Impact of Free Space Consolidation On Windows File System Performance

White Paper created by
Balder Technology Group, Inc.

August 2008

Table of Contents

Executive Summary

It has long been recognized that file fragmentation is detrimental to Windows® system performance. Several previous studies have examined the impact of file fragmentation on overall performance of file systems[1] and the negative impact fragmentation has on the total cost of ownership (TCO) for IT departments[2].

While it is intuitive that fragmented free space will lead to new file fragmentation, this study looks at the effect of free space consolidation on system performance in general and disk performance in particular on the mobile laptop environment where disks are slower and smaller and defragmentation may not be under centralized control.

Based on our extensive lab testing, we conclude the following:

- Free space consolidation improves disk and operating system performance.
- Free space consolidation is a critical aspect of disk defragmentation.
- Free space consolidation is as important as defragmentation of files for a disk defragmenter.
- If proper attention is not paid to free space consolidation, disk defragmentation is likely to not be effective.
- If free space is not adequately consolidated, fragmentation will occur, and even newly-created files will be fragmented.
- PerfectDisk® 2008's emphasis on free space consolidation provides performance improvements beyond those of Diskeeper® 2008 and the built-in defragmenter. Indeed, the largest chunk of contiguous free space ended up being **smaller** after Diskeeper finished defragging.

Balder Technology Group Background

David Goebel, the president and chief executive officer of Balder Technology Group, Inc., has been intimately involved with the NTFS file system since its original design in 1991. As a Microsoft employee and member of the four-person Windows NT® file system team from 1990 to 1995, Mr. Goebel worked on the architecture, design, and implementation of NTFS[3].

Since leaving Microsoft® and founding Balder Technology Group in 1995, he has developed file system and other kernel mode software for various clients, including Microsoft®. The video of his talk on file system synchronization at Microsoft's only file system conference is featured in Microsoft's Windows Server® 2003 Installable File System Development Kit[4]. He built the foundation for the ntifs.h header file in the original IFS kit so that it would auto-generate for future releases.

Mr. Goebel helped design the SIS or 'Single Instance Store' storage architecture first introduced in Windows 2000 Datacenter Edition, resulting in a professional paper[5] and a patent issued to Microsoft[6], with Mr. Goebel listed as a co-author and co-inventor respectively.

Microsoft sent him along with a Microsoft employee to Romania in 2005 to perform technical due diligence on its acquisition of GeCAD anti-virus technology[7] and write Microsoft's anti-virus file system filter driver, MpFilter.sys, which was the core component of its OneCare™ release.

In 2006 and 2007 Mr. Goebel helped design and authored the new exFAT.sys file system released in Vista SP1 and Server 2008.

Other notable contributions include staffing all nineteen of Microsoft's file system plugfests[8] and the file system development labs before them.

Testing Overview

Fragmentation Background

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, as each noncontiguous read incurs a disk seek[9] (approximately 10 msec on today's disks, or about 30 million wasted processor cycles on a 3GHz processor). 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.

Products Used in Tests

To test the effect of free space consolidation, two third party disk defragmentation products that have starkly different perspectives on free space management were chosen: Raxco Software's PerfectDisk® and Diskeeper Corporation's Diskeeper®, as well as the built in defragger which had significant changes starting in Windows Vista®. This selection was made due to the behavior of the respective products with respect to free space management, their publicly stated approaches to free space management and consolidation, and recognition that they are the two leaders in the defrag market space. The products used in the testing were Diskeeper 2008 Pro Premier and PerfectDisk Pro 2008, both of which were available to the general public at the time of the testing.

Diskeeper states "Free space consolidation might be important if you have to create one gigantic contiguous file, but it has no effect on performance."[10] Raxco claims "In a single pass, PerfectDisk 2008 creates the largest piece of contiguous free space possible. This results in less fragmentation occurring when new files are created and also improves write file access time, in addition to read file access time, which is all typical defragmenters focus on."[11] Microsoft has also weighed in, writing "…and while it's good to have free space, it's not good if it's fragmented. Free space fragmentation refers to file space that's broken into small pieces, rather than joined together. This type of fragmentation results in slowed performance."[12]

Operating System

We were eager to test on the newly released Windows Vista to discover what impact vast kernel and file system changes have made since our previous testing on Windows 2000 and Server 2003. In addition, the built-in defragger has seen significant changes and it was compared as well. We used Vista SP1 for all testing.

Testing Background

To conduct our tests we used an industry standard performance benchmarking suite and a custom tool that would actually measure physical disk accesses per file request. For benchmarking, we selected the BAPCo SYSMark 2007 v1.03. SYSMark installs a variety of applications and automates them in order to measure real world performance and responsiveness. BAPCo (Business Applications Performance Corporation) is a non-profit industry consortium and its benchmarking tools are used by Ziff Davis Media among other computer press. See http://bapco.com/products/sysmark2007preview for additional details.

When the file system receives a read/write request that spans a fragmented region of a file, it must split the single input/output request packet (IRP) into multiple associated request packets that read or write the scattered data. A custom kernel mode tool was developed by the Balder Technology Group that filters the requests between the file system and the disk collecting statistics on fragmented reads and writes. By measuring both the number of requests sent to the file system and the number of resulting requests sent to the disk, we can determine the number of extraneous, or wasted, disk requests and thus wasted seeks. The tool developed was predicated on the *diskperf.sys* sample in Microsoft's Windows Driver Kit (WDK).

In addition, a user mode tool was developed that could make a sector by sector copy of a partition to a file and then later restore that image back to the partition. This tool was used to guarantee that each defrag run started with the verbatim partition image.

Test Equipment and Methods

The following system configuration was used in our evaluations.

Table 1: System Configuration[13]

Table 1: System Configuration			
Make/Model	**CPU**	**Memory**	**Disk**
Compaq/HP Presario F700 laptop	AMD Athlon™ 64 X2 1.6HGz	1 GB	WD 120GB drive-Model WD1200BEVS ATA device

SYSmark only runs on a boot partition, which was formatted to NTFS as Vista requires. The partition was filled with random length files leaving 30GB free as required by SYSmark. A custom tool was used to fragment the resulting collection of files. SYSmark was installed and the partition was sector by sector imaged so it could be restored to provide an identical starting point for an accurate comparison of the respective products for each defragmentation run.

To help achieve consistent results, the Windows Search service was disabled and the network was unplugged so that the machines would be truly idle when not defragging or running the benchmark.

Defragmentation Methods

For PerfectDisk and Diskeeper, their recommended approach was used to defragment the volumes, respectively. For PerfectDisk, this was SMARTPlacement™[14] mode online and an offline pass. For Diskeeper, this was using Invisitasking® which runs in the background and defragments during system idle time. Invisitasking presented the unique problem of knowing exactly when it was "done". Since the system was otherwise idle, we would wait until the Diskeeper processes had stopped accruing any user or kernel CPU time for several minutes. A manual analysis then recommended an offline pass, which was done.

For the built in defragger, its command line interface was used: 'defrag c: -w -v'.

Test Results

Fragmentation Statistics

After fragmentation, the file and metadata contained 270,604 excess fragments.[15] The volume had 9.65 gigabytes of free space which was in 80,559 pieces, the largest being 198 MB.

Table 2: Post-Defragmentation Results

	Excess File Fragments	Free Space Fragments	Largest Contiguous Free Space
After PerfectDisk	17	2001	5.24 GB
After Diskeeper	112,742	67,443	8.5 MB
After Built-In	11,539	37,978	28.3 MB

Diskeeper decreased the total excess file fragments, but it also shrunk the largest chunk of free space by more than an order of magnitude.

Disk Drive Performance Analysis

In order to establish a performance baseline we ran SYSmark on the original fragmented disk and established base scores. The idea here was to establish a norm with the base fragmented drive and compare this to the scores SYSmark provided after defragmentation with the commercial products and built-in tool.

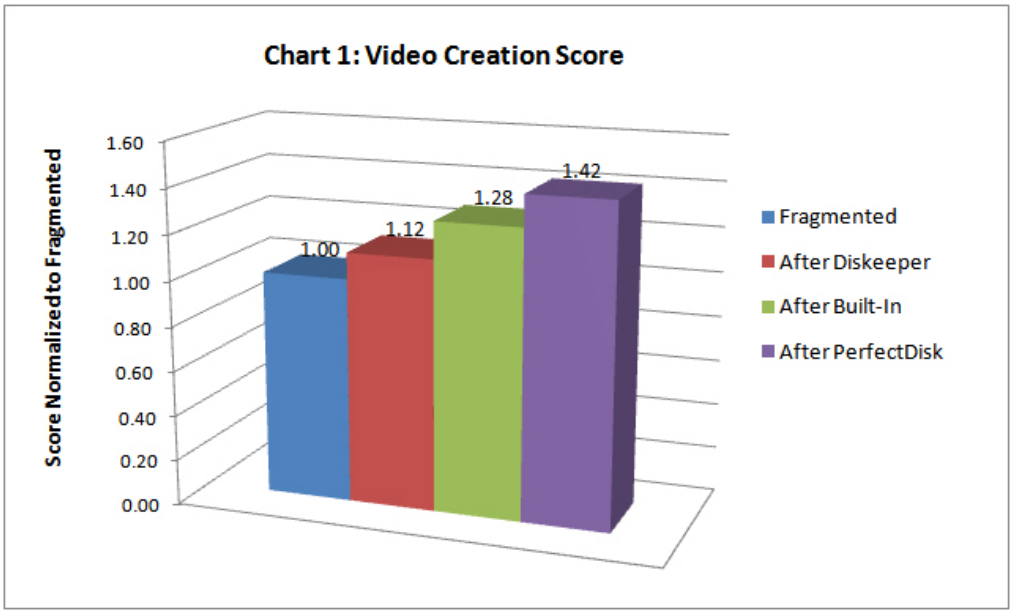
SYSmark results are split into four categories (E-Learning, Video Creation, Productivity, 3D), each with a collection of representative Apps. BAPCo's web site gives details on the apps involved in each category.[16] In general the Video Creation category, due to its manipulation of large video files, best represents disk performance and is the number we will use for our graphs. Inversely the 3D test is almost exclusively a measure of graphics performance. In order to publish results, BAPCo requires three consecutive runs of the test with variance of less than 10% between runs. The numbers presented here are the "Official Rating" averages.

Higher scores indicate better performance. The important thing to remember about SYSmark scores is that a ***higher score is better***.

Table 3: SYSmark Result on Compaq/HP Laptop

Partition State	E-Learning	Video Creation	Productivity	3D	SYSmark Score
Fragmented	55	43	60	74	57
After PerfectDisk	61	61	66	77	66
After Diskeeper	57	48	64	76	61
After Built-In	56	55	66	76	63

When we normalized the base fragmented partition to be 1.0, we see the following performance chart for Video Creation:



These results lead us to the following conclusion:

- PerfectDisk resulted in higher performance than either Diskeeper or the built-in defragger.

Disk Access Analysis

The seek data acquired during the test run tracks the final performance data.

The table summarizes the total number of seeks performed on the baseline disk when running SYSmark, as well as the total number of seeks on the drive defragmented by PerfectDisk, Diskeeper, and the built-in defragger.

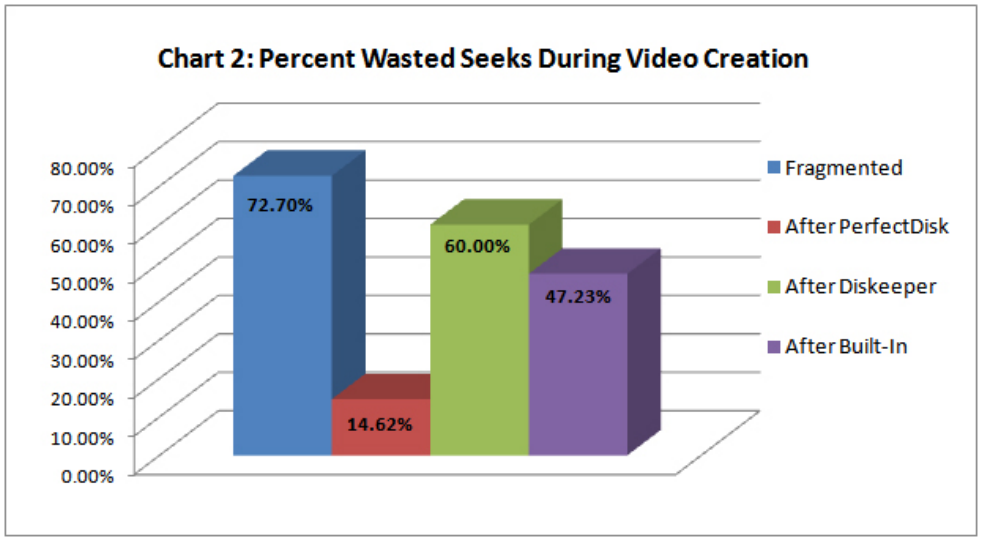
Table 4: Wasted Seek Analysis

Partition State	Total Number of I/O Requests Sent to the File System[17]	Total Number of Resulting Disk Accesses/Seeks	Net Wasted Seeks When Running SYSmark	Percent Net Wasted Seeks When Running SYSmark
Fragmented	1,320,686	2,090,649	769,963	58.30%
After PerfectDisk	1,434,454	1,616,847	182,393	12.72%
After Diskeeper	1,341,262	1,961,954	620,692	46.28%
After Built-In	1,411,613	1,931,395	519,782	36.82%

Further refining this to just include the Video Creation data in order to focus on disk performance, we see the following results.

Table 5: Wasted Seek Analysis for Video Creation Test Only

Partition State	Total Number of I/O Requests Sent to the File System	Total Number of Resulting Disk Accesses/Seeks	Net Wasted Seeks When Running SYSmark	Percent Net Wasted Seeks When Running SYSmark
Fragmented	494,104	853,341	359,237	72.70%
After PerfectDisk	534,795	612,957	78,162	14.62%
After Diskeeper	485,547	776,881	291,334	60.00%
After Built-In	512,624	754,720	242,096	47.23%



From this data it is clear that in this environment PerfectDisk resulted in the fewest wasted seeks.

Summary

Consolidating free space is singularly effective in postponing refragmentation of a file system. When a file is copied to a volume or a new file created and its size set, Windows is successful in using contiguous free space when it is available. If sufficient contiguous free space is not available the file is created already fragmented.

When running the relatively small disk footprint SYSmark test, PerfectDisk 2008 consolidated free space and scored up to 27% better than Diskeeper 2008, and incurred as few as one quarter the number of wasted disk seeks as Diskeeper 2008.

Based on our extensive lab testing, we conclude the following:

- Free space consolidation increases disk and operating system performance.
- Free space consolidation is a critical aspect of disk defragmentation.
- Free space consolidation is as important as defragmentation of files for a disk defragger.
- If proper attention is not paid to free space consolidation, disk defragmentation is likely to not be effective.
- If free space is not adequately consolidated, fragmentation will occur, and even newly-created files will be fragmented.
- PerfectDisk 2008's emphasis on free space consolidation provides performance improvements beyond those of Diskeeper 2008 and the built-in defragmenter. Indeed, the largest chunk of contiguous free space ended up being more than an order of magnitude smaller **after** Diskeeper 2008 finished defragging.

(Footnotes)

[1] http://www.raxco.com/products/perfectdisk2k/whitepapers/ohio_sb_ss.pdf
[2] http://www.raxco.com/new/idcstudy.pdf
[3] Helen Custer, Inside the Windows NT File System (Redmond: Microsoft Press, 1994), pp. vii
[4] Microsoft Part # X09-46661.
[5] William J. Bolosky, Scott Corbin, David Goebel, and John R. Douceur. "Single Instance Storage in Windows™ 2000," In Proceedings of the 4th USENIX Windows Systems Symposium, pp. 13— 24, August 2000.
[6] Bolosky, Douceur, Cutshall, Rashid, Myrhvold and Goebel. "Single Instance Store for File Systems". US Patent number 6,477,544. Nov 5, 2002.
[7] Fisher, Dennis, "Microsoft Signs Anti-Virus Deal." EWeek. Ziff David Media. June 16, 2003. http://www.eweek.com/c/a/Past-News/Microsoft-Signs-AntiVirus-Deal
[8] Plugfests are an approximately biannual week-long event held at Microsoft to test the interoperation of third party file systems and file system filter drivers.
[9] A seek is the physical movement of the disk read/write head
[10] http://support.diskeeper.com/support/diskeeperfaqs.aspx?Page=5&Subpage=2&cust=1&RId=1&CId=1&SId=1 (click "Analysis and Defragmentation" then "Why doesn't Diskeeper result in a fragmentation display showing all the files in one place and all the spaces consolidation into another place?")
[11] http://www.raxco.com/small_medium/small_perfectdisk_professional.cfm
[12] http://technet.microsoft.com/en-us/library/bb742585.aspx
[13] The hardware was identical for all tests.
[14] Davy, William. "Method for eliminating file fragmentation and reducing average seek times in a magnetic disk media environment." US Patent number 5,398,142. March 14, 1995
[15] 'excess fragments' is the difference between how many unique extents are allocated on the volume, and how many there would be if all the files on the volume were contiguous. The goal is zero.
[16] http://www.bapco.com/techdocs/SYSmark2007Preview_WhitePaper.pdf
[17] This is the number of non-cached I/O requests received by the file system. Cached I/O requests are not directly included as they don't directly cause a disk request, however if there is a cache miss, a non-cached read to satisfy the page fault will be sent to the file system, and at that time included in the total.