

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 (server-to-Internet) and east–west (server-to-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.

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 not new. But the reality of achieving micro-segmentation is new, made possible for the first time with network virtualization and security distributed in 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. Perimeter firewalls still perform an important function as the first line of defense for a network, but as discussed earlier, their inherent weakness in a perimeter-centric strategy means they should not be the foundation for data center security alone.

Next, micro-segmentation does not introduce a performance bottleneck. The firewalling performance of the VMware NSX platform, for example, is impressive. NSX delivers near line-rate firewall throughput. This means that if your physical hosts have dual 10 GbE interfaces, NSX can deliver almost 20 Gbps of firewalling throughput per host. And because the firewall capability within NSX is distributed in the hypervisor, every time another host is added to the data center, another 20 Gbps of firewall throughput capacity is also added!

Micro-segmentation isn't possible with existing tools and technologies either, 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).

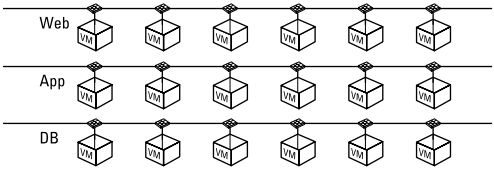


Figure 2-2: Intelligent grouping of data center workloads in a multitier environment that is completely independent of the underlying physical network is only possible with micro-segmentation.

This level of granularity would potentially require deploying thousands of firewalls (physical, virtual, or both) in a large environment— which isn't feasible, financially or operationally, for most organizations. Although micro-segmentation effectively enables the deployment and centralized management, automation, and orchestration of potentially 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 or host based firewalls installed on individual workloads themselves. There are two problems with this approach: lack of isolation and operational overhead. A common technique employed by attackers is to disable security agents or firewalls inside a workload, when they've gained access to it. Installing agents or firewalls on individual workloads also creates an operational nightmare for security teams, often requiring manual installation and frequent updating. Micro-segmentation is isolated from the attack surface, inside the hypervisor itself. It's also baked into the platform and distributed throughout the environment, eliminating the need to install and update agents.

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

Data centers today are receiving tremendous benefits from compute and storage virtualization solutions that consolidate infrastructure resources, reduce operational complexity, and dynamically align and scale their application infrastructure in response to business priorities. However, data center networking and security remain rigid, complex, proprietary, and closed to virtualization innovation — a barrier to realizing the full potential of the software-defined data center.

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.

Moving networking and security to software creates a common platform that delivers the operational model of a virtual machine (the ability to create, snapshot, delete, and restore) to the network.

The current hardware-centric approach to networking has resulted in slow, manual, error-prone provisioning of networking and security services to support application deployment. Network operators are dependent on command-line interfaces (CLIs) and scripting to manipulate a multitude of virtual LANs (VLANs), routing configurations, firewall rules, load balancer policies, 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 single monolithic application stacks that could be provisioned in isolation on a physical server with a static security policy tied to it. Today's applications are distributed and demand workload mobility, speed to market, and accessibility in a way that simply did not exist before. But the current hardware-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, firewall rules, and so forth. This process is not only slow and complex, but also one that is fundamentally limited (for example, 4,094 VLANs). Organizations often resort to expensive overprovisioning of compute capacity for each cluster or pod in order to achieve isolation or compliance objectives, 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).

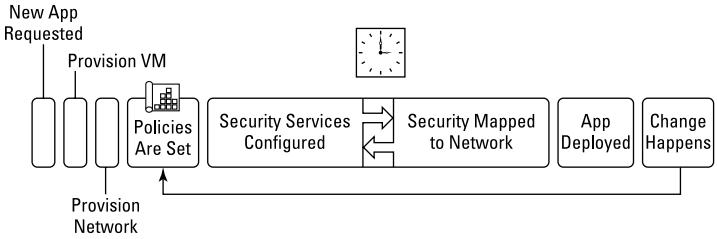


Figure 3-1: Rigid security policies mapped to network topology can't keep up with the pace of business changes.

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 concepts to the entire data center. The SDDC transforms data center economics and business agility through abstraction and automation, while still leveraging existing physical compute, network, and storage infrastructure investments.

In much the same way that server virtualization programmatically creates, snapshots, deletes, and restores virtual machines (VMs), network virtualization programmatically creates, snapshots, deletes, and restores 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 underlying physical network.

Network virtualization is a completely nondisruptive solution that can be deployed on any IP network, including traditional data center network architectures or more modern fabric architectures.

With compute virtualization, a software abstraction layer (server hypervisor) reproduces the 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).

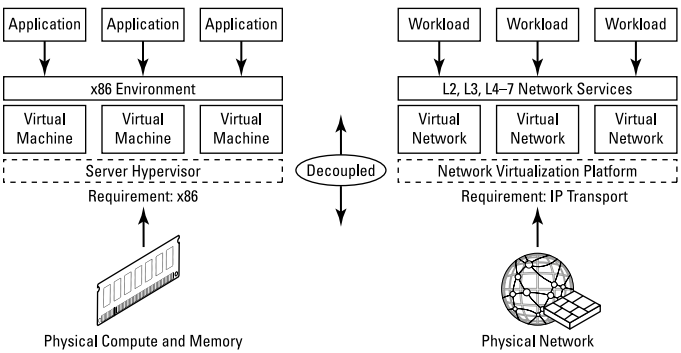


Figure 3-2: Compute and network virtualization.

With network virtualization, the functional equivalent of a "network hypervisor" reproduces the attributes of a data center network — 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 inserted and chained together, in any combination or order, to deliver the most effective security controls possible (see Chapter 2 to learn more about this capability).

Additionally, network virtualization enables you to distribute these same 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 also allows you to scale your capacity with the addition of each new host. For example, as you add a new host to scale your data center workload capacity, you're also adding 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 approaches, 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 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 construct that presents logical compute services to an application, a virtual network is a software construct 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 already present in the hypervisor and the network services are pushed to them automatically, to effectively deliver a platform — or "network hypervisor" — for the creation of virtual networks (see Figure 3-3).

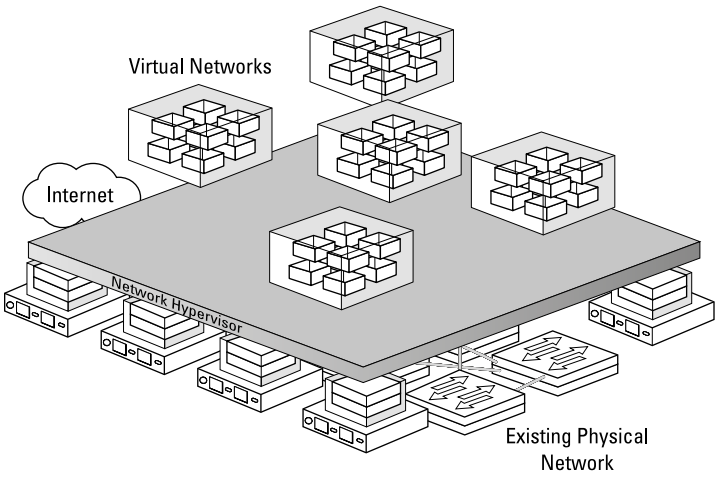


Figure 3-3: The "network hypervisor."

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).

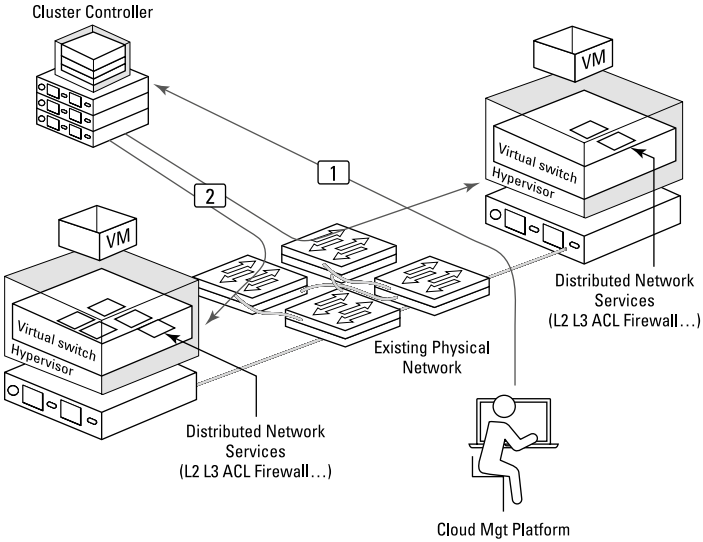


Figure 3-4: Virtual network provisioning.

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.

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.

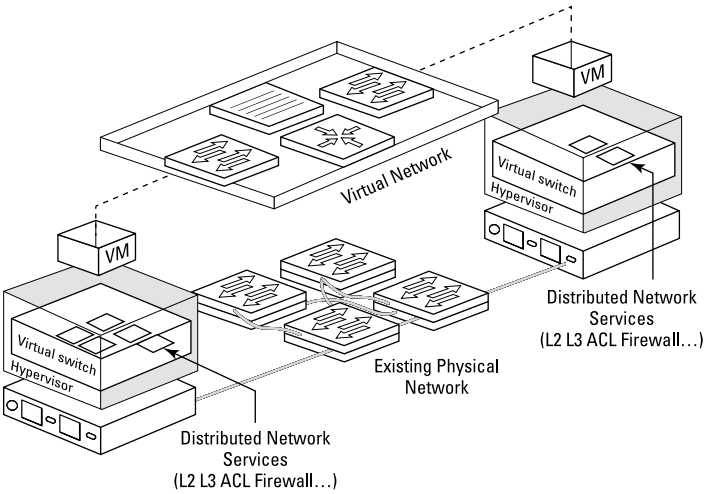


Figure 3-5: The virtual network, from the workload's perspective (logical).

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. 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.

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.

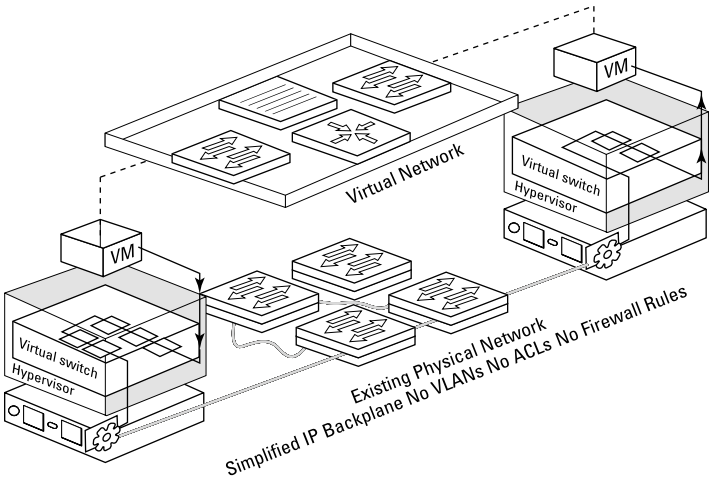


Figure 3-6: The virtual network, from the network's perspective (physical).

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 key for network virtualization: centralized control with granular security 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 networking components that provide the IP backplane.

Unlike software-defined networking (SDN), which is essentially a special type of hardware managed centrally by software, network virtualization technology is entirely hardware agnostic and truly decouples network services 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 runs in controller nodes and is the central control point within a virtualized network. The controllers maintain information about all virtual machines, hosts, distributed local routers, logical switches, and VXLANs (explained later in this chapter). The controller nodes do not have any dataplane traffic passing through them, and 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 through a GUI or CLI. The management plane may be integrated with a cloud management platform for orchestrating the deployment 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, even over layer 3 boundaries, irrespective of the physical network design.
- **Routing:** Routing between IP subnets can be done in the logical space without traffic going out to a 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 virtual network interface–level and like routing, is performed in the hypervisor kernel. 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 virtual and physical spaces.

The *data plane* transports the network traffic. In a virtualized network, data plane functions are 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 layer 2 overlay networks, running on top of existing physical IP infrastructure, without the need to re-architect any of the data center network.

The data plane also supports comprehensive traffic management, visibility, monitoring, and troubleshooting within a virtual network using features such as port mirroring, NetFlow and IP Flow Information Export (IPFIX), configuration backup and restore, quality of service (QoS), Link Aggregation Control Protocol (LACP), and more.

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 in software at layer 2 (bridging) or layer 3 (routing), or by integrating the network virtualization platform with a physical hardware switch, also known as a hardware Virtual Tunnel Endpoint (VTEP).

Encapsulation

Encapsulation (or "overlay") protocols decouple networks in the logical space from the physical network infrastructure. Devices connected to logical networks can leverage network functions — including switching, routing, firewalling, and load balancing — independently from how the underlying physical infrastructure is configured. The physical network effectively becomes an IP backplane used to simply 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 (or even across multiple data centers), 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 4,094. This number, while it may seem large for typical deployments, is becoming a serious bottleneck for most cloud providers and even large enterprises.

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, using some techniques similar to VLAN but supporting up to 16 million logical networks.
- **Network Virtualization using Generic Routing Encapsulation (NVGRE)** uses Generic Routing Encapsulation (GRE) to tunnel layer 2 packets over layer 3 networks.
- **Generic Network Virtualization Encapsulation (Geneve)** specifies a data plane schema and decouples encapsulation from the control plane to provide flexibility across different deployment scenarios.
- **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.

Not sure which encapsulation protocol to use? Don't worry, they can all coexist on the same network and essentially perform the same function. Use any encapsulation protocol your particular network virtualization technology supports.

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 4,094 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).

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.

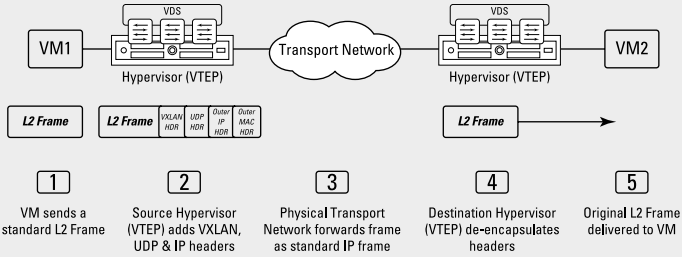
The following figure describes (at a high level) the steps required to establish layer 2 communication between VMs leveraging VXLAN overlay functionality:

1. VM1 originates a frame destined to the VM2 part of the same layer 2 logical segment (IP subnet).

(continued)

(continued)

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, deencapsulates it, and identifies the layer 2 segment to which it belongs.
5. The frame is delivered to VM2.



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

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

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 trends such as virtualization, containerization, 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 because they're forced to use tools and solutions that haven't kept pace with the demands of the business and other areas of IT.

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

Creating Security Policies for Modern Application Environments

Network virtualization provides the capability to micro-segment the data center to provide a robust security posture by intelligently grouping workloads based on their attributes (such as machine name, OS, tags, and so on) rather than static network constructs (such as IP address, port, protocol, and so on) and applying appropriate security policies to them automatically.

There are various models for creating security policy with network virtualization, as shown in Figure 4-1.

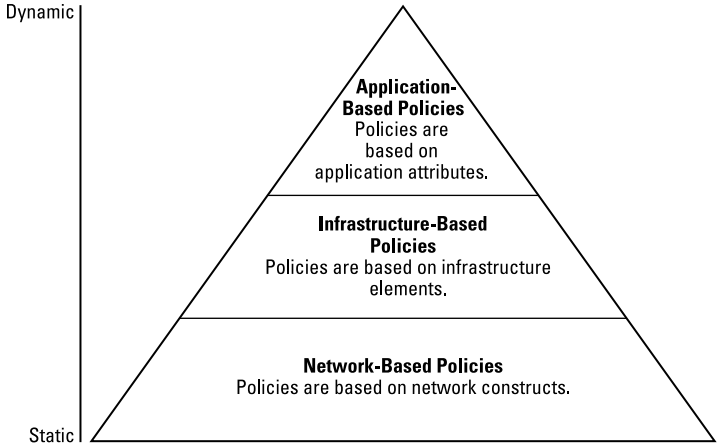


Figure 4-1: Network-, infrastructure-, and application-based policies. Network-based policies are intended for static environments. In more dynamic environments, policies need to evolve with the dynamic nature of the applications.

Network-based policies

Network-based security policies group workloads 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 topology to deploy network-based policies. There is a high probability of security rule sprawl as grouping based on workload attributes are not used. In addition, stale rules are often left in firewalls when workloads are decommissioned, as rules are not directly tied to the workload. This method of grouping works fine if you have a relatively static environment.

In dynamic environments, such as clouds leveraging automation or those with self-service capabilities, where you're adding/deleting virtual machines (VMs) and application topologies change at a rapid rate, network-based grouping approaches may not be suitable because there can be significant delay between provisioning a VM and the manual creation of network-based policies. This is true also for environments with a high amount of workload mobility (for example, where VMs are migrated dynamically to better distribute resources).

Infrastructure-based policies

Infrastructure-based policies group infrastructure elements such as hosts, clusters, logical switches, 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 security policies applied to the VLAN. 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 network, security, and application teams to understand logical and physical boundaries in the network.

If there are no physical or logical boundaries in your data center or cloud environment, then an infrastructure-based policy approach isn't feasible. You also need to be cognizant of where applications can be physically 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, a better approach is one that is application-based, is automatically deployed, and can move with the application.

Application-based policies

Application-based policies group data center elements based on attributes of the workloads, such as the machine name (for example, VMs that have a name starting with "Web-"), application environment (for example, all VMs tagged as "Production_Zone"), or operating system (for example, all VMs running Windows). The advantage of this approach is that the security posture of the application is not tied to either network constructs or the physical infrastructure. Security policies can move with the application irrespective of network changes or infrastructure boundaries and policy templates can be created and automatically used when new workload instances are deployed.

To implement application-based policies, the security team only needs to be aware of the application that they are trying to secure and don't need a deep understanding of the network or data center topology. 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 security policy approach enables automated cloud and self-service IT models. Concise and reusable security rules and templates can be created without knowledge of the underlying network topology.

Infrastructure- and application-based policies provide the best security in dynamic 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 installing and manually configuring legacy physical network hardware (see Chapter 6).

The powerful orchestration capabilities in a network virtualization platform programmatically distribute network and security services in lock step with VMs. Organizations use these

capabilities to standardize and maintain predefined templates that consist of the network and security topologies and services.

For example, an administrator can create a template for a multitier application for development purposes. The environment can then be provisioned for 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 policies. These automation capabilities reduce operational expense, and speed up IT service delivery to ultimately accelerate time to market.

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 when security workflows are automated. Today's adversaries have 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, with security policy applied to individual workloads in the data center and advanced security service insertion.

For example, micro-segmentation enables security teams to enforce a security policy that provides both high-performance and robust security for a multi-tiered application. 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 enforce a more restrictive security policy to the affected components in order to prevent further exploitation of the application or any other applications in the data center. The newly applied policy might 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.

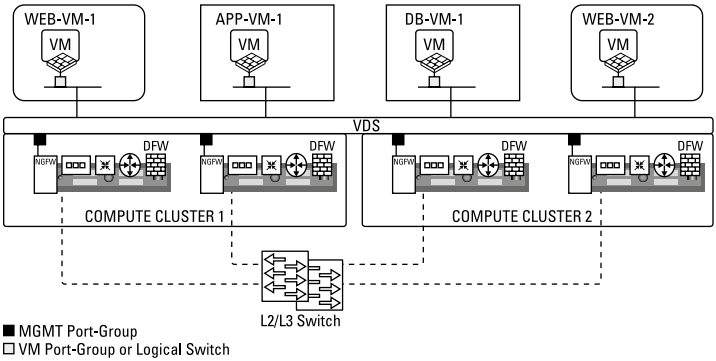


Figure 4-2: Physical view of micro-segmentation in a multi-tiered application.

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.

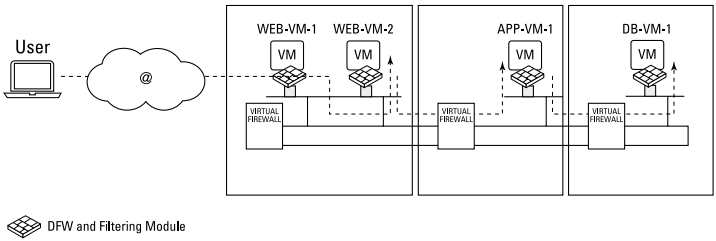


Figure 4-3: Logical view of micro-segmentation in a multi-tiered application.

Firewalling Tens of Thousands of Workloads with a Single Logical Firewall

Finally, a network virtualization platform makes it possible to manage literally thousands of distributed firewalls embedded in each hypervisor, as if they were a single firewall, managed

from a single "pane of glass." 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 automatically — to protect the data center *inside* the perimeter. Put another way, network virtualization enables distributed security policy enforcement with centralized management.

Chapter 5

Getting Started with Micro-segmentation

In This Chapter

- Planning for micro-segmentation
- Deploying micro-segmentation
- Looking at additional security use cases enabled by network virtualization

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 new software-defined data center (SDDC) built on network virtualization with micro-segmentation or adding micro-segmentation capabilities to an existing data center, a network virtualization platform enables organizations to build a data center that is truly optimized for security.

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

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

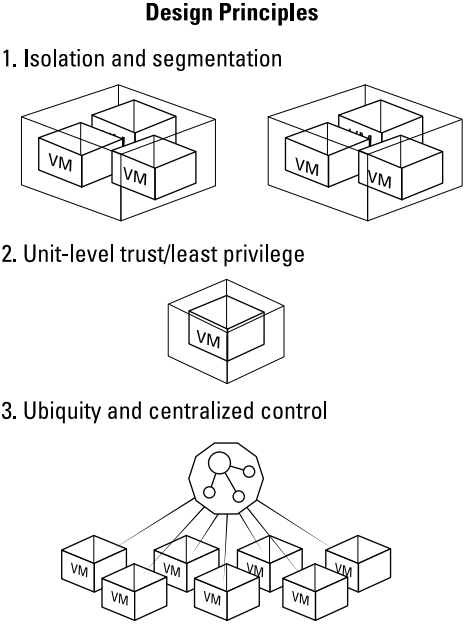


Figure 5-1: Micro-segmentation, a new model for network security.

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

Legacy data center designs build a network around rigid hardware constructs and direct traffic through predetermined communication paths and security choke points. Network virtualization can be built on any networking hardware, turning the physical network into a simple pool of IP transport capacity — 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 and their traffic communication patterns. Finally, create a policy model that is aligned with the individual security needs of each application and workload.

Determine network flows

Understanding how network traffic flows into, out of, and within the data center is an important first step in planning for micro-segmentation. This process often uncovers inefficiencies in traffic flows that can be optimized with micro-segmentation, reducing application latency and load on the network.

Begin your traffic analysis by reviewing existing firewall rules on your perimeter firewalls and identifying north–south versus east–west traffic. Flow monitoring protocols, such as IPFIX (NetFlow), can help you collect and analyze these traffic flows. Commercial monitoring tools are also available to simplify the flow monitoring process, providing a deeper level of analysis and enabling correlation of flows with your existing firewall policies (see the sidebar "Intelligent operations for network virtualization and micro-segmentation").

Flow patterns that are hairpinned (see Chapter 1) are typically indicative of east–west traffic flows. Analyzing existing network flows and firewall rules helps build an understanding of how to reduce hairpinned traffic with more efficient firewall controls, implemented in the hypervisor.

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 that should be created in the micro-segmentation security model.

Flow patterns provide insight into the relationships that exist between workloads within your data center. For example, you can see how each of your workloads interacts with other workloads, with shared IT services, other applications, 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.

Intelligent operations for network virtualization and micro-segmentation

The VMware vRealize Network Insight platform provides a complete solution for planning, configuring, and troubleshooting network virtualization. It is purpose-built for VMware NSX and provides planning and security policy recommendations for implementing micro-segmentation.

In addition, vRealize Network Insight fits into the overall vRealize family of products to provide intelligent operations for the entire software-defined data center across compute, storage, network, and applications.

Here are some of the key use cases for vRealize Network Insight.

Planning for micro-segmentation deployments

- Comprehensive NetFlow assessment and analysis to model security groups and firewall rules
- Recommendations to make micro-segmentation easier to deploy
- Continuous monitoring and audit compliance postures of the distributed firewalls over time
- Optimization of network performance across virtual and physical networks
- Physical to virtual network topology mapping
- Performance optimization across overlay and underlay networks

Ensuring virtual network health and troubleshooting

- Intuitive user interface (UI) and natural language search to quickly pinpoint issues
- Best practice and compliance checking
- Logging of analytics for accelerated troubleshooting via integration with vRealize Log Insight

Create and apply the policy model

Micro-segmentation can be implemented nondisruptively and doesn't need to be enabled for the entire data center at once. A common implementation strategy is to apply micro-segmentation one application at a time. This enables you to build your micro-segmentation security policy at your own pace, without disrupting critical business operations.

To enable a unit-level, zero-trust model of micro-segmentation, start with a "default block" policy model, where no communication between workloads is allowed for the application that you're focusing on. In other words, start with every door and drawer in the bank vault locked. Based on your analysis of the traffic flow patterns and the relationships between workloads, 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 for a particular application. Then block any inappropriate communication channels that are subsequently identified, to eliminate traffic between those micro-segments.

As workload and application/user/data requirements 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

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

The following sections highlight three such use cases: securing server-to-server traffic, the creation of DMZs anywhere, and securing end-user environments.

Securing server-to-server traffic

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 that occurs, 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 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.

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, enforced in the hypervisor layer, 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 to applications from anywhere, IT needs the ability to create a DMZ anywhere!

A data center network DMZ traditionally required separate server and networking hardware isolated from the rest of the data center, leading to underutilized pools of resources and increased costs. This physically separate network model also required complex manual configurations of networking and security policies, drastically slowing operations in the data center.

Micro-segmentation enables security controls to be assigned to the individual VM workload. Thus, the network perimeter and DMZ are no longer defined by the separation of physical infrastructure. Instead, a DMZ can be defined for each unique application, in software, on a common pool of hardware resources.

Secure user environments

Many organizations have deployed a virtual desktop infrastructure (VDI) to leverage virtualization technologies to improve their end-user computing experience (see Figure 5-2). Micro-segmentation enables these organizations to extend the same micro-segmentations concepts to the virtual desktop — and even to mobile devices themselves — including the following:

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

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.

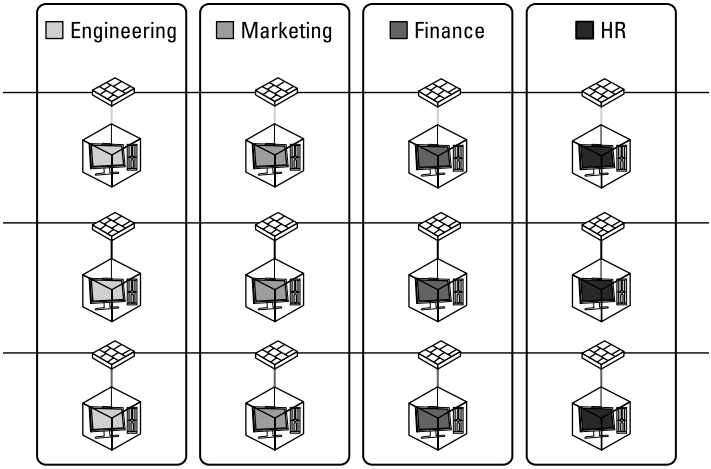


Figure 5-2: Micro-segmentation in a VDI environment.

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, outsourcing hotline support, and more

As noted in Chapter 1, the average cost of a data breach globally has been estimated at almost $3.62 million in 2017, and several high-profile breaches in recent years have far exceeded $100 million in losses.

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 network virtualization to provision security services with the same agility, speed, and control as VMs for computing.

With micro-segmentation, enterprises can provision security services for applications in a matter of seconds. Application teams can access self-service provisioning systems to deploy their own secure environments — 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 downtime — worse yet, security holes.

See Chapter 2 for a complete discussion on how the growth of east–west traffic has created 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 for organizations.

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 and increases latency in the network.

Network virtualization and micro-segmentation enable direct east–west communication between server workloads, while still being protected by a firewall (embedded in the hypervisor).

- Significantly reduces east–west traffic hops for better application performance (without traffic even hitting the wire for VMs on the same physical host)
- Eliminates inefficient *hairpinning* (forcing east–west traffic through physical firewalls), which creates choke points and backhauls unnecessary traffic
- 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

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 network security capabilities can leverage micro-segmentation to distribute, enable, and insert advanced network security services in a virtualized network context. Network virtualization distributes network services directly to the virtual network interface (vNIC) of individual VM workloads. This creates a logical pipeline where network and security services can be applied to virtual network traffic for complete threat protection at a granular level.

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 special hardware or any configuration changes to be made to the underlying physical network, they can transparently coexist on the physical network — with as much or as little micro-segmentation of existing application workloads as desired.

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 switch hardware — deliver the ability to seamlessly interconnect virtual and physical networks. These can be used, for example, to securely connect legacy virtual LANs (VLANs) and physical server workloads to virtual networks.

Organizations are using micro-segmentation and network virtualization to bridge and simplify data centers without disruption. Micro-segmentation works with traditional three-tier network architectures or modern fabric architectures. The result is a common platform with the same logical networking, security, and management model. Organizations 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 multi-tier 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.

Reduce Capital Expenditures

Deploying additional physical firewalls to control increasing volumes of east–west traffic inside the data center and implement micro-segmentation is cost prohibitive for most organizations. Additionally, the sheer number of devices needed and the effort required to set up and manage such 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, resulting in significant savings in for data centers. Figure 6-1 illustrates this use case for a typical data center.

Environment and Capacity	
Number of VMs	2,500
VMs per CPU (Consolidation Ratio)	5
CPUs per Host	2
Total Hosts	250
Average East–West Traffic per VM (Gbps)	0.3
% of East–West Traffic Needing Micro-Segmentation	80%
Total Firewall Throughput Required (Gbps)	600
Total Hardware Firewalls Required (20 Gbps Each x2 for HA)	60
Cost for Hardware Solution	
20 Gbps Hardware Firewall List Cost	$135,000
Total Hardware Firewall Cost (But Operationally Infeasible)	$8,100,000
Cost for NSX Solution	
Total NSX License Cost	$2,247,500
CapEx Savings with NSX	**$5,852,500**
	72%

Figure 6-1: Micro-segmentation eliminates the need for additional physical firewalls.

Lower Operating Expenses

Micro-segmentation dramatically reduces the manual effort and time required for security tasks, including moves/adds/changes, scaling, and troubleshooting/remediation.

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 6-2 shows, micro-segmentation dramatically speeds up 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 on average. 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 on average. 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 6-2: 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 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, and will continue to discover innovative uses and applications for this truly game-changing technology.

Micro-segmentation is the 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. 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 about the problem with traditional data center security — and how to fix it with micro-segmentation***
- ***Make Zero Trust in the data center a reality — with 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***
- ***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!

Matt De Vincentis is a Group Product Marketing Manager for Networking and Security at VMware. He has a Master's of Business Administration and a Master's of Networking and Systems from Charles Sturt University in Australia. Matt currently lives with his family in Cupertino, California.

FOR DUMMIES® A Wiley Brand

Also available as an e-book

ISBN: 978-1-119-44854-9
Not for resale

WILEY END USER LICENSE AGREEMENT

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