

BEST PRACTICES

Deploying Microsoft Exchange in VMware Infrastructure



Table of Contents

Introduction 3

Exchange on VMware Infrastructure 4

Exchange Best Practices 4

Server Hosts..... 4

Virtual Machines 6

Storage..... 7

Implementation 9

Case Study 11

Appendix 14

Deploying Microsoft Exchange in VMware Infrastructure

Introduction

Microsoft Exchange Server is the most widely used messaging platform in the world. It is used for approximately 31 percent of all corporate mailboxes as of March 2006, according to a study by the Radicati Group (http://download.microsoft.com/download/e/8/a/e8a154bf-cc35-4340-bd26-6265cdb06b6e/market_share_March06.pdf). Due to the importance of messaging within organizations, architects strive to design messaging solutions that are highly available and quickly recoverable. With Exchange, there are many methods of achieving these design goals. Unfortunately, most of these methods are complicated and expensive.

At the same time, many organizations have invested heavily in new virtual infrastructures based on VMware Infrastructure technologies. This platform offers fairly simple and straightforward mechanisms for the messaging architect to provide higher availability and far easier recovery while actually saving money, rather than spending significantly on more traditional solutions. VMware Infrastructure offers these and many other advantages including:

- Higher availability — VMware HA has been shown to work with Exchange Server. With HA, if a host fails, the host's virtual machines are brought back online by another host in the cluster. In most cases (and in all of our testing), the Exchange server is able to recover and comes back online very quickly.
- Simplified local recovery — When you use ESX Server, what was previously a physical server becomes a set of files. These files can be backed up and restored in order to recover the server if a host fails or a system becomes corrupted. The backup is often performed via a SAN-based copy or snap technology that allows for very quick recovery of the server on the same or alternate hardware.
- Facilitated disaster recovery — Disaster recovery of Exchange has always been expensive and difficult. Most organizations that have implemented disaster recovery have kept hot recovery servers in the disaster recovery site and used the Exchange disaster recovery installation options to recover the Exchange systems, followed by tape restores for the data. Cutting-edge organizations use integrated SAN technologies to create SAN snapshots or copy backups of the data that can be restored quickly. Disaster recovery of systems using ESX Server snapshots helps accelerate system recovery, allowing you to begin data recovery much sooner.

- More efficient use of resources — It is common for architects to design systems with redundant components, not because the load requires multiple servers, but in order to provide high availability. Exchange services that benefit from redundant servers include bridgeheads, gateways, and Outlook Web Access servers. VMware Infrastructure provides a much more cost-effective means to provide this redundancy.

An important point is that Exchange, although it is more complicated, should be treated just as you would any other enterprise application: the decision to virtualize should be based on an analysis of performance data, not a preconceived expectation of performance. In many organizations, Exchange running in a virtual environment meets those performance tests.

The rest of this paper will help you determine the best design for Exchange Server virtual machines in your environment. It discusses best practices for running Exchange 2003 in a VMware Infrastructure environment, based upon experience that consulting firm RapidApp has gained while working with its customers and on the firm's internal testing. Some of the recommendations presented here might apply to other versions of Exchange. However, if you are using a different version, you should verify each item for your configuration. Note that Exchange 2007 has a significantly different architecture and design, so specific recommendations made in this paper must be evaluated with particular care. However, you can apply the general principles of this analysis to your environment.

Although this paper discusses techniques for determining the resources you should provision for a virtual Exchange infrastructure, it does not make specific sizing recommendations. Provisioning information is addressed in more detail in other documents that are dedicated to the topic. In particular, Exchange 2003 is a very storage-intensive application, and you should use the standard best practices for issues such as sector alignment for the storage subsystem and determination of spindle count along with the recommendations in this document.

Exchange on VMware Infrastructure

Organizations that are designing and implementing VMware Infrastructure have been wary of moving their Exchange mailbox servers to VMware Infrastructure. The most often cited reasons for this are:

- Exchange imposes resource demands that are too heavy for a virtualized environment.
- Exchange is critical to the organization and, therefore, saving money by consolidating it is unnecessary.
- There is simply too much messaging data to handle using virtual servers.

With the proper design, these should not be issues that stop your organization from deploying Exchange within a virtual Infrastructure where you will receive all the benefits that virtualization has to offer. These objections can be addressed in turn by considering the following points:

- Many Exchange servers are extremely under-utilized, maybe even more than servers used for other purposes, because architects tend to over-engineer Exchange solutions. This paper examines several sources of data on Exchange performance that show moderate to low use of processor, disk, memory, and network resources. With growing use of newer dual-core and quad-core servers, the underutilization of hardware resources will become more dramatic. Furthermore, with the dynamic server migration capability that VMware Infrastructure provides, you are actually better able to handle peak loads.
- Exchange recovery and disaster recovery can be complicated. Installing Exchange on VMware Infrastructure simplifies these processes and, at the same time, shortens recovery times dramatically. Disaster recovery can also be expensive, but with virtualization, the cost for both hardware and resources is reduced dramatically.
- Organizations have been moving Exchange storage off local server storage onto the SAN for many years now. Once on the SAN, Exchange storage can be addressed in the exact same manner by enterprises using ESX Server as it is if they use physical servers. Corporations can continue to use the same tools for backup and recovery that are used now.

In order to set a baseline for Exchange virtual machine system requirements, we performed the following activities:

- We performed simulations using the Microsoft Exchange Load Simulator 2003 (LoadSim) utility in our lab.
- We reviewed VMware Capacity Planner data to examine Exchange Server performance.
- We examined real statistics from a moderate-size Exchange deployment with 10,000 mailboxes.

The results of these activities are detailed in the Appendix and show that Exchange is similar to most applications — a good candidate for virtualization. In fact, even though most organizations are not targeting their messaging systems as the place to save money through consolidation, there are many very good reasons to virtualize Exchange to obtain better service for users and organizations. Not all Exchange environments are amenable to virtualization, but for those that are, the benefits are significant.

Exchange Best Practices

In the following sections, we define four areas of best practices:

- **Servers:** the design and configuration of servers running ESX Server 3
- **Virtual machines:** the configuration and deployment of Exchange virtual machines
- **Storage:** the design and configuration of the back-end storage system, a particularly critical area for a messaging system.
- **Implementation:** the process for planning, testing, and implementing the virtualized Exchange environment. This includes the steps an organization should take to determine whether Exchange should be virtualized and, subsequently, the process necessary to design the environment and deploy Exchange on VMware Infrastructure.

It is important to be familiar with all of these sections before you make decisions about how to configure your Exchange virtual machines.

Note: In Microsoft terminology, a front-end server is a server running either Exchange Server 2003 or Exchange 2000 Server software that is specially configured to accept requests from clients and proxy them to the appropriate back-end server for processing. A back-end server is a server with a standard configuration. This paper refers to components of Exchange based on their functionality, such as: SMTP, bridgehead, and Outlook Web Access, which run on a front-end server, and mailbox, which runs on a back-end server.

Server Hosts

This section outlines best practices for hosts. Although some recommendations are Exchange-specific, most are design goals that should benefit any virtualized host infrastructure.

Redundancy

Redundancy is always of primary importance. If you are expecting high availability, and it is a requirement in your current design, all major components should be redundant. A general rule is that ESX Server hosts support 20–30 virtual machines on quad-processor systems and 8–13 on dual-processor systems.

In order to ensure that you meet your organization's availability targets, consider providing the following for your ESX Server hosts:

- Redundant HBAs — Do not use a single dual-port card. A minimum of two cards is required for redundancy.
- Redundant power supplies.
- Redundant network connections to each required VLAN. Consider port trunking if many VLANs must be supported.
- Mirrored ESX Server operating system drives.
- Redundant core network switches with the redundant network adapters connected to separate switches.
- Redundant SAN switches with HBA connections to separate switches.

If you do not have redundancy in your current configuration, and it is not within your budget, do the best you can, but communicate the risks to your organization

Virtual Machine Mobility

In addition to configuring ESX Server hosts to allow VMotion of virtual machines among the hosts, you should deploy the hosts in clusters in which all hosts are similarly configured so Exchange-related virtual machines can run equally well on any host. By enabling virtual machine mobility through VMotion, you give your solution the following benefits:

- VMware DRS (Distributed Resource Scheduler) and VMware HA – VMware DRS tracks the performance of virtual machines and, depending on configuration, recommends target hosts for best performance or actually migrates hosts based on policy. VMware HA automatically restarts virtual machines that run on hosts that experience a failure — for example, if a motherboard fails or the host panics.
- You can perform hardware maintenance without affecting availability. You can migrate virtual machines to other hosts before maintenance, then migrate them back after you finish the maintenance.
- You can allocate resources in advance for planned peaks. If you know that a virtual machine will require more processor resources for a process that occurs at the end of the week, such as database maintenance, you can migrate that virtual machine to an underutilized host for that period of time. This allows you to maintain service levels during the task. When the peak need has passed, you can migrate the virtual machine back to its normal host. You can even script and schedule this process.
- You can migrate virtual machines to a host you are not using or to an underutilized host for troubleshooting. This allows you to isolate a virtual machine that is experiencing an application issue.

Hardware Resources

The overall capacity of your hosts must be great enough to provide resources for all of the virtual machines you plan to run there and must provide room for the variability of the overall system. In RapidApp's practice, most of our designs — which are all determined by our clients with direction from our architects — have used a major brand of four-processor host with from 24GB to 48GB of RAM. The amount of RAM is determined by the expected average RAM per virtual machine multiplied by the expected number of virtual machines per host. RAM and processor have typically been the bottleneck for the implementations we have performed. For example, if you are using new dual-core processors and expect to support 25–32 virtual machines per host with each virtual machine using between 1GB and 1.5GB of RAM, you would design for $(16 \times 1) + (16 \times 1.5)$ or 40GB of RAM. You could choose to round up to 48GB, or you might want to reduce the cost and use 32GB. Although the latter choice limits the number of virtual machines on the host, you might determine that it is better to spend the saved money on another server.

ESX Server 3 has the ability to run virtual machines with up to four virtual processors. Exchange is multithreaded, and can often take advantage of two processors. However, in order to avoid performance penalties due to scheduling conflicts, it is recommended that the number of physical processors exceed the maximum number of virtual processors on a single virtual machine. Therefore, if you intend to use two-way Virtual SMP virtual machines for Exchange, you should plan for servers with at least four processor cores.

In contrast to CPU and memory, network and HBA configurations are usually governed by security and redundancy concerns rather than capacity. In our analysis, network bandwidth is not a significant factor in Exchange Server performance, since in most large-scale deployments, hosts are provided gigabit connectivity

You must provide adequate excess capacity within a cluster. The system should be able to handle a server failure with no performance degradation. If it takes a long time in your organization to have hardware repaired, you should provision still more capacity. Remember: since significant savings result from deploying virtualization in the first place, this is not the place to cut corners.

Quality of Hardware

Use quality servers and components in your design. Again, you achieve substantial reductions in the number of physical servers deployed in the environment, so don't settle for lesser quality server hardware and SAN infrastructure components. Ensure that all hardware is on the VMware hardware compatibility list and that the hardware meets your other criteria for performance, scalability, and availability.

Virtual Machines

The following sections outline technical considerations for deploying Exchange within ESX Server virtual machines.

Virtual Machine Timekeeping

Because of auditing regulations due to Sarbanes-Oxley laws, the time of day on all messaging-related systems must be correct so that timestamps on messages and message-related events are accurate. Since much of operating system timekeeping depends upon a count of CPU cycles, and because virtual machines which do not require CPU cycles do not get them, time in a virtual machine needs to be synchronized explicitly.

The best way to maintain proper time on all virtual machines is to turn on VMware Tools time synchronization within each virtual machine, then implement the NTP daemon from within the service console on every ESX Server host and point to an external stratum 1 NTP source or a corporate time source that synchronizes to an external stratum 1 NTP source. This will allow the virtual machines to keep time based on the underlying ESX Server host and the ESX Server host to keep proper time using an authoritative external source. The use of NTP is described in VMware knowledge base article 1339, and the overall issue of time accuracy in virtual machines is discussed in VMware knowledge base article 1318.

Note that the use of Windows Time service to synchronize time in a virtual machine is not recommended. Since this service is based on the behavior of a physical machine and is not aware of the unusual clock behavior of a virtual machine, it does not always synchronize accurately. However, if you have a requirement to run the Windows Time service or other non-VMware software to synchronize a virtual machine's clock, be aware that you should never run more than one clock synchronization tool in the same machine at the same time. The different tools are not aware of each other and are likely to operate at cross purposes, making the clock erratic. Therefore, if you choose to let the Windows Time service set the time in the guest operating system, turn off VMware Tools time synchronization.

Memory Sizing

Because Exchange 2003 is a 32-bit application, it is not capable of utilizing more than 4GB of RAM in a virtual machine. However, the Exchange Server store service always takes most of the available memory up to this limit, making it difficult to analyze how much memory is truly required for the application on systems that have a substantial amount of memory installed. The factors that affect memory usage are: number of users, number of mailboxes, and average size of mailboxes. Based on our experience, start with 1GB of memory for each type of Exchange server that you virtualize. If performance is not as expected, monitor the active memory defined in Virtual Center and the Page File/% Usage counter from Perfmon. If active

memory continually approaches granted memory, you should check the page file to see if it is also being heavily utilized. If both of these metrics are high, increase memory to 1.5GB, and see if performance improves. If not there may be some other issue.

Virtual Processors

Unless a virtual machine is running a mailbox server, configure it with a single virtual processor. For applications that do not need more than a single processor, configuring for more than one results in lower efficiency.

For mailbox servers with less than 500 mailboxes, start with a single virtual processor. Monitor performance using VirtualCenter, and if processing consumes nearly a whole processor on average, reconfigure the virtual machine to use two virtual processors. For mailbox servers with more than 500 users, start with two processors. There should be no need for more than two processors.

Network

The virtual machine obtains network redundancy and gigabit connections provided by the host. If your implementation does not provide gigabit bandwidth to the host, you may need to analyze the aggregate requirements of your virtual machines to ensure that you have adequate bandwidth for the Exchange virtual machines.

Co-Location of Virtual Machines

In order to minimize risk within the overall design, it is prudent to ensure that you spread virtual machines across hosts by function. In other words, when you are deciding which host runs which virtual machine, you should always separate virtual machines of the same type. This means you should not place two mailbox servers on the same host. No two antivirus virtual machines should run on the same host, etc. However, it makes sense to mix these functional servers on the same hosts. Therefore, on host A, you could run a mailbox server, an SMTP gateway, an Outlook Web Access server, and an antispam server. If you are using VMware DRS, it is important to configure mailbox server virtual machines so that they cannot be run on the same host, using the affinity rules. This ensures that your collocation strategy is maintained within the VMware DRS policies.

Shares and Resource Allocation

If both shares and resource allocation are set to the defaults, then ESX Server is allowed to handle scheduling appropriately. This ensures fair utilization of resources by all the virtual machines on the host. It is important to monitor resource utilization to ensure that the hosts are not overutilized, thus causing the virtual machines to compete for resources. Setting the shares higher for messaging-related virtual machines can

help to ensure that these are given top priority in case there is resource contention. You can go further and set resource reservations for these virtual machines so that they are guaranteed some minimum amount of resources — for example, memory. Enabling DRS to recommend changes will also help you understand the load on servers in the cluster.

VMware DRS and VMware HA

Initially, it is probably desirable to run VMware DRS in manual mode. In manual mode, VMware DRS recommends changes, then you can choose if and when to implement them. When you feel comfortable with the VMware DRS recommendations, you can configure virtual machines to use automatic mode.

During a host failure, VMware HA restarts all of the affected virtual machines on other hosts in the cluster. In the case of the Exchange mailbox servers the database is flagged as “dirty” because there was not a normal shutdown and the system automatically performs a recovery when it comes back up. Since VMware HA uses the CPU and memory reservation instead of actual usage to decide failover destinations, those hosts on which they were placed could be heavily loaded, while other hosts are comparatively lightly loaded.

VMware DRS subsequently determines the optimal location in the cluster for the restarted virtual machines. If VMware DRS is in manual mode, you must keep your virtual machine collocation strategy intact by moving virtual machines to appropriate servers, keeping virtual machines that you want separated on different hosts. However, if you have automated VMware DRS, you can configure virtual machine affinity rules so that, for instance, two mailbox servers are never allowed to run on the same host (see *Co-Location of Virtual Machines* on page 6).

You can still use Microsoft Cluster Services (MSCS) to provide high availability for Exchange, if you prefer. You can cluster virtual machines running Exchange on ESX Server just as you would physical machines, and VMware supports the use of MSCS. However, even if you cluster virtual machines using MSCS, there is significant additional cost, including, setup, maintenance, licensing, and possibly additional hardware.

Virtual Machine Templates

Templates enable you to keep build consistency between systems. This is very useful for organizations that have branch offices or remote campuses and want to ensure that the configuration of all Exchange deployments is kept synchronized. It also enables the exact configuration used in development and test to be deployed into production, which aids in support and troubleshooting. If you have a large deployment, you may want to build specific templates based on server type including

mailbox, bridgehead or gateway, and Outlook Web Access. You should build in operating system and Exchange prerequisites including

- Base operating system
- Latest service packs
- Antivirus and management agents
- VMware Tools
- Extensions: NNTP, SMTP, IIS, WWW Publishing Service, .NET framework and ASP .NET
- DCOM enabled
- Multiprocessor HAL (if using Virtual SMP)

Remember that you should always use Sysprep for these templates and should not install Exchange itself in the templates.

Storage

Because storage is such an essential component of an Exchange deployment, these sections address this particular area in detail.

Disk

This section discusses system partitions, data partitions and store sizing. It assumes that for critical applications, VMotion is required and, therefore, recommends only solutions that utilize shared storage. This storage can be on any supported platform. Always check the latest hardware compatibility list for storage devices under consideration.

ESX Server 3.0 supports connections to storage using iSCSI, NFS or Fibre Channel as depicted in Figure 1 on page 8. All of these can work well so long as you follow certain guidelines when you deploy them. If you use iSCSI, you should consider either a dedicated VLAN or a physically separate switched network infrastructure, just as you would when setting up a Fibre Channel infrastructure. This ensures a high level of performance for your ESX Server hosts. The network should be routable to the network or VLANs supporting Exchange client connections to simplify iSCSI setup.

You can run iSCSI within the virtual machines (label 2 in Figure 1) or on the hosts (label 1). This makes the configuration options very flexible. Both options work well, although more effort is required to set up each virtual machine to attach to separate LUNs via iSCSI. On the other hand, if you think that you may want to move a virtual machine between clusters, attaching via iSCSI inside the virtual machine allows you to do this without configuration changes. Remember that if iSCSI is used inside the virtual machine, disk traffic goes through the LAN connections. You can configure a separate LAN or VLAN (label 1) and virtual switch to ensure traffic routes through the iSCSI dedicated network adapters rather than through the public network (label 3).

System LUNs

In order to provide separation for controlling growth as well as enabling isolation for backups, consider the following guidelines for creating virtual disk (.vmdk) files:

- Configure system partitions as .vmdk files on your standard LUNs. Make sure to leave sufficient space for growth when applying service packs. With Windows Server 2003, 10–12GB for the system partition is a good starting point.
- Create a separate .vmdk for applications including Exchange. Usually 4GB is ample for this. This application partition can hold MTA and SMTP queues as it is SAN-attached and, therefore, redundant and high-performance.
- If you plan to snap and replicate the system partitions, you may want to consider creating a separate LUN for page files, creating a small .vmdk of appropriate size on this LUN to hold the page file. This way, the page file can be excluded from snapshots and replication. If the page file is split out, the system partition can be smaller.

Data LUNs

Raw device mappings (RDM), .vmdk-based partitions, and iSCSI connections in the virtual machine all work well. You have more flexibility with iSCSI and RDMs, because you can use third-party SAN applications with these just as you would with a physical machine, and it is possible to grow these volumes dynamically. However, .vmdk files offer the advantage of encapsulation, which simplifies the overall design.

The type of storage you choose depends strongly on your backup and recovery design. If you plan to configure snapshots of the data using SAN frame functionality, you should follow the vendors' recommendations. Normally, snapshots require

separate LUNs for each information store and set of logs. That means that if there are four stores in a storage group, there are five LUNs associated with that storage group: one for the logs and one for each store. Also, as with an Exchange deployment on physical infrastructure, you must determine the spindle count and capacity based on expected usage.

Store Size/LUN Size

This section pertains more to Exchange design than VMware Infrastructure, but it still merits a discussion. In Exchange, the size of information stores is usually determined by the system recovery time objective and the time required to perform offline maintenance or recovery. If your recovery time objective is four hours, you must size an information store to be equal to the amount of time needed to recover the information store and the processing time to run utilities **ESEUTIL (/P and /D)** and **ISINTEG** for that store, if necessary. If you use snapshots, restore time is very quick. However, it may still take an hour or more to process each 18–20GB using **ESEUTIL** on the latest fast processors.

One thing to remember is that if you use a snapshot technology to back up Exchange data, you must also perform a consistency check following the backup. Most organizations that do this run the checksum (**ESEUTIL .exe /k**) nightly, during the overnight hours. (If you use normal tape backup programs, this is not a requirement.) The check is required because when you snap a volume, the database is not checked for corruption. The normal recovery step from corruption is to restore the last good database. If the corruption is not caught in time, the recovery point is so far in the past you cannot recover.

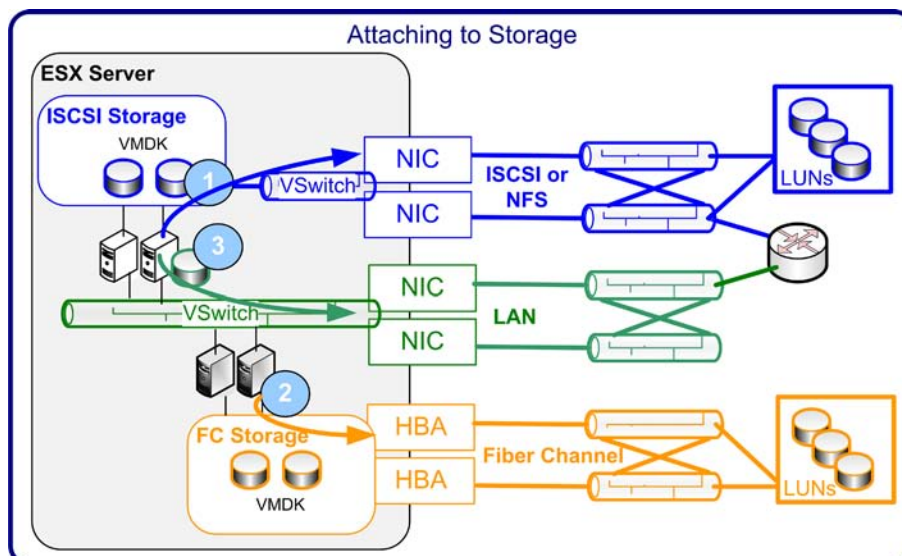


Figure 1 — Attaching to storage

Implementation

In the previous section provided suggestions for configuring hosts, virtual machines, and storage to meet your service level objectives. This section outlines the processes you should consider to ensure a solution that is engineered well and that functions well.

Analyze Current Systems Performance

- Perform data capture on your existing systems. Monitor processor, disk, memory, and network statistics over a period of time that makes sense for your business. In Perfmon for mailbox servers, Outlook Web Access, bridgeheads, and gateways, the relevant metrics are:

| Object | Counter |
|-------------------|---|
| Processor | % processor time |
| Physical disk | Disk bytes/sec and IOPs (To calculate IOPs, see http://www.microsoft.com/technet/prodtechnol/exchange/Guides/StoragePerformance/f4dc0eaf-3d21-4650-83b7-86526694f05a.mspx?mfr=true) |
| Memory | Committed bytes |
| Network interface | Bytes total/sec and IOPs |

Table 1 — System performance metrics

- Analyze the data to understand the peaks — when they occur and how long they last

Review Vendor Support

Review vendor support for virtualization with your team and management. Obtain consensus on the strategy before moving forward. You should consider the following points when deciding your plan for support:

- Microsoft supports Exchange on third-party virtualization platforms for customers with Premier contracts.
- Microsoft may ask you to recreate the problem on a nonvirtualized platform
- Microsoft will ask customers who do not have Premier contracts to recreate the problem on a nonvirtualized platform.

For more information on this subject, see Microsoft knowledge base articles 320220 and 897615.

Size the Solution

Size your solution so you feel comfortable with the load while using standard configurations for the ESX Server hosts. Ensure that you take into account all requirements your organization may have for virtualization. In other words, do not design a separate system for Exchange. Exchange should fit into your overall virtualization strategy as another application in your total set of application services.

- The solution can be made up of dual- or quad-processor servers. For large organizations, the most cost-effective platform generally is a quad-processor platform, especially for configurations utilizing virtual machines with more than one virtual processor. You can use larger systems if that better matches your architectural standards, though usually the cost increases with these solutions.
- Ensure that there is enough spare capacity in the server farm to account for a hardware failure while maintaining full performance.
- Ensure that you can meet your security and network requirements. If you plan to virtualize all servers, you may need access to the internal network and the DMZ. If you want to mix test and production, you may also need access to the test and development environment.
- Assume that you will not run more than one of the same server type — for example, mailbox, bridgehead, or gateway — on the same host system, though you certainly could mix these types. You should plan for redundancy in nonmailbox servers spread over different hardware in case a server fails.
- Ensure that you test using the same applications you will have in production. For instance, Research in Motion Blackberry services can add a substantial disk I/O load and a moderate processor load to your Exchange servers. If you perform testing without this component, you miss a significant portion of the true load. In the case of Blackberry, Research in Motion has published a document on performance considerations for adding Blackberry services to your environment. This document, *Performance Characteristics: BlackBerry Enterprise Server Version 4.0 for Microsoft Exchange*, is available on the Research in Motion Web site.
- Follow the redundancy guidelines discussed elsewhere in this document.
- If you can afford only a single quad-processor host, using three dual-processor hosts would provide more redundancy than would a single quad-processor host at a similar cost.

Test the Solution in the Lab

Test your solution in the lab using Microsoft tools such as Microsoft Exchange Load Simulator 2003 (LoadSim) and Microsoft Exchange Server Stress and Performance 2003 (ESP). Try to match your users' habits. Much of this information is available from your current systems using the following Exchange performance counters:

- Average messages sent per day
- Average messages received per day
- Average size of attachments, etc.

Document Processes and Configurations

Create process and configuration documentation prior to the pilot. Use the lab to prove out and engineer the processes. These may include:

- Backup and recovery including VMware HA
- Disaster recovery
- Provisioning
- Running **ESEUTIL**
- Reorganization
- Recovery
- Consistency or checksum

Pilot Test the Solution

Add a new server using VMware Infrastructure and move a small number of IT user mailboxes to the new server. Perform backups and run normal processes until you feel comfortable with the design. Next, increase the pilot using business users who are friendly to IT. Subsequently, select a representative group to test. Finally, plan for full-scale deployment.

Full Migration

Although it is possible to virtualize an existing Exchange server using VMware P2V Assistant, most organizations choose to use the MoveMailbox migration method to perform a migration. MoveMailbox is a function within Exchange that allows administrators to transfer a mailbox from one server to another. When you perform this function, your users do not need to make any changes in the configuration of their Outlook clients. This method has benefits and drawbacks. The main benefits are:

- An incremental approach allows administrators to monitor performance and decide on the mailbox load that works best for a given deployment. In other words, as you are migrating, you may decide to stop at 500 mailboxes on a server or 1000. If you use the P2V Assistant approach, you have to match the number of mailboxes you had on a physical server, and the migration is all or nothing.
- A MoveMailbox migration is gradual. If you have a problem, it should turn up early and will not affect as many users.
- Each mailbox is migrated independently, and if there are too many errors, the migration is rolled back with no service outage.

The main drawbacks are:

- When you use MoveMailbox, single instancing in the information stores is lost. Therefore, the information stores grow and use more storage. This increase in storage requirements is resolved over time, but it can add as much as 30–40 percent to the size of the stores during the migration.
- Using MoveMailbox can take quite a while depending on the other technologies used in your Exchange environment. For example, if snapshot backups are used in your organization, you may need to substantially limit the number of migrations performed per day in order to limit the size of change in your information stores, allowing you to stay within your snapshot volume size limitations.

If you choose to move to VMware Infrastructure and undertake a major version upgrade (for instance, upgrading to Exchange 2007) at the same time, you must make migration choices similar to those you would make for a physical server solution.

Case Study

The case study described in this section brings the concepts discussed in this paper together. RapidApp worked with a client to develop a new architecture for its organization based on VMware Infrastructure, Exchange, and Microsoft Active Directory. The architecture described here is similar to that solution and illustrates the alternatives presented in this paper. The original design was for Exchange 2007, The example used here shows Exchange 2003 to reflect currently shipping products. The concepts described here will work with Exchange 2007 when it is released.

The company in the design has three offices with 200 to 300 users in each. Most of the company's approximately 20 other offices have fewer than 25 users.

The organization decided to create two main datacenters, one in the North and one in the South. The larger offices are connected to the datacenter with Gigabit Ethernet. The remaining offices are connected with a multiprotocol label switching (MPLS) network with multimegabit connectivity.

VMware Infrastructure was chosen as the virtualization platform for the company. Each datacenter will have five ESX Server hosts configured as follows:

- 4 dual-core processors
- 32GB DDR2 400MHz (16 x 2GB) dual-ranked DIMMs
- 2 73GB U320 SCSI hard drives
- 2 QLogic 4GB HBA cards
- 2 on-board Ethernet adapters
- 3 Intel Pro 1000MT dual-port Gigabit Ethernet adapters

The company plans to use a SAN that offers both Fiber Channel and iSCSI connectivity. Exchange data will be backed up using an online backup process that backs up to disk using Exchange backup APIs. Exchange virtual machines will be backed up using SAN snapshots. Both types of backups will be replicated to the alternate datacenter. Due to the amount of replication required for this solution, the page files will be stored in separate .vmdk files, so that they will not be replicated.

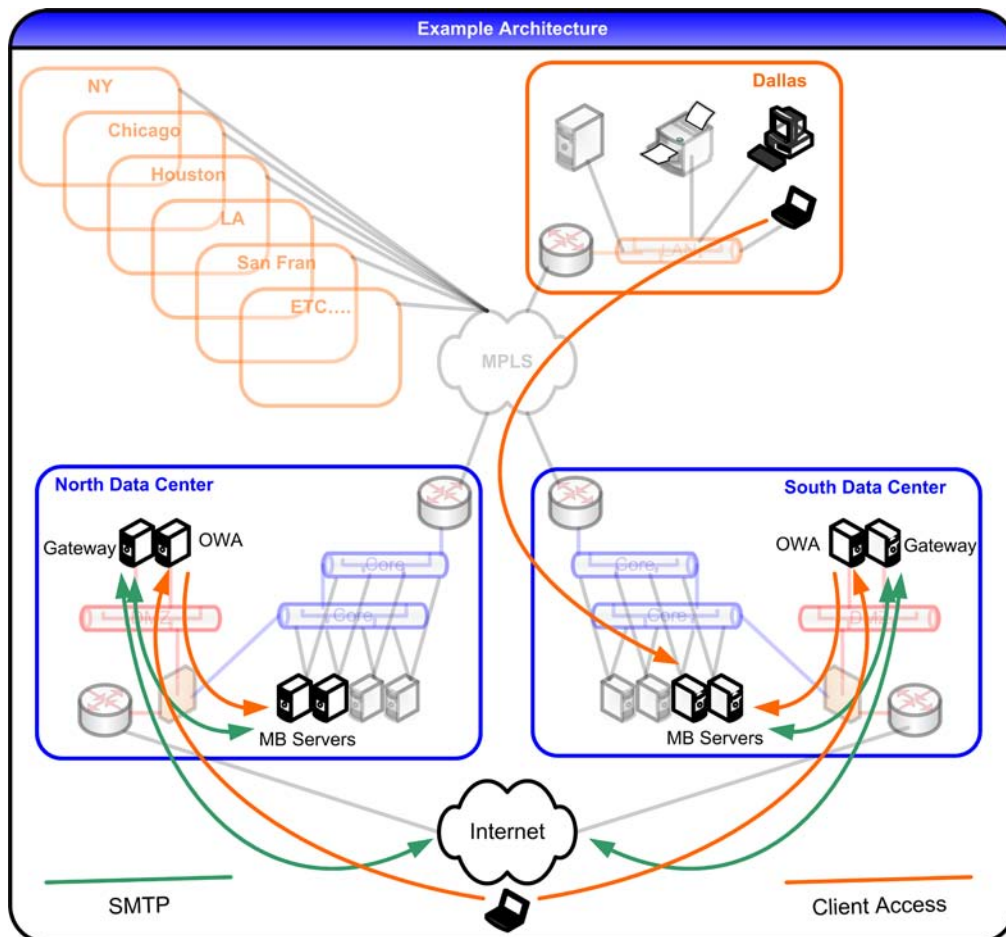


Figure 2 — Example architecture

The configuration of virtual disks (.vmdk files) and LUNs is shown in Figure 3. Information stores will be limited to 60GB per store. Average mailbox size is 500MB, so there should be approximately 100 mailboxes per information store. Initially, there will be six information stores across two storage groups. The configuration will use .vmdk files for the logs and information stores — one for each storage group and one for each information store. The .vmdk files will be spread over four LUNs in order to lessen the risk from a LUN failure. Logs will be placed on the same LUN that holds the system partition .vmdk files, because if the system partition LUN goes down, the logs will be unusable, anyway. As mentioned above, the page file will be in a separate .vmdk file on the same LUN, as well.

The design will support SAN replication of systems. Disaster recovery will be provided for each datacenter by the alternate datacenter. Therefore, the virtual machines for applications

housed in the South Datacenter will be replicated to the North Datacenter and vice versa. If a disaster recovery event is declared, nonessential virtual machines will be shut down to provide capacity for essential virtual machines.

Redundant network adapters are deployed to create a total of five separate networks as shown in Figure 4:

- Core
- DMZ
- iSCSI
- Service console
- VMotion

Mailbox and global catalog servers are placed on virtual switches connecting to the internal core network, and Outlook Web Access and gateway virtual machines are connected to the DMZ. The other three networks exist to provide support for operating various VMware Infrastructure components as well as for access to storage.

Exchange servers themselves are hosted only within the two datacenters. All clients will use cached-mode clients remotely. Users will also use Outlook Anywhere, both internally and when traveling. Each datacenter will have two mailbox servers, two global catalog servers, one Outlook Web Access server, and one SMTP gateway. The company uses external services for

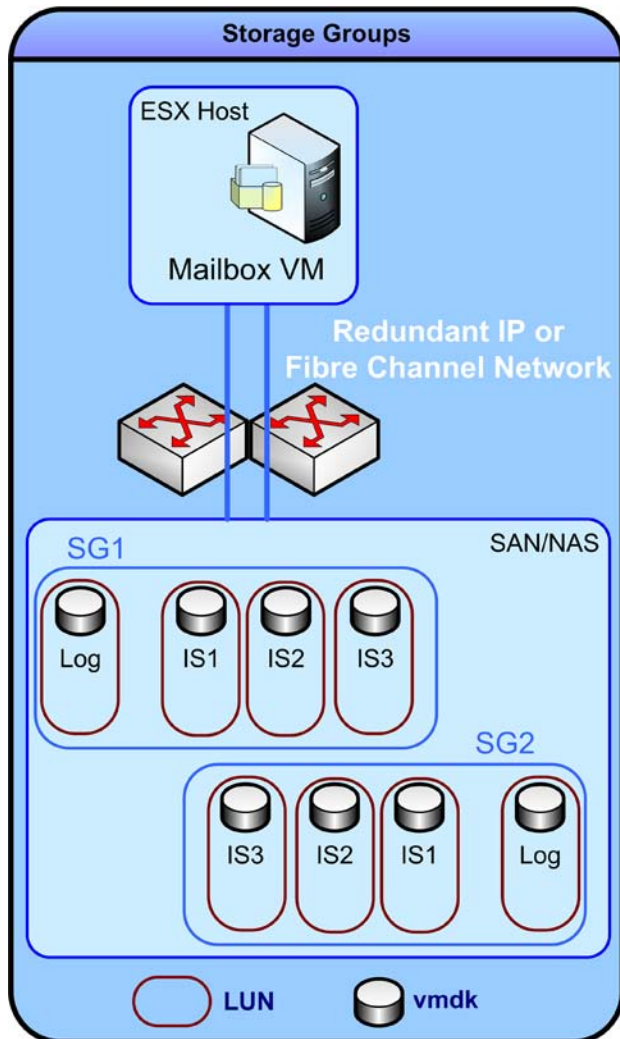


Figure 3 — Configuration of storage groups

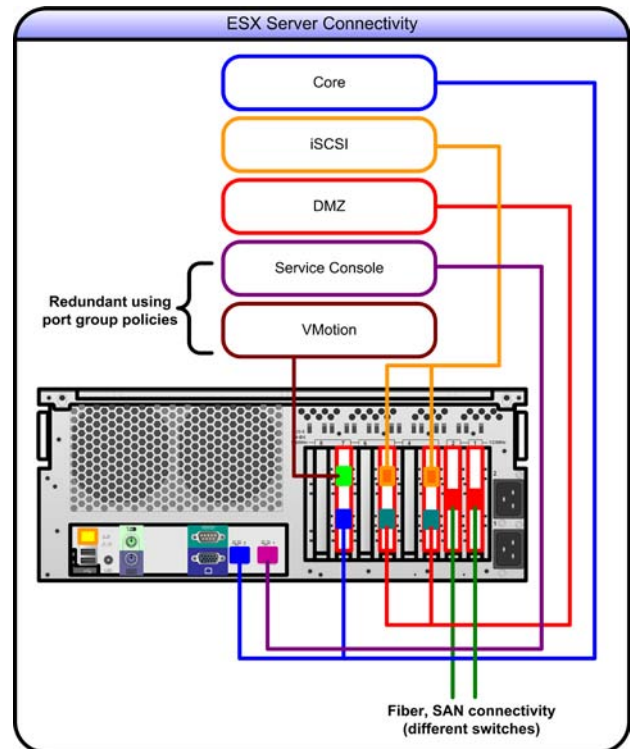


Figure 4 — Networks connected to the ESX Server system

mail scrubbing and routing, therefore, the SMTP gateways will only send and receive from the service provider. The Outlook Web Access and gateway servers are made redundant across the two sites. If, for any reason, one of these servers fails at one site, traffic will be sent to the corresponding server in the other. If mailbox growth is greater than expected, a third mailbox server may be introduced into the model. The location of virtual machines in the datacenter is depicted in Figure 5. Note that many other application virtual machines run alongside Exchange in this environment.

One of the five servers is referred to as a “swing” server. Swing servers are used to determine baseline performance of new virtual machines before placing them permanently on production hosts. This makes it possible to find the bottlenecks of the new virtual machine before deciding what ESX Server hosts could run it. Use of a swing server also makes it easier to decide which virtual machines can be collocated on the same ESX Server host, based on load, seasonality, and other performance metrics. It also engenders confidence in the administrative staff, because the virtual machines will be on servers with much lower server-to-processor ratios from the start.

The swing server also acts as the recovery server for the farm. If any one host fails, the swing server can mount the virtual machines for the failed host, bringing all the affected services back online quickly. This decreases the overall capacity of the farm, but it ensures greater availability of the virtual machines.

This infrastructure is expected to host no fewer than 100 virtual machines in each datacenter. Each server will support mixed loads including test, QA, DMZ, and internal production applications. If it is necessary to troubleshoot performance of any one of the application virtual machines, it will be migrated to the swing server and diagnosed there.

Overall, this design provides an almost ideal fit for the client’s requirements, providing a reliable, recoverable, and cost-effective computing infrastructure for the organization.

In conclusion, if you have already set up VMware Infrastructure, you should strongly consider moving your entire Exchange service to this platform. If you do not already have VMware Infrastructure in place, the advantages outlined in this paper should provide a basis for you and your organization to consider deploying it.

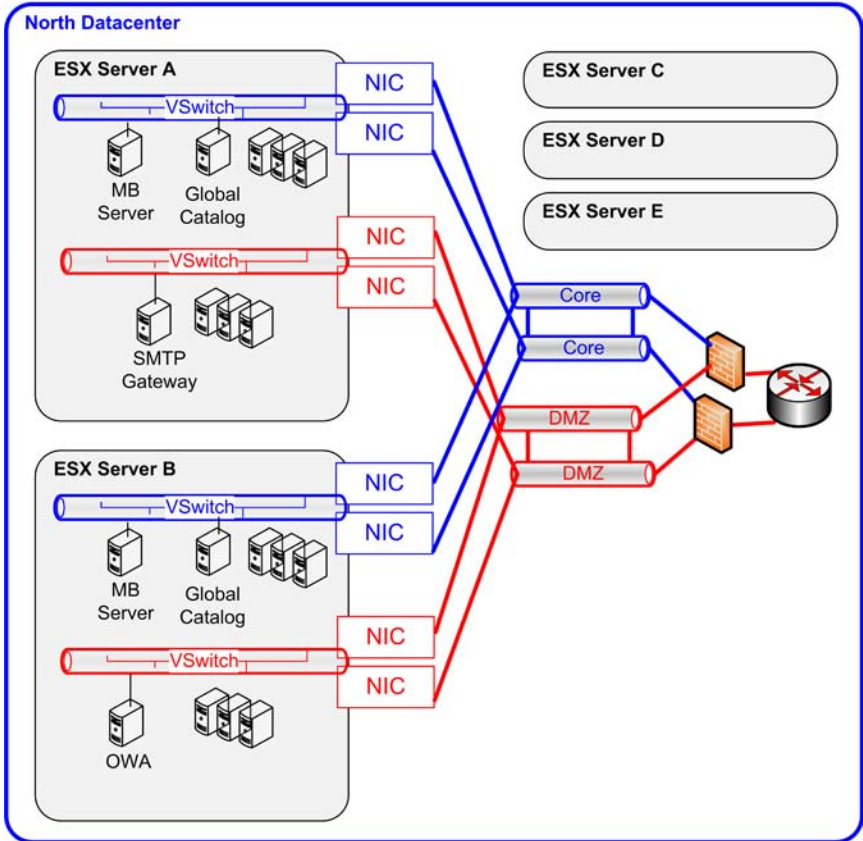


Figure 5 — Architecture of the north datacenter

Appendix

The findings explored here are intended to show that virtualization is indeed a worthwhile option for Exchange. You should not use these statistics to size your specific implementation. Each implementation has many different factors that make it unique. You should use this information both to understand the capabilities of virtualization and to set a baseline for your thinking. With this in mind, you can perform your own tests and create your own strategy.

This section examines Exchange server performance statistics in three ways:

- Using LoadSim analysis in the lab
- Using informal representative client data
- Using simple Capacity Planner statistics

All these sources demonstrate that Exchange Server is very likely a good candidate for virtualization in your environment, given that, as with many applications, hardware resource utilization in most organizations proves to be moderate to low.

Load Simulator Analysis

We tested using Microsoft LoadSim in the lab. We tested with 500 users in a heavy-load scenario based on Microsoft's LoadSim scale.

The lab was configured as follows:

Two Dell PowerEdge 2900 Servers configured with:

- 2 Intel dual-core Xeon 5060 3.0GHz processors
- 3 Gigabit Ethernet adapters
- Low-end iSCSI NAS storage solution using a Compaq ProLiant DL 360 with SCSI storage
- 5GB RAM
- ESX Server 3.0.1

The virtual machines were configured with:

- Dual processors
- 768MB memory
- Gigabit virtual Ethernet adapter
- 54GB iSCSI LUN connected in guest

We started the simulations and let them run for about two hours, then monitored them for about 90 minutes to capture the results provided in this analysis. The tables below summarize the results of our testing.

Table 2 shows each metric's average values.

| Metric | 500 Users Heavy |
|--------------------------|-----------------|
| Processor (MHz) | 2217 |
| Physical Disk (KB/s) | 502 |
| Memory (MB) | 562 |
| Network Interface (KB/s) | 2159 |

Table 2 — Average results

Table 3 shows each metric's maximum values.

| Metric | 500 Users Heavy |
|--------------------------|-----------------|
| Processor (MHz) | 3493 |
| Physical Disk (KB/s) | 3568 |
| Memory (MB) | 658 |
| Network Interface (KB/s) | 3680 |

Table 3 — Maximum results

The processor averages show just over 2200MHz for 500 users. These values are based on 3GHz dual-core processors. So, with 500 users, the configuration uses less than half of a physical processor for the simulation.

You can see from the Table 3 that the peak values were higher, but not significantly so. Even with peaks, the 500-user simulation used just a little over half of a physical processor.

All of the other critical metrics were similarly low compared with the power of these machines.

For an example of these performance metrics, see figures 6–9, which show statistics captured using VirtualCenter 2.0.1 on ESX Server 3.0.1.

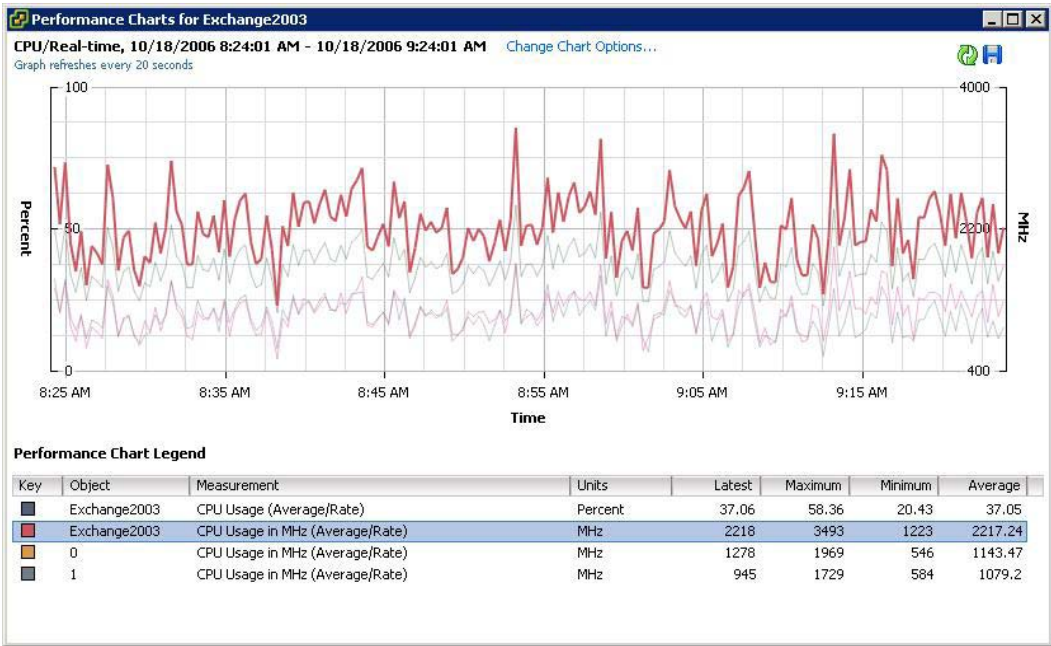


Figure 6 — CPU performance with 500 heavy users

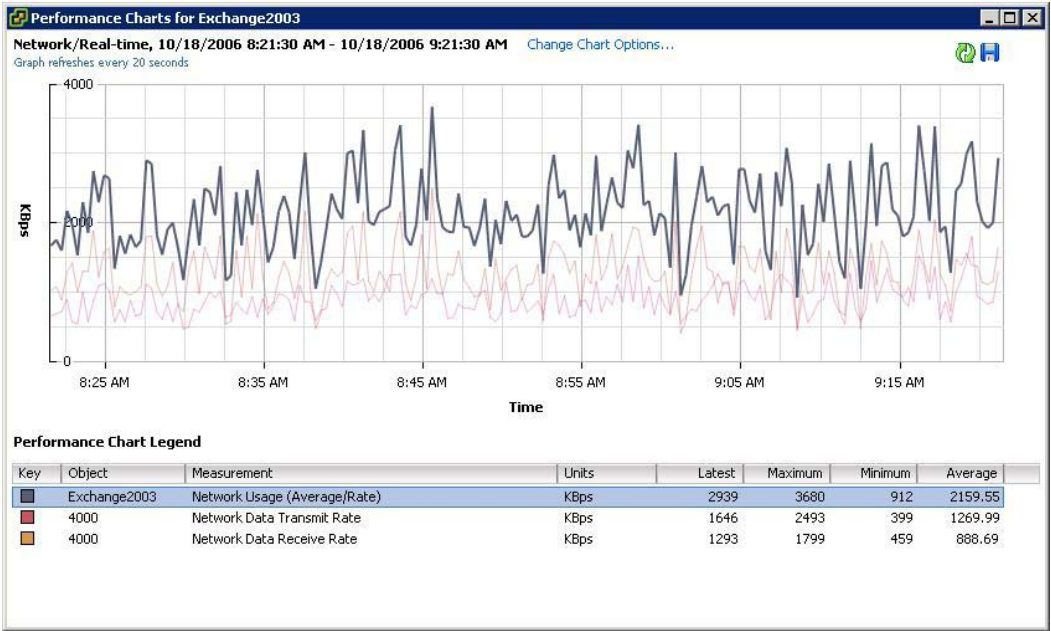


Figure 7 — Network performance with 500 heavy users

VMWARE BEST PRACTICES

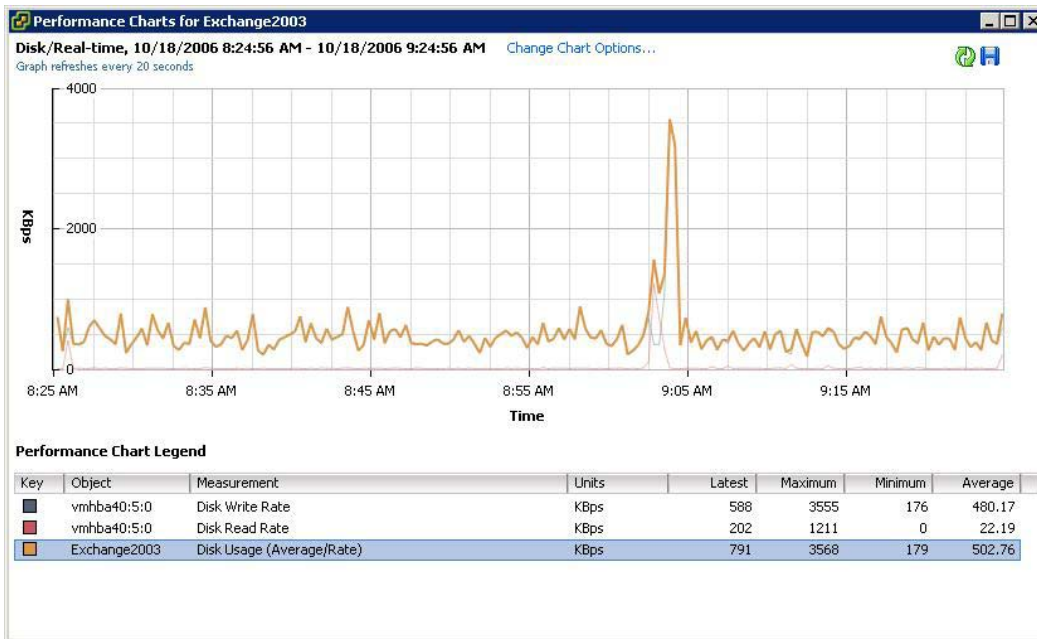


Figure 8 — Disk performance with 500 heavy users



Figure 9 — Memory performance with 500 heavy users

Capacity Planner Data

VMware provides a service called VMware Capacity Planner to many of its customers, both directly and through partners. This service captures performance data from servers, then uploads the data to a centralized database at VMware. Clients can then analyze their own data and print reports on server performance metrics from a Capacity Planner Web site. Among other statistics, the Web site shows aggregate performance metrics categorized by application. Table 4 shows information that has been captured for Exchange server as of Oct. 30, 2006. These statistics are for servers with Exchange installed. It does not differentiate between types of Exchange servers.

| CPU's | Total GHz | RAM Used (MB) | % CPU | Paging./Sec | Disk Busy | Disk Bytes | Disk IOs/Sec |
|-------|-----------|---------------|-------|-------------|-----------|------------|--------------|
| 2.00 | 4.793 | 590 | 6.5 | 19 | 9.7 | 609K | 57.41 |

Table 4 — Capacity Planner aggregate statistics for Microsoft Exchange servers

The average number of processors on Exchange servers in the database is two, with an average of almost 4.8GHz per server. Of that 4.8GHz, on average, only 6.5 percent was utilized. Average RAM used was 590MB and, as seen above, disk performance was low, as well.

The types of organizations that provide metrics to this database are extremely varied and include small companies with 100 users and large organizations with thousands of users. The breakdown of types of servers in the data is illustrated in the table and graph in Figure 10. Remember, this data is based on over 10,000 servers in the database.

Note: As with most applications on modern servers, this distribution is weighted significantly to less than 10 percent processor utilization.

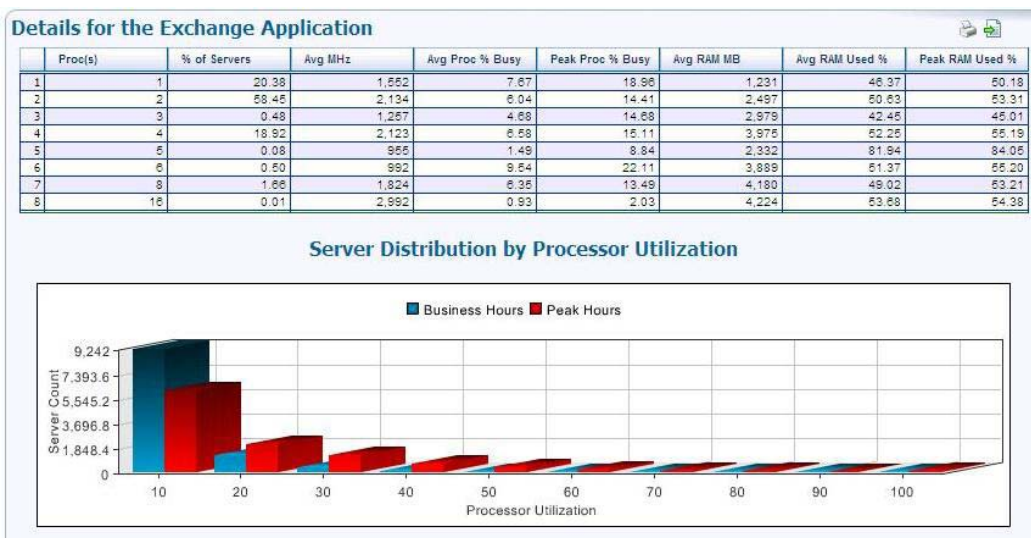


Figure 10 — Comparative performance statistics from VMware Capacity Planner

Client Example: Exchange Performance in a Nonvirtual Environment

In RapidApp's experience, most large organizations use quad processor servers for Exchange mailbox servers with 1000 to 3000 mailboxes, and dual processor servers for mailbox servers with 50 to 1000 mailboxes. Organizations normally use dual-processor servers for bridgeheads and gateways. As seen from the Capacity Planner data, smaller organizations may use smaller servers for these component servers.

In order to provide some real world performance data for Exchange, we asked one of our clients if they would provide data from two types of servers, chosen at random, in their environment; a mailbox server and a bridgehead server. These servers were representative of their environment which has seven mailbox servers of similar size and six bridgehead servers configured exactly the same connecting three main hub sites with two bridgeheads each. The results show that both are not heavily utilized.

| Metric | | Bridgehead | Mailbox Server |
|-------------------|------------------|------------|----------------|
| Processor | % Processor Time | 8.13% | 15.58% |
| Physical Disk | Disk KBytes/Sec | 3481 KB/s | 4065 KB/s |
| Memory | Committed Bytes | 1.47 GB | 2.67 GB |
| Network Interface | KBytes Total/Sec | 382 KB/s | 379KB/s |

Table 5 — Performance in a nonvirtual environment

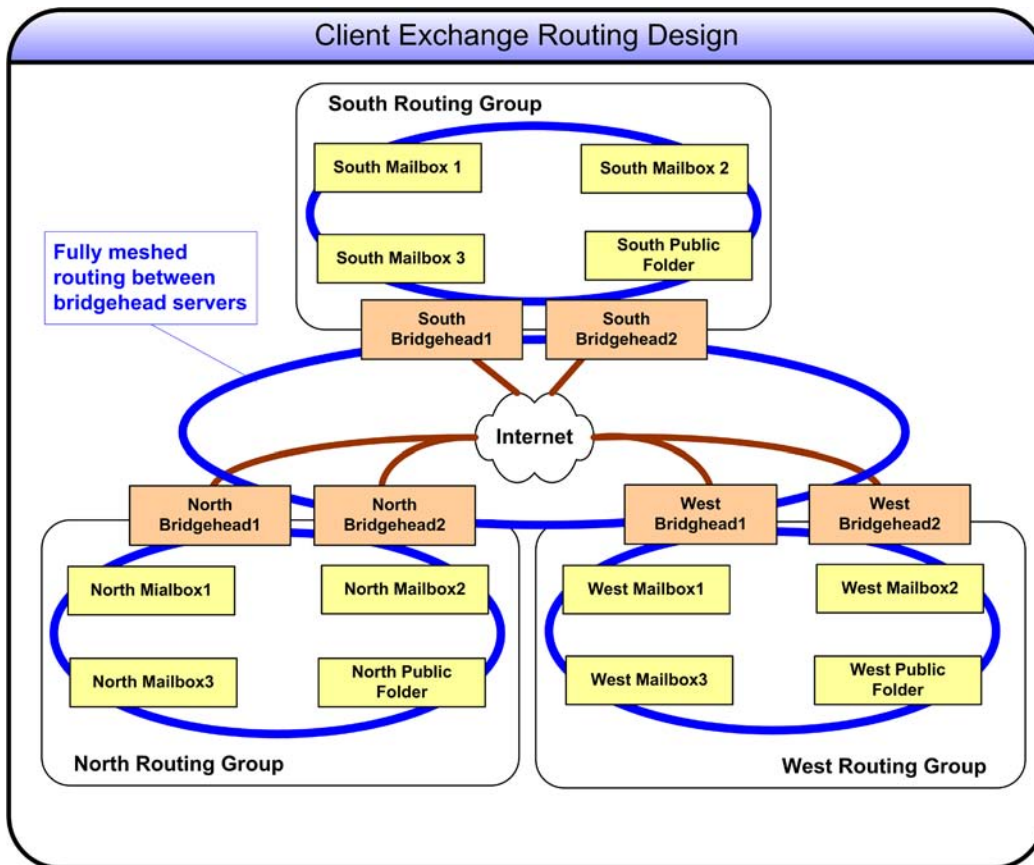


Figure 11 — Client Exchange routing design

These results are shown in more detail below. The configuration of these servers is as follows:

Mailbox Server

The mailbox server was configured with:

- 4 Intel Xeon 2.6GHz processors
- 4GB of RAM
- High-end iSCSI attached NAS storage with gigabit connections
- Gigabit network adapter connectivity

The charts on the following pages illustrate this server’s performance. This server hosts approximately 1600 Mailboxes. It is deployed with Exchange 2003 Enterprise.

CPU — As seen in Figure 12, average CPU utilization over the course of a typical afternoon from 1 p.m. to 5 p.m. was only 4.6 percent. Peak utilization during this time was less than 20 percent and, therefore, less than a single processor.

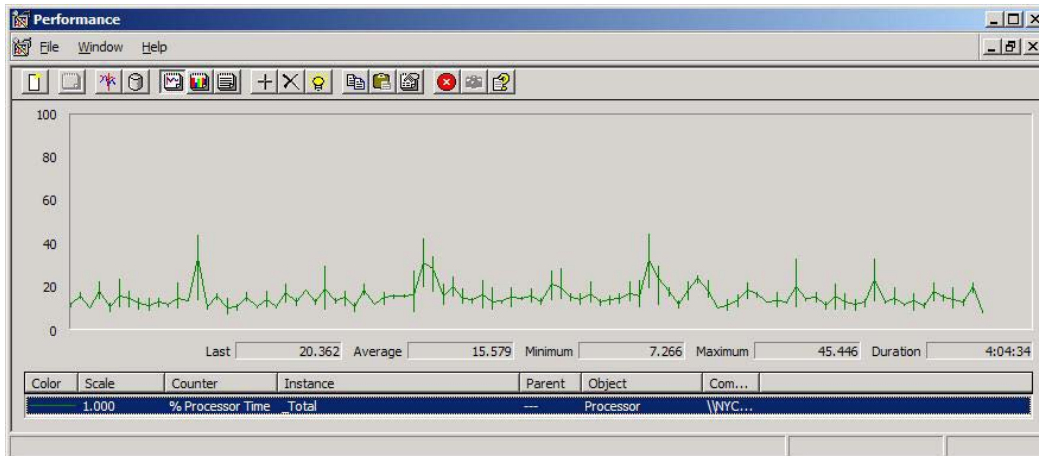


Figure 12 — CPU performance on an Exchange mailbox server

Disk — As seen in Figure 13, average disk bytes/sec was a little over 4MB/sec.

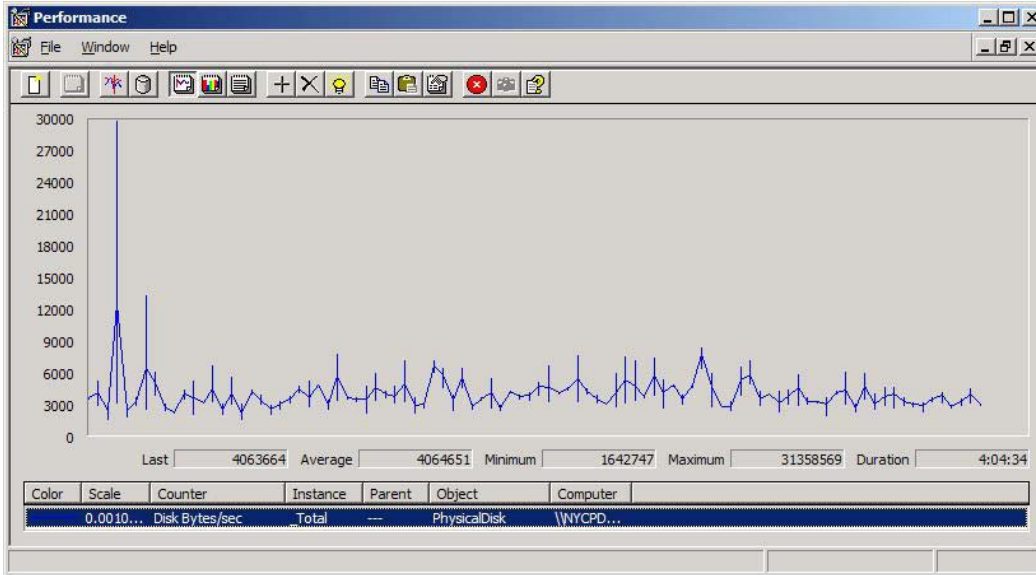


Figure 13 — Disk performance on an Exchange mailbox server

Network Interface — As seen in Figure 14, average network utilization was less than 870KB/sec.

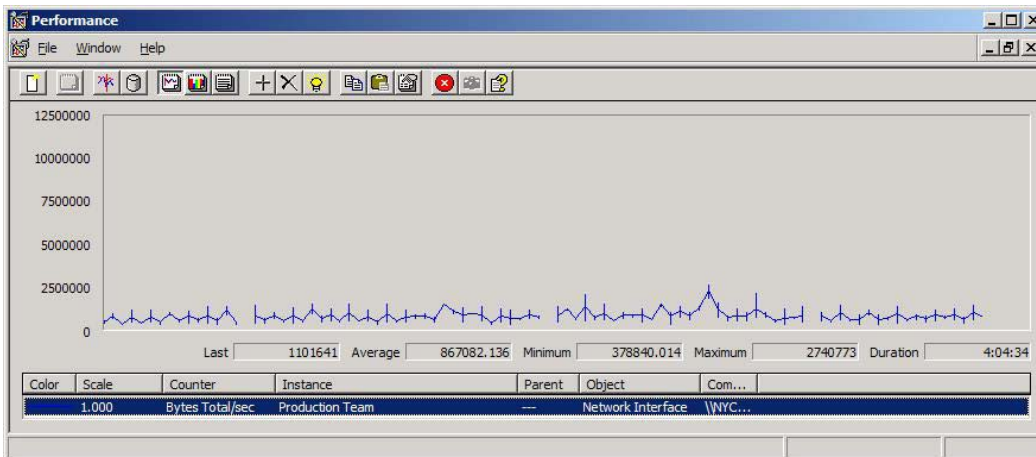


Figure 14 — Network performance on an Exchange mailbox server

Memory — As seen in Figure 15, average committed bytes of memory was 2.7GB. Remember that Exchange will utilize a large amount of memory if it is available but it is not always necessary for adequate performance as seen in the LoadSim testing.

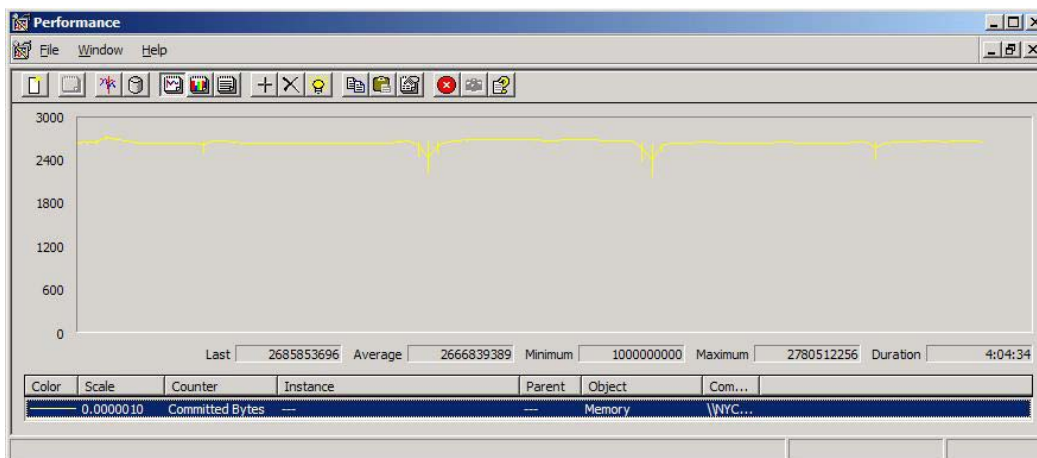


Figure 15 — Memory performance on an Exchange mailbox server

Bridgehead Server

The bridgehead server was configured with:

- 2 Intel Xeon 2.8GHz processors
- 3GB of RAM
- Locally attached SCSI storage
- Gigabit network adapter connectivity

The following charts illustrate this server’s performance. This bridgehead is one of two in a site with approximately 3,000 users.

CPU — As seen in Figure 16, average CPU utilization over the course of a typical morning from 8 a.m. to 11 a.m. was only 8.1 percent. Peak utilization during this time was less than 25 percent and, therefore, less than half a single processor of this dual-processor server.

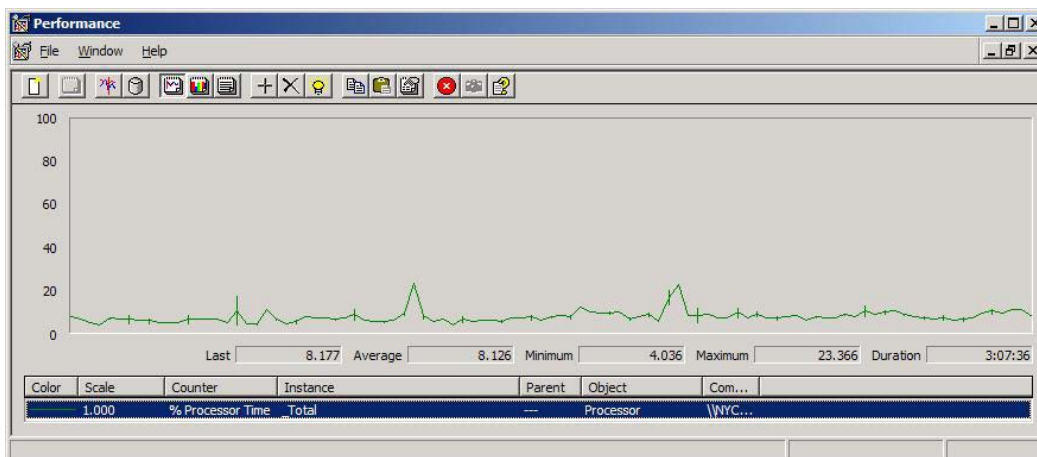


Figure 16 — CPU performance on an Exchange bridgehead server

Disk — As seen in Figure 17, disk utilization is also very low with an average of 348B/sec and a maximum of 189KB/sec.

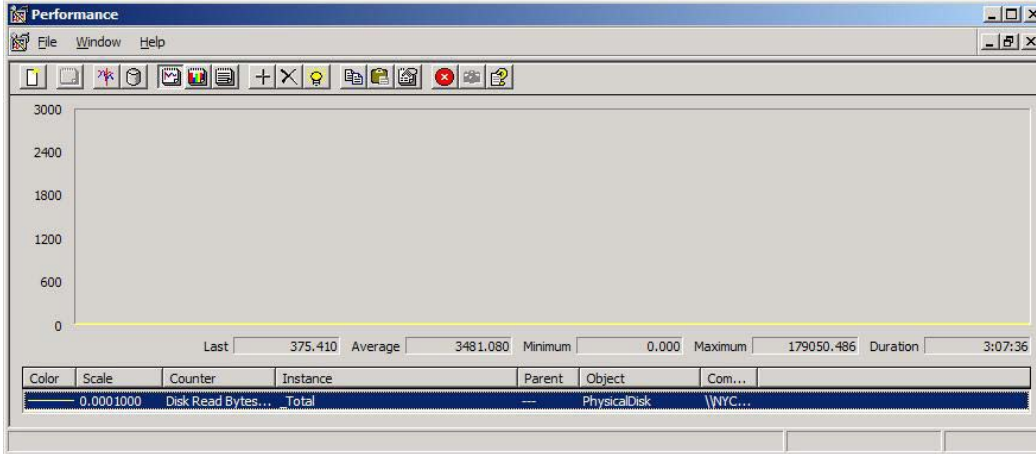


Figure 17 — Disk performance on an Exchange bridgehead server

Network Interface — As seen in Figure 18, network utilization over this time period was 382KB/sec.

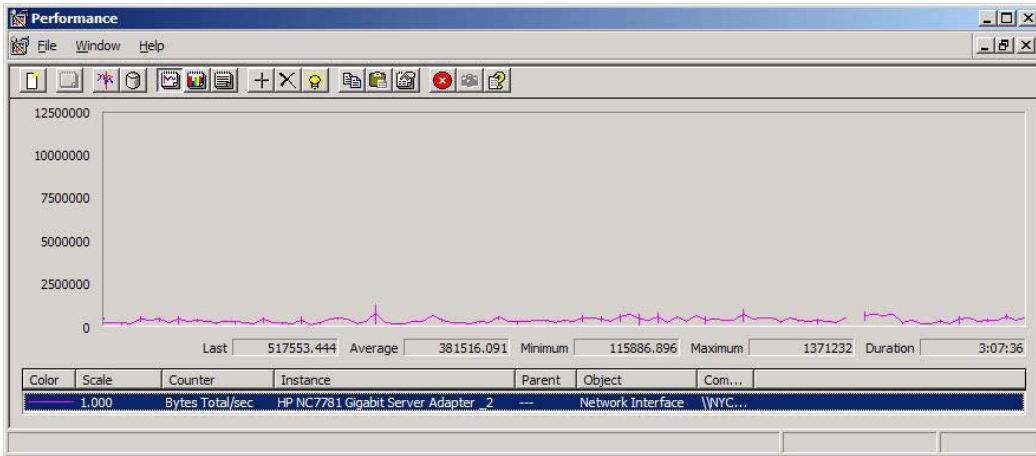


Figure 18 — Network performance on an Exchange bridgehead server

Memory — As seen in Figure 19, memory was the only metric on this server that could be considered substantial. The average committed memory was 1.46GB with a maximum of 1.49GB. Exchange is an application that will take up most available memory, but this system could still run well with less memory if necessary.

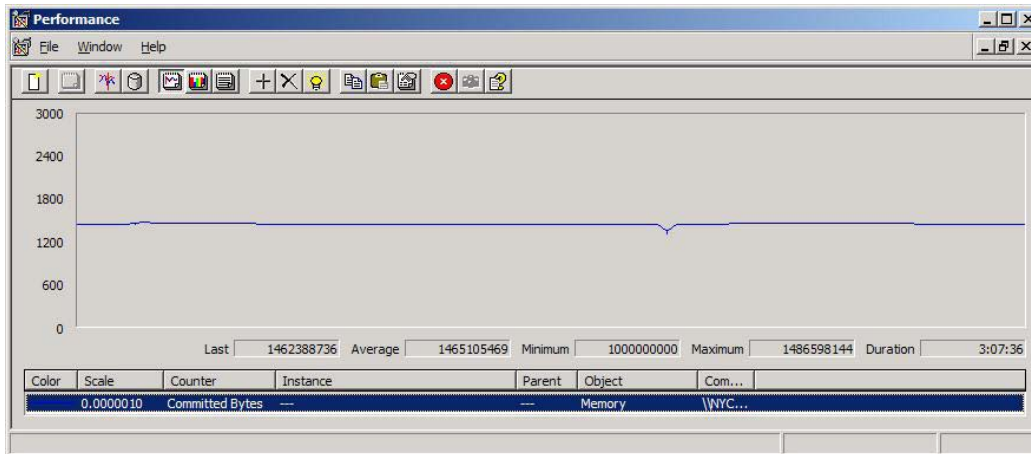


Figure 19 — Memory performance on an Exchange bridgehead server

About the Authors

Robert Zylowski is the Director of Consulting for RapidApp. He has over 19 years of Information Technology experience. He specializes in infrastructure architecture and design development as well as high-level technology management. He is an expert in the design, implementation, and management of multi-platform infrastructures. He has been a messaging architect and expert for over 15 years and has been working with Microsoft Exchange Server since its introduction in 1996. At RapidApp, he has been providing clients with the architecture and planning of ESX technologies for almost three years. His experience spans all major computing platforms and network technologies.

Zylowski earned an MBA from the Simon School at the University of Rochester in Rochester, NY and has experience managing both financial and operational systems.

Several RapidApp staff members helped with the development of this paper. Guidance with many technical questions and answers was provided by Ron Oglesby, RapidApp's Director of Technical Architecture, and Robert Wilson, one of RapidApp's Senior Technical Architects.

Charu Chaubal is a Technical Marketing Manager at VMware. He provided guidance, editing, and ideas.

Revision: 20070108 Item: BP-010-ISV-01-01



VMware, Inc. 3145 Porter Drive Palo Alto CA 94304 USA Tel 650-475-5000 Fax 650-475-5001 www.vmware.com
© xxxx VMware, Inc. All rights reserved. Protected by one or more of U.S. Patent Nos. 6,397,242, 6,496,847, 6,704,925, 6,711,672, 6,725,289, 6,735,601, 6,785,866, 6,789,156, 6,795,966, 6,880,022, 6,944,699, 6,961,806, 6,961,941, 7,069,413, 7,082,598, 7,089,377, 7,111,086, 7,111,145 and 7,117,481; patents pending.
VMware, the VMware "boxes" logo and design, Virtual SMP and VMotion are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

