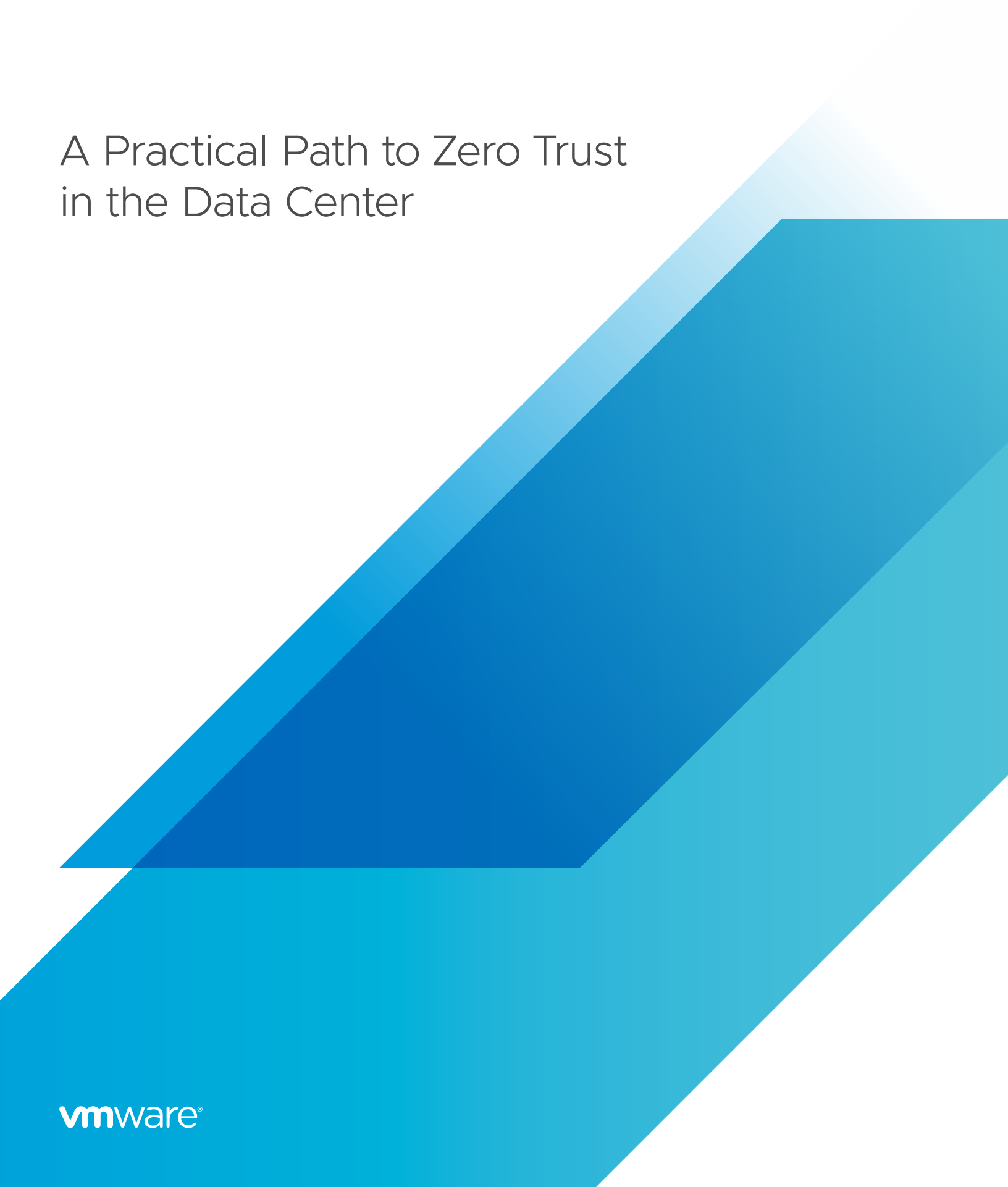
WHITE PAPER–DECEMBER 2020

A Practical Path to Zero Trust in the Data Center



vmware®

Introduction

Securing the enterprise has never been easy. Even in the old days, when IT infrastructure was small and self-contained and cyber threats were few, it was not enough just to insert a few edge firewalls to protect an organization from outside attack. By implicitly trusting everyone on the inside, a perimeter approach assumed that all internal users would act responsibly, no user IDs would be compromised, and everyone could be trusted. This, in turn, allowed malicious or compromised insiders to wreak havoc, steal sensitive data, and bring organizations to their knees.

Today, the network is increasingly complex, distributed, and remote: large percentages of the workforce are doing their work remotely, employee-owned (BYOD) devices are commonly used for work-related activities, and cloud-based applications, services, and data are becoming the norm. All the while, bad actors invent new and even more malicious ways of attacking weak spots in organizational security. The old methods of protecting an increasingly diffuse and porous network are not enough.

Increasingly, organizations are adopting the concept of zero trust. For the data center, this means, by default, trusting no entity on the network and distrusting all traffic unless a security policy explicitly allows it. According to NIST[1], "Zero trust security models assume that an attacker is present in the environment." Further, "a zero trust architecture is … designed to prevent data breaches and limit internal lateral movement." (For a broader overview of zero trust see [2])

Lateral movement is key. While older, edge (perimeter) firewall-based methods concentrated on north-south traffic (coming from the Internet into the organization and going back out), the internal movement was invisible. Hackers who succeeded in penetrating the network were free to move about the data center, escalating privileges, and gaining undesired access to sensitive assets. With zero trust, all east-west network traffic is forced to undergo inspection so that anomalous or unwanted behavior can be spotted and stopped before damage is done. Several approaches have been tried to achieve zero trust in the past, but with varying degrees of success.

EDGE FIREWALLS

- Can't enforce workload-level security policies
- Can't cost-effectively inspect all east-west traffic
- Don't have full workload visibility

Traditional Firewalls Fall Short

Firewalls would appear to be ideally suited to achieving zero trust. Many organizations already use firewalls to segment their end-user networks from their data center networks, in response to the fact that hackers may find it easier to infiltrate end-user devices and from there get a toehold into the data center. When organizations started segmenting end-user networks from the data center, edge firewalls were the only tool available. Whether physical or virtual appliances, they were repurposed to serve as internal firewalls to segment the network. However, there are three main problems with this approach.

First, edge firewalls don't explicitly consider application architecture in their design. Some might recognize specific workload types, but they remain blind to the relationship between workloads and applications. Thus, they cannot deal with modern, modularized applications. Today an application may comprise multiple workload types, microservices, and containers that run in the data center or in the public cloud. An edge firewall repurposed as an internal firewall can only block traffic based on ports, protocols, IP addresses, or the workload type. But to be truly effective, an internal firewall needs to enforce policies holistically at the level of individual workloads within an application.

Second, an edge firewall used as an internal firewall can be a costly proposition. Such firewalls quickly run into capacity problems if trying to inspect all east-west traffic, creating the need for multiple firewalls that must be periodically upgraded to deal with traffic increases. As a result, most organizations choose to inspect only a small amount of east-west traffic, if at all.

Third, and perhaps most important, to monitor east-west traffic and enforce granular policies, the internal firewall needs to have full visibility down to the workload level, enabling it to automatically determine the communication patterns between workloads and microservices. Traditional edge firewalls deployed internally lack this visibility. Often, they are not in the path of east-west traffic. They lack the logic to cleanly distinguish traffic flows between one application's workloads from those of another even when they are. This makes it challenging—if not almost impossible—to create and enforce rules at the workload or individual traffic flow level.

Micro-segmentation Orchestrators Have Their Own Issues

Enterprise edge firewalls repurposed as internal firewalls presented a set of insurmountable obstacles. *Micro-segmentation* services came into being to overcome the weaknesses of this approach in protecting applications and their workloads at a granular level. The typical micro-segmentation solution consists of two components: an agent and an orchestrator. The agent resides on the workload (e.g., inside the virtual machine or the physical server hosting a database) and enforces security policies distributed to it by the orchestrator. Micro-segmentation allows for very granular control over the interaction between workloads in a micro-segment while restricting an attacker's visibility over attackable assets and lateral movement across the network.

The biggest problem with most micro-segmentation solutions for zero trust is that the agent is actually a user-level process that utilizes the host operating system's firewall to program security policies into the OS kernel. This means they cannot add security capabilities beyond what is offered by the operating system (OS)—typically maintained by a third party such as Red Hat or Microsoft[3].

For example, suppose the security team wants to implement a security policy focused on users and applications. Say, one that allows users from the finance group to access finance applications, but not human resources applications. Because the host OS firewall can only operate on IP addresses, port numbers, and protocol identifiers, the team cannot create such a policy. This limits their ability to express security policies for applications and users.

MICRO-SEGMENTATION ORCHESTRATORS:

- Are limited by the host operating system
- Can't deploy IDS/IPS threat controls as a second layer of defense
- May increase the complexity of the network, not reduce it

Another weakness of the host OS further diminishes the utility of micro-segmentation for zero trust: most host operating systems don't implement intrusion detection/prevention systems (IDS/IPS). These are software or network hardware solutions deployed to analyze live traffic as it passes through the network, detecting threats that have slipped through other defenses. This means there is no way to deploy IDS/IPS threat controls as a second layer of protection within permitted traffic. Security teams faced with this limitation are forced to either purchase and deploy a specialized IDS/IPS appliance or maintain an internal firewall with IDS/IPS—in addition to the microservices orchestrator. Clearly, this does nothing to simplify infrastructure or reduce costs.

Distributed Internal Firewalls—The Best of Both Worlds

Traditional edge firewalls, deployed internally, provide some advantages for organizations moving down the path to zero trust. So too do micro-segmentation orchestrators. A distributed *internal firewall* that judiciously combines the best characteristics of both results in a far superior approach. The distributed internal firewall overcomes the weaknesses presented by both lack of visibility and granular control in edge firewalls and limitations in the host OS; it likewise can simplify network architectures and reduce overall costs.

The VMware NSX Service-defined Firewall was explicitly designed to be an internal firewall, implemented entirely in software[4] and satisfying the critical requirements for internal firewalling in the data center. With an access control engine—a classic stateful firewall that also recognizes applications and users/groups—and an analytics engine plus a threat control engine to provide advanced threat control (such as IDS/IPS), the *Service-defined Firewall* provides full topology visibility, advanced analytics, and a simplified security architecture (see Figure 1).

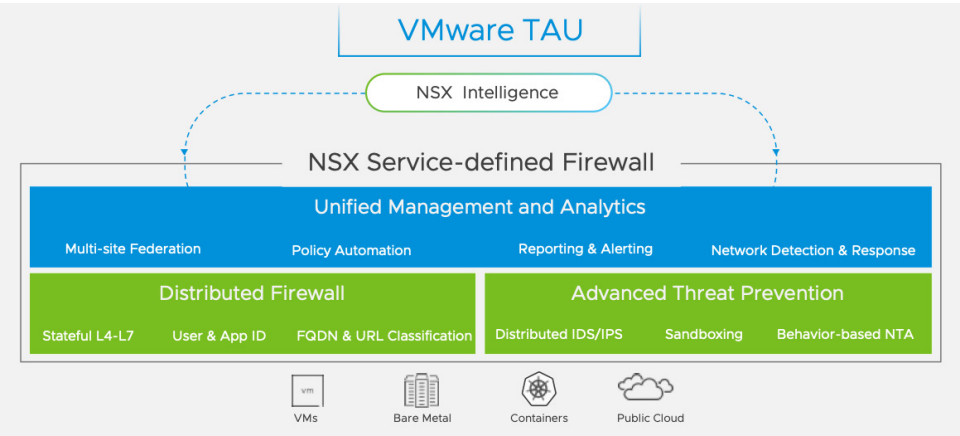


Figure 1: VMware NSX Service-defined Firewall

The Service-defined Firewall provides granular enforcement at the level of individual workloads within an application. It can inspect all east-west traffic for access control, threat control (via its built-in IDS/IPS functionality), and analytics. When the distributed internal firewall runs on the physical servers hosting the workload, close to the origin or destination of east-west traffic, its capacity can grow or shrink with the servers rather than being constrained by the inspection capacity available in a centralized firewall appliance.

Centralized management enables security teams to manage east-west network security from a single portal that presents a unified view of the topology down to the workload level. Based on communication patterns and observed behavior of applications, workloads, and microservices, the Service-defined Firewall can make security policy recommendations and check that the results match the intended outcomes.

By associating security policies with workloads, lifecycle management is simplified: policies can be constructed for a workload before it is even created, ensuring protection from the minute it comes into being. If the workload is moved or decommissioned, the team does not need to scramble to define or remove policies. And security tags can be associated with workload properties (OS type, VM name, user identities, etc.), making security policies easier to express and maintain.

Best of all, by combining full network firewall capabilities with threat controls such as IDS/IPS, there is no need for separate micro-segmentation and firewall products. This, in turn, reduces security architecture complexity.

A Step-by-Step Process for Zero Trust in the Data Center

Securing the data center and achieving zero trust will undoubtedly take some time and effort – both of which are scarce commodities today, especially for security teams. That's why it's best to break down the journey into discrete, phased steps. Not only does this let the team show early success, but it also helps build expertise along the way, which accelerates later stages.

The distributed internal firewall is the foundation for the journey, providing full topology visibility, advanced analytics, and simplified security architecture to help security teams through these five steps.

Step 1: Macro-segment the network

Most organizations know how to do rough segmentation of the network based on existing firewall configurations or documentation. Armed with this information, use the *Service-defined Firewall* to segment the network at a coarse level, isolating and securing zones (e.g., development, test, and production) from one another (see Figure 2). With no need to redesign the network or even make network address changes, this step simplifies the security architecture and accelerates time-to-value.

This simple step prevents attackers from moving between zones, limiting the damage they could potentially cause. It provides a flexible solution with the ability to quickly expand the number of zones as needed. Just as importantly, it allows the security team to show early progress in its zero trust journey.

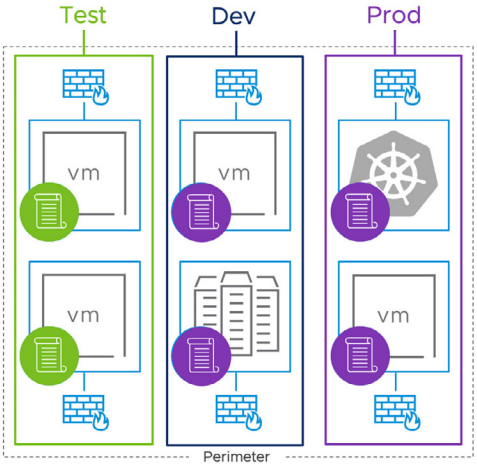


Figure 2: Macro-segmentation of the network

Step 2: Obtain topology visibility

The adage "You can't control what you can't see" is never truer than with east-west traffic. The Service-defined Firewall provides full topology visibility down to the workload level, enabling security teams to see the communication patterns and behavior of applications and their constituent workloads and microservices (see Figure 3). The comprehensive map of application topography gives data center-wide visibility on all applications. Built-in machine learning helps the team understand application behavior and traffic flows, making later steps easier to complete.

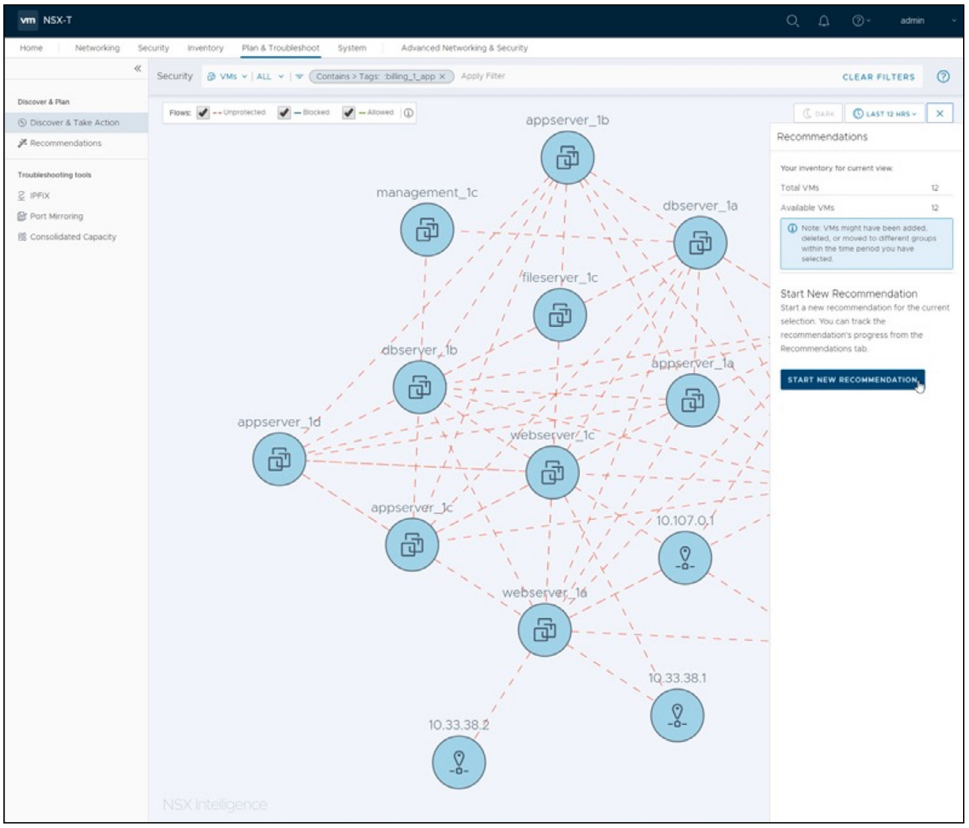


Figure 3: Network topology visibility

Step 3: Micro-segment a few well-known applications

Choose a small number of business-critical applications that are well-understood and well-documented as the first applications to be micro-segmented. *Micro-segmentation* isolates and protects them, applying layered security controls that are specific to the applications. It further prevents lateral movement into and out of the segment where the application runs (see Figure 4).

Many organizations start with their virtual desktop infrastructure (VDI) or other shared services like Active Directory or DNS servers. The VDI environment is often chosen as an early micro-segmentation target because it can expose the data center to threats that stem from end-user security violations. The Service-defined Firewall micro-segments VDI by isolating desktop zones from sensitive data center assets with just a few user-group specific security policies. It can also analyze behavioral patterns to make security policy recommendations, and its centralized management ensures policy consistency across the environment.

Step 3 begins reducing the attack surface by isolating critical applications from other data center assets, enabling user- and application-specific access controls, and mitigating lateral movement.

Step 4: Turn on threat control capabilities

Enhancing granular, workload-level security controls for critical applications is the next step. The Service-defined Firewall provides additional granular controls through its threat control capabilities, such as distributed intrusion detection/prevention (IDS/IPS). Because

this is available out of the box, there's no need to add, maintain, or manage another device on the network. The security team simply turns on the IDS/IPS functionality to provide enhanced granular security. Not only does the Service-defined Firewall IDS/IPS protect the organization by detecting traffic patterns that could indicate an attack, but it also helps with compliance, fulfilling the IDS/IPS requirements for regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and the Payment Card Industry Data Security Standard (PCI-DSS).

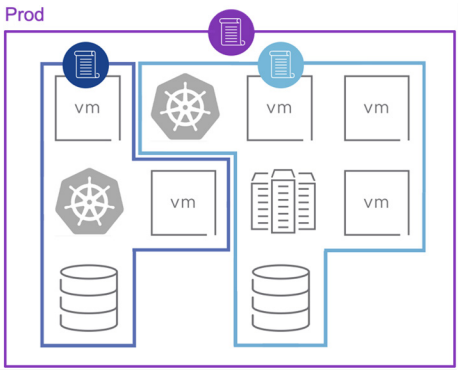


Figure 4: Micro-segmentation

Step 5: Micro-segment all applications to achieve zero trust in the data center

The team, armed with the experience gained in earlier steps, can now secure all critical applications in the data center. It's best to tackle related applications, such as those used by specific end-user groups: e.g., finance apps, then HR apps, then developer apps, etc. These applications are likely to have similar constraints and policies, so addressing them together should lead to operational efficiencies.

Here, the *Service-defined Firewall* can prove invaluable by providing visibility into east-west traffic and increasing the team's awareness of how applications behave. The Service-defined Firewall can automatically generate policy recommendations and deploy updated policies, constantly monitoring traffic flows to ensure policy compliance.

As the team extends the Service-defined Firewall to inspect all east-west traffic, it reduces blind spots, allowing it to detect and block lateral movement early and limit the damage.

Conclusion

Zero Trust initiatives are increasingly funded with the intent of protecting the data center by providing fine-grained control over the network, remote users, BYOD and cloud-based assets, and applications. For this approach to work, network security must happen at the workload level, isolating workloads from one another and securing them individually. A *distributed internal firewall* specifically built for zero trust is the best way to move quickly down the path to enhanced security.

Key things to look for when moving toward Zero Trust are easy segmentation of applications (to remove lateral movement of threats), reduced cost and complexity, and enhanced team efficiency. The *VMware Service-defined Firewall* is a perfect match for all three requirements:

- A comprehensive view of network topology and in-depth information about application behavior enables the team to quickly micro-segment even poorly understood applications
- Built-in IDS/IPS and automated network and security configurations keep costs low, removing the need for additional devices and even making it possible to retire some legacy tools
- Centralized management provides the team with single-pane-of-glass security management, which improves operational efficiency while effectively mitigating risk

Learn more about how distributed internal firewalls function and how organizations successfully deploy them for zero trust initiatives in our Internal *Firewalls for Dummies Guide*[5].

References

1. Zero Trust Architecture: SP 800-207. Retrieved from *https://csrc.nist.gov/publications/detail/sp/800-207/final*
2. How to Get from Here to Zero Trust. Forrester. Retrieved from *https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/products/nsx/vmware-zero-trust-spotlight.pdf*
3. The Four Barriers to Micro-segmentation. Retrieved from *https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/products/nsx/vmware-wp-four-barrs-micro-segmntatn-uslet-Final.pdf*
4. Five Critical Requirements for Firewalling in the Data Center. Retrieved from *https://www.vmware.com/learn/492966_REG.html*
5. Internal Firewalls for Dummies Guide. Retrieved from *https://www.vmware.com/content/microsites/learn/en/656351_REG.html*

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com