WHITE PAPER

Tips and Tricks for Implementing Infrastructure Services on ESX Server

vmware®

foedus

Contents

Overview

VMware Infrastructure 3 is the industry's first full Infrastructure virtualization suite that empowers enterprises and small businesses alike to transform, manage and optimize their IT systems infrastructure through virtualization. VMware Infrastructure 3 delivers comprehensive virtualization, management, resource optimization, application availability and operational automation capabilities in an integrated offering. VMware Infrastructure 3 includes the following products:

- ESX Server
- VMFS
- Virtual SMP
- VirtualCenter
- VMotion
- DRS
- HA
- Consolidated Backup

The purpose of this paper is to provide advice on leveraging key features of VMware Infrastructure for deployment of infrastructure services for system administrators who are new to VMware Virtual Infrastructure. This paper is intended for system administrators who are new to VMware Infrastructure. The recommendations are applicable to small or mid size companies. Recommendations and examples will illustrate what can be accomplished using the VMware Infrastructure platform.

When new and disruptive technology such as the VMware Infrastructure enters an IT environment, perhaps the greatest challenge faced by IT staff is learning to think in new ways about server and workload management. People are generally risk averse and this can lead to reactive thinking when some hurdle or problem arises within the Virtual Infrastructure. Obtaining the greatest reward from Virtual Infrastructure largely depends on a commitment to virtualization as an architectural principle. Resolving problems within the context of tools provided by the virtual infrastructure will help getting the most out of virtual infrastructure. Too often system administrators conclude prematurely that a particular server or application cannot be virtualized. This can lead to deployments that do not provide the full return on investment that is possible leveraging the power of VMware Infrastructure.

This paper will assist the system administrator facing this new and potentially daunting technology by providing new ways to think, plan and execute on the virtual platform. This paper will provide an understanding of virtual workloads together with practical tips and tricks regarding the use of VMware Infrastructure.

Running Infrastructure Services on VMware Infrastructure

Virtualization's Impact on the Budgeting Process

Server virtualization is a new paradigm for server management. Physical servers are often purchased individually and tied to a unique budget line item that must be justified on its own

merits. Virtual servers are essentially purchased in bulk based on capacity planning considerations. This opens opportunities for IT to experiment with new solutions using excess capacity on the ESX Server platform without obtaining specific budgetary justification prior to the exploratory process. This "under the radar" solution development can unleash the creativity of the IT staff, facilitating the testing and validating of solutions that may never have otherwise seen the light of day.

Unleashing Your Creative Potential

The ability to deploy a server in minutes using virtual machine templates offers a low risk way to validate products before presenting lengthy proposals to management. The typical budgetary processes within mid to larger companies forces IT staff to submit proposals that are the most likely to meet with financial approval. This limits the discovery process and often prevents IT staff from pursuing solutions, which, though viable and low cost, cannot be validated without access to servers on which to test them. With access to virtual infrastructure, these servers now stand ready just a few clicks away. Whether validating new solutions, furthering education and skills, or creating a world class virtual lab, virtualization technology will changes the way IT is managed. IT can now provide an increased range of services and unleash the full creative potential of the system-administrator team.

Put on your Architect's Hat

The gradual accumulation of physical servers in the typical IT environment can result in the lack of data-center wide capacity planning and less than optimal resource management. Physical servers silo applications and minimize interactions between the CPU resources of different applications. Virtualization brings a new level of interaction by pooling all hardware resources including CPUs. This opens up new opportunities for efficient use of hardware but brings new tuning challenges. Investing in virtual infrastructure is an architecture decision. It implies a commitment to finding ways of managing resources that are compatible with the virtual framework whenever possible. In order to maintain any architecture, changes and sometimes sacrifices must be made at the level of individual components for the benefit and integrity of the whole. To see the greatest return on investment for virtual infrastructure, a system administrator must acknowledge and become comfortable with the necessity of adopting designs and processes that are optimal at an architecture level even if they sometimes go against accepted wisdom in the world of physical server management. Examples of such instances will be provided throughout this document.

Choose to Virtualize Services

Organizations that make a commitment to virtual infrastructure eventually reach a decision to virtualize services whenever possible, and deploy physical servers only by specific justification. Many infrastructure services today are provided by dedicated hardware appliances, dedicated servers, or other isolated solutions. Offering as many of these services from a common platform reduces the number of management processes necessary to maintain the environment. Offering the majority of IT services from a virtual platform creates the opportunity to form unified solutions and strategies to common problems such as backup, disaster recovery, business continuity, upgrade management, resource estimation, and many more. Start with a bias towards running solutions on the virtual platform. Learning to favor solutions that meet the business need and are compatible with the virtual platform is important in achieving the fullest benefits of the technology.

Optimizing Virtual Workloads

Workloads in the Physical World

When considering deploying infrastructure services, it is important to have a clear understanding of some basic differences between physical and virtual workloads. In the world of physical servers, workloads exist in isolated physical silos. Workloads are optimal when they maximize utilization of available resources within these silos. This encourages applications to take as much of the CPU, memory, disk and network resources as possible since there is an assumption that other applications will similarly be contained in their own silos. A physical file server with a single CPU running at an average of 40% is considered a better use of resources than a server with two CPUs where each runs at 20% on average. This reflects the bias towards vertical workloads that is common in the world of physical server provisioning. By *vertical* we mean maximum utilization within a resource container such as CPU, disk and memory of a given server or cluster.

Workloads within Virtual Infrastructure

Virtual infrastructure lends itself to the opposite paradigm. In general, large workloads should be spread out *horizontally* across multiple virtual CPUs where possible so as to maximize VMware Infrastructure's ability to balance the entire platform. The reason for this way of distributing workloads goes to the heart of virtualization technology. The ESX Server scheduler is always scanning across the available physical CPUs on an ESX Server host looking for CPUs with available headroom to service additional virtual CPUs that are requesting time. In a busy ESX Server, smaller chunks of CPU loan have more places they can be placed than large demanding chunks. A smaller CPU demand can more easily added to physical CPUs without causing other virtual machines on that same CPU to be rescheduled elsewhere and without other virtual machines waiting longer than previous for CPU cycles. Consider a scenario where a physical CPU servicing four virtual machines each requesting 20% of its time. The assigned CPUs for these virtual machines is able to give all Virtual machines the cycles they seek and still have 20% of its time to spare. Adding a fifth VM to this CPU that requests 15% CPU time will be possible with minimal impact on the other Virtual machines. Adding a fifth VM requiring 35% CPU time will cause the CPU to be overcommitted by at least 15%. The scheduler will then determine whether to move one of the five Virtual machines to another physical CPU or, on a busy ESX Server host, to distribute the 15% overbooking across all five Virtual machines by scheduling fewer CPU cycles. This may cause virtual machine performance to slow on this CPU. While VMware Infrastructure can certainly handle heavy loads with minimal CPU overhead, breaking workloads into smaller chunks will give the scheduler the most flexibility in keeping the platform balanced and responsive. An Example of redistributing workloads into smaller chunks might include breaking one very busy file server into two or more file server, each targeting a different set of content.

The Benefits of CPU Dense ESX Server Hosts

Extending this principle, ESX Server installations with a greater number of physical CPUs offer a greater chance of servicing competing workloads optimally. The chance that the scheduler can find room for a particular workload without much reshuffling of virtual machines will always be better when the scheduler has more CPUs across which it can search for idle time. For this reason, it will generally be better to purchase two four-way ESX Server licenses than to purchase four two-way machines. Similarly, two eight-way servers will provide more scheduling flexibility than four four-way servers. Cost considerations will obviously weigh into such decisions but in thinking about optimizing virtual workloads, focus on:

- creating lighter workloads, breaking up large CPU demands into several smaller chunks when possible.
- distributing your workloads across the greatest number of physical CPUs.

This gives the ESX Server scheduler the greatest flexibility in servicing the workloads quickly and efficiently.

Virtual SMP and CPU Dense ESX Server Hosts

The positive effect of additional CPUs in ESX Server hosts is dramatically illustrated when scheduling virtual machines with multiple virtual CPUs (VCPUs). With the release of ESX Server 3.0, four-way SMP(symmetric multiprocessing) virtual machines are now supported. In order for the scheduler to allocate time to an SMP virtual machine, it must be able to find available CPU resources for each VCPU at the same time. If a virtual machine with two CPU's wants 40% for each of its VCPUs, the scheduler will try to locate two physical CPUs with adequate headroom at the same time. Depending on the CPU shares assigned, the scheduler can schedule virtual machines on CPUs without adequate headroom provided that other virtual machines on these CPUs can be de-prioritized. For the purpose of this discussion, we will assume for a moment that all virtual machines have equal shares.

The Math behind Virtual SMP Scheduling

Scheduling a two-VCPU machine on a two-way physical ESX Server hosts provides only one possible allocation for scheduling the virtual machine. The number of possible scheduling opportunities for a two-VCPU machine on a four-way or eight-way physical ESX Server host is described by combinatorial mathematics using the formula N! / (R!(N-R)!) where N=the number of physical CPUs on the ESX Server host and R=the number of VCPUs on the machine being scheduled.[1] A two-VCPU virtual machine running on a four-way ESX Server host provides (4! / (2! (4-2)!) which is (4*3*2 / (2*2)) or **6** scheduling possibilities. For those unfamiliar with combinatory mathematics, X! is calculated as X(X-1)(X-2)(X-3)…. (X- (X-1)). For example 5! = 5*4*3*2*1.

Using these calculations, a two-VCPU virtual machine on an eight-way ESX Server host has (8! / (2! (8-2)!) which is (40320 / (2*720)) or **28** scheduling possibilities. This is more than four times the possibilities a four-way ESX Server host can provide. Four-vCPU machines demonstrate this principle even more forcefully. A four-vCPU machine scheduled on a four-way physical ESX Server host provides only one possibility to the scheduler whereas a four-VCPU virtual machine on an eight-CPU ESX Server host will yield (8! / (4!(8-4)!) or **70** scheduling possibilities, but running a four-vCPU machine on a sixteen-way ESX Server host will yield (16! / (4!(16-4)!) which is (20922789888000 / (24*479001600) or **1820** scheduling possibilities. That means that the scheduler has **1820** unique ways in which it can place the four-vCPU workload on the ESX Server host. Doubling the physical CPU count from eight to sixteen results in **26** times the scheduling flexibility for the four-way virtual machines. Running a four-way virtual machine on a Host with four times the number of physical processors (16-way ESX Server host) provides over **six** times more flexibility than we saw with running a two-way VM on a Host with four times the number of physical processors (8-way ESX Server host).

[1] See *http://en.wikipedia.org/wiki/Combinatorics* for additional information on the history and theory behind Cominatorics.

These examples demonstrate the exponential benefits of CPU-dense ESX Server hosts in cases where ESX Server 3.0 four-way SMP capabilities will be utilized. The larger scheduling search space provided by CPU-dense hosts gives the scheduler the best chance of finding the right combination of CPUs to service complex workloads immediately and with minimal impact to other virtual machines. Using CPU-dense host machines together with distributing virtual workloads horizontally across multiple vCPUs ensures optimal performance and utilization of the virtual platform.

Distributed Resource Scheduler

With the introduction of ESX Server 3.0, VMware Infrastructure 3, VMware has introduced VMware DRS (Distributed Resource Scheduler). DRS dynamically allocates and balances computing capacity across collections of hardware resources aggregated into logical resource pools. VMware DRS continuously monitors utilization across resource pools and intelligently allocates available resources among virtual machines. VMware DRS aligns computing resources with business goals while ensuring flexibility and efficient utilization of hardware resources. Should a particular ESX Server host become overloaded with CPU, memory or disk activity, designated virtual machines can be moved to ESX Server hosts better able to handle the load. This provides another tool that will assist in a variety of ways but especially with servicing heavier workloads. Virtual workloads that do not readily lend themselves to being split across multiple VCPUs can be serviced without compromising the lighter workloads by moving them to a designated "Heavy Load" ESX Server host or groups of hosts. These ESX Server hosts will have a lower ratio of Virtual machines to physical CPUs.

Migrating Infrastructure Workloads to VMware Infrastructure

As workloads are migrated from physical servers onto the virtual platform, system-administrators have the responsibility of making workloads designed for the physical world function within the virtual world. Too often decisions about an application's inability to be virtualized are made without adequate analysis of ways in which the workload can be redistributed horizontally. Tools that examine whether a physical server can be migrated to a virtual machine may miss the mark by limiting analysis to the constraints of the source server. If physical source servers are heavily loaded or demand high CPU loads once converted, evaluate how the source workload can be broken up into logical chunks and distributed across the virtual platform in new configurations. Physical to virtual(P2V) tools such as VMware's P2V Assistant should be engaged to minimize the time involved in moving infrastructure services to VMware Infrastructure. P2V tools are a quick and convenient way to avoid the many manual steps of rebuilding a server. After P2V migrations, allocate infrastructure virtual machines to groups of hardware resources known as resource pools based on common performance expectations or business unit groupings. Infrastructure services can aggressively be migrated to VMware Infrastructure provided that system administrators are conscious of dividing workloads into smaller chunks and fully leverage the new capabilities and products such as resource pools, DRS, and VirtualSMP.

Setting up IP and Directory Services

Startup and Shutdown Sequence

Virtual Machines can readily be used to provide basic IP and directory services on the network. Providing IP services such as DNS and DHCP are critical to other functions on the network. These services are typically light to moderate workloads and should be placed on virtual machines that

boot up first when an ESX Server host comes online. Starting virtual machines that rely on DNS or Active Directory before these servers are operational will cause many problems. ESX Server provides for a timer between when a virtual machine comes online and when ESX Server starts the next virtual machine in the startup sequence. Use this feature to ensure that IP services are fully functional before starting other virtual machines. Use this principle in reverse for power shutdown events. Set the shutdown sequence so that IP services go down last.

Plan for Redundancy from the Beginning

For those companies with Microsoft centric networks, consider setting up two virtual machines each configured for DHCP, DNS, and Active Directory. Make sure that DHCP is only active on one virtual machine per VLAN. Each virtual machine proving IP services should ideally run on separate ESX Server hosts. If there is a problem with one ESX Server host, Active Directory and DNS services will stay online. Running two virtual machines each with DHCP, DNS and Active Directory should be adequate for companies of up to 500 people.

High Availability with VMware HA

Since there can be only one DHCP device per VLAN, a problem on the ESX Server host running your virtual machine with DHCP enabled could cause DHCP to become unavailable on the network. To counteract this and similar predicaments, VMware HA was introduced with VMware Infrastructure 3 and VirtualCenter 2. VMware HA provides easy to use, cost effective high availability for applications running in virtual machines. VMware HA minimizes downtime and IT service disruption while eliminating the need for dedicated stand-by hardware. In this example, HA would be leveraged to automatically restart the virtual machine providing DHCP on another functioning ESX Server host. This will minimize interruption of services for DHCP and other IP services you wish to restart immediately in the event of an ESX Server host failure.

Hardware Maintenance

In the event that hardware maintenance is required, leverage VMware VMotion™ to move the running virtual machines to another server. VMotion enables the live migration of running virtual machines from one physical server to another with zero downtime, continuous service availability, and complete transaction integrity. Use VMotion to move Active Directory virtual machines within production infrastructure to other ESX Server installations so that Active Directory replication can continue uninterrupted. Similarly, VMotion DHCP and DNS if you have installed them on separate Virtual machines. Ensuring the replication of data between Active Directory and DNS machines lowers the risk of an anomaly occurring when you bring your ESX Server host back online. When maintenance is complete, redistribute the load appropriately using VMotion, and repeat the process as necessary.

Scaling IP Services

Scaling up to larger designs, separate the Active Directory workloads from the DNS and DHCP workloads so each workload can be tuned individually. User environments with many Active Directory driven policies will want to ensure that the ESX Server %Ready time as viewed in the ESXTOP utility stays under 10. Active Directory machines that are not responsive enough can

slow the user login process. VMware Infrastructure 3 introduces the concept of Resource Pools. A resource pool is a collection of hardware resources including processor, memory, disk and networking that is aggregated by VMware Infrastructure into a unified logical resource that can be allocated to virtual machines on-demand. Resource pools abstract the underlying heterogeneous hardware and present uniform resources to virtual machines. Virtual machines in a resource pool become agnostic on which particular physical server they are running at any given point in time. As a result, available resources can be dynamically and intelligently allocated among virtual machines based on pre-defined rules that reflect business needs and changing priorities. In larger installations of Active Directory consider placing all the Active Directory servers into the same resource pool and controlling their access to VMware Infrastructure resources through a resource pool policy.

Time Synch Issues on ESX Server

Virtualization introduces new issues in the area of timekeeping. Because virtual machines work by time-sharing host physical hardware, a virtual machine cannot exactly duplicate the timing behavior of a physical machine. VMware virtual machines use several techniques to minimize and conceal differences in timing behavior, but the differences can still sometimes cause timekeeping inaccuracies and other problems in guest software. When virtual machines are not scheduled for execution on a physical CPU, the guest OS clock is suspended, and to the exterior world, the clock appears to drift. Over time, this can cause problems with Active Directory servers that require close clock synchronization. Best practices around time keeping involve setting ESX Server hosts to synchronize with external time servers using NTP protocol. This will keep ESX Server hosts in synch with each other and ensure that VMotion operations do not cause sudden time disruptions on virtual machines that synchronize with the ESX Server host clock. Next, be sure to select the checkbox in VMTools to synchronize the guest OS clock with the underlying ESX Server clock. This ensures timely correction of any time drift. For a more thorough exploration of this important topic, see VMware's whitepaper titled "Timekeeping in Virtual Machines"[2]

Migrating File and Print Services to VMware Infrastructure

Accessing Workloads

File servers can pose special tuning challenges in a virtual environment. Physical file servers often combine content serving different groups and business functions with widely differing performance requirements attached to each workload stream. A physical file server with a high average usage may initially be considered a poor candidate for virtualization. A good approach to take in such cases is to divide out the separate workloads and place them onto separate smaller file servers. This allows the ESX Server scheduler to spread the workloads over a larger number of virtual processors.

[2] "Timekeeping in Virtual Machines" can be downloaded from the Technical Papers sections in the documentation category of www.vmware.com or directly at http://www.vmware.com/pdf/vmware_timekeeping.pdf

Dividing Workloads from Physical Servers

By way of example, imagine a company converting to virtual infrastructure that is struggling to provide the right tuning parameters for their demanding file server requirements. Their physical server holds thousands of small files that are frequently searched and accessed by a business critical application. Slow access to these files will cause the application to freeze the user's application screen. The same server hosts many customer files in the form of Adobe PDF and Microsoft Office files. These are larger but less frequently accessed. The third data stream consists of all the end user "My Documents" and redirected Desktop files that resided on the same file server. Lastly, almost 600 Gigabytes of infrequently accessed scanned documents are hosted on the same file server.

Post Migration Problems

In this example, all data is initially migrated to a single virtual machine. After initially poor performance, resource shares on the virtual machine are increased to a very high level. The server still performs poorly. Adding a 2nd virtual CPU helps somewhat but the server still appears so busy that the ESX Server scheduler seldom views it as idle and the overall ESX Server host usage level stayed artificially high. Eventually the company divides this file server into three separate single virtual CPU file servers. All the line of business content is placed on one server (Server A), the user's personal and desktop docs were placed on another (Server B), and the customer files and scan archive placed on a third (Server C).

Segmentation of Workloads across Virtual Machines

Through experiments and monitoring with the esxtop utility, the company determines that their application performs well as long as Server A's %Ready metric in esxtop stays under 5. Resource shares are increased to a level where this low %Ready is obtained. Server B holds the documents that users are frequently working on in Microsoft Office applications. Through experiments, it is determined that Excel and Word load documents in under a second if the %Ready stays less than 10, while document load times could take 2-4 seconds if %Ready exceeds 10 and as much as 6 seconds if it goes over 20. Server B is tuned to load docs in about a second and Server C is tuned less aggressively to load documents in the 2-4 second range. After tuning these workloads separately, the company's ESX Server host usage drops significantly and the schedule correctly reports idle time for each file server in the esxtop utility. In this example, each workload is tuned appropriately to its business purpose resulting in an improved overall ESX Server host tuning and improved responsiveness for end-users.

Hosting Print Services on VMware Infrastructure

Print servers can be very busy machines with high loads at times but in most cases they do not require the highest levels of performance. By the time users walk over to the network printer, the print job will be there waiting for them. Print servers should receive a smaller number of resource shares. Let print servers take cycles when higher priority machines do not need them.

Just as file servers can be divided into logical work streams and placed on separate virtual machines, so too print servers can be segmented into printer groupings to bring down their

average CPU utilization to something under 25%. Provisioning servers takes little time and making the effort to create lighter workloads for VMware Infrastructure will help the whole virtual environment run better.

Print servers can generate a lot of bulky network traffic that can potentially degrade network performance of applications such as Citrix which are very latency sensitive. If you are mixing these types of network traffic on the same ESX Server host, consider putting these two classes of applications on separate Virtual Switches that are tied to different physical NICS. This will assist in reducing network congestion for the latency sensitive applications.

File & Print Summary

When migrating file servers, divide the workloads into logical groupings, and create one or more file servers for each workload, tuning each separately. Do not think in terms of whether a server can be migrated but rather what type of performance expectations should be tied to each workload stream.

Proxy and Firewall Services

Setting up a Virtual Proxy Server

Another area where virtual infrastructure can add value to an organization is in the security space. Smaller companies may not see the need or be able to justify expenditures on dedicated proxy servers for example. Proxy servers can provide significant performance acceleration, often up to 25% or more reduction in bandwidth utilization and much faster retrieval of frequently used websites. In addition, they provide the ability to block undesirable websites and track usage of Internet sites.

By deploying an open source product such as IPCOP inside a virtual machine, companies have a quick and compact solution that provides immediate benefit. IPCOP and other similar virtual appliances offer versions of Squid Proxy, which is mature high performance proxy code that performs very well in a compact framework. IPCOP is but one example of readily available open source platforms that can be readily transformed into virtual appliances offering valuable infrastructure services without introducing additional stand alone devices to the environment.

Hosting a Virtual DMZ

IPCOP can also be used to host an entire DMZ on VMware Infrastructure. If making web servers, email servers or any Internet exposed server available on the Internet from your corporate site is on the agenda, a DMZ should be created. A DMZ (Demilitarized Zone), takes the military metaphor and creates firewall containment around the web servers serving the outside world. In deploying a DMZ, it is assumed that web servers might be compromised by an unknown exploit. It is the function of the DMZ firewall to prevent these servers from being used as springboards into the vulnerable LAN environment by only allowing predefined network communication between them and other more sensitive servers.

IPCOP can be obtained through IPCOP.org or the VMware Virtual appliance library[3]. A growing company in New Jersey recently used IPCOP to host a busy DMZ with over ten servers and to provide proxy services to a user community of over 300 people. IPCOP runs on a hardened Linux OS and only takes about 256MB of Ram and 2Gigs of disk space. Proxy and DMZ workloads seldom pushed the IPCOP firewall past 15% of a single VCPU. IPCOP has a red, green and orange NIC. The red NIC is meant to sit on the public Internet, the green NIC will be assigned an address on the LAN, and the orange NIC is used for the DMZ address space… usually 10.x.x.x.(See Figure 1 below)

Setting up DMZ Virtual Switches

To setup up IPCOP as a DMZ firewall, create two virtual switches on one or more ESX Server hosts named DMZ-EXT and DMZ-INT. Plug the red NIC of IPCOP into DMZ-EXT and the orange NIC into DMZ-INT. Plug the green NIC into whatever virtual switch is associated with the internal LAN address space. DMZ-INT will function as the switch servicing DMZ virtual machines. DMZ-EXT will be used to send packets out to the WAN router or perimeter firewall. Internal servers will communicate with servers in the virtual DMZ via the green NIC, so add a route for sending data into the DMZ on the LAN router. The route will point to the gateway address of the DMZ firewall's green NIC in order to reach the DMZ subnet. If you already have hardware based perimeter firewalls, assign a transit network to the red NIC that is dedicated to sending packets between your DMZ-EXT's virtual switch and a dedicated port on the WAN router or perimeter firewall. Plug the DMZ-EXT's associated physical NIC on each ESX Server host into a common physical switch or VLAN segment in order to logically isolate DMZ traffic from other network segments.

While the DMZ-INT virtual switch could be configured as an isolated switch on installation with only one ESX Server host, it is better to associate this virtual switch with a physical NIC on each ESX Server in cases where the virtual platform has multiple ESX Server hosts. VMotion requires servers to be plugged into virtual switches with associated physical NICS. VMotion of VM's with multiple NICS such as IPCOP requires that all NICS are plugged into virtual switches with associated physical NICS. Configured properly, all component virtual machines and the DMZ firewall itself will be able to move to other ESX Server hosts during a VMotion without disrupting DMZ security or proxy services. Installations with only two ESX Server hosts servicing the DMZ can simply connect the DMZ-INT switch's associated physical NIC on each ESX Server with a crossover cable. A dedicated mini-switch or VLAN can connect more than two ESX Server hosts so that DMZ elements can be moved around as needed among a group of ESX Server hosts.

Integrating the Virtual DMZ with Upstream Firewalls and Routers

Create one or more mapped IP addresses on the perimeter firewall and map them through to the IP addresses assigned to IPCOP's red NIC. In cases where a WAN router sits behind one or more perimeter firewalls, pass the packets from the mapped IPs to the WAN router and set static routes in the WAN router that send packets along to the DMZ via IPCOP's red NIC. It is good practice to

[3] *http://www.vmwarez.com/2006/01/ipcop-virtual-machine-new-esx-version.html* to download or see VMware site at *http://www.vmware.com/vmtn/appliances/directory/9*

dedicate a switch or VLAN for transport of data between the WAN router port and the physical NIC on each ESX Server associated with DMZ-EXT.

The Benefits of a Virtual DMZ

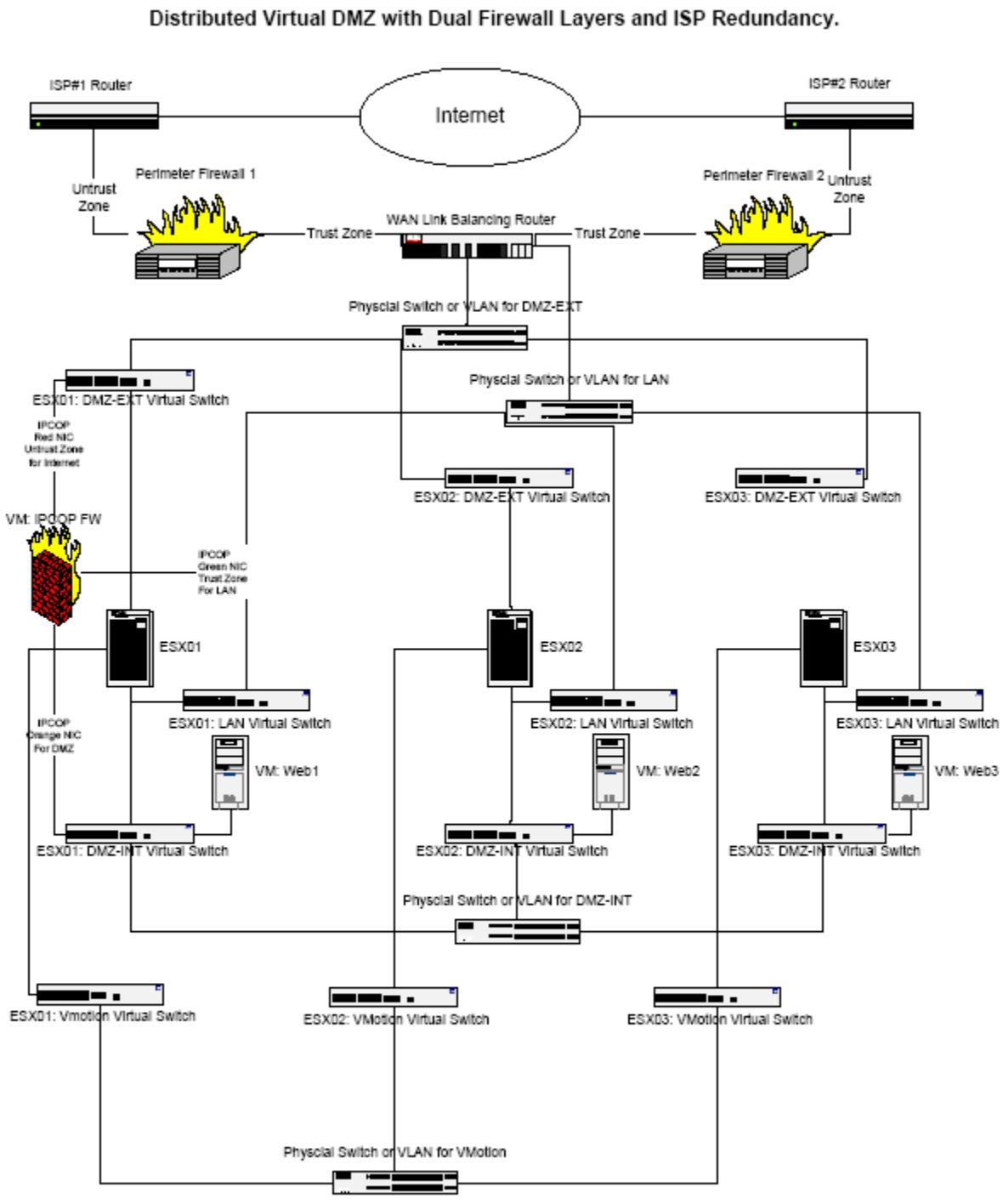
The benefits of creating a virtual DMZ include amazing flexibility to move all components of the DMZ including the DMZ firewall itself to different ESX Server hosts during maintenance windows. Imagine a DMZ with IPCOP as the controlling firewall and ten virtual web servers in the DMZ. In this example, as shown in Figure 1 below, all DMZ related virtual machines including the firewall and DMZ member servers are hosted on an eight-way server named ESX01. In order to perform a BIOS upgrade on ESX01 without disrupting service, VMotion can be used to move these business critical servers to other ESX Server machines. However sister hosts ESX02 and ESX03 are four-way ESX Server hosts and only have spare capacity for approximately six or seven extra virtual machines each. With the virtual DMZ depicted in figure 1, the IPCOP firewall and DMZ web servers can be moved to ESX02, the remaining five DMZ web servers can be moved to ESX03. This reshuffling of virtual machines is completely transparent to the Internet clients using DMZ web servers, takes just a few minutes, and maintains the identical security configuration as when the whole DMZ is hosted on ESX01. The two firewall layers between the public internet and the DMZ member servers help protect against exploits that might capitalize on known vulnerabilities of a particular firewall platform.

This type of virtual DMZ configuration was deployed with great success at a growing mortgage company in the northeast. The flexibility helped in achieving almost two years of uptime for DMZ servers. Results like these are more difficult and expensive to achieve in a physical DMZ design where various components cannot be taken offline without creating a service outage. The added features of proxy acceleration and Internet tracking helped this company contain bandwidth expenditures without spending more for dedicated physical servers. Using virtual infrastructure to host your infrastructure services will provide a completely new level of flexibility and sophistication at a modest price.

Conclusion

This paper has emphasized some of the changed thinking that is required in deploying and managing infrastructure services on the VMware Infrastructure. Methods for creating and distributing workloads across one or more lightly loaded virtual CPUs have been given together with advice on combining virtual SMP with single virtual CPU workloads. Try to keep workloads light, distribute them across appropriately sized ESX Server hosts and leverage DRS to balance workloads across ESX Server hosts as necessary. Plan for virtual machine redundancy across ESX Server hosts and use VMware HA to ensure continuous availability of services. Hosting infrastructure services on the ESX Server platform offers new levels of flexibility, reduced downtime, and the benefit of unifying the way IT resources are managed. Making the tradeoffs necessary to host infrastructure services together with other types of servers on a common virtual platform will pay strategic and financial dividends in reduced complexity and improved manageability of the IT environment.

Figure 1


Distributed Virtual DMZ with Dual Firewall Layers and ISP Redundancy.
ISP#1 Router
Internet
ISP#2 Router
Untrust Zone
Perimeter Firewall 1
WAN Link Balancing Router
Trust Zone
Trust Zone
Perimeter Firewall 2
Untrust Zone
Physcial Switch or VLAN for DMZ-EXT
Physcial Switch or VLAN for LAN
ESX01: DMZ-EXT Virtual Switch
IPCOP Red NIC Untrust Zone for Internet
ESX02: DMZ-EXT Virtual Switch
ESX03: DMZ-EXT Virtual Switch
VM: IPCOP FW
IPCOP Green NIC Trust Zone For LAN
ESX01
ESX02
ESX03
ESX01: LAN Virtual Switch
ESX02: LAN Virtual Switch
ESX03: LAN Virtual Switch
IPCOP Orange NIC For DMZ
VM: Web1
VM: Web2
VM: Web3
ESX01: DMZ-INT Virtual Switch
ESX02: DMZ-INT Virtual Switch
ESX03: DMZ-INT Virtual Switch
Physcial Switch or VLAN for DMZ-INT
ESX01: Vmotion Virtual Switch
ESX02: VMotion Virtual Switch
ESX03: VMotion Virtual Switch
Physcial Switch or VLAN for VMotion

About the Author

Daniel Beveridge is a Senior Systems Engineer with Foedus, a leading Virtual Infrastructure Services company in the Northeast. Daniel has three years experience managing a mid-size company's entire server infrastructure on ESX Server, and prior to that, fifteen years experience in project and relationship management, extensive system and software architecture design, and implementation of infrastructure solutions.

VMware, Inc. 3145 Porter Drive Palo Alto CA 94304 USA Tel 650-475-5000 Fax 650-475-5001 www.vmware.com

vmware®

foedus