

# STORAGE SWITZERLAND

## IMPROVING VMWARE STORAGE I/O PERFORMANCE BY REALIGNING PARTITIONS



Storage Switzerland, LLC

George Crump, Senior Analyst

The virtual infrastructure is an environment where the old ways of just throwing hardware at a storage performance problem are no longer viable. Rather, what's needed is an understanding of the exact source of the problem to assure that foundational issues are taken care of first. A great example of this is the proper alignment of VM storage partition tables to their respective virtual disk or physical LUNs. Skipping this core step can limit other performance hardware solutions to far less than their full potential.

What is Misalignment?

To get to the physical location of data a virtual machine typically goes through three layers of disk abstraction. The first is a VMDK formatted with its own file system, Windows NTFS in most cases. The second layer is the virtual machine file system layer (VMFS) and the third is the Logical Unit Number (LUN) provided by the storage system. Another option for disk image storage is to have the VMDK reside on an NFS share instead of a LUN. In this case alignment issues can still occur but only the VMDK

needs to be aligned. The alignment issue is primarily a legacy storage issue as storage vendors and virtualization vendors have taken steps to address this, but there is some clean up work to be done on legacy systems.

These three layers (or two in the NFS case) should be in alignment to ensure maximum performance. Typically, most older operating systems (Windows 2003 and XP) are not aligned by default when they are virtualized. In other words the block boundaries of all three layers won't be at the same physical address, causing writes to overlap adjacent blocks. The third layer of VMFS confuses the older operating systems and results in this misalignment. The problem is that everything will still function properly but not at optimum levels. As the VM density increases on the host and the number of hosts connecting to the shared storage infrastructure also increases the performance overhead of this misalignment becomes more apparent.

What is the Impact?

Much has been written about the problems with misalignment, including documents by VMware, Microsoft, EMC, NetApp and many other storage and virtualization vendors. The performance impact can be as little as 20% in sequential large block workloads and as much as 100% in small block, random I/O workloads. Basically, the greater the number of reads and, especially writes generated, the more performance will deteriorate when a partition is misaligned. In short, unneeded reads and writes are caused by a read or a write having to cross multiple block boundaries within each layer, making the storage controller do extra work in repositioning drive heads to multiple locations.

This extra work also impacts the performance of the other VMs that are accessing that shared storage system, even if they themselves are aligned, because the controller is processing the extra reads and writes of the misaligned partitions. While modern versions of operating systems and VMware's new conversion utilities will correct alignment problems in the future, many current virtualized infrastructures may have hundreds of misaligned LUNS and they need a realistic solution to the problem today.

### The Challenge with Re-Aligning Partitions

While there are an increasing number of tools available that help with partition realignment, the reality is that most administrators are hesitant to take action to fix those partitions. The reason is that the time and processing involved in remediation is considerable. One remediation path could be to create a new VM with a correctly aligned template. Then the misaligned VM would be taken offline and restored to the new correctly aligned VM.

This restoration process, depending on the capacity requirements of the VM, can take from 30 seconds to 4 hours per VM. No matter the case, downtime will be required. Again, while it's happening, the virtual machine, and of course the application that it supports, need to be

taken down. This is also a manual VM-by-VM process, with each conversion effort being individually monitored for successful completion. Since the performance problem is going to first appear on small block I/O VMs this means that critical applications like databases and email are going to be the most likely candidates for conversion.

### Real-World Alignment

Again, while there are several sources for software solutions that will detect misaligned partitions these are only slightly helpful given the manual conversion process described earlier, which can take hours to complete. Instead, administrators will want to be more 'surgical' in their implementation of the solution. What they'll want to do is assess which VMs are being most impacted by misalignment and fix them first. For example VMs with low I/O activity or high sequential I/O activity are only impacted 20%, typically, and could be made a lower priority. VMs that have high small-block random I/O, like databases and email applications, should be addressed first, since this would improve their performance and the efficiency of the storage devices they are using. This could, in turn, provide a ripple effect, benefiting the entire infrastructure.

Products like [NetApp OnCommand Insight Balance](#) can show which VMs are impacted most by misalignment, among a host of other detailed virtual infrastructure management capabilities. Not only does this solution identify the affected disk partitions, it also provides a ranking of those most impacted. This enables the administrator to create a prioritized list of VMs for correction based on biggest payoff for the downtime that the conversion process will generate. Of course OnCommand Insight Balance can also be used to document the performance improvement that the downtime investment provided. The administrator may also decide that some VMs don't have a performance profile that makes this mitigation worthwhile. Either way, the administrator, armed with the facts, can make the right decision given the available downtime window.

## Future Mitigation and Best Practice

While the alignment process is being conducted, and even when it completes, steps should be taken to make sure there are no future alignment problems. As stated earlier, most modern operating systems are now aligned as a default. Also, in VMware vSphere 5 the VM converter function will now properly align a new VM as it's being converted from a physical server during a P2V process. The same processes can be used to correct VMs through a V2V process. For environments that are not yet on VMware 5 and where older operating systems need to be virtualized, best practice is to use a pre-aligned template and install the new guest operating system. The final step is to keep monitoring the virtual environment for any case of misalignment. Early detection is ideal before the virtual machine grows in size and importance.

Eventually, mis-alignment issues may cease to be a concern. Now, however, it can be a real problem, especially for customers that have pushed well beyond the evaluation phase of a virtualization rollout. While much has been written about the problem and how to fix it, the real world challenges of converting to a new aligned partition still make it a major undertaking. With many data centers now in the hundreds to thousands of virtualized servers, a more intelligent option of detection is needed. Solutions like OnCommand Insight Balance offers the VMware administrator the ability to clearly see which VMs will benefit the most by realignment and to begin making those intelligent choices.

### **About Storage Switzerland**

Storage Switzerland is an analyst firm focused on the virtualization and storage marketplaces. For more information please visit our web site: <http://www.storage-switzerland.com>

*Copyright © 2011 Storage Switzerland, Inc. - All rights reserved*