vmware®

VMware ESX Server 2

Using vmkusage to Isolate Performance Problems

The web-based `vmkusage` utility displays historical graphs that show physical server, ESX Server system and virtual machine statistics. By default, these graphs show the most recent, daily and weekly data. The `vmkusage` utility gathers the statistics it uses to generated the graphs from the ESX Server machine's `/proc` nodes.

Samples are taken every minute and the graphs are regenerated every five minutes with the average of one-minute statistics over five-minute intervals.

The web pages generated by the `vmkusage` utility show CPU utilization, memory utilization, network use and disk use (both I/O operations and bandwidth). There are system aggregate, per-virtual machine, storage and networking-focused views of the data. You can use this information to isolate performance problems or observe longer-term trends for capacity planning.

The `vmkusage` utility's disk usage and execution speed are non-intrusive. Typically, disk usage is around .5MB per virtual machine and execution speed is 180 to 200 ms per minute for data collection, and three to five seconds every five minutes for graph generation.

The `vmkusage` utility is designed to be run as a cron job which is installed with the `vmkusagectl` auxiliary utility. It can also be run from the ESX Server console command line or from a script. This paper describes running `vmkusage` as a cron job.

The following sections describe the process of using `vmkusage` to troubleshoot a suspected performance problem.

Starting vmkusage

The `vmkusage` utility is installed by default beginning with ESX Server 2.1, but is not started until you use the `vmkusagectl install` command.

Note: If you have not upgraded to ESX Server 2.1 or later, see the VMware Knowledge Base article 1078 for instructions on how to download, install and start `vmkusage`.

Perform the following steps to start `vmkusage` as a cron job.

1. Log on to an ESX Server console as root.

2. Run the `vmkusagectl install` command.

The syntax for `vmkusagectl` is:

```
vmkusagectl [options]
```

The available options include:

Option	Description
install	Starts `vmkusage` and removes any existing `vmkusage` databases.
uninstall	Stops `vmkusage`.
cleandb	Removes any existing `vmkusage` databases.
support	Creates a compressed tar file (tar ball) containing the `vmkusage` database, graphs and HTML files. You can un-tar this file on any machine and view the graphs in a Web browser because the links are relative.
save	Saves the `vmkusage` database to a file.
hyper	Collects statistics every ten seconds and generates graphs every five minutes. The default is to collect statistics every minute and generate graphs every five minutes. **Note:** Hyper mode consumes additional CPU resources because of increased sampling rate and should be used only to diagnose problems, not in production.
*	Prints `vmkusagectl` usage to the screen.

Using vmkusage

To view the performance history of an ESX Server machine, open a Web browser and enter the URL of the ESX Server machine, using the format `http://<ESXservername>/vmkusage`. For example, entering the following sample URL:

```
http://esxserver1.svc.com/vmkusage
```

Causes its **vmkusage** Web page displaying an overview of the system to appear.
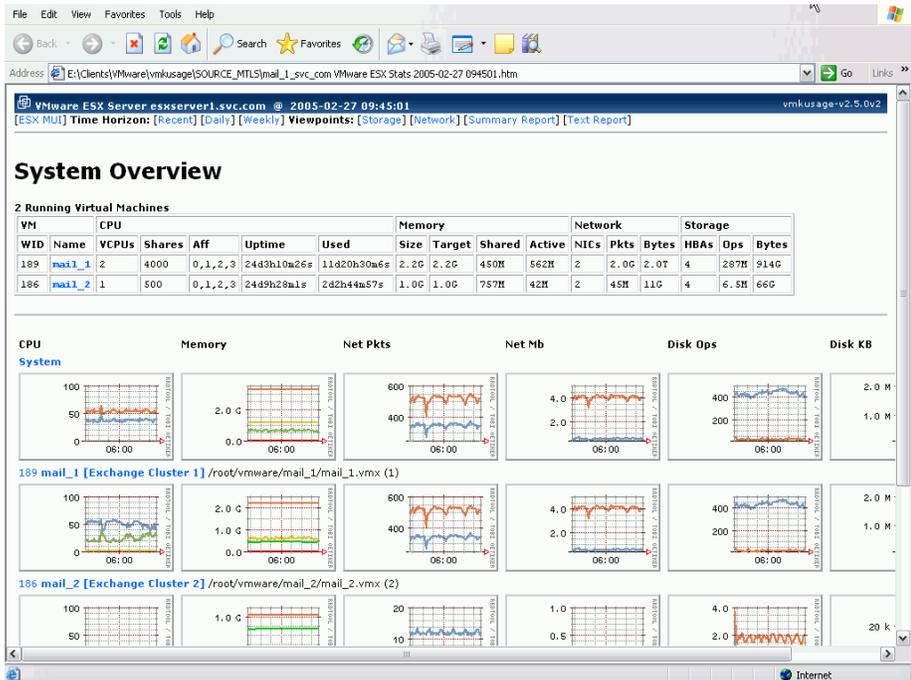
The first row of graphs, labeled **System**, shows total system resource utilization. The subsequent rows of graphs. labeled with their respective virtual machine's WID and name, show per-virtual machine resource utilization. The virtual machine labels are links to pages that show their graphs in greater detail. In this example, if you click the first virtual machine's label, the following page appears:

Analyzing CPU Usage

This section describes how to analyze system CPU loading, percentage of individual CPU use and individual virtual machine CPU use.

The following is an example of a **Recent** historical graph for a dual-CPU virtual machine.
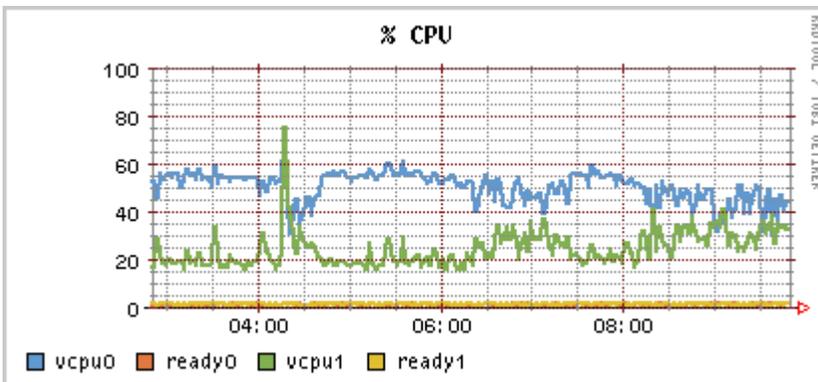
The CPU time is split into %CPU for VCPU0 (as a percentage of a single CPU) and %CPU for VCPU1.

Examine **ready0** and **ready1** for the percentage of time that the virtual machine was ready but could not get scheduled to run on a physical CPU. Under normal operating conditions, this value should remain under 5%.

Examine **vcpu0** and **vcpu1** for the percentage of physical CPU resources used by a VCPU. If the physical CPU resources are running at full capacity, you can use these percentages to identify a virtual machine that is using a large amount of physical CPU resources. If a VCPU consistently saturates the physical CPU resources, possible fixes include:

- Decreasing the number of virtual machines running on the ESX Server machine.
- Moving the virtual machine to a different ESX Server machine that has more physical CPU resources.

  Note: Possible causes of application problems for a VCPU that consistently saturates the physical CPU resources include:

  - A screen saver is running (you should turn off screen savers on virtual machines).
  - The application has a software bug.
  - The application might not be a candidate for virtualization.

Analyzing Memory Usage

This section describes how to analyze memory usage.

The `vmkusage` utility displays the total amount of memory in use. The maximum of the `vmkusage` plot is the total amount of memory currently allocated on the host.

The following figure is an example of a **Recent** historical graph for a dual-CPU virtual machine. The graph shows memory usage where size is the maximum amount of memory available to the virtual machine.
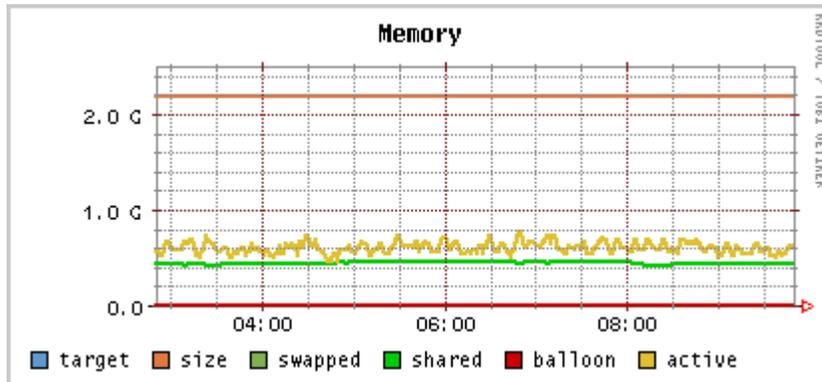
The memory graphs in `vmkusage` represent the memory the virtual machines are consuming, not the amount of physical memory (after sharing) used.

The descriptions of the graph lines are:

| Graph Line	Description
target	Target size of how much memory should be allocated to the virtual machine.
size	Current amount of memory the virtual machine has (size should approach target, and they are often equal).
swapped	Amount of swap space used by the virtual machine.

vmware®

Graph Line	Description
shared	Amount of memory this virtual machine has marked shared, and can be shared either with itself or other virtual machines.
balloon	Amount of memory ballooned out by the ESX Server machine.
active	Amount of its allocated memory that the virtual machine actively used.

For example, when the ESX Server machine is short on memory, you usually see from the graphs that the balloon driver starts using a lot of memory. If you are hitting swap and not seeing the balloon driver using a lot of memory, you should check that the VMware tools are installed correctly on your guest operating systems.
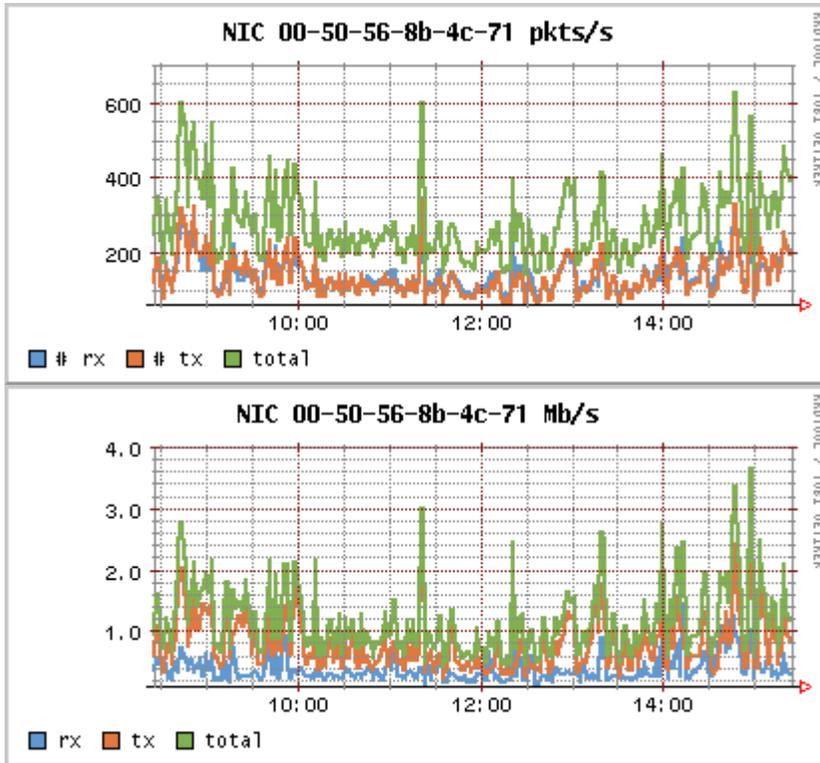
Assessing Disk and Network Usage

This section describes how to analyze disk and network usage.

The following figures are examples of a **Recent** historical graph for a dual-CPU virtual machine.
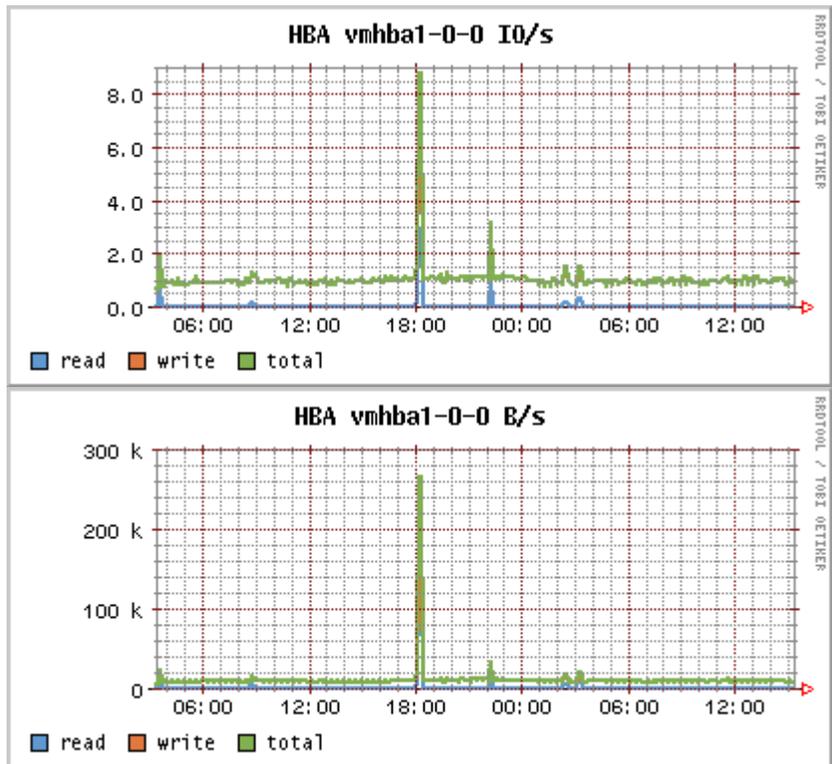
You see a graph for each vmhba showing both number of I/Os and bytes per second that are read and written.

You see a graph for each NIC showing both the number of packets and the MBytes per second that are transmitted and received.

Examine the disk and NIC activity. If the throughput either is not meeting expectations or is approaching the maximum capacity of the hardware device, there is a potential for a performance bottleneck.