Windows® IT Pro

The Impact of Disk Fragmentation on Servers

By David Chernicoff

Published: May 2009

The Impact of Disk Fragmentation on Servers

Contents

Diskeeper corporation

Testing Server Disk Defragmentation

IT professionals responsible for server hardware well understand the value that professional grade disk defragmentation software brings to their servers. Storage servers can experience high levels of disk thrashing (the constant writing and rewriting of small amounts of data) caused from excessive file fragmentation.

Problems in delivering services to users however are difficult to directly trace to server fragmentation issues. Network and application issues have a much more visible impact on the performance of network-based services, especially when problems with those functions are encountered. But with the negative impact on ROI that network performance problems cause, IT pros would be ill-advised to overlook the advantages that assuring the optimization of the underlying hardware infrastructure can bring. Optimal disk performance translates into better ROI. Testing will bear this out.

We tested the impact of server disk defragmentation by looking at common tasks that network servers, both physical and virtual, encounter, ranging from maintenance tasks such as server backup and anti-virus scans, to basic knowledge worker tasks involving opening files stored on the host server and virtual machines, and manipulating email. We also looked at tasks that are more taxing on the server, such as database queries, index creation, and bulk updates. Each test was performed as the sole task on the server.

When considering the results of our testing keep in mind that a production environment will see significantly heavier server use, which results in much greater potential for ongoing disk fragmentation. In your production environment with dozens, if not hundreds, of users touching your server storage simultaneously, your disk fragmentation can become severe in a very short time. Preventing this fragmentation from affecting server performance is an ongoing process.

The Testing Environment

For our benchmark tests we used an HP ProLiant DL380 G5 equipped with dual quad-core 2.83 GHz Xeon processors, each with a 2x6MB L2 cache, 16 GB of RAM and seven 72 GB 10,000 RPM SCSI drives attached to an HP Smart Array P400 controller that has a 256 MB cache and that supports both serial-attached SCSI and SATA drives. The volumes we tested against were 30 GB, 80 GB, and 175 GB. We used a 500 GB 7200 RPM locally attached SATA drive for backup only. The server operating system was Microsoft Windows Server 2008 Enterprise; the application server software installed in VHDs was Microsoft SQL Server 2008 and Microsoft Exchange Server 2007. All server software was updated with service packs, patches and hotfixes current as of February 2009. The disk defragmentation software used was Diskeeper Server.

The seven SCSI drives attached to the array controller were configured as two physical drives. We used the first physical drive, comprised of two drives configured as a RAID 0 stripe set for maximum performance, for the installation of the operating system and all related files. We configured the remaining five drives as a RAID 5 stripe set to be representative of the type of hardware storage configuration found in most business environments. We performed all applications, VHDs, and tests on the RAID 5 stripe set. The volume size was dependent upon the test level.

As an example of the effect fragmentation can have, the screen capture in Figure 1 shows the Diskeeper fragmentation analysis of a severely fragmented disk. The severe fragmentation documented here will have a negative impact on storage performance.

We tested three levels of fragmentation, described herein as low, medium, and high. We used the Diskeeper Diskcrusher fragmentation utility to create fragmented files and directories. We ran all tests a minimum of three times with the results reported here being the average of all test runs.

	Low	Medium	High
Number of files	101,652	1,220,660	2,087,158
Avg. Number of Fragments per File	3.21	1.69	2.30
Number of Fragmented Files	99,074	613,221	1,994,117
Number of Excess Fragments	225,216	840,076	3,005,400
Percent Fragmented – Volume	40%	50%	84%
Percent Fragmented – Data	51%	58%	91%
Free Space	22%	15%	15%

Table 1: Fragmented disk test configurations

As shown in Table 1 the level of fragmentation and the number of affected files increases with each testing tier. The level of fragmentation you'll encounter in production environments is dependent upon the level of use and types of applications the server deals with. In all likelihood, if your server storage levels are consistently exceeding 75 percent or so, you've begun aging data off of the servers or you're planning to add additional storage. While fragmentation isn't a direct result of reduced capacity, the chances for fragmentation increase as free storage space decreases and the operating system is forced to write data into an ever-increasing number of non-contiguous spaces.

By using an automated defragmentation process, the same disk volume sees absolutely minimal fragmentation even though it is in continual use by applications and users (Figure 2).

Figure 1: Fragmentation map of a heavily fragmented disk

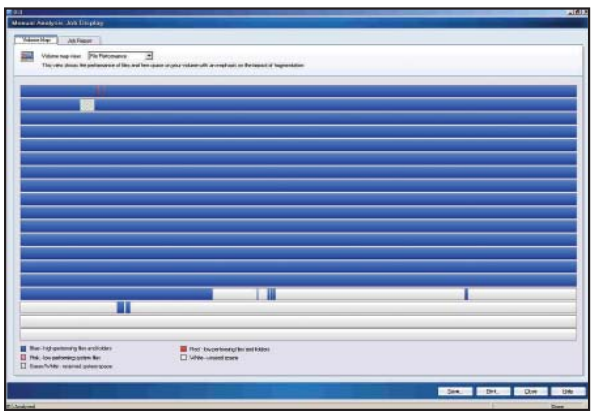

Figure 2: Fragmentation map after automated defragmentation by Diskeeper.

We ran each set of tests for three iterations, and then defragmented the storage using Diskeeper to reduce or eliminate the disk fragmentation. We repeated each test (also for three iterations) and averaged the results. In the following test descriptions and analysis, the comparisons are all before and after defragmentation at each specific fragmentation level tier. We did not do cross-tier comparisons. All test times are reported in seconds.

	Low	Medium	High
Number of files	101,652	1,220,660	2,087,158
Percent Fragmented – Volume	0	0	0
Percent Fragmented – Data	0	0	0
Avg. Number of Fragments per File	0	0	0
Number of Fragmented Files	0	1	1
Number of Excess Fragments	0	2	4
Free Space	22%	15%	15%

Table 2: State of fragmentation after Diskeeper has been run

The Tests

In our first set of tests we look at common server tasks that are likely to be affected by disk fragmentation. These tasks are all primarily storage related; that is, the performance of the storage media will have a primary impact on the performance of these tasks.

File Copy

In the file copy test, a folder containing 5 GB worth of files and sub-directories was copied from the test volume to the boot volume of the server. To minimize variables, the copy was done locally, not across the network. We timed the test using a stopwatch. This is one of the most basic tasks done with server data and, in a severely fragmented environment, showed some of the most significant performance improvements.

File Copy Tests (measured in seconds)

Low – Fragmented	44
Low – Defragmented	39
Medium – Fragmented	72
Medium – Defragmented	60
High – Fragmented	97
High – Defragmented	54

The basic task of moving data from one location to another on the server shows that a fragmented disk has a major negative impact on the file copy. Even the lightly fragmented low-level test showed an improvement in copy time of over 11 percent, while the copy that was done from the very highly fragmented drive improved in time by almost 45 percent. Given how common the file copying task is the benefit is clear. Defragmented disks are a significant time saver for common user tasks.

While the limiting factor in doing a file copy from the server to the client might be the available network bandwidth, as technologies such as Gigabit Ethernet become more common, the base limiting factor will be how fast the operating system can feed data to the network request, which is directly impacted by fragmentation of the data on the local drive.

Document Open

In this test, a 100-page Microsoft Word document was opened from the server to a Windows XP client running Microsoft Office 2007. The size of the document was 3.3 MB.

Document Open Tests (measured in seconds)

Low – Fragmented	11.7
Low – Defragmented	10
Medium – Fragmented	12.7
Medium – Defragmented	10.7
High – Fragmented	14.7
High – Defragmented	10.3

Our test results showed performance improvements of upwards of 30 percent. In the case of any file load from server to client the performance improvement will be determined by just how badly fragmented is the file located on the server. In our tests, the file was clearly badly fragmented, significantly so at the highest level of fragmentation testing. To prevent this type of file fragmentation, the best methodology is an ongoing background file defragmentation process, the benefits of which are clearly demonstrated by this test. And given how often this type of task is performed in most business environments, the value of the defragmentation cannot be understated. As shown in this and the File Copy test, basic data manipulation is much faster on defragmented storage.

Backup
In the first test, we backed up the test volume using disk-to-disk backup as supported by Windows Server Backup, which is a component of Windows Server 2008. Backup was done using the VSS copy method, which is designed to work with other backup tools that would require that the archive and backup information in the files remain unmodified. We backed up to a SATA-attached dedicated hard drive that was reformatted between tests. Timing was done using the backup application.

Backup Tests (measured in seconds)

Low – Fragmented	1193
Low – Defragmented	1130
Medium – Fragmented	2787
Medium – Defragmented	2300
High – Fragmented	6960
High – Defragmented	6620

While different backup tools will be differently affected by disk fragmentation, our tests showed one simple fact; defragmented disks back up faster. Individual runs demonstrated performance improvements of up to 20 percent with our test data set and the built-in Windows Server backup. Our least effective test result, a large data backup that can represent a significant amount of time, still showed an improvement of 5 percent. Our highest report results, which averaged a 17 percent reduction in backup time, shows that reducing or eliminating disk fragmentation prior to backup will allow larger amounts of data to be backed up, especially if time is a constraint in your backup process. If backup is run as a background application, reduced fragmentation will allow for lower resource consumption necessary for the backup process, minimizing further the impact of the backup on active users of the storage.

The single, consistent result that appears in all of our tests is that defragmented server drives using Diskeeper deliver better performance.

Anti-Virus Scan
For the AV scan test, we performed a complete scan of the test volume using the Kapersky Lab AntiVirus Version 6 Windows Server software, current as of February 2009. The default configuration of the AV software was used with only the test volume selected for scanning. Timing was done using the AV application.

Anti-Virus Scan Tests (measured in seconds)

Low – Fragmented	256
Low – Defragmented	238
Medium – Fragmented	1485
Medium – Defragmented	1359
High – Fragmented	4428
High – Defragmented	4004

Many factors will have an impact on the speed of a complete anti-virus scan of your storage. The way the scanner works, the total number of files that need to be scanned, the size of the files, and the fragmentation level of the disk all have a direct impact on the length of the AV scan process. In our tests with the Kapersky Lab AV solution, the disk defragmentation resulted in upwards of a 10 percent performance improvement—with the improvement being more significant as the test drives increased in size, number of test files, and fragmentation.

VHD Start
This test measured the amount of time it took to launch the saved test virtual machine. The VM was launched from a saved state and timing stopped when the Hypervisor manager reported that the VM was successfully started.

VHD Start Tests (measured in seconds)

Low – Fragmented	62.3
Low – Defragmented	51
Medium – Fragmented	60.7
Medium – Defragmented	58
High – Fragmented	55.3
High – Defragmented	47

With as much as a 17 percent improvement in the start time of the test virtual machine, the effects of fragmentation on the VHD are clear. This fragmentation will also impact the performance of the VM itself, because all of the additional I/O necessary to read from a severely fragmented VHD will reduce the performance of the virtual computing environment. Fragmentation must also be watched if your VMs are configured with the dynamic disk option, which allows the virtual machine to grow the size of its storage as necessary. This means that as the size of the VHD grows it will continue to fragment into the available space on the hard drive. Making sure that the host machine hard disk is regularly defragmented and managed will improve the performance of virtual machines running on the host and allow for the use of dynamic disk allocation within the VM without danger of disk performance issues.

Even with significant free space of the disk, as shown by the white space in the fragmentation map (Figure 3), major fragmentation can still occur even without VHD test volume.

Figure 3: Fragmentation map of VHD volume

VHD Save

This test measured the length of time required to save the test virtual machine. From the Hypervisor manager, the running machine was saved and timing stopped when the manager reported the save complete.

VHD Save Tests (measured in seconds)

Low – Fragmented	365.3
Low – Defragmented	271.7
Medium – Fragmented	409.3
Medium – Defragmented	402
High – Fragmented	447.7
High – Defragmented	390.3

With test results indicating as much as a 25 percent performance improvement after defragmentation, the VHD Save tests show quite clearly the effect of writing a very large file to a fragmented hard drive. The more fragments on the drive the less likely it will be that a large file can be written contiguously. And in the world of virtualization, large files are the standard, and the need to be able to read and write those files with a minimum of fragmentation is a requirement to meet the basic ROI needs of the enterprise.

Automated background defragmentation results in a major reduction in fragmentation even with an active VHD (Figure 4). Regular use of the background defragmenter will continue to minimize fragmentation.

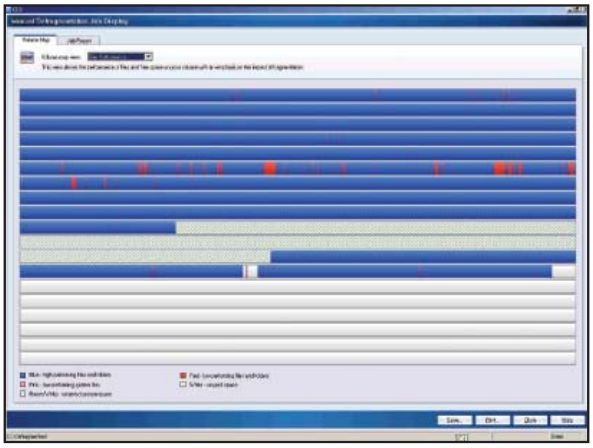

Figure 4: Fragmentation map after automated defragmentation by Diskeeper.

Server Application Tests

In the server application tests we looked at the impact of fragmented storage on server-based applications. Other factors will have an impact on the overall performance of these applications; optimizing storage strategies, including defragmentation, reduces the impact of storage performance on the overall application performance.

Exchange Test One

In this first Exchange test, the client, a Windows XP Professional Workstation running Office 2007, uses Outlook to open 100 messages from the server. One hundred messages are highlighted then opened simultaneously. Timing starts when the open is launched and stops when all of the messages have been opened and console control returns.

Exchange Test One (measured in seconds)

Low – Fragmented	7.7
Low – Defragmented	7
Medium – Fragmented	10.7
Medium – Defragmented	8.6
High – Fragmented	18.4
High – Defragmented	11.6

While the impact of server fragmentation gets significantly greater as the disk becomes more fragmented, even the common lower levels of fragmentation will have a large impact on user response time when you consider that hundreds of users may be accessing the data store at the same time. Delayed response time for email users is a generator of a large percentage of help desk calls, and implementing a defragmentation strategy can help to solve the problem. As our tests show, allowing the data to become seriously fragmented can have a major negative impact on the

Exchange user experience with a 40 percent reduction in performance in our highly fragmented test environment. Good defragmentation strategies result in fewer help desk calls.

Exchange Test Two
In this test, the contents of an existing folder were moved to a new folder. Time to complete was measured from the client side.

Exchange Test Two (measured in seconds)

Low – Fragmented	9
Low – Defragmented	8
Medium – Fragmented	13.8
Medium – Defragmented	9
High – Fragmented	24.9
High – Defragmented	12.3

A new folder was created and the contents of the Inbox were moved to the new folder. With our heavily fragmented test environment showing a greater than 50 percent performance improvement after defragmentation it's clear that this test was extremely sensitive to higher levels of fragmentation on the server. If users are often found reorganizing the data in the Exchange mailbox, the impact of fragmentation can be quite severe.

SQL Server Bulk Insert
We tested SQL Server 2008 with a bulk insert of 50,000 rows of data. The bulk insert is often the fastest method of getting data into a SQL Server database.

SQL Server Bulk Insert Tests (measured in seconds)

Low – Fragmented	22.1
Low – Defragmented	20.9
Medium – Fragmented	31
Medium – Defragmented	25
High – Fragmented	53.3
High – Defragmented	33.4

As has been seen with the Exchange tests, a highly fragmented database structure can have a severe negative impact on loading and extracting data from server applications, with our test showing a performance improvement of 40 percent in the most heavily fragmented environment. Because Microsoft offers APIs for moving open files, defragmentation software is able to safely work on database files without risk of data loss or corruption. Loading data into a defragmented environment not only improves load times but reduces the amount of disk thrashing necessary to manipulate the data and the amount of work that is necessary to later defragment the database.

Table Key Creation (measured in seconds)

	Table 1	**Table 2**
Low – Fragmented	12.5	15.9
Low – Defragmented	12	14.9
Medium – Fragmented	14.1	18.23
Medium – Defragmented	12.4	17.1
High – Fragmented	25.5	32.4
High – Defragmented	20.6	25.3
	Table 3	**Table 4**
Low – Fragmented	26	35.4
Low – Defragmented	24.2	33
Medium – Fragmented	32.3	49.1
Medium – Defragmented	30.4	43.8
High – Fragmented	51	68.8
High – Defragmented	46.7	61.3

In this test each table was opened, a field was selected as the primary key, and the change was saved. The table key creation times are directly related to how much data SQL Server had to touch, and the level of fragmentation that had to be dealt with. SQL Server 2008 does a very good job of managing its databases, but defragmentation shows appreciable improvement in the performance of tasks such as this with a performance improvement of over 11 percent in the most fragmented environments.

With the SQL queries, the two tests differ primarily in the amount of data that SQL Server returns in response to the query. The tests depict the effects of manipulating the data on a fragmented drive with peak performance improvements of approximately 18 percent.

SQL Query 1 – Simple (measured in seconds)

Low – Fragmented	23.9
Low – Defragmented	22.3
Medium – Fragmented	28.2
Medium – Defragmented	24.8
High – Fragmented	43.5
High – Defragmented	33

SQL Query 2 – Complex (measured in seconds)

Low – Fragmented	35.3
Low – Defragmented	33.3
Medium – Fragmented	41.5
Medium – Defragmented	38.5
High – Fragmented	61.3
High – Defragmented	50.8

Conclusion

The single, consistent result that appears in all of our tests is that defragmented server drives using Diskeeper deliver better performance.

Every application that touches the hard drive will benefit from a good tool that defragments and manages the files on your servers.

Almost every role filled by Windows servers in your computing environment will benefit from the use of disk defragmentation software. The simplest file and print services delivery requires a significant amount of disk I/O and will easily benefit from file defragmentation. As our simple tests show, even Exchange and SQL Servers benefit from defragmentation; reading and writing data with either application simply works better when the files are not fragmented. The result is improved performance.

Throwing more storage resources (hardware) at a problem should be the last resort, because it only masks the potential problems that intelligent disk defragmentation addresses.

Our test results showed performance improvements of upwards of 30 percent.

Quicker response time in databases and mail servers means that more time is spent getting work done, rather than waiting for information to be delivered.

Diskeeper is the only true server defragmentation software that runs silently in the background, continually improving performance.

With the current economic and business environment, maximizing ROI becomes even more critical. Adding Diskeeper to your server toolkit gives you the ability to get the maximum speed from your storage subsystems of your existing hardware.

David Chernicoff is a technology consultant with a focus on the mid-market space, *Windows IT Pro* Senior Contributing Editor, founding Technical Director for *PC Week Labs* (now *eWeek*), former Lab Director for *Windows NT Magazine/Windows 2000 Magazine* (now *Windows IT Pro*) and formerly Chief Technology Officer for a network management tools ISV. David has been writing computer-related feature and product reviews for more than 20 years and is coauthor of a number of operating system books, ranging from the *Windows NT Workstation: Professional Reference* (New Riders Publishing), to the *Microsoft Windows XP Power Toolkit* (Microsoft Press), as well as over a dozen eBooks on topics ranging from network switching topologies to production FAX technology.