

Is Preventing Fragmentation Needed in Today's Environment?

A New Storage Age

For years computing demands have been pushing along Moore's Law. That "law" is directly related to CPU computing power and states that the number of transistors per CPU doubles over certain time intervals (12, 18, or 24 months - depending on when or where you heard it). However, in the new millennium the growth of data storage has been so rapid it is even exceeding the industry standard growth index laid forth by Moore's Law (any way you define it).

The exponential growth of storage requirements is driven by the Information Age, the public's unquenchable thirst for information and the increased complexity and size of applications, operating systems and data.

Additionally another huge influence on the IT world has been government regulations such as Sarbanes-Oxley, the Patriot Act, Health Insurance Portability and Accountability Act (HIPAA) as well as the Gramm-Leach-Bliley act. These regulations often govern documenting certain business transactions and the retention and security of all relevant data. Many in the industry are required to store everything for periods as long as 7 years to maintain regulatory compliance. Deleting files has become taboo.

A booming storage industry has grown in response to this new age. Hard drive manufacturers increase storage capacities at rates that often keep pace with transistor count increases. Advanced storage solutions such as Cloud Storage, Storage Area Networks (SAN), Network Attached Storage (NAS), Storage Management (e.g. deduplication, thin provisioning) software, and e-mail archiving solutions continue to fill the headlines.

New Maintenance Practices Needed

In this new climate, it's important to re-examine old beliefs and ensure our storage and systems management practices are sufficient for today's environment.

Take for example a modern enterprise SAN environment. Here scheduled processing of files is increasingly more impractical, given the sheer amount (terabytes+) of files to manage. Alternative designs that function in real time only on file *changes* (at the block) level are increasingly more commonplace. SAN snapshots are one such example of continuous data protection used to augment, or sometimes applied in lieu of, scheduled file backup.

Information Technology needs to continually adapt to the changing demands of the business. Storage management is a crucial topic, and the need to maximize the efficiency of storage is now more vital than ever before.

Disk Fragmentation is More Prevalent Than Ever

Today the number of files stored on volumes is much greater than times past. The increase in the number of files not only necessitates larger storage capacities, but due to inherent fragmentation problems, puts a burden on local disk file systems to keep files stored contiguously.

As files are deleted, the space the deleted files formerly occupied will be randomly spread across the disk(s). These non-contiguous segments of available space encourage new files to be created in places where they can't be written contiguously.

As a general rule, more files equal more fragmentation problems.

Another fragmentation issue is the increasing size of files. The typical Word or PowerPoint® document is bigger than ever. Additionally the use of video and graphic files have become commonplace and these files have grown to massive proportions. Bigger files have an obvious connection to increased file fragmentation.

Another general rule: bigger files equal more fragmentation problems.

With the exponential growth of storage, managing one's backup window becomes a major challenge when designing storage architectures and setting backup practices. Handling disk fragmentation is vital to managing backup windows when file level backups are performed. Many system administrators are battling the ever-expanding backup window. In fact, it is not uncommon for file-based backup times to exceed 24 hours, driving storage managers to seek out continuous data protection, data segregation, deduplication, expensive hardware, and other strategies, lest they risk data loss. Recent studies have shown that defragmenting before backups are performed can decrease backup times by up to 69%¹, often making a spiraling issue more manageable. At the very least, it provides breathing room en route to permanent solutions.

Not Just a Server Issue

One might mistakenly consider that since user files and data are stored on servers in the typical enterprise client-server environment, disk fragmentation doesn't occur fast enough to warrant frequent defragmentation jobs on desktops. Nothing could be further from the truth.

When managing storage devices in a client-server environment it's important to consider the files that are temporarily created on users' local hard disks, and files that are backed up locally by commonly-used applications. Applications such as Microsoft Outlook, web browsers and many others create and use files in the background. These background files are very often heavily fragmented and since the user is running applications that have to operate with fragmented files, the user is really feeling the performance degradation.

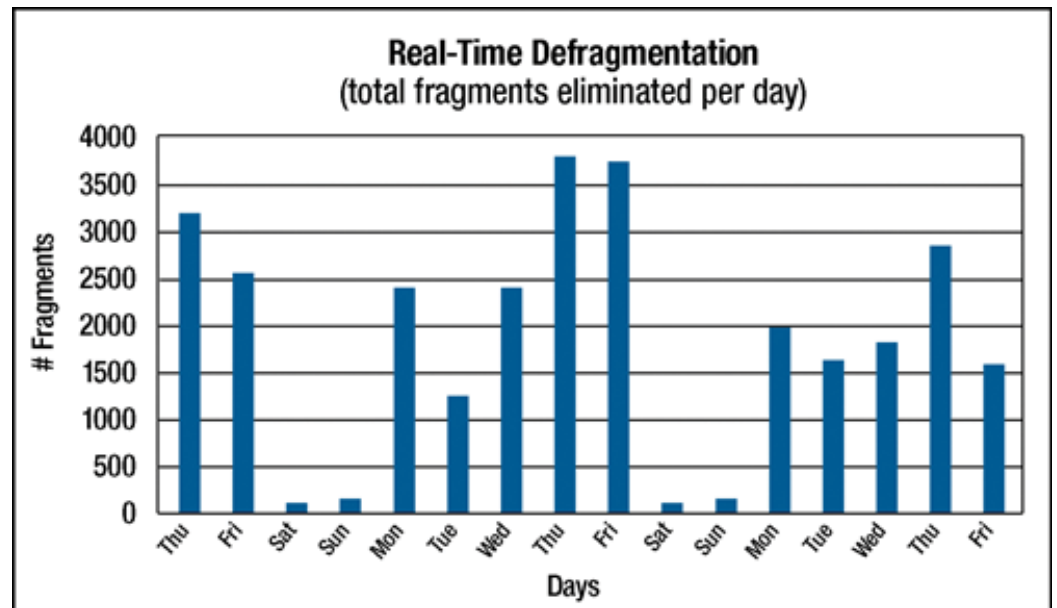
1. Improving Backup Performance with Defragmentation (A business application study), Diskeeper Corporation

The proliferation of wireless technology and improvements in security have fueled a new mobile workforce. Corporate culture more readily accepts work-from-home employees. Laptop sales now commonly outpace that of the desktop counterparts. The workforce, for all the consolidation and centralization of servers, is becoming more widely distributed in the workstation user population. For those segments, data distribution and/or synchronization are an increasing reality.

Real World Test Results on Fragmentation Levels

An experiment was performed using a desktop running Windows Vista in a typical Monday through Friday business environment which utilizes a file server to store user documents. For this test, Diskeeper was installed and allowed to perform reactive defragmentation (i.e. defragment data that is written in a fragmented state). The desktop user went about his normal job-related activities for 2 weeks².

Normal operations for the user included, Internet browsing (Mozilla Firefox), e-mail (Microsoft Outlook w/offline store), word processing (Microsoft Word), spreadsheet (Microsoft Excel), and design (Adobe Photoshop, Microsoft Visio).



The most heavily fragmented files included a 2.5 GB Outlook file (this OST file was locally stored) and several System Restore files. The majority of fragmented files were part of the user's profile, affecting logon/logoff.

A similar previously published two-week experiment was conducted on Windows XP, where the user used only Word and Internet Explorer (significantly less activity). The results showed accumulating fragmentation topping 4000 fragments over the same period.

2. Diskeeper had been used to fully clean up fragmentation on this system up until two days prior (Tuesday) to the beginning of this test and had left the machine with fewer than 100 fragments. The test was then begun on a Thursday, after the Vista built-in defragmenter had performed its default scheduled defragmentation job the night before. The Vista defragmenter still left 1245 fragments after its run, which was meant to address only the prior two days of fragmentation buildup (since Diskeeper had been removed).

As demonstrated in the chart, fragmentation levels rise quickly on the desktop, resulting in performance being degraded each day. Cumulative weekly buildup reaches more than 12,000 fragments each week³. That is fragmentation that a weekly defragmentation job such as on Windows Vista and Windows 7, does not address.

Even though this computer stores only very few files locally, fragmentation will slow system performance and hurt user productivity. Worse yet, as fragmentation levels increase for larger files that are not addressed, they will get up into levels where one can start experiencing reliability problems. Had this been a computer on which a great deal of data is locally stored (perhaps an attorney, field engineer, or salesman's notebook) the degree of fragmentation would most certainly be higher, and cause an even more significant impact to productivity.

Evolution and Intelligent Design

Over the years the Diskeeper system performance solution has evolved from simple fixed schedules to heuristic scheduling to real-time defragmentation. However, when fragmentation occurs, the system is wasting precious I/O resources by writing non-contiguous files to scattered free spaces across the disk. The best strategy is, deviating from the evolutionary path, a truly *revolutionary* one; prevent the problem from ever happening in the first place.

The exclusive IntelliWrite™ technology does just that. Its design *prevents* fragmentation (up to 85% of it) from being written to the hard drive by writing the files intelligently. The benefits are readily evident; continuous peak performance, and no administrative overhead for IT departments.

Consider again the above test case. Instead of 3500 fragments a day, a system with IntelliWrite technology generates only 500 or so fragments. Those 500 extents are quickly corrected a few minutes after the fact with additional advanced real-time technology also exclusive to Diskeeper.

Maintaining System Uptime

System administrators using the manual disk defragmenter built into Windows XP/2003 must wait for lengthy defrag jobs, which use up enough system resources that it must be done off-line or after hours. The Task Scheduler can be used to schedule jobs, but that itself generates massive management overhead, and it is still unlikely to solve the issue of fragmentation. Windows Vista/7/2008/R2 all provide essentially the same basic built-in pre-scheduled defragmenter. As the study shows, fragmentation increases at a phenomenal rate. Far beyond the rate previously experienced with earlier operating systems, hence it is understandable that Microsoft has made efforts to address this with their built-in utility in newer version of Windows. The weekly buildup tops 15,000 fragments. As is also shown, fragmentation is

3. Note these tests can easily be re-created and verified by installing trial software and viewing the Diskeeper History graphs.

never truly eliminated in the weekly pre-scheduled job and actually accumulates from one week to the next. The problem continues to exist, and it could be argued that on a day-to-day basis, the issue has actually worsened.

On the other end of the spectrum is Diskeeper, the market-leading solution, with over a decade of system performance innovations for Windows file systems. Diskeeper invented defragmentation for Windows in the mid 1990's, and maintains innovative leadership to this day as the *only* solution to prevent fragmentation.

Conclusion

In today's environment of bigger disks storing not only larger files but more files than ever before, the effects of fragmentation worsen markedly with each day's use. To eliminate the window of performance loss (between defragmentation cycles), eliminate wasted I/O resources and maximize file write performance, fragmentation should be prevented.