VMware ESX Server 2

Best Practices Using VMware Virtual SMP

VMware® Virtual SMP™ is an add-on software module that allows VMware ESX Server to run virtual machines using multiple virtual CPUs. Multi-threaded applications such as databases and mail servers, and other applications that use multiple processes, can support greater loads when deployed using symmetric multi-processing (SMP) virtual machines, as compared to virtual machines with a single virtual CPU. This white paper describes best practices for deploying Virtual SMP and achieving the highest levels of application performance and throughput.

This paper includes the following topics:

This white paper is intended for advanced ESX Server administrators who wish to deploy Virtual SMP and need to understand the factors and trade-offs that may affect their choices.

Introduction to Virtual SMP Technology

Virtual SMP allows an ESX Server virtual machine to have access to more than one CPU. These virtual machines implement a standard shared-memory SMP computer architecture and are compatible with guest operating systems such as Windows® Server 2003 and Red Hat™ Enterprise Linux.

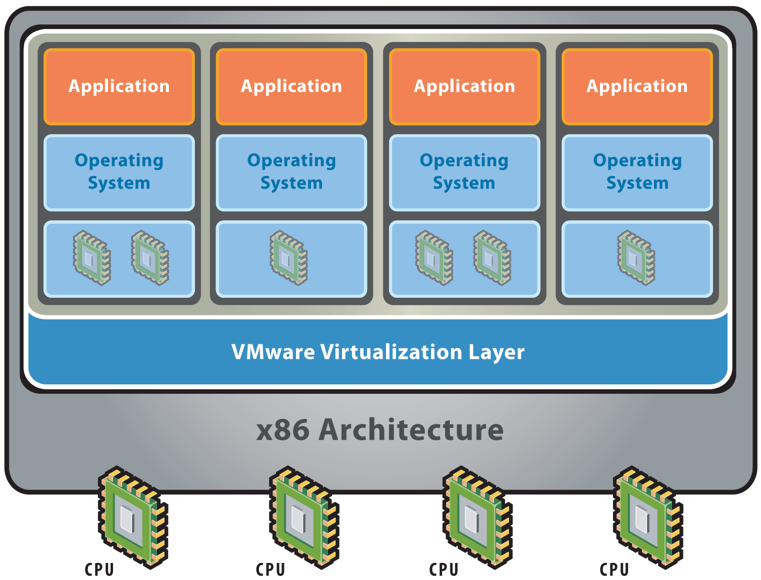

Figure 1: Virtual SMP Computer Architecture

By default, virtual machines are allocated one CPU each (called a virtual CPU). However, Virtual SMP allows virtual machines to have access to more than one CPU; that is, virtual machines can be allocated two virtual CPUs.

Virtual SMP and ESX Server give users the flexibility to deploy virtual machines with multiple virtual CPUs for applications spawning multiple processes and multi-threaded applications that can take advantage of additional virtual CPUs. On a single physical ESX Server machine, you can also mix and match Virtual SMP virtual machines and uniprocessor virtual machines.

Note: Even without Virtual SMP enabled, ESX Server can use multiple physical processors in a physical SMP system. With such a system, multiple physical processors can be used to run multiple virtual machines. However, without Virtual SMP, each virtual machine can have only one virtual CPU.

SMP in the Physical World

In the physical world, SMP describes the most common type of computer system architecture that uses more than one CPU. In these types of systems, all processors must share a common architecture and all resources such as memory, disks, buses, and I/O adapters are shared among all the processors. (This arrangement is also called tight coupling.) Only one copy of an operating system runs on all of the CPUs. Any CPU can execute any program or task (process or thread), whether it is a part of the operating system or part of a user application. In these scenarios, the operating system must be designed to support SMP and the operating system is responsible for balancing the processes among all the available CPUs as evenly as it can. The operating system is also responsible for ensuring synchronization of the data that needs to be modified by multiple threads or processes.

Scalability of SMP

In most cases, adding more CPUs to an SMP system does not increase throughput in direct proportion to the new resources because workloads cannot always take advantage efficiently of multiple CPUs. There is also some overhead involved in sharing resources and scheduling processes. For example, a four-CPU SMP system is not four times as productive as a single-CPU system. The efficiency of a multiple-CPU SMP system, versus a single-CPU system, is defined by the workload's ability to scale, or the workload's scaling ratio. This scaling ratio varies for different workloads.

Although single-threaded workloads utilize only one CPU at a time, the operating system can dynamically migrate a thread or process between physical CPUs in a multi-CPU system. In addition, system processes can be scheduled on other available CPUs. However, CPUs in the SMP system may still compete for memory, bus cycles, and other system resources. Newer operating systems such as Windows Server 2003 and SUSE™ Linux Enterprise Server scale more efficiently on systems with multiple CPUs compared to older operating system versions.

SMP systems are also affected by the need for locking and synchronization of resources. Before a task can modify a shared data item, it must ensure that no other task will change the data item. This is usually done by means of a lock. While a process or thread is waiting to obtain a lock, it is not productive. Further, while a thread is waiting for the lock, some of its cache lines may be replaced. Thus, when the thread is scheduled again, it may experience higher memory latency.

The operating system's kernel contains many shared data items, so it must perform synchronization internally. Synchronization delays can occur even in an application program that does not share data with other programs because the kernel services have to serialize shared kernel data. In addition to lock contention, path length also increases because more code is being executed.

Advantages of SMP

The main advantage of deploying an SMP system is the ability to use multiple processors simultaneously to execute different tasks constituting a program, thereby increasing the throughput (for example, the number of transactions per second), compared to a single-CPU system.

Only workloads that support parallelization (including multiple processes or multiple threads that can run in parallel) can benefit from SMP. Single-threaded workloads can be scheduled to only one CPU at the time; and thus, cannot take advantage of additional CPUs being available.

On the other hand, some modern applications with significant computational components have built in multi-threaded structures and a high scalability ratio. Good examples of the latter are Microsoft® SQL Server and Microsoft Exchange. Microsoft recommends deploying these applications on SMP systems with at least two CPUs.

For many applications that do support parallelization, it is possible to implement either a scale-up approach (deploying the applications on larger servers) or a scale-out approach (deploying a larger number of smaller servers and distributing the application load among them). The scale-out option presents additional challenges compared to the scale-up approach because it means having to manage additional servers. It also creates the additional overhead of application partitioning and transaction arbitration. In virtual infrastructure, the trade-offs between scaling up and scaling out approaches are somewhat different and are discussed later in this paper.

vmware®

Scheduling of Virtual CPUs in Virtual SMP

For SMP-enabled virtual machines, the virtual processors from the same virtual machine are co-scheduled. That is, if physical processors are available, the virtual processors are mapped one-to-one onto physical processors that are run simultaneously. In other words, if one virtual CPU in the virtual machine is running, a second virtual CPU is co-scheduled so that they execute nearly synchronously.

Co-scheduling ensures that the guest operating system behaves as if it were running on a dedicated server with the same number of processors. In some cases, though, guest operating systems have timing constraints that require operations on other processors to complete within a certain amount of time. So, if a virtual CPU's execution becomes too skewed from the other co-scheduled virtual CPU, ESX Server de-schedules both virtual CPUs.

Co-scheduling also improves performance by facilitating prompt communication and synchronization between the processors. For example, if one processor is spinning on a lock waiting for the other processor to release it, co-scheduling the processors allows the spinning processor to proceed as soon as the lock is released by the other processor—without the additional delay of waiting for the other processor to get scheduled. If the virtual CPU is halted, ESX Server does not need to co-schedule it. (The halted CPU is considered scheduled for the purposes of co-scheduling and does not continue consuming physical CPU resources.)

Bare metal virtualization architectures like that of ESX Server are the most conducive to implementing virtual SMP systems because the ESX Server virtualization layer has complete control of CPU scheduling. Hosted virtualization architectures (in which the virtualization application is hosted by a standard operating system that is not virtualization aware) do not have the same advantage. In those cases, scheduling of multiple virtual processors would be controlled by the kernel of the host operating system and the processors would not have the ability to be co-scheduled.

Co-scheduling can however lead to the problem of processor fragmentation on two-way systems. Consider the case where a single-processor virtual machine is running and a two-processor virtual machine is ready to run. One physical CPU is idle, but ESX Server is not able to run the SMP virtual machine with two active (non-halted) virtual CPUs because ESX Server needs two available physical processors. Therefore, a physical CPU may be left idle.

Note: The situation just described is not a problem for hyper-threaded (HT) systems. Hyper-threading is a feature of Intel® Pentium 4 and Xeon processors that allows two independent program threads to execute simultaneously. ESX Server v2.1 and later releases support hyper-threading on processors that have this feature. Hyper-threading does not provide the same performance as a two-core processor but does allow better utilization of processor resources for certain workload combinations. (For more information on hyper-threading, see the white paper available at *www.vmware.com/pdf/esx21_hyperthreading.pdf*.)

For example, with hyper-threading enabled, VMware ESX Server can dedicate one processor (with two logical CPUs) to the SMP virtual machine and another processor to the single-processor virtual machine (running on one logical CPU, with the other halted), thus fully utilizing the system's resources. This increased utilization may provide substantial performance benefits for many workloads using a mixture of SMP and single-processor CPU virtual machines.

Virtual SMP Best Practices

Virtual machines in an SMP system should be configured to use multiple CPUs only if they are running applications that are multi-threaded or implemented to use multiple processes. Also, as mentioned earlier, single-threaded workloads cannot make use of the second virtual processor in a virtual machine. Depending on the guest operating system scheduling behavior, the second virtual CPU may still consume resources and reduce the flexibility of the ESX Server scheduler, without enhancing application performance.

This best practice recommendation differs from the common practice of configuring physical enterprise servers with at least two CPUs. While in typical business environments such a practice normally enables standardization and the possibility of system reuse, in virtual infrastructure, where provisioning is virtually free, having the second CPU just in case, does not carry similar benefits. In addition, virtual infrastructure changes the nature of the trade-off between scaling up and scaling out your platforms. In many cases, ESX Server running a combination of SMP and single-virtual CPU virtual machines can utilize the physical CPUs most efficiently.

When considering whether to deploy multiple, single-CPU virtual machines running the same application or to deploy a single SMP virtual machine, you need to understand the cost of the associated application partitioning as well as the cost of managing multiple virtual machines. Typically, the hardware cost for running two single-CPU virtual machines and the hardware cost for running a single SMP virtual machine are roughly equivalent.

To redeploy some of the workloads that currently run on dual-processor physical system platforms to single-CPU virtual machines, you can use VMware P2V Assistant or third-party tools to perform the conversion. During the conversion process, the P2V Assistant replaces the original SMP kernel and hardware abstraction layer (HAL) with a uniprocessor (UP) kernel and HAL in the virtual machine. This procedure works best if the HAL of the original physical system is an ACPI (advanced configuration and power interface) HAL. While P2V Assistant also attempts to perform the conversion on systems with a standard HAL, this procedure is not recommended.

For certain Windows operating systems, if you wish to change existing SMP virtual machines to UP virtual machines, it is possible do so by changing the number of virtual processors to one (as shown in Figure 3) and modifying the HAL with appropriate validation from VMware support. Please note that, for Windows 2000, you may want to consult the following Microsoft Knowledge Base article to understand how to modify the HAL:

support.microsoft.com/default.aspx?scid=kb;EN-US;237556.

To downgrade an SMP system to one virtual processor, you need to use the ACPI Uniprocessor HAL.

Note: Manual replacement of an ACPI HAL with a standard HAL should never be attempted.

If hyper-threading is available on the physical processor platform on which ESX Server is installed, for scheduling efficiency, it's recommended that you select the HT-sharing **any** setting (the default) which permits ESX Server to schedule virtual CPUs from this virtual machine on the same processor package as any other virtual CPU. This allows the system to fully exploit hyper-threading and is the best choice for a majority of applications.

When sizing systems that run ESX Server, you should take into consideration the additional memory requirements associated with SMP virtual machines. For guidelines on allocating memory to ESX Server, refer to the *ESX Server Administration Guide* found at:

www.vmware.com/support/esx2/doc/esx20admin_res-mem-sizing-intro.html.

For more information about troubleshooting Virtual SMP and optimization options, please refer to the VMware knowledge base (articles such as #1077 and #1730) at the following location:

www.vmware.com/support/kb/enduser/std_alp.php

Configuring Virtual SMP Virtual Machines

The Virtual SMP module installs with ESX Server. Virtual SMP is not available with VMware GSX Server or with VMware Workstation.

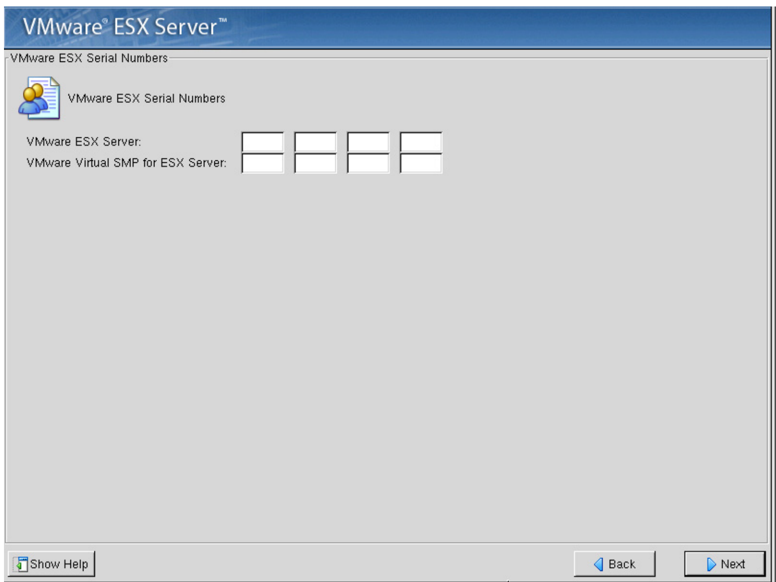

Figure 2: Installing Virtual SMP

As shown in Figure 2, when prompted by the ESX Server installation wizard, enter both the ESX Server serial number and the Virtual SMP serial number.

To add a virtual SMP module to an existing ESX Server installation, you can add Virtual SMP license keys from the ESX Server Management User Interface (MUI).

To create a new SMP virtual machine for an ESX Server installation where Virtual SMP is enabled, select the two CPU option in the virtual machine configuration wizard. Note that you are presented with this option only when selecting to configure a custom virtual machine.

You can check whether or not a virtual machine is an SMP virtual machine by reviewing the virtual hardware configuration, as shown in Figure 3.

vmware®

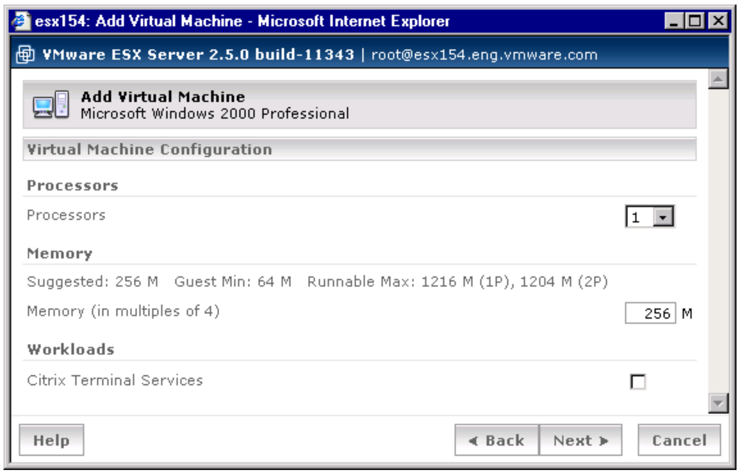

Figure 3: Configuring a Number of Virtual CPUs on a Virtual Machine

You can assess resource utilization of an SMP virtual machine with a number of VMware performance monitoring tools, such as VirtualCenter, `vmkusage` and `esxtop`. For more information on the use of these tools, go to:

www.vmware.com/pdf/esx2_using_vmkusage.pdf

www.vmware.com/pdf/esx2_using_esxtop.pdf

For example, `esxtop` displays information about the state of the physical server running ESX Server. The `esxtop` tool lists CPU and memory utilization for each individual virtual machine and for each individual virtual CPU. These CPU and memory statistics let you monitor the resource utilization for each of your virtual machines.

When you use the `esxtop` tool, two virtual CPU IDs matched to the same WID indicates that those two virtual CPUs are configured into an SMP virtual machine.

The Future of Virtual SMP

Certain pending technology developments are likely to significantly affect Virtual SMP.

For example, the introduction of multi-core processor designs will allow for more efficient scheduling. Both Intel and AMD® started shipping dual-core x86 architecture CPUs in 2005. Essentially, a multi-core processor is a set of two independent processing cores, each with its own cache, placed in a single socket. ESX Server running on a system using multi-core processors perceives each of the execution cores as a discrete processor. In contrast to Hyper-Threading, which is limited to using a single core's existing execution resources, a multi-core capability provides two or more complete sets of execution resources to increase computer system throughput.

Another future technology that is likely to affect Virtual SMP deployments is the four-way Virtual SMP by VMware. First announced at VMworld in 2004, four-way SMP is expected to optimize the deployment of highly parallel applications such as databases running on top of a virtual infrastructure. Four-way Virtual SMP should significantly expand the capability of ESX Server to handle the biggest SMP-aware workloads in virtual machines.

Conclusions

VMware® Virtual SMP™ is an add-on software module that allows VMware ESX Server to run virtual machines with multiple virtual CPUs. Multi-threaded applications such as databases, mail servers, and other applications that use multiple processes can support more load when deployed in an SMP virtual machines than when deployed in a virtual machine with a single virtual CPU.

Virtual SMP can provide a significant advantage for multi-threaded applications and applications using multiple processes in execution. However, since not all applications are able to take advantage of a second or multiple CPUs, SMP virtual machines should not be provisioned by default. For each virtual machine, system planners should carefully analyze the trade-off between possible increases in performance and throughput gained through individual virtual machines and the number of virtual machines supported by physical hardware.

The best practice for deploying Virtual SMP depends on a number of performance and configuration factors, some of which are similar to physical SMP. You should analyze your specific applications to understand if deploying them in SMP virtual machines will improve throughput. In addition, you can achieve more efficient use of the underlying hardware by deploying a mixture of single-CPU and SMP virtual machines to ESX Server platform systems.

Some key points regarding Virtual SMP deployment to take away from this white paper are:

- Virtual SMP allows a virtual machine to access two or more CPUs.
- Up to two physical processors can be consumed when you run a two-way virtual machine.
- Only allocate two or more virtual CPUs to a virtual machine if the operating system and the application can truly take advantage of all the virtual CPUs. Otherwise, physical processor resources may be consumed with no application performance benefit and, as a result, other virtual machines on the same physical machine will be penalized.

VMware, Inc. 3145 Porter Drive Palo Alto, CA 94304 www.vmware.com