Wilson IT

tel -0870 486 1305
fax -0870 486 1304

www.wilson-it.com
enquiries@wilson-it.com

NTFS cluster size performance differences on VMWare hosts/guests

By Jonathan Wilson

Before rolling out our new VMWare server infrastructure I decided to do some testing to see if changing the NTFS cluster size on a VMWare host would make any difference to the speed of a VMWare guest. Logic would dictate that as the files are very large in size (gigabytes) a larger cluster size would be better.

Unfortunately there are no benchmarks anywhere and people on various forums either said to use large 64k clusters because VMWare files are 2GB+ or that it wouldn't really make any difference. No one could point to any actual data or real world testing and VMWare itself doesn't seem to have any recommendations in its white papers on performance tuning (which I found surprising). So before creating a stable production environment I decided to run some tests on a new server to see what happens.

I setup a single RAID 5 array with four disks and then created a 40GB boot partition (4K Clusters) onto which I installed a fresh copy of Windows 2003 Enterprise SP1. I then created an extended partition on the rest of the free space and within that created 2 logical 40GB drives, one formatted with the default cluster size of 4K and the other with a 64K cluster size. VMWare was installed on the boot partition and two identical Virtual Machines were placed on the logical drives (J: and K:).

SiSoft Sandra 2005 SR1 was installed inside each VM to benchmark file performance and also installed on the main server itself as a reference.

Setup details:-

Dell PE2900, 1x Quad Core 1.6Ghz Xeon 5310, 4GB 667Mhz FB RAM, PERC 5/i 256MB RAID Controller (256MB Cache with "read ahead normal" & "Write Back" cache settings) Stripe size 64K, 4x 250GB 7.2k SATA disks in RAID 5 configuration.

Host OS setup - Windows 2003 Enterprise Edition with SP1 - NO roles defined - NO Extra services, latest Dell drivers, VMWare Server 1.0.1 build 29996, SiSoft Sandra 2005.2.10.50

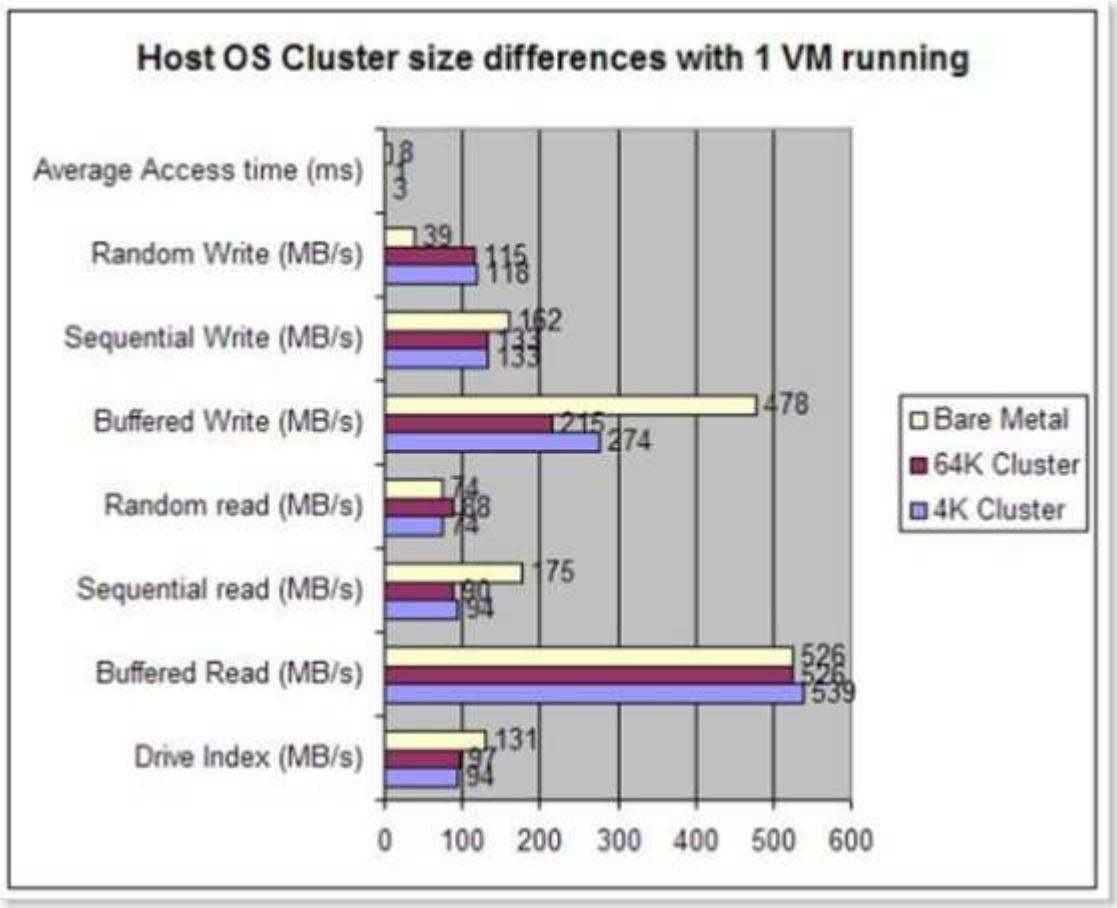
Virtual Machine OS setup- Windows 2003 Enterprise Edition SP1 - NO roles defined - NO Extra Services, VMWare tools, SiSoft Sandra 2005.2.10.50

*Notes:- VM Tests were run 3 times and then averaged, 'average access time' results varied wildly on the VM's and sometimes didn't appear in SiSoft so you should *ignore them*.

The Bare metal tests were only run twice (and averaged) and the same results are used on ALL graphs merely as a reference. In general the results were more consistent than the VM tests. They are there solely as a reference point to see the difference between bare metal and a VM.

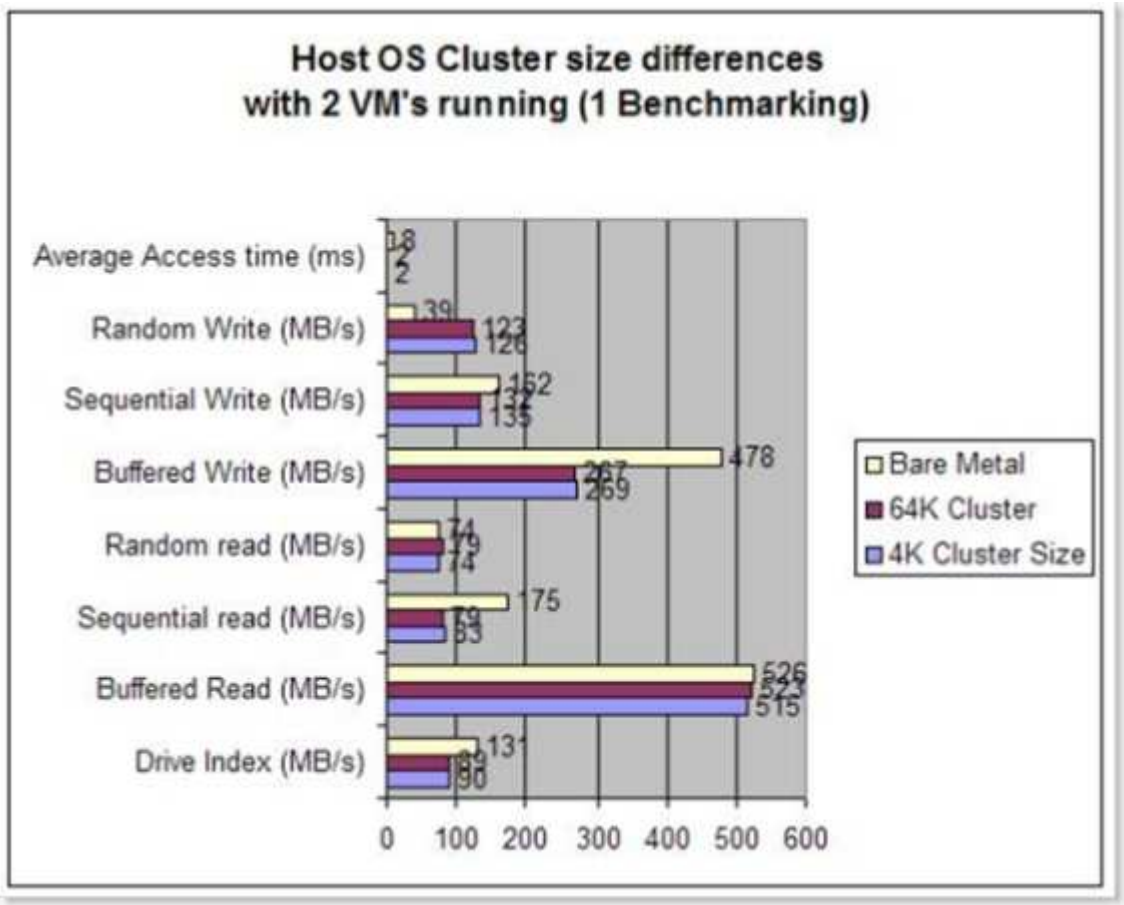
IMPORTANT: Please remember, the bare metals tests were run without anything else running, are duplicated on all graphs and can only be directly compared to a Virtual Machine in the first graph.

Results as follows:-

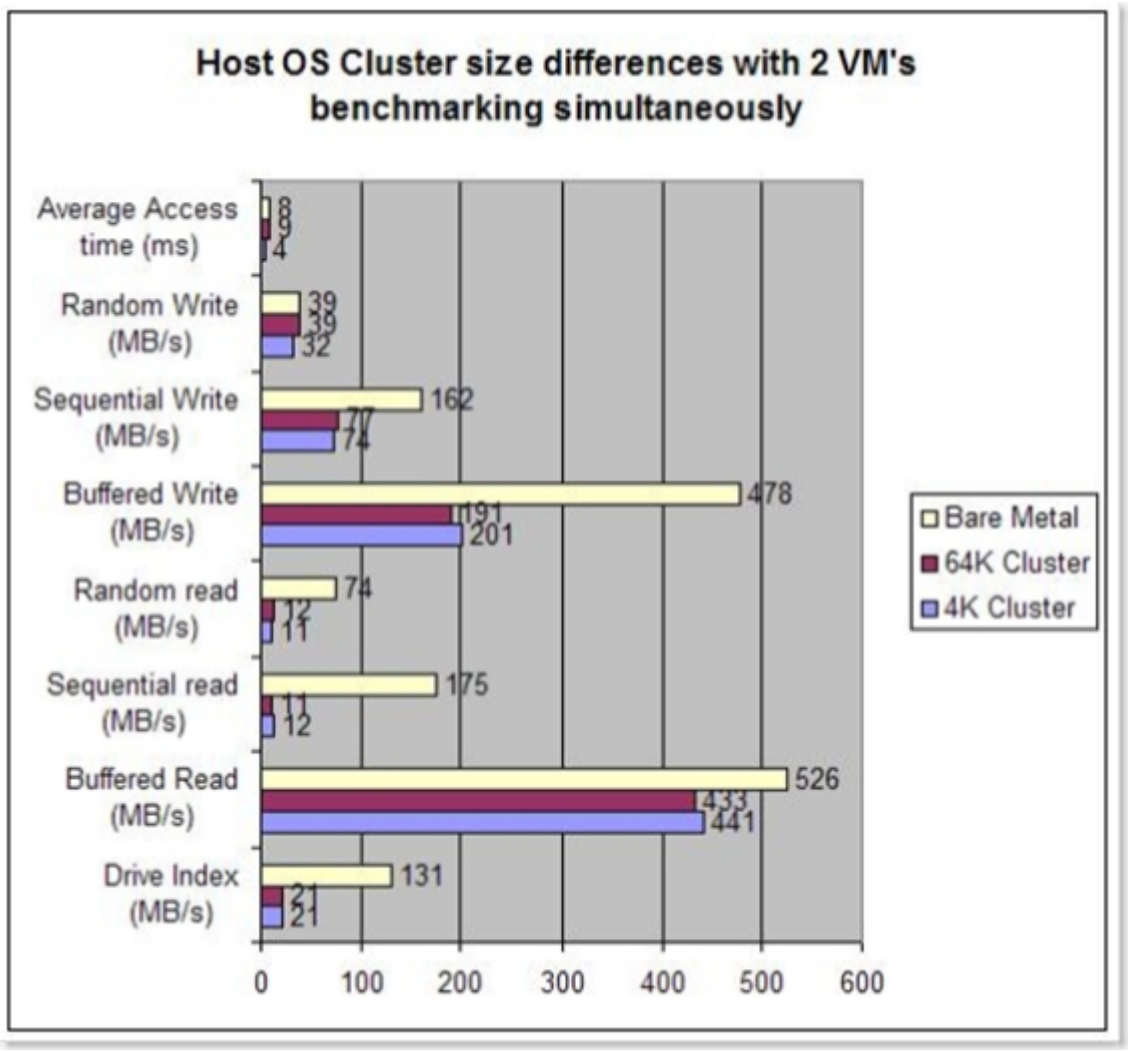


*** The 4K Cluster 'buffered write' seen above is an anomaly, the three results were 240, 351 and 231 so clearly the 2nd run was way out and should be dismissed.**

As you can see straight away there is barely any difference between cluster sizes and overall file system performance holds up incredibly well compared to bare metal, especially 'sequential writes'. The one area where it does fall down quite a bit is 'sequential reads' meaning if your users regularly pull lots or large files off your server there is no substitute for bare metal.



With one VM sitting idle and the other benchmarking there is very little difference in performance. Again, the difference between 4K and 64K clusters is minimal and could quite easily be put down to benchmarking variables.



Again, there is very little difference between the two cluster sizes, the differences are almost too small to be noteworthy. Please note that although performance has dropped off considerably here this is due to the benchmarks we are running and is not indicative of every day usage, these results should ONLY be used to compare cluster size differences.

Conclusion

First off SiSoft Sandra results varied quite a bit between each run so if you did enough runs you would probably find that the results between cluster sizes would be even closer.

In general though I'm amazed at how good VMWare is, seeing as this is a virtualised environment it's very impressive. Generally speaking the biggest hit comes in buffered writes and sequential reads so I suggest leaving Write cache on your RAID controller to "Write Back" which is faster but less secure in a power cut. If you have a decent UPS which allows your server to shutdown gracefully (and you've tested this!) you should be fine.

So, 64K or 4K cluster size? Well, from the results above I'd give the nod to the default 4K clusters, there's barely anything in it but it does appear fractionally quicker in sequential reads (the important bit) and you also get the nice warm feeling knowing that your defragmenter of choice will be compatible with 4K clusters.