RAXCO SOFTWARE

Third-Party White Paper

Impact of Free Space Consolidation On Windows File System Performance

Property of
Balder Technology Group, Inc.

May 2004

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]. Microsoft, the maker of the Windows file systems, and industry analysts concur on the problem. In February 2003, Gartner issued a Research Note[3] in which the authors made this observation about the effects of the Windows built-in defragmenter and its multi-pass defragmentation engine on the remaining free space on the disk:

"The Windows 2000 Server built-in defragmentation tool is a multi-pass defragmenter that must be run over and over to defragment the disk, especially when defragmenting very large disks with heavy fragmentation and limited free space. As such, multi-pass defragmenters characteristically fragment the remaining free space on the disk, which accelerates fragmentation later. It is recommended that a third-party single-pass server defragmentation tool be implemented instead."

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 servers and workstations.

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 servers and workstations.

Defragmentation, seemingly a straightforward concept, has actually evolved considerably over the past few years. For example, vendors came out with different approaches to the problem, including an attempt to stop fragmentation from occurring, utilizing thresholds to minimize how often defragmentation runs are necessitated, and more. Free space consolidation has received increased interest over the last few years.

This paper details a study done by the Balder Technology Group to determine the impact of free space consolidation on Windows file system performance. It attempts to answer the question whether free space consolidation matters. For all the talk about it from vendors – is the effort and expenditure worth it.

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®'s emphasis on free space consolidation provides performance improvements beyond those of Diskeeper® and the built-in defragmenter. Indeed, their inability to effectively consolidate free space can be detrimental to system performance.

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 [4].

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[5]. 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 for Windows 2000 Datacenter Edition, resulting in a professional paper[6] and a patent issued to Microsoft[7], with Mr. Goebel listed as a co-author and co-inventor respectively.

Other notable contributions include staffing all eleven of Microsoft's file system plugfests[8] and the file system development labs before them. Microsoft sent him along with a Microsoft employee to Romania last year to perform technical due diligence on its acquisition of GeCAD anti-virus technology[9] and is currently writing Microsoft's anti-virus file system filter driver to interface with that technology.

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[10] (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 disk defragmentation products that have starkly different perspectives on free space management were chosen: Raxco Software's PerfectDisk® and Executive Software's Diskeeper®.

To test the effect of free space consolidation, two disk defragmentation products that have starkly different perspectives on free space management were chosen: Raxco Software's PerfectDisk® and Executive Software's Diskeeper®. 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 Version 8.0 and PerfectDisk Version 6.0, both of which were available to the general public at the time of the testing.

Executive has stated "Free space consolidation might be important if you have to create one gigantic contiguous file, but it has no effect on performance."[11] Raxco claims "PerfectDisk's one pass defragmentation technology results in fast and efficient defragmentation and consolidation of free space, eliminating the need to run over and over again like other defragmenters."[12] 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."[13]

Testing Background

To conduct our tests we needed two tools: a performance benchmarking tool and something that would actually measure physical disk accesses. We selected the Veritest WinBench benchmarking suite for the performance testing (WinBench actually uses WinStone data to measure disk performance. WinStone is another Veritest benchmark that installs and runs a

number of Windows applications like: MS Excel, MS Word, MS Front Page, MS Access, MS PowerPoint, Norton Anti-Virus, WinZip, Lotus Notes, MS Project, and Netscape Communicator.). The benchmark uses these applications to create, edit, and delete various file types. WinBench simulates the reads/writes that would have occurred with the WinStone applications to benchmark disk activity in a typical environment. WinBench and WinStone are recognized standards for this type of testing. See www.veritest.com for additional details on WinBench and WinStone.

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 Driver Development Kit (DDK).

Test Equipment and Methods

The two following system configurations were used in our evaluations.

Table 1: System Configuration[14]

Computer	Operating System	Disk
HP/Compaq 8500, 8x700MHz Intel® Pentium® III Xeon™ processors, 4GB RAM	Windows Server 2003	Compaq StorageWorks 4314R, 120 GB, attached via a Compaq 5304 array controller card
AMD Athlon™ Workstation 2.1GHz, 1GB RAM	Windows 2000 SP4	Promise FastTrak 133 ATA RAID0, 80 GB

All disks were formatted with NTFS. Microsoft tools were used to populate the disk with random length files. A custom tool was used to fragment the resulting collection of files. The disk was imaged so it could be restored to provide an identical starting point for an accurate comparison of the respective products for each defragmentation run.

High-End Server Test Results

Starting Statistics

The Compaq StorageWorks disk array was used exclusively for test data (the HP 8500 had a separate 18GB system drive) and

formatted with 4k clusters and populated with 40 gigabytes of randomly sized files. After fragmentation, the file and metadata contained 253,170 excess fragments.[15] The volume had 72 gigabytes of free space (60%) that was in 87,327 pieces.

The disk was then defragmented with Diskeeper in its Improved Free Space defragmentation method. In accordance with Executive Software's recommendation, we first ran an online defragmentation to defragment the data files, and this completed in 2 hours 26 minutes. This was followed with an offline defragmentation to defragment the Master File Table, and the metadata files. The offline defragmentation completed in 2 hours 42 minutes, which did not include the required system reboot, for a total defragmentation time of 5 hours 8 minutes. After defragmentation with Diskeeper, there were zero excess file fragments; however, the free space was still in 72,720 pieces. This constitutes a 17% reduction in free space fragmentation.

The disk was restored to its original state and then defragmented with PerfectDisk. Raxco Software recommends running the offline defragmentation first, and this completed in 5 minutes. We then ran the online defragmentation which completed in 1 hour 3 minutes for a total defragmentation time of 1 hour 8 minutes. After defragmentation with PerfectDisk, there were also zero excess file fragments, and the remaining free space was in 2 pieces. This constitutes a 99.998% reduction in free space fragmentation.

Table 2: High-End Server Summary Table

	Excess File Fragments	Free Space Fragments	Duration
Before Defrag Run	253,170	87,327	
After Diskeeper Run	0	72,720	5 hr. 8 mins.
After PerfectDisk Run	0	2	1 hr. 8 mins.

Disk Drive Performance Analysis

In order to establish a performance baseline we ran WinBench 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 WinBench provided after defragmentation with the commercial products.

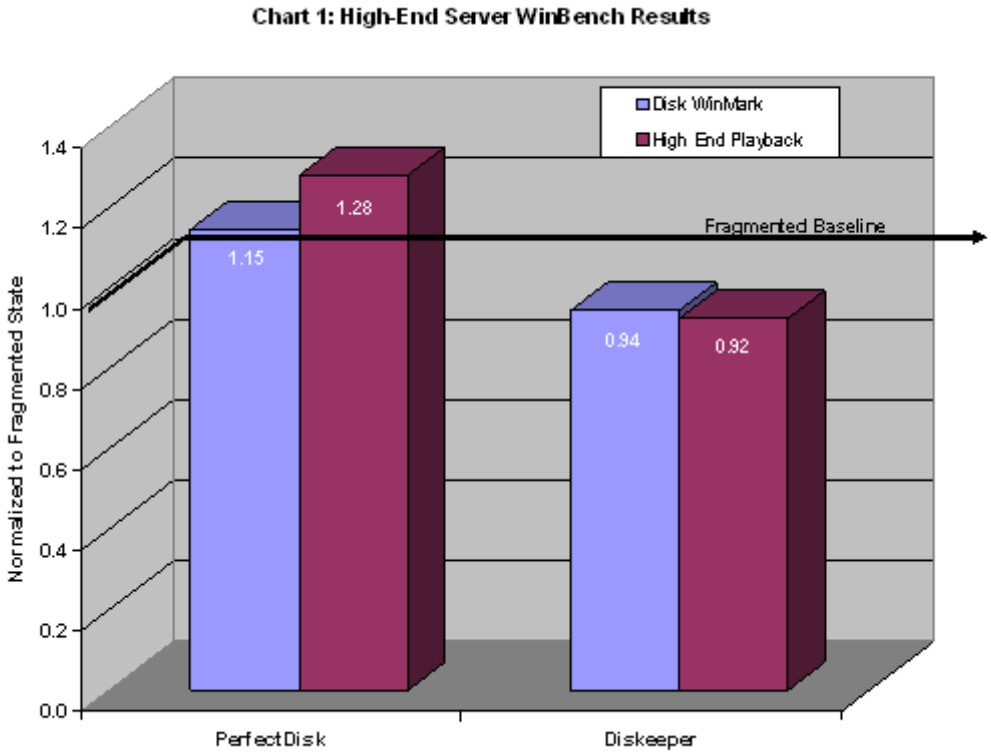
We then ran it on the partitions after the respective defragmentation passes. WinBench results are given as two numbers: a Business Disk WinMark score and a High End Disk Overall Score. Both are in thousands of bytes per second. Higher scores indicate better drive performance. The important thing to remember about WinBench scores is that a ***higher score is better***.

For our test on Windows Server 2003 with the 120GB Compaq disk array we got the following WinBench scores.

Table 3: WinBench Result on High-End Server

	Business Disk WinMark (Disk throughput in thousands of bytes per second)	High-End Disk WinBench Overall (Disk throughput in thousands of bytes per second)
Base Fragmented Disk	6,740	42,100
After PerfectDisk	7,720	53,900
After Diskeeper	6,360	38,900

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



The Business Disk WinMark and High End Playback scores show that Diskeeper underperformed the base fragmented partition by 5.6% and 7.6% respectively. Reducing the total number of free space fragments to 72,720 pieces did not improve the benchmark scores.

Conversely, the Business Disk WinMark and High End Playback scores show that PerfectDisk outperformed the base fragmented partition by 14.5% and 28.0% respectively. Since both products defragmented all the files, the consolidated free space and eventual location of individual files on disk were the only differentiators. PerfectDisk uses a patented technology called SMART Placementä[16] to determine the optimal target location for files. This technology uses last modify time to classify files into placement zones on disk. However, since the disk was principally populated with random files that had very similar last modify times, and more importantly weren't accessed during the WinBench test runs, the affect of SmartPlacement on this particular test scenario can be considered vastly outweighed by the consolidated free space. The conclusion is that consolidating the free space into 2 pieces did improve the benchmark scores.

To summarize, from a baseline fragmented disk drive, performance decreased after running Diskeeper, with the accompanying 72,720 free space fragments remaining; performance increased after running PerfectDisk, with the accompanying 2 free space fragments remaining.

These results lead us to the following conclusion:

· Defragging without free space consolidation does ***not*** reliably increase system performance.

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 WinBench, as well as the total number of seeks on the drives defragmented by PerfectDisk and Diskeeper.

Table 4: Wasted Seek Analysis on Compaq Array

SERVER	Total Number of I/O Requests Sent to the File System[17]	Total Number of Resulting Disk Accesses / Seeks	Net Wasted Seeks When Running WinBench
Fragmented	58,844	63,038	4,194 (7.12%)
PerfectDisk	58,517	59,469	952 (1.63%)
Diskeeper	58,955	65,227	6,272 (10.64%)

On the partition defragmented with Diskeeper, the total wasted seeks are 49.5% **greater** than experienced on the base fragmented partition. In other words, Diskeeper underperformed the baseline fragmented partition and resulted in a substantial increase in wasted seeks.

On the partition defragmented with PerfectDisk, 1.63% of the seeks were unnecessary. The total wasted seeks were over 77.3% **less** than experienced on the base fragmented partition. In other words, PerfectDisk resulted in far fewer wasted seeks than were experienced with the baseline fragmented partition.

From this data it is clear that in this environment consolidated free space had a profound effect on disk performance.

From this data it is clear that in this environment consolidated free space had a profound effect on disk performance.

Typical Workstation Test Results

Starting Statistics

The Promise FastTrack RAID 0 80 GB drive was split into two 40 GB partitions. One partition was used for system files and the other was used exclusively for test data. Each partition was formatted with 4k clusters and the test partition was populated with 37 gigabytes of randomly sized files. After fragmentation, the file and metadata contained 71,172 excess fragments. The volume had 1.378 GB of free space (3.4%) that was in 4,933 pieces.

The partition was then defragmented with Diskeeper with its Improved Free Space method in accordance with the vendor's recommendations. The online defragmentation took 55 minutes. After the initial defragmentation with Diskeeper, there were still 26,522 excess file fragments, and there were 14,109 fragments of free space. This was followed by an offline defragmentation that took 6 minutes 24 seconds. We then let Diskeeper make three more passes on the disk, since the Improved Free Space method requires multiple passes to consolidate free space.[18] The three additional passes took another 58 minutes, for a total defragmentation time of 1 hour 59 minutes. After the fourth pass completed there were 24,523 excess file fragments and the free space was in 13,022 fragments.

In this test, Diskeeper decreased the total excess file fragments by 65.5%, but it **increased** the free space fragmentation by 164%. This is the kind of negative behavior Gartner recognized with the built-in defragmenters multi-pass defrag engine.

The disk was restored to its original state and then defragmented with PerfectDisk in accordance with the vendor's recommendations. The offline defragmentation completed in 1 minute. The online defragmentation completed in 2 hours 55 minutes for a total defragmentation time of 2 hours 56 minutes. After defragmentation with PerfectDisk, there were zero excess file fragments, and the free space was in 22 pieces.

Table 5: Workstation Summary Table

	Excess File Fragments	Free Space Fragments	Duration
Before Defrag Run	71,172	4,933	
After 4 Diskeeper Runs	24,523	13,022	1 hr. 59 mins.
After 1 PerfectDisk Run	0	22	2 hr. 56 mins.

Disk Drive Performance Analysis

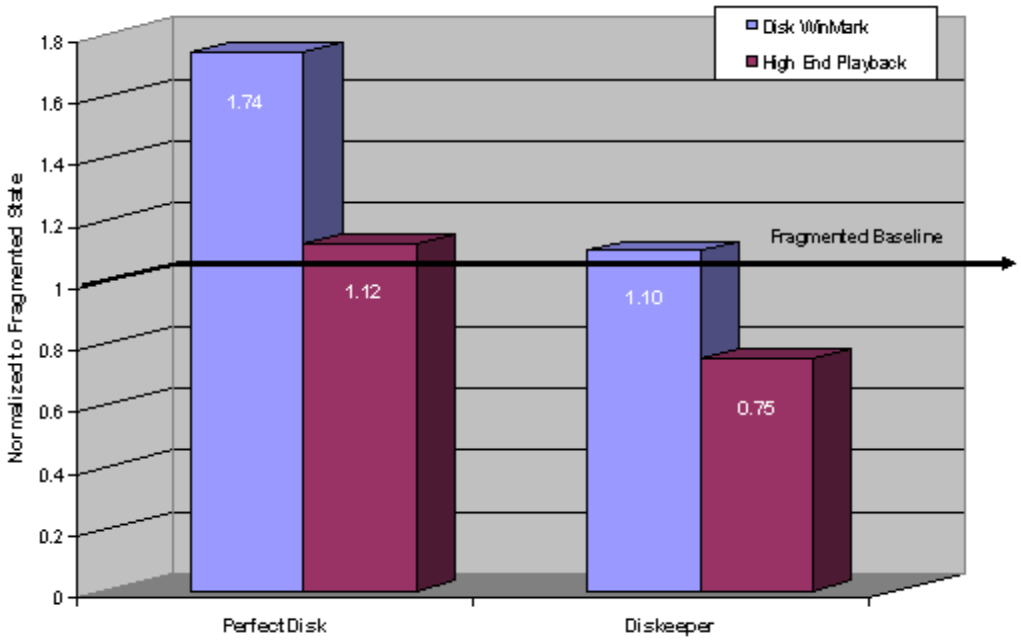
Once again we ran WinBench on the original fragmented partition to establish base scores and then on the respective defragmented partitions with the following results. Again, these results are in thousands of bytes per second, so the larger the number, the better the result.

Table 6: WinBench Result on AMD Workstation

	Business Disk WinMark (Disk throughput in thousands of bytes per second)	High-End Disk WinBench Overall (Disk throughput in thousands of bytes per second)
Fragmented	4,900	18,200
After PerfectDisk	8,540	20,400
After Diskeeper	5,410	13,700

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

Chart 2: AMD Workstation WinBench Results



The Business Disk WinMark score shows Diskeeper outperformed the base partition by 10.4%. The High End Playback score shows Diskeeper underperformed the base partition by 24.7%.

The PerfectDisk defragmented partition outperformed the base fragmented partition by 74.3% and 12.1% respectively.

Disk Access Analysis

Once again there is a correlation between the relative amount of fragmented free space and the number of Wasted Seeks.

Table 7: Wasted Seek Analysis on AMD Workstation

WORKSTATION	Total Number of I/O Requests Sent to the File System	Total Number of Resulting Disk Accesses / Seeks	Net Wasted Seeks While Running WinBench
Fragmented	56,745	58,458	1,713 (3.02%)
PerfectDisk	55,508	56,340	832 (1.50%)
Diskeeper	57,123	63,253	6,130 (10.73%)

On the partition defragmented with Diskeeper, the percentage of seeks which were unnecessary was 10.7%. The total wasted seeks are 255% **greater** than experienced on the base fragmented partition. In other words, Diskeeper underperformed the baseline fragmented partition and resulted in a substantial increase in wasted seeks.

On the partition defragmented with PerfectDisk, 1.50% of seeks were wasted. The total wasted seeks were 51.4% **less** than experienced on the base fragmented partition. In other words, PerfectDisk resulted in far fewer wasted seeks than were experienced with the baseline fragmented partition.

From this testing on the workstation, it is again clear that in this environment, consolidated free space had a profound effect on disk performance.

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.

Consolidating free space is singularly effective in postponing refragmentation of a file system.

When running the relatively small disk footprint WinBench test, PerfectDisk consolidated free space and scored up to 50% better than Diskeeper, and incurred almost an order of magnitude fewer wasted disk seeks. Indeed, in some cases, both the WinBench score and seek data were better **before** Diskeeper ran, as it further fragmented free space. This finding is consistent with the observation Gartner made in its assessment of the built-in Windows defragmenter and multi-pass defragmentation engines.

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's emphasis on free space consolidation provides performance improvements beyond those of Diskeeper and the built-in defragmenter. Indeed, the inability of Diskeeper and the built-in defragmenter to effectively consolidate free space can be detrimental to system performance.

(Footnotes)

[1] http://www.raxco.com/products/perfectdisk2k/whitepapers/ohio_sb_ss.pdf
[2] http://www.idc.com/en_US/templates/viewTOC.jhtml?containerId=29945
[3] Gartner Research Note Windows 2000 Server, February 16, 2003 Mary Hubley and Mary Ann Richardson
[4] Helen Custer, Inside the Windows NT File System (Redmond: Microsoft Press, 1994), pp. vii
[5] Microsoft Part # X09-46661.
[6] William J. Bolosky, Scott Corbin, David Goebel, and John R. Douceur. "Single Instance Storage in Windows ä 2000," In Proceedings of the 4th USENIX Windows Systems Symposium, pp. 13 — 24, August 2000.
[7] Bolosky, Douceur, Cutshall, Rashid, Myrhvold and Goebel. "Single Instance Store for File Systems". US Patent number 6,477,544. Nov 5, 2002.
[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] Fisher, Dennis, "Microsoft to Buy Antivirus Company." EWeek. Ziff David Media. June 10, 2003. http://www.eweek.com/article2/0,1759,1504728,00.asp
[10] A seek is the physical movement of the disk read/write head
[11] http://www.execsoft.com/eletter/previous/eletter.asp?F=2003070202.html#2
[12] http://222.raxco.com/products/perfectdisk2k/
[13] http://www.microsoft.com/windows2000/techinfo/administration/fileandprint/defrag.asp
[14] The hardware was identical for all tests.
[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] 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
[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.
[18] From the Diskeeper help file under "Using Different Defragmentation Methods." Improved free space consolidation – free space consolidation improvement will be gradual as the option is used over time.