Demystifying CPU Ready (%RDY) as a Performance Metric

Don't Trust Available CPU

By David M. Davis

Contents

Introduction

CPU Ready is a critical statistic that VM admins need to know, understand and monitor in their vSphere infrastructure. However, it is often a silent killer of virtual machine performance because it's either overlooked or misunderstood. This white paper is a deep dive into CPU Ready in an effort to demystify CPU Ready and explain why CPU Ready is a very important performance metric. Specifically, this white paper discusses the following topics:

- CPU Ready overview
- Why VM admins should not rely solely upon CPU Usage for understanding CPU performance
- Common CPU Ready misconceptions
- How much CPU Ready is "normal" and what causes CPU Ready
- How to resolve CPU Ready issues

What is CPU Ready?

The term "CPU Ready" is somewhat of a misnomer. It could be easily thought that this metric is how much CPU is ready for use and to have high amounts of "CPU Ready" could be a good thing. However, the more CPU Ready you have, the worse off your vSphere infrastructure is and the more your applications are suffering.

The official VMware definition of CPU Ready is:

> *The % of time a "world" is ready to run and awaiting the CPU Scheduler for approval. In vSphere a "world" is a process. The larger CPU Ready is, the more your virtual machines (and applications) aren't doing what they are supposed to do.*

To put it another way, a "world," in the CPU scheduler run queue, is waiting for the CPU scheduler to let it run on the physical CPU and the "world" is a virtual CPU that the virtual machines use to process application data. How long that vCPU waits to be scheduled to run on the physical CPU core, on the server, is the CPU Ready time.

What Causes High CPU Ready?

While the cause of high CPU usage can seem easy to identify, the cause of CPU Ready can be a little more elusive. The two most common causes of high CPU Ready are high CPU oversubscription and setting CPU limits.

Here is a full list of possible causes of high CPU Ready values:

CPU Oversubscription – The most common cause of high CPU Ready is oversubscribing the number of physical CPUs on the host with too many vCPUs being active. Keep in mind, that it is very normal and safe to have more vCPUs allocated than pCPUs, but the higher the ratio, the more difficult it is for the ESXi scheduler to perform its tasks without effecting performance.

At what point is the ESXi scheduler "juggling too many balls"? There is no universally agreed upon number, but some guidelines for vCPU to pCPU ratio are:
- 1:1 to 1:3 is no problem
- 1:3 to 1:5 may begin to cause performance degredation
- 1:5 or greater is often going to cause a problem

In Figure 1, the host has a single quad core CPU with hyperthreading enabled. Thus, the host has 8 logical to work with. However, as Figure 2 demonstrates, the VMs on this host have a total of 13 vCPUs allocated. If all of those VMs were powered on, using the logical core value, this host would have an 8:13 vCPU to pCPU ratio, putting this host well under the danger zone of a 1:5 ratio. Using the aforementioned guidelines, somewhere between 24 (1:3 ratio) and 40 (1:5 ratio) vCPUs, CPU Ready may become unacceptable.

Figure 1

Figure 2

CPU Limits – Putting limits on virtual machine CPU is not a best practice and should only be used in certain situations. CPU limits can easily cause high CPU Ready values. In esxtop there is another counter called %MLMTD which is the percentage of time the VM is ready to run but isn't scheduled because it would violate the CPU Limit set. Since %MLMTD is added to CPU Ready time, this simple formula can be used: %RDY – %MLMTD. Normally, %MLMTD should be 0.000% and limits should not be used in order to prevent these kinds of CPU Ready issues.
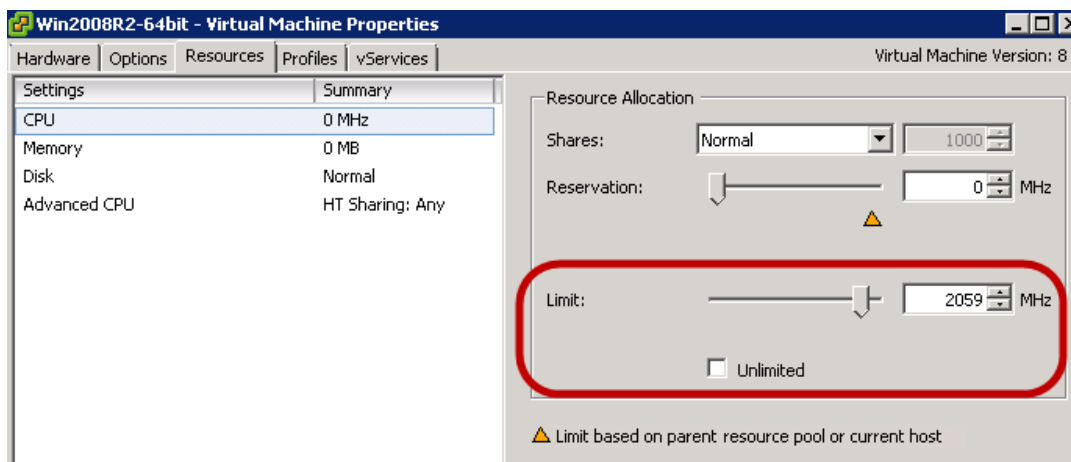

Figure 3

CPU Affinity – Utilizing CPU affinity can be dangerous because you are making decisions for the ESXi CPU scheduler without knowing the full picture. While utilizing CPU affinity on a single VM for a test probably isn't a problem, utilizing CPU affinity across numerous VMs should not be done and will likely result in high CPU Ready values. It is also important to keep in mind that a VM moved with vMotion loses its CPU affinity value even when moved back to the original host.

Fault Tolerance (FT) – In extreme scenarios, if the fault tolerance (FT) network between the primary and the replica host can't keep up with the volume of changes to the protected VM that VM will actually have its CPU throttled down. In other words, the CPU %MLMTD will be increased by ESXi to slow down the CPU so that FT can keep up with the changes. Of course, this usage of %MLMTD will result in high CPU Ready and this can be difficult to track as it will happen sporadically as the FT network is congested and relieved. Best practices dictate that there should be a dedicated, high-bandwidth, FT network between primary and secondary hosts to prevent these types of problems.

Common CPU Ready Misconceptions

There are a couple of common CPU Ready misconceptions. The first is hyperthreading and how it impacts performance. The ESXi scheduler considers a CPU with hyperthreading enabled as a full core. However, the hyperthreaded core doesn't offer 100% performance capacity. It would not be accurate to consider a single CPU with hyperthreading as providing the performance of two full CPU cores. In other words, with hyperthreaded CPUs only offering 75% the performance of a real CPU core and it is important to be very conservative when calculating the CPU oversubscription on hosts with CPU hyperthreading in use.

Second, a large amount of available CPU MHz and GHz on an ESXi host does not necessarily mean that you are operating without CPU Ready issues. Active usage doesn't cover how many cores are being used by virtual machines at any given time or preventing other VMs from being able to do so. Likewise, a VM's CPU usage and CPU Ready metrics are not directly correlated. A VM can have a serious issue with CPU Ready when its usage doesn't appear to be that high. In order to have a full picture of your CPU performance, you need to look at both CPU Ready and CPU usage.

A common question about CPU Ready is whether or not it can be relieved by using VMware's Distributed Resource Scheduler (DRS). The short answer is "no," DRS won't help with CPU Ready issues. The reason for this is that DRS does not take CPU scheduling into account. DRS just monitors CPU and memory usage when determining where to place virtual machines and does not measure pCPU contention between allocated vCPUs.

Where to Look for CPU Ready

To check how much CPU Ready is being experienced in your environment, it's most effective if you look at CPU Ready on a per-vCPU basis, at the VM level. As shown in Figure 4, this can be done using the vSphere Client by going to "Chart Options" while viewing performance graphs and selecting CPU Ready as a viewable metric.
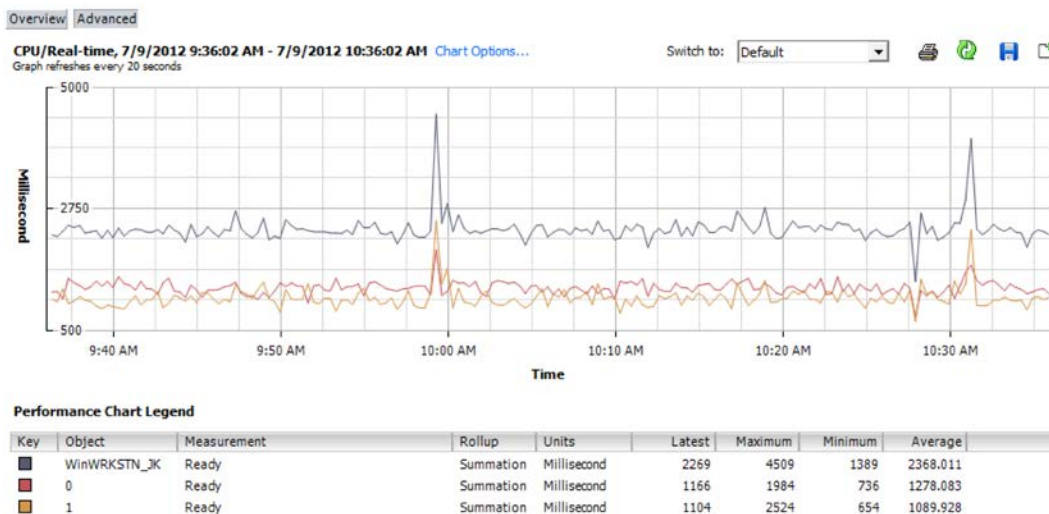

Figure 4

In the vSphere Client, measuring CPU Ready at a host level can be tricky and often impossible. Looking at Figure 5, you can see how CPU Ready is displayed for hosts; it's a summation of all CPU Ready across all vCPUs on all VMs. This would make it very hard to understand what a "bad" value would be because it would depend on too many variables that would differ from host to host and environment to environment.
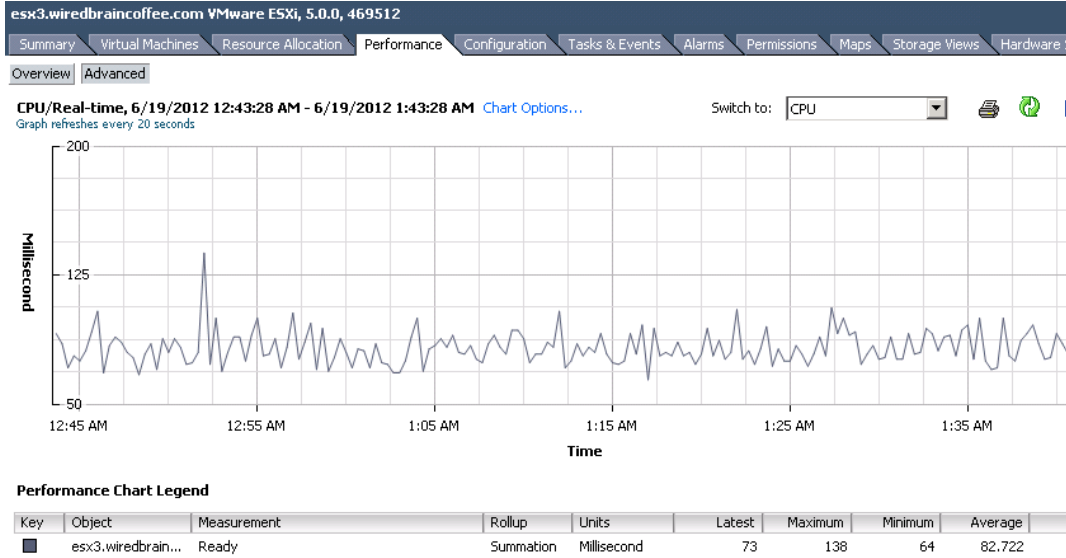

Figure 5

Using ESXTOP to measure CPU Ready on a host can be slightly more effective since it breaks down the %RDY on a per-VM basis; however, the number of vCPUs allocated to each of those VMs is still a variable that must be account-ed for. The next section will detail why this is.
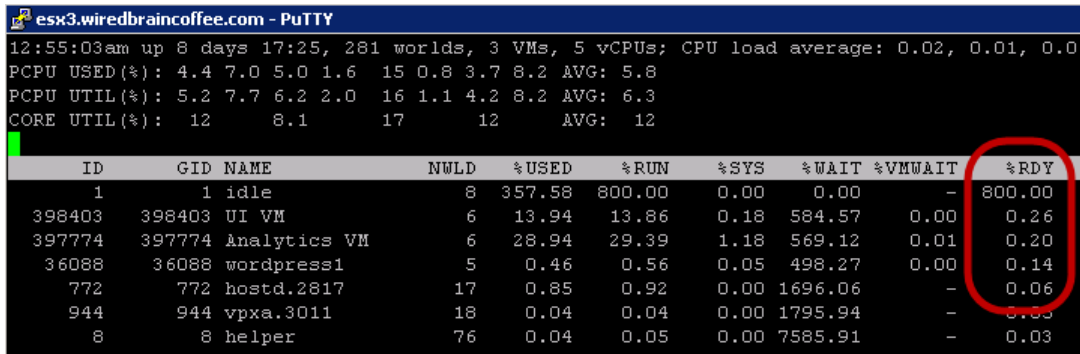

Figure 6

How Much CPU Ready is "Normal"?

While it is easy to look at CPU usage and understand that you are using 25% of 100% total capacity, it is a little more difficult to understand what is normal versus bad when looking at CPU Ready.

For starters, CPU Ready is measured in the vSphere Client is measure in milliseconds (ms). This needs to be reconciled with VMware's best practice guidelines that indicate it is best to keep your VMs below 5% CPU Ready per vCPU. When trying to make this reconciliation, it's important to note that vSphere performance graphs are based upon 20-second data points. Using Figure 7 as an example, to convert this to a percentage value, you have to take what is reported (2173 ms) and divide by 20 seconds (20000 ms) to arrive at a CPU Ready value as a %. In the case of this graph, 2173/20000 = 0.10865, or 10.865%, which is twice the 5% guideline.
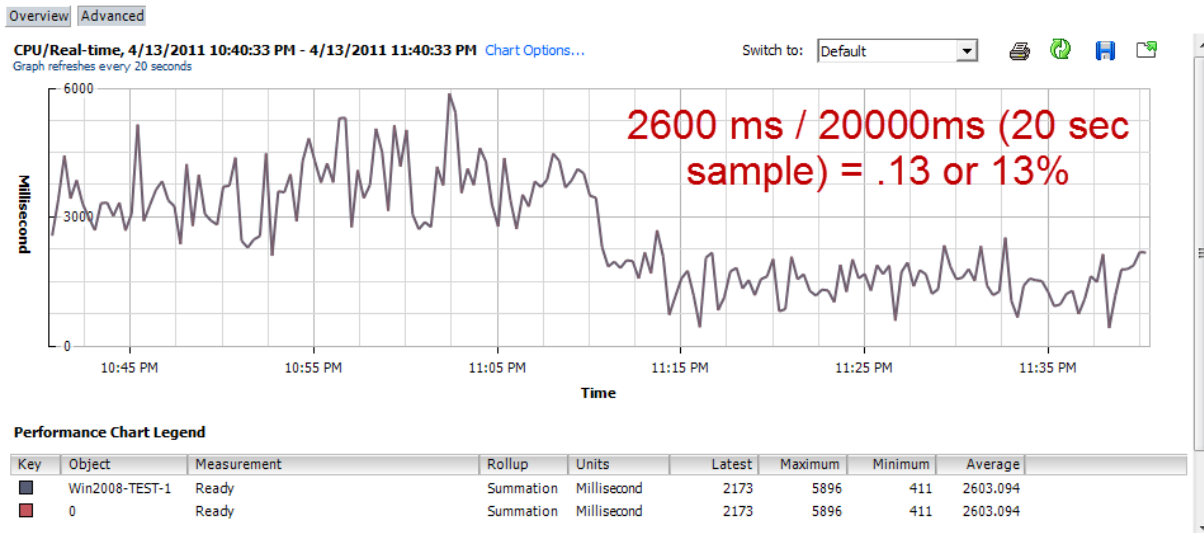

Figure 7

There are a couple of other considerations to make when measuring CPU Ready. First, know that an environment will always have some % of ready time. No virtual CPU has 100% access to the physical CPU as the scheduler takes vCPUs on and off the pCPU. In fact, every 25ms the ESXi CPU scheduler evaluates whether or not a vCPU world should be on or off a CPU core. And secondly, what is considered a problematic CPU Ready value might differ from the aforementioned guidelines depending on the role of the virtual machine. Certain types of applications – particularly CPU intensive ones - will be much less tolerant to higher CPU Ready than others.

Five Ways to Fix CPU Ready Issues

How do you fix your CPU Ready issues when you find them? Here are 5 ways to get CPU Ready resolved quickly (listed in order or importance):

1. **Proper sizing of virtual machines and hosts** – Ensure that virtual machines aren't significantly oversubscribing the pCPU of your ESXi hosts with too many vCPUs. Additionally, if hosts are truly reaching their limits in terms of pCPU usage then consider adding additional pCPUs to the hosts or adding additional hosts to your cluster.

2. **Don't use CPU limits** – Limits cause CPU Ready issues and should only be used for short-term testing or troubleshooting.

3. **Don't use CPU affinity rules** – CPU affinity rules can cause CPU Ready issues and, like limits, should only be used for short term testing or troubleshooting.

4. **Use best practices with Fault Tolerance (FT)** – Ensure that your FT implementation has its own FT network to communicate changes and isn't congested so as not to prevent CPU Ready issues.

5. **Understand what DRS does and doesn't do** – DRS doesn't help with CPU Ready and so ensure that DRS doesn't actually cause CPU Ready issues.

Automate CPU Ready Issue Resolution

VKernel's vOPS product line incorporates CPU Ready detection and resolutions into a solution that automates the steps described in this paper. The end result is a listing of all VMs experiencing these kinds of problems (as shown in the screenshot below) with drill-down information that offers specific recommendations to fix an issue. Some of these resolutions also be implemented with one-click automation. A free 30 day free trial of vOPS Server Standard can be downloaded and installed as a vSphere, Hyper-V or Red Hat Enterprise Virtualization (RHEV) VM in 20 minutes, and will immediately show any CPU Ready issues existing in an environment.

Conclusion

Understanding and monitoring CPU Ready is essential to maintaining a health virtual environment. The information in this white paper provides VM admins the guidance and tools needed to use it as a performance metric and take the mystery out of CPU Ready.

For more on CPU Ready, go to the VKernel website and listen to the "Demystifying CPU Ready as a Performance Metric" podcast with VKernel Senior System Engineers Makis Koilakoudis and Jonathan Klick. You can also find white papers, podcasts, blogs and other CPU and vCPU resources on VKernel's vCPU & CPU Management Resources page.

About the Author

David Davis is the author of the best-selling VMware vSphere video training library from TrainSignal. He has written hundreds of virtualization articles on the Web, is a vExpert, VCP, VCAP-DCA, and CCIE #9369 with more than 18 years of enterprise IT experience. His personal Website is VMwareVideos.com.

About the Sponsor

VKernel, a division of Quest Software (NASDAQ: QSFT), is the number one provider of virtualization management products for virtualized data centers and cloud environments. The company's powerful, easy-to-use and affordable products simplify the complex and critical tasks of planning, monitoring and predicting capacity utilization and bottlenecks. Used by over 50,000 system administrators, the products have proven their ability to maximize capacity utilization, reduce virtualization costs and improve application performance.