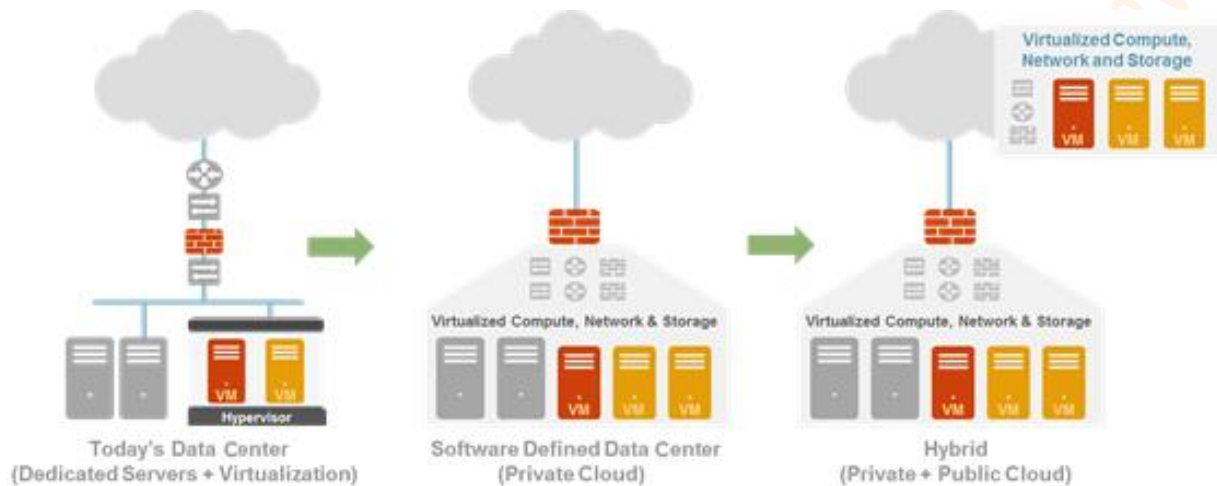


Hybrid Data Center Security

Data centers are rapidly evolving from a traditional, closed environment with static, hardware-based computing resources to an environment in which traditional and cloud computing technologies are mixed (see Figure 3-9).

Figure 3-9 Data centers are evolving to include a mix of hardware and cloud computing technologies.



The benefit of moving toward a cloud computing model – private, public, or hybrid – is that it improves operational efficiencies and lowers capital expenditures:

Optimizes existing hardware resources. Instead of using a “one server, one application” model, you can run multiple virtual applications on a single physical server, which means that organizations can leverage their existing hardware infrastructure by running more applications within the same system, provided that sufficient compute and memory resources exist on the system.

Reduces data center costs. Reduction of the server hardware “box” count not only reduces the physical infrastructure real estate but also reduces data center costs for power, cooling, and rack space, among others.

Increases operational flexibility. Through the dynamic nature of virtual machine (VM) provisioning, applications can be delivered more quickly than they can through the traditional method of purchasing them, “racking/stacking,” cabling, and so on. This operational flexibility helps improve the agility of the IT organization.

Maximizes efficiency of data center resources. Because applications can experience asynchronous or bursty demand loads, virtualization provides a more efficient way to address resource contention issues and maximize server use. It also provides a better way to address server maintenance and backup challenges. For example, IT staff can migrate VMs to other virtualized servers or hypervisors while performing hardware or software upgrades.

Traditional data security solution weaknesses

Traditional data center security solutions exhibit the same weaknesses found when they are deployed at a perimeter gateway on the physical network: They make their initial positive control network access decisions based on port, using stateful inspection, and then they make a series of sequential, negative control decisions using bolted-on feature sets. This approach has several problems:

Limited visibility and control. The “ports first” focus of traditional data security solutions limits their ability to see all traffic on all ports, which means that evasive or encrypted applications, and any corresponding threats that may or may not use standard ports can evade detection. For example, many data center applications (such as Microsoft Lync, Active Directory, and SharePoint) use a wide range of contiguous ports to function properly. You must therefore open all those ports first, exposing those same ports to other applications or cyberthreats.

No concept of unknown traffic. Unknown traffic is high risk but represents only a relatively small amount of traffic on every network. Unknown traffic can be a custom application, an unidentified commercial off-the-shelf application, or a threat. The common practice of blocking all unknown traffic may cripple your business. Allowing it all is highly risky. You need to be able to systematically manage unknown traffic using native policy management tools to reduce your organizational security risks.

Multiple policies, no policy reconciliation tools. Sequential traffic analysis (stateful inspection, application control, intrusion prevention system (IPS), anti-malware, etc.) in traditional data center security solutions requires a corresponding security policy or profile, often using multiple management tools. The result is that your security policies become convoluted as you build and manage a firewall policy with source, destination, user, port, and action; an application control policy with similar rules; and any other threat prevention rules required. Multiple security policies that mix positive (firewall) and negative (application control, IPS, and anti-malware) control models can cause security holes by missing traffic and/or not identifying the traffic. This situation is made worse when there are no policy reconciliation tools.

Cumbersome security policy update process. Existing security solutions in the data center do not address the dynamic nature of your cloud environment, because your policies have difficulty contending with the numerous dynamic changes that are common in virtual data centers. In a virtual data center, VM application servers often move from one physical host to another, so your security policies must adapt to changing network conditions.

Many cloud security offerings are merely virtualized versions of port- and protocol-based security appliances with the same inadequacies as their physical counterparts.

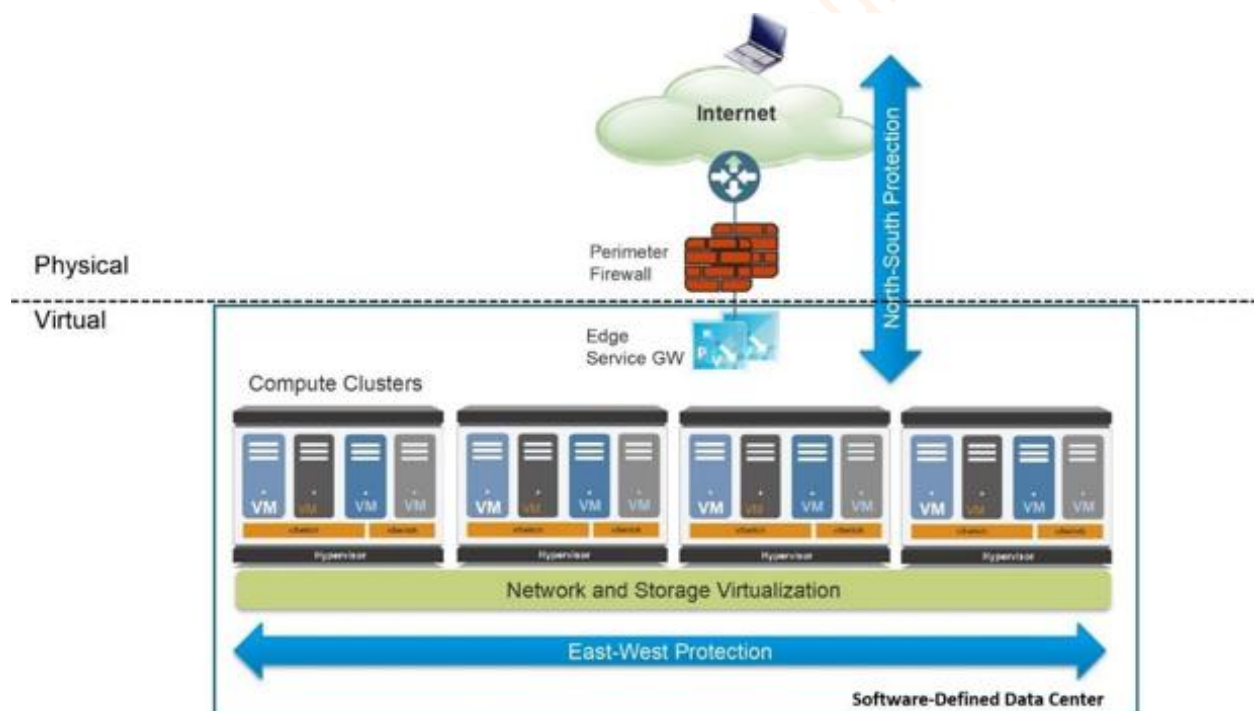
East-west traffic protection

In a virtual data center (private cloud), there are two different types of traffic, each of which is secured in a different manner (see Figure 3-10):

North-south refers to data packets that move in and out of the virtualized environment from the host network or a corresponding traditional data center. North-south traffic is secured by one or more physical form factor perimeter edge firewalls. The edge firewall is usually a high-throughput appliance working in high availability active/passive (or active/active) mode to increase resiliency. It controls all the traffic reaching into the data center and authorizes only allowed and “clean” packets to flow into the virtualized environment.

East-west refers to data packets moving between virtual workloads entirely within the private cloud. East-west traffic is protected by a local, virtualized firewall instantiated on each hypervisor. East-west firewalls are inserted transparently into the application infrastructure and do not necessitate a redesign of the logical topology.

Figure 3-10 Typical virtual data center design architecture

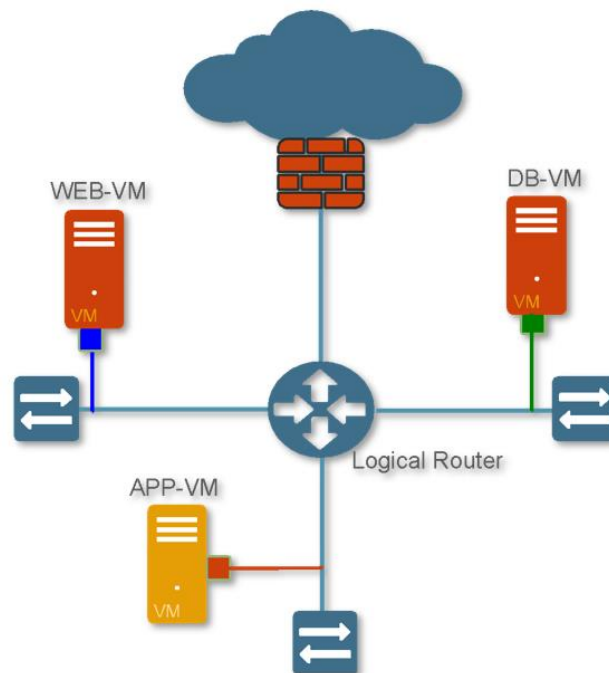


The compute cluster is the building block for hosting the application infrastructure and provides the necessary resources in terms of compute, storage, networking, and security. Compute clusters can be interconnected using OSI model Layer 2 (Data Link) or Layer 3 (Network) technologies such as virtual LAN (VLAN), virtual extensible LAN (VXLAN), or Internet Protocol (IP), thus providing a domain extension for workload capacity. Innovations in the virtualization space allow VMs to move freely in this private cloud while preserving compute, storage, networking, and security characteristics and postures.

Organizations usually implement security to protect traffic flowing north-south, but this approach is insufficient for protecting east-west traffic within a private cloud. To improve their security posture, enterprises must protect against threats across the entire network, both north-south and east-west.

One common practice in a private cloud is to isolate VMs into different tiers. Isolation provides clear delineation of application functions and allows a security team to easily implement security policies. Isolation is achieved using logical network attributes (such as a VLAN or a VXLAN) or logical software constructs (such as security groups). Figure 3-11 displays a simple three-tier application that is composed of a WEB-VM as the front end, an APP-VM as the application, and a DB-VM providing database services.

Figure 3-11 *Three-tier application hosted in a virtual data center*



An attacker has multiple options to steal data from the DB-VM. The first option is to initiate an SQL injection attack by sending HTTP requests containing normalized SQL commands that target an application vulnerability. The second option is to compromise the WEB-VM (using vulnerability exploits) and then move laterally to the APP-VM, initiating a brute-force attack to retrieve the SQL admin password.

After the DB-VM is compromised, the attacker can hide sensitive data extraction by using techniques such as DNS tunneling or by moving data across the network with NetBIOS and then off the network via FTP. In fact, attackers using applications commonly found on nearly every network have virtually unlimited options for stealing critical data in this environment. Infiltration into the environment and exfiltration of critical data can be completely transparent and undetected because the data is carried over legitimate protocols (such as HTTP and DNS) that are used for normal business activities.

Virtual data center security best practices dictate a combination of north-south and east-west protection. East-west protection provides the following benefits:

- Authorizes only allowed applications to flow inside the data center, between VMs
- Reduces lateral threat movement when a front-end workload has been compromised (the attacker breaches the front-end server by using a misconfigured application or unpatched exploit)
- Stops known and unknown threats that are sourced internally within the data center
- Protects against data theft by leveraging data and file filtering capability and blocking anti-spyware communications to the external world

An added benefit of using virtual firewalls for east-west protection is the unprecedented traffic and threat visibility that the virtualized security device can now provide. After traffic logs and threat logs are turned on, VM-to-VM communications and malicious attacks become visible. This virtual data center awareness allows security teams to optimize policies and enforce cyberthreat protection (for example, IPS, anti-malware, file blocking, data filtering, and DoS protection) where needed.

Security in hybrid data centers

The following approach to security in the evolving data center – from traditional three-tier architectures to virtual data centers and to the cloud – aligns with practical realities, such as the need to leverage existing best practices and technology investments, and the likelihood that most organizations will transform their data centers incrementally.

This approach consists of four phases:

Consolidating servers within trust levels. Organizations often consolidate servers within the same trust level into a single virtual computing environment: either one physical host or a cluster of physical hosts. Intra-host communications are generally minimal and inconsequential. As a matter of routine, most traffic is directed “off box” to users and systems residing at different trust levels. When intra-host communications do happen, the absence of protective safeguards between these virtualized systems is also consistent with the organization’s security posture for non-virtualized systems. Live migration features are typically used to enable transfer of VMs only to hosts supporting workloads within the same subnet. Security solutions should incorporate a robust virtual systems capability in which a single instance of the associated countermeasures can be partitioned into multiple logical instances, each with its own policy, management, and event domains. This virtual systems capability enables a single physical device to be used to simultaneously meet the unique requirements of multiple VMs or groups of VMs. Controlling and protecting inter-host traffic with physical network security appliances that are properly positioned and configured is the primary security focus.

Consolidating servers across trust levels. Workloads with different trust levels often coexist on the same physical host or cluster of physical hosts. Intra-host communications are limited, and live migration features are used to enable transfer of VMs only to hosts that are on the same subnet and that are configured identically with regard to routing of VM-to-VM traffic. Intra-host communication paths are intentionally not configured between VMs with different trust levels. Instead, all traffic is forced off box through a default gateway – such as a physical network security appliance – before it is allowed to proceed to the destination VM. Typically, this off-box routing can be accomplished by configuring separate virtual switches with separate physical network interface cards (NICs) for the VMs at each distinct trust level. As a best practice for virtualization, you should minimize the combination of workloads with different trust levels on the same server. Live migrations of VMs also should be restricted to servers supporting workloads within the same trust levels and within the same subnet. Over time, and in particular as workloads move to the cloud, maintenance of segmentation based on trust levels becomes more challenging.

Selective network security virtualization. Intra-host communications and live migrations are architected at this phase. All intra-host communication paths are strictly controlled to ensure that traffic between VMs at different trust levels is intermediated either by an on-box, virtual security appliance or by an off-box, physical security appliance. Long-distance live migrations (for example, between data centers) are enabled by a combination of native live migration features with external solutions that address associated networking and performance challenges. The intense processing requirements of solutions such as next-generation firewall virtual appliances will ensure that purpose-built physical appliances continue to play a key role in the virtual data center. However, virtual instances are ideally suited for scenarios where countermeasures need to migrate along with the workloads they control and protect.

Dynamic computing fabric. Conventional, static computing environments are transformed into dynamic fabrics (private or hybrid clouds) where underlying resources such as network devices, storage, and servers can be fluidly engaged in whatever combination best meets the needs of the organization at any given point in time. Intra-host communication and live migrations are unrestricted. This phase requires networking and security solutions that not only can be virtualized but are also virtualization-aware and can dynamically adjust as necessary to address communication and protection requirements, respectively. Classification, inspection, and control mechanisms in virtualization-aware security solutions must not be dependent on physical and fixed Network layer attributes. In general, higher-layer attributes such as application, user, and content identification are the basis not only for how countermeasures deliver protection but also for how they dynamically adjust to account for whatever combination of workloads and computing resources exist in their sphere of influence. Associated security management applications also need to be capable of orchestrating the activities of physical and virtual instances of countermeasures first with each other and then with other infrastructure components. This capability is necessary to ensure that adequate protection is optimally delivered in situations where workloads are frequently migrating across data center hosts.