



Global Knowledge®

Expert Reference Series of White Papers

Performance and Troubleshooting with esxtop

Performance and Troubleshooting with esxtop

Steve Baca VCP, VCI, VCAP, Global Knowledge Instructor

Introduction

This paper introduces and gives examples of how the **esxtop** utility can help address performance issues. First, we will discuss the history of **esxtop** and show several different methods that can be utilized to start the monitoring tool. Next, we will discuss how to use **esxtop** by using interactive commands that can be typed in while **esxtop** is running. Finally, we will look at how to use **esxtop** is given by looking at how to interpret CPU data utilizing the **esxtop** utility.

History

The **esxtop** command is a tool based upon the old UNIX command-line tool called **top** that continuously updates every five seconds, displaying a snapshot of the processes running on an ESXi host. The **top** program has been around since the mid-1980s and has been ported to many different versions of UNIX and Linux. Originally, VMware ported a version of the UNIX **top** program and customized it to gather statistics for the ESX host, the standard **top** program was included in the service console as well. When VMware changed the direction of its hypervisor and removed the service console, **esxtop** continued to be a useable command-line utility within the ESXi hypervisor, which runs a proprietary version of UNIX. VMware also modified **esxtop** to run remotely and called it **resxtop**. The remote **resxtop** runs within the vCLI, and allows the user to remotely connect to an ESXi host and run **esxtop**.

esxtop/resxtop

The **resxtop** command is used when you want to run **esxtop** remotely from the vSphere command-line interface (CLI) using vCLI, usually within the vMA. The **resxtop** utility is referred to as remote **esxtop** and offers a secure method to run scripts across multiple ESXi hosts and virtual machines. This paper concentrates on how to use **esxtop**, since once **resxtop** is started all of the counters and fields are the same.

Using esxtop in Batch Mode

The **esxtop** command can also be run in batch mode, which allows statistics to be collected and saved into a file, then played back at a later point in time. The data can be read using the Windows **Perfmon** utility or Microsoft Excel. To start running **esxtop** in batch mode use the following syntax.

```
# esxtop -a -b > outputfile.csv
```

-a show all of the statistics

-b stands for batch mode

> outputfile.csv redirect the output to the file and the file must end with **.csv**

To stop processing in batch mode do **Ctrl+C**.

Using esxtop in Interactive Mode

By default, **esxtop** runs in interactive mode, which initially begins by typing in **esxtop** at the command line. Depending on what system you are running on, you might have to set the **terminfo** database to **xterm**.

```
# export TERM=xterm
```

```
# esxtop
```

Once you launch **esxtop** you will see a default screen (Figure 1), I included callout descriptions to some of the main host statistics and fields. The **esxtop** output can show more information than you will need for the performance or troubleshooting problem that you are addressing. There are also interactive commands that can be issued to customize the display, which will be shown in Figure 3. Figure 1 is an example of the output generated from **esxtop** or **resxtop**. There are several screens that can be viewed. The default screen is always the CPU view as shown in the screen shot Figure 1, and the screen refreshes every five seconds by default. The **esxtop** displays statistics based on worlds. A world can be defined as schedulable entity, and other operating systems would call it a process. Each virtual machine will have multiple worlds running based on several factors. There will be one world for each of the vCPUs running on the VM. There will be a world for the VM's MKS, and a world for the virtual machine monitor (VMM) of the world.

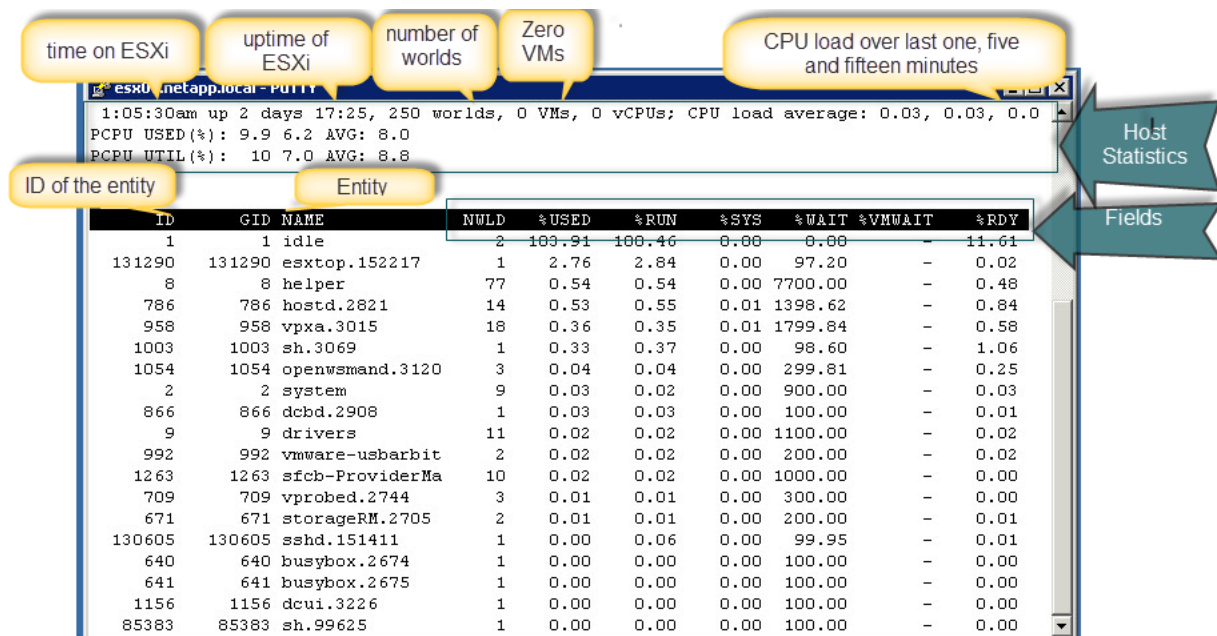


Figure 1. Esxtop outlining main statistics and showing location of fields

Screen Views with the esxtop Utility

The default view when **esxtop** is launched is going to show information for the CPU. You can change the screen view by simply typing in a corresponding letter for the view that you are interested in inspecting. Here is the list of views that you can switch to by simply typing in the letter associated with the view.

- **c**: CPU view which is the default view
- **m**: Memory view
- **n**: Network view
- **d**: Disk adapter view
- **u**: Disk device view
- **v**: Disk VM view
- **i**: Interrupts
- **p**: Power management

For example, if you want to switch from looking at the CPU view information to looking at the memory view, simply type in the letter **m** to make the switch. Figure 2 shows the memory view.

```

esx01.netapp.local - PuTTY
2:06:14am up 3 days 18:26, 249 worlds, 0 VMs, 0 vCPUs; MEM overcommit avg: 0.00, 0.00, 0.00
PMEM /MB: 6143 total: 805 vmk, 103 other, 5234 free
VMKMEM/MB: 6115 managed: 326 minfree, 2112 rsvd, 4002 ursvd, high state
PSHARE/MB: 18 shared, 18 common: 0 saving
SWAP /MB: 0 curr, 0 rclmtgt: 0.00 r/s, 0.00 w/s
ZIP /MB: 0 zipped, 0 saved
MEMCTL/MB: 0 curr, 0 target, 0 max

  GID NAME                MEMSZ  GRANT  S2TGT  TCHD  TCHD W  SWCUR  SWTGT  S
-----
786 hostd.2821            54.77  44.41  0.00   44.41  44.41   0.00   0.00
958 vpxa.3015             22.14  14.49  0.00   14.49  14.49   0.00   0.00
1261 sfcb-ProviderMa     14.90  12.39  0.00   12.39  12.39   0.00   0.00
1263 sfcb-ProviderMa     10.07   6.79  0.00    6.79   6.79   0.00   0.00
932 vobd.2975             8.94   1.23  0.00    1.23   1.23   0.00   0.00
604 vmsyslogd.2620        6.68   4.53  0.00    4.53   4.53   0.00   0.00
603 vmsyslogd.2619        4.62   3.71  0.00    3.71   3.71   0.00   0.00
1362 sfcb-ProviderMa      3.74   1.50  0.00    1.50   1.50   0.00   0.00
1257 sfcb-ProviderMa      3.68   1.71  0.00    1.71   1.71   0.00   0.00
1156 dcui.3226            3.48   1.05  0.00    1.05   1.05   0.00   0.00
611 vmkeventd.2640        3.01   1.18  0.00    1.18   1.18   0.00   0.00
1262 sfcb-ProviderMa      2.79   0.67  0.00    0.67   0.67   0.00   0.00
671 storageRM.2705        2.73   1.36  0.00    1.36   1.36   0.00   0.00
709 vprobed.2744          2.56   0.74  0.00    0.74   0.74   0.00   0.00
85383 sh.99625            2.53   0.12  0.00    0.12   0.12   0.00   0.00

```

Figure 2. Default esxtop screen when first started

Help Screen

To learn more about other options you can choose, type in **h** to get the help view for **esxtop**.

```

esx01.netapp.local - PuTTY
space - update display
h or ? - help; show this text
q - quit

Interactive commands are:

fF      Add or remove fields
oO      Change the order of displayed fields
s       Set the delay in seconds between updates
#       Set the number of instances to display
W       Write configuration file ~/.esxtop50rc
k       Kill a world
e       Expand/Rollup Cpu Statistics
V       View only VM instances
L       Change the length of the NAME field
l       Limit display to a single group

Sort by:
      U:%USED      R:%RDY      N:GID
Switch display:
      c:cpu        i:interrupt      m:memory        n:network
      d:disk adapter  u:disk device    v:disk VM       p:power mgmt

Hit any key to continue:

```

Figure 3. Displays the help screen interactive commands

Calculating Performance Counters

The performance counters are calculated in different ways. The counters or statistics types can be a **Rate**, **Delta**, or **Absolute** value. **CPU Ready** is a Delta, which is the change from the previous interval. As an example, some counters are calculated as the delta between two successive snapshots or intervals. The **%Used** is a good example of a Delta.

%Used = (Total CPU used time at the second snapshot – Total CPU used time at the first snapshot) / time elapsed between snapshots

To help understand the **esxtop** output it helps to define fields and counters that you are viewing.

World – Is a schedulable entity

ID – World Identifier

GID – World Group Identifier

NWLD – Number of Worlds for an entity

CPU Load Average – is the mean of CPU loads in 1 minute, 5 minute, and 15 minutes, base on 6 second samples.

Interpreting CPU Activity using Esxtop Utility

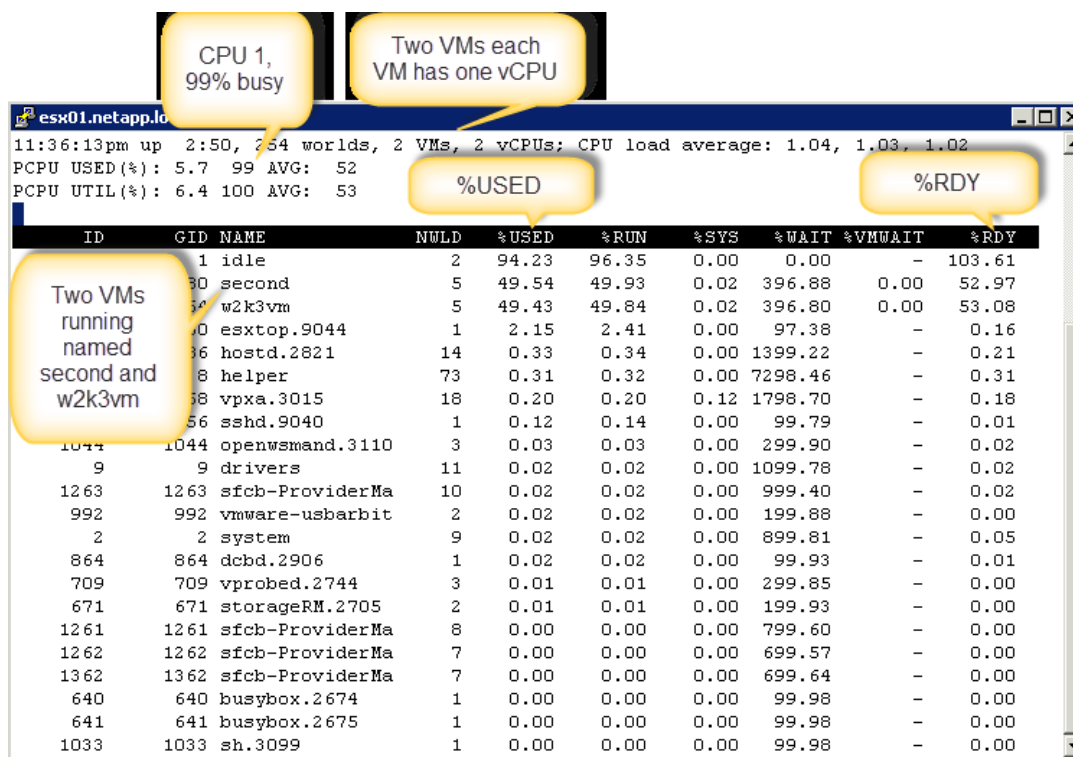


Figure 4. Displays the CPU screen with VMs running

Figure 4 shows CPU activity for the ESXi host and there are two VMs running on the system named **second** and **w2k3vm**. In order to create contention on the CPU both VMs have a CPU affinity set for CPU 1 and are running a math application in a loop, which is generating 99% busy. If you look at the **%USED** for both VMs, they are both running at a little more than 49%, since they are competing equally for the same PCPU. Another interesting field that is used for performance monitoring when it comes to CPU issues is the **%RDY** field. The **%RDY** field is the percentage of time that the world was ready to run, but was waiting for its turn. In the example above, the two VMs, second and w2k3vm, have a **%RDY** time a little greater than 50%, which is extremely high. Normally, I become concerned if I see a steady value greater than 10%. If the **%RDY** is greater than 10%, I would look to see if **%MLMTD** is high as well. If **%MLMTD** is high, it would signify that a CPU Limit has been set on the VM and needs to be investigated. In addition, there is a field called **%WAIT** that shows wait and idling time together.

CPU Statistics

PCPU USED% – CPU utilization per physical CPU (includes logical CPUs)

%USED – CPU Utilization. The percentage physical CPU time accounted to the world.

The formula is: **%USED = %RUN + %SYS - %OVERLP**

It is possible that the **%USED** of a world can be greater than 100%, if the system service runs on a different PCPU for this world.

If the **%USED** of a VM is high, that means the VM is using lots of CPU resources, which can be normal.

%RDY – The percentage of time the world was ready to run, but was not provided the CPU resources. A world in a run queue is waiting for the CPU scheduler to let it run on a PCPU. If **%RDY** of a VM is high, it means the VM is possibly under resource contention. Check **%MLMTD** as well. If **%MLMTD** is high, you may raise the CPU Limit setting for the VM. If **%RDY - %MLMTD** is high, the VM is under CPU contention.

%MLMTD – The percentage of time the world was ready to run but deliberately was not scheduled because that would violate the CPU Limit setting. What does it mean if **%MLMTD** of a VM is high, the VM cannot run because of the CPU limit setting.

%SYS – The percentage of time spent on the ESXi VMKernel running process interrupts and other system services on behalf of the world.

%IDLE – The percentage of time the vCPU world is in an idle loop.

%CSTP – The percentage of time the vCPUs of a VM are spent in the co-stopped state, waiting to be co-started.

%SWPWT – The percentage of time the world is waiting for the ESXi’s VMKernel to swap memory. If **%SWPWT** is high, then the VM is swapping memory.

%RUN – The percentage of total scheduled time for the world to run. If **%RUN** of a VM is high, the VM is using lots of CPU resources, but does not necessarily mean the VM is under resource constraint.

%WAIT – The percentage of time the world spent in the wait or idle state. This **%WAIT** is the total wait time, the world is waiting for some VMKernel resource. The **%WAIT** time can be high because there are many worlds waiting for events to happen, and the total wait time can be high due to the large number of worlds waiting on events.

Summary

The **esxtop** utility provides detailed performance data for an ESXi host. This real-time data gives the system administrator information that aids in detecting performance issues. To better interpret **esxtop** data, it helps to understand how to setup the **esxtop** view with the appropriate fields. When dealing with CPU performance problems for a VM, one of the first fields to observe is **%RDY**. If this field is larger than 10%, it could mean that you have more requests for CPU processing than resources available. Thus, **%RDY** time is the best indicator of possible CPU performance issues.

Learn More

To learn more about how you can improve productivity, enhance efficiency, and sharpen your competitive edge, Global Knowledge suggests the following courses:

[VMware vSphere: Fast Track \[V5.1\]](#)

[VMware vSphere: Optimize and Scale \[V5.1\]](#)

Visit www.globalknowledge.com or call **1-800-COURSES (1-800-268-7737)** to speak with a Global Knowledge training advisor.

About the Author

Steve Baca has been working in the Information Technology field for more than 15 years, after graduating from the University of Nebraska with a Bachelors degree in Computer Science and Mathematics. After spending time programming and doing Systems Administration, Steve has been doing technical training for VMware, Netapp, Sun Microsystems, and Symantec.