

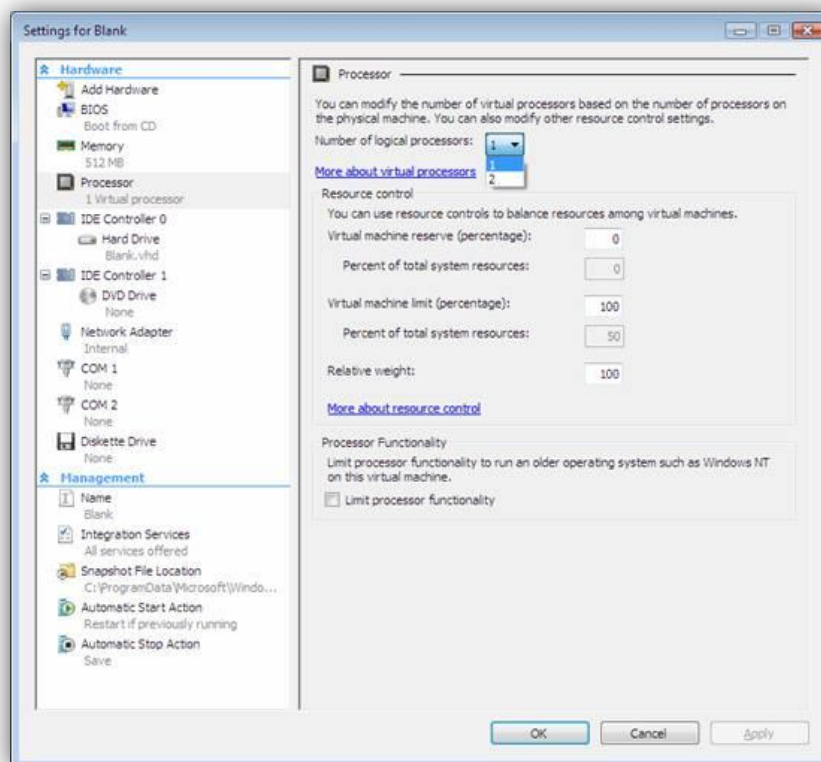
CONTROLLING PROCESSOR RESOURCES IN HYPER-V GUESTS

Windows Server 2008 Hyper-V allows up to four virtual processors in a virtual machine and allows you to configure options on how those virtual processors are balanced across virtual machines. This article describes the options for processor resource control and how to use them.

Hyper-V uses physical processors and cores to provide virtual processors to virtual machines. Each virtual machine starts with a single virtual processor, but you can increase virtual processors to 2 or 4 per virtual machine. Virtual processors are actually threads in the parent partition running on a physical processor. As each virtual machine is powered on, a separate thread is created for each virtual processor in the virtual machine. Each thread can be scheduled by the **Virtual Machine Manager (VMM)** on separate physical core or processor.

Configuring the number of processors in a virtual machine is done from the virtual machine settings dialog. While you can view the number of virtual processors when the virtual machine is running, you cannot change the number of available processors until the virtual machine is in a powered off state. **Figure 1** shows the settings dialog with the Processor hardware node selected, you can see the right hand side has the ability to select the number of processors, in this case 1 or 2 processors are available.

This is determined by the number of cores that are available in the parent partition. To be given an option to select 4 processors in a virtual machine the parent partition must have 4 cores (*or 4 processors if not multiple cores*).



RESOURCE CONTROL

Virtualization allows you to over subscribe the processing limits of the physical hardware. I could only have 4 cores on the physical server, but create and have more than 4 virtual machines running. Each virtual machine would be sharing a virtual processor in the parent partition, but with no limits imposed a virtual machine could consume an entire core in the machine.

For example, if I allocate a virtual machine 1 processor on a machine that has a single quad core processor, that virtual machine will have one thread that can consume the equivalent of an entire core or processing time. If it has 2 processors configured it can consume two cores. If you configure the virtual machine to have 4 cores, it could attempt to consume all the processing power of the server and starve the other virtual machines.

The **Virtual Machine Manager (VMM)** manages the scheduling of threads of all the running virtual machines. By default it attempts to balance the processing evenly across all the cores in the physical machine to get a load balanced processor distribution. As discussed above, it is possible for a virtual machine to consume entire processor cores and starve other threads running on a core.

While **Hyper-V** does not have processor affinity (*the ability for the admin to specify the processor core that a thread runs on*), you have the ability to tell **VMM** that you want to apply resource control on a per virtual machine basis to set limits on the virtual machines ability to starve other virtual machines on a single core.

Hyper-V accomplished resource control in different ways:

- Setting a reserve on processing resources using a percentage
- Setting a maximum on processing resources using a percentage
- Setting a relative weight of the virtual machine to others in the system

VIRTUAL MACHINE RESERVE

The virtual machine reserve allows you to specify the percentage of the assigned virtual processor that this virtual machine will be guaranteed on the physical host. This value can range from 1-100% and is relative to the number of processors assigned to the virtual machine. For example, if a physical host has 4 cores and you assign a single processor to a virtual machine, that virtual machine can potentially consume up to an entire processor core of processing time, but it is not guaranteed that the processing time is available at all times.

By setting the **reserve value to 100%**, a virtual machine will be reserved the equivalent of an entire processor core. If that virtual machine sits idle at 10% most of the time, the other 90% of the processing time is still unavailable to any other virtual machine. Using *reserve resource control will limit the amount of available virtual processor resources that can be shared on a Hyper-V host* and therefore limit the number of concurrent virtual machines you can power on. If you have 20 virtual machines configured with a single processor on a host with 4 cores and you have each of them set to a reserve of 100%, you can only power on 4 virtual machines.

Reservation should only be used if you want to guarantee a virtual machine processing power. This is typically used on virtual machines that you know will require lots of processing power and has spikes where having a guarantee is extremely important.

VIRTUAL MACHINE LIMIT

Virtual Machine limits are the opposite of virtual machine reserve, it allows you to specify the maximum amount of processing power that a virtual machine can consume. This value can range from 100-1% and is relative to the number of processors assigned to the virtual machine. For example, if a physical host has 4 cores and you assign a single processor to a virtual machine, that virtual machine can potentially consume up to an entire processor core of processing time, but there is no limit, so the virtual machine can consume the entire core. By setting the limit value to 10%, that virtual machine will be limited to a maximum of 10% of an entire processor core.

If that virtual machine has a spike in processing it will be limited to no more than 10% of the core and therefore will suffer in performance. Using limit resource control will limit the amount of available virtual processor resources that can be consumed on a **Hyper-V** host and allow you to control the number of concurrent virtual machines you can power on and have defined amounts of processing power per virtual machine.

If you have 20 virtual machines configured with a single processor on a host with 4 cores and you have each of them set to a limit of 10%, you have only consumed the equivalent of two processor cores and can still power on another 20 virtual machines (*assuming you have enough memory and disk resources*). This also means that if you power on only a single virtual machine it can never consume more than 10% of a single core, so you are performance limiting the virtual machine even when you have excess processing capacity.

Processor limits should only be used if you want to limit a virtual machine processing power. This is typically used by web service providers who want the ability to maximize the number of virtual machines on a host, but provide a specific level of service.

RELATIVE WEIGHT

Relative weight allows you to specify a virtual machine has processing priority without applying a specific limit or reserve. This value can range from 0-10000.

Relative weight is used to make a determination of who should get processing resources when multiple requests are being made at a time.

For example, by default if you have 4 virtual machines running, all 4 will get equal sharing of the available processing power because they all have the same relative weight. If you have a machine that is more important than the others and want to give that virtual machines request for processing power priority over others, you can assign that VM a higher weight than the other virtual machines.

This means that if a virtual machine with a higher weight needs resources, it gets them, but if it is not using them, other virtual machines can use them. This is all still limited by the number of processors assigned to the virtual machine. If you assign a virtual machine a single virtual processor and a high **relative weight**, it can still only consume a maximum of a single core.