The Performance Impact of Windows Guests Defragmentation and Free Space Consolidation on VMware ESX

Presented by

RAXCO SOFTWARE

April 2010

Table of Contents

Executive Summary

It has long been recognized that file and free space fragmentation are detrimental to Windows® system performance. There are numerous articles by third party authors and in Microsoft's Technet knowledgebase about the negative impact file and free space fragmentation has on read/write performance and the need to perform regular defragmentation. File fragmentation increases the time it takes to read a file, while free space fragmentation increases the time it takes to write a file. In both instances, the extra workload increases the demand for CPU and memory while issuing extraneous IO to the disk. This IO is often the cause of resource bottlenecks.

This paper details the results of testing done to determine the effectiveness of defragmenting virtual servers. With server virtualization, an organization can deploy several virtual instances of Windows Server on a single physical machine. The combined workload of these machines must share the resources of their physical host. Given the well-documented track record of the Windows file system to fragment files and free space; we set out to determine if fragmentation inside Windows guests had any impact on performance.

The testing, conducted on a VMware ESX cluster, indicated that defragmentation and free space consolidation of Windows server guests has a positive effect on the performance of the both the Windows guest and the host. Based on the metrics detailed in this report we conclude:

- File defragmentation and free space consolidation combine to significantly reduce the total number of IO that need to traverse the virtual storage stack
- Free space consolidation can significantly improve the number of large writes to the disk improving throughput
- File defragmentation and free space consolidation dramatically reduce disk latency by orders of magnitude
- File defragmentation and free space consolidation of Windows guests in a VMware environment reduces the total workload on the hypervisor
- Free space consolidation improved write throughput in the Windows guest and on the ESX platform

Test Equipment and Methods

The following system configuration was used in the testing.

ESX server configuration

ESX Version:	3.5.0 Update 1
Motherboard:	Intel S5000PSL
CPU Type:	Intel(R) Xeon(R) CPU E5345 @ 2.33GHz
Number of CPUs:	2
Cores per CPU:	4
Logical Processors:	8
Memory:	4 GB

Storage Configuration

RAID controller:	Adaptec RAID 3805
Number of Drives:	4
Drive Type:	WD1001FALS 1TB 7200 RPM 32MB Cache
Total Capacity:	4.0 TB
Number of LUNS:	2
LUN 1 RAID level:	5
LUN 1 Capacity:	2.00 TB
LUN 1 Partitions:	1
LUN 1 Name:	IOTesting
LUN 2 RAID level:	5
LUN 2 Capacity:	744.75 GB
LUN 2 Partitions:	1

VM Configuration

Number of VMs:	2
Operating System:	Windows Server 2008 R2 (64-bit)
Memory:	2GB
Number of CPUs:	2
SCSI Controller:	LSI Logic (no SCSI bus sharing)
Number of Disks:	1
Size of Disk:	50 GB
Provisioning Type:	Thick
Backing Datastore:	IOTesting
Virtual Memory:	none (pagefile disabled)
Network:	disabled

Testing Overview

Fragmentation in a Virtual Environment

File fragmentation is a function of how the file system allocates space to a file. To create a file, the NTFS file system looks to the $Bitmap file to determine where space is available. The $Bitmap file identifies which logical clusters are in use and which ones are free. If the file system cannot allocate space for the entire file in a contiguous string of logical clusters, the file will be fragmented. When a read or write request is received for that file, the Master File Table is accessed and it provides the starting logical cluster number (LCN) and the run length for each fragment needed to satisfy the requested read range. The more fragments there are, the longer it takes to read the file. If a read request spans 10 fragments, the file system needs to report the 10 starting LCN's and run lengths to the disk controller.

The same behavior described above occurs inside each virtualized Windows server. File and free space fragmentation occur within each VMDK and impose a performance penalty on the system. File and free space fragmentation is relevant in a virtual environment because the finite resources of the host must be shared with other virtual machines. If file and free space fragmentation create a resource bottleneck on one virtual machine, the rest are going to be deprived access to those resources.

Products Used in Tests

To test the effect of file defragmentation and free space consolidation we used Raxco Software's PerfectDisk. PerfectDisk is unique in that it is the only disk defragmentation solution that consolidates the free space on the disk into the largest possible contiguous chunk. Defragmenting files improves read access time, but consolidating free space improves write access and slows the re-fragmentation of the disk.

Testing Procedure

To conduct our tests we needed to be able to collect performance metrics from the VMware environment. We used the *vscsiStats* which are collected in the VMware Monitor. The vscsiStats collect information on every IO coming through the system and sorts them into various categories for performance analysis. The versatility of the vscsiStats provides several different metrics we can use to determine the benefits of file defragmentation and free space consolidation with PerfectDisk.

All disks were formatted with NTFS. The disks were populated with ISO images and other random length files. A custom tool was used to fragment the resulting collection of files. The fragmented disk was imaged so it could be restored to provide an identical starting point in subsequent tests. The fragmented VMDK was cloned so we had two identical but separate test disks.

The first step in the test was to defragment the cloned VMDK with PerfectDisk. The fragmented disk was designated VM1 and the PerfectDisk disk was designated VM2. All extraneous activity on the ESX cluster was shut down to ensure the vscsiStats counters were only counting IO related to the test activities. The vscsiStats were enabled on VM1 and we installed Microsoft Office. Upon completion of the installation the vscsiStats on VM1 were stopped. We then installed MS Office on VM2 while collection the vscsiStats on that device.

Upon completion of the MS Office installation and data collection the VM2 disk was again defragmented with PerfectDisk and the same process was followed while we installed MS SQL Server which is a larger package. The pre and post disk status for VM1 and VM2 is shown below.

VM1-Fragmented Disk Pre-Post Software Installation Details

	Pre- Software Installation	Post-Software Installation
File System	NTFS	NTFS
Bytes/Cluster	4096	4096
File Fragmentation	29.9%	28.7%
Directory Fragmentation	0.2%	0.6%
Free Space Fragmentation	100%	100%
Metadata Fragmentation	10.7%	12.2%
Excess File Fragments	378723	428174
Excess Directory Fragments	87	525
Excess Metadata Fragments	65	79
Total Number of Files	122322	135676
Total Number of Directories	33099	35783

VM2- PerfectDisk Pre-Post Software Installation Details

	Pre- Software Installation	Post-Software Installation
File System	NTFS	NTFS
Bytes/Cluster	4096	4096
File Fragmentation	0%	0.2%
Directory Fragmentation	0%	0.3%
Free Space Fragmentation	0.5%	0.1%
Metadata Fragmentation	0%	2%
Excess File Fragments	3	1869
Excess Directory Fragments	0	307
Excess Metadata Fragments	0	13
Total Number of Files	122239	135592
Total Number of Directories	33099	36096

MS Office Installation Statistics

In a VMware environment it is well understood that more IOP's are bad, and fewer IOPS are good. The first vscsiStat we looked at was total IO and its breakdown into read IO and write IO for the MS Office installation. In Table 1 we see the total IO count was 7,263 IO's higher on the VM1 (fragmented disk). As you would expect for a software installation, almost all of the IO reduction benefit came in the form of writes. Since identical software was installed on identical disks this shows defragmentation and free space consolidation reduces IO.

	VM1-Fragmented	VM2 PerfectDisk	Percent Improvement
Total IO Count	32314	25051	22.4
Read IO Count	2886	2509	13.0
Write IO Count	29428	22542	23.3

Table 1. Total IO Counts-MS Office Test

More telling is the distribution of these IO by their size. On the VM1 disk, all of its files and 21GB of free space are fragmented all over the disk. When installing software, the fragmented free space will have a profound effect since it limits the maximum IO size the system can perform. The vscsiStats sorts each IO into one of 18 buckets ranging in size from 512bytes to >524K. The

two largest buckets are 524K and >524K. Table 2 shows the total number of IO to each of the two largest buckets. Writes of these buckets would be the most desirable.

	VM1-Fragmented	VM2-PerfectDisk
Total IO Equal to 524K	1280	299
Total IO > 524K	143	1576
Read IO Equal to 524K	0	0
Read IO >524K	48	39
Write IO Equal to 524K	1280	299
Write IO >524K	95	1537

Table 2. IO Distribution by Size –MS Office Test

The results shows that the PerfectDisk disk with consolidated free space was able to perform **16 times** as many write IO greater than 524K than the same disk with fragmented free space.

The third metric we looked at is latency. This is the amount of time it takes the system to complete an IO. Again, vscsiStats sort each IO into one of 11 buckets ranging from 1 microsecond, to greater than 100,000 microseconds (100ms). To provide a point of reference, a fast IO would be anything faster than 15000 microseconds (15ms), and a slow IO is anything over 15ms. Table 3 shows the total IO distribution taking longer than 15ms.

	30ms	50ms	100ms	>100ms
Total IO -VM1	1037	871	1198	6145
Total 10-VM2	730	158	120	187

Table 3. IO Distribution by Latency-MS Office Test

These results show the PerfectDisk disk significantly reduced the total number of slower IO. In fact, the VM2 disk did **23 times fewer** IO of 100ms or greater than the VM1 disk. This data is complementary to Table 2 which shows the same disk doing a greater number of large IO. If the system can do larger IO then the latency should be less.

MS SQL Installation Statistics

Following the completion of the MS Office installation we defragmented the VM2 disk again and then installed MS SQL on both VM1 and VM2 while collecting the vscsiStats on both disks. The MS SQL installation is a larger package with more files and larger files than the MS Office installation, which means more work for the file system. In Table 4 we see that the VM1 disk performed significantly more IO.

	VM1-Fragmented	VM2- PerfectDisk	Percent Improvement
Total IO Count	166412	105620	36.5
Read IO Count	44230	37988	14.3
Write IO Count	122182	67632	44.6

Table 4. Total IO Count-MS SQL Test

The total IO count on VM1 was 60,792 greater on VM1. The PerfectDisk disk (VM2), with consolidated free space, made it possible for the file system to perform more writes in a single IO. The total write IO count was 54,550 less on the VM2 disk, almost half as many writes as the fragmented disk. The consolidated free space on VM2 also means this disk was able to do a greater number of large writes. This is supported by the data in Table 5.

	VM1-Fragmented	VM2-PerfectDisk
Total IO Equal to 524K	2298	1031
Total IO > 524K	179	4065
Read IO Equal to 524K	35	9
Read IO >524K	50	22
Write IO Equal to 524K	2263	1022
Write IO >524K	129	4043

Table 5. IO Distribution by Size –MS SQL Test

The results show that the PerfectDisk disk with consolidated free space was able to perform **31 times** as many write IO greater than 524K than the same disk with fragmented free space.

In the MS SQL test we again measured the disk latency on the two disks and the PerfectDisk disk reported fewer IO taking longer than 15ms to complete.

	30ms	50ms	100ms	>100ms
Total IO -VM1	8236	3794	5081	16712
Total 10-VM2	3193	1085	631	448

Table 6. IO Distribution by Latency- MS SQL Test

We see a significant reduction in the number of slower IO on the PerfectDisk VM2 disk. The fragmented VM1 disk did **2.5 times** as many IO taking 30ms as the VM2 disk. When you look at the IO taking more than 100ms or longer, the VM1 disk did **20 times** more slow IO than the VM2 disk. This demonstrates file defragmentation speeds up disk reads and free space consolidation speeds up disk writes. The ability to perform fewer and larger IO means the total IO is more efficient and IO complete faster.

Summary

Performance improvements from file defragmentation and especially free space consolidation are commonplace on physical Windows servers. Under the NTFS file system, normal file activity like creation, extension, truncation and deletion, contribute to both file and free space fragmentation. The net effect of both kinds of fragmentation is increased IO which in turn increases the demand for CPU and memory.

In a virtual environment multiple virtual machines share the resources of their physical host. A virtualized Windows server guest behaves just like a physical server. All of the file system workings are exactly the same. Files and free space inside the vmdk fragment and create an extra IO load that increases the demand for CPU and memory resources from the ESX host. The purpose of this testing was to determine if the effects of fragmentation had measureable effect on resources in a virtual environment. The VMware vscsiStats are collected at the monitor level just above the VMware kernel and they intercept every IO coming through the virtual stack. If defragmentation and free space consolidation were going to make a difference, the IO metrics collected by the vscsiStats would show it.

The testing conducted was simple in that it compared the IO counts, IO size and latency on disks where we installed software. Software installation is write-intensive so it is a good test for the efficacy of free space consolidation. The VMware vscsiStats provided compelling data to support the argument for the defragmentation of virtual servers.

Based on the testing we conclude the following:

- Free space consolidation improves disk write performance. The installation of MS Office and MS SQL Server showed disk write improvements of 22.3% and 44.6% respectively when comparing the fragmented disk to the disk where PerfectDisk disk.
- File defragmentation and free space consolidation improved over IO performance. The total IO counts improved 22.4% and 36.5% respectively on the benchmarked disks.

- Free space consolidation improved total throughput. During the software installations the PerfectDisk disk was able to perform more IO's of greater than 524K. Large IO on the PerfectDisk disk was 16 times and 31 times greater respectively than the fragmented disk.
- File defragmentation and free space consolidation combine to improve disk latency. Using 15ms or faster as the definition of a fast I0, we saw the fragmented disk on the MS Office install performed 9,251 IO greater than 15ms, while the PerfectDisk disk only did 1195, an 87% improvement. On the MS SQL install, the fragmented disk performed 33,823 IO over 15ms and 49% of these were over 100ms. The PerfectDisk disk did 5,357 IO over 15ms IO and only 8% were over 100ms.
- File defragmentation and free space consolidation improve productivity. While the vscsiStats do not time the work done, we did note installation time for MS Office was about 10 minutes on the fragmented disk and 5 minutes on the PerfectDisk disk. The MS SQL installation was approximately 45 minutes on the fragmented disk and 30 minutes on the PerfectDisk disk.