

NTFS/FAT Disk Defragmentation in RAID and SAN Environments



This document explains how disk defragmentation really works and how it can deliver significant benefits to RAID or SAN/NAS storage.

There is a common misconception that disk defragmentation is not necessary in RAID or SAN/NAS storage environments. This document explains how disk defragmentation really works and how it can deliver significant benefits to RAID or SAN/NAS storage.

The idea that disk defragmentation software is not necessary on RAID or SAN storage arises from a misunderstanding of how disk defragmentation software works and how it interacts with the disk controller software in these environments. At the core of this issue is the common perception that disk defragmentation software directly shuffles around disk clusters. That belief is incorrect.

The Windows file systems, FAT and NTFS, “know” absolutely nothing about the configuration of the attached storage. They do not know if a disk is IDE, SCSI, RAID, or SAN. The only thing the file system knows is what kind of partition it is, NTFS or FAT, and how big it is. When a disk is formatted, the file system enumerates the drive into logical disk clusters starting with Logical Cluster Number (LCN) zero up to LCN NNNNNNNN, depending on the size of the disk. As file space is allocated and deleted, the file system updates its logical map of the LCN’s.

When an application requests a file, the file system queries the Master File Table (for NTFS) and obtains the starting LCN and the run length for the file. If the file is logically contiguous, the file system need only do a single logical I/O to access the file. If the file is in more than one logical piece, the file is fragmented. In this case, the file system must locate the starting LCN and run length for every piece of the file. It is not uncommon for “write-active” files to be in thousands of fragments. To access these files the file system must perform thousands of logical I/O’s using both time and resources.

Disk defragmentation software defragments the LCN’s. The Microsoft Defrag API lets defrag software employ different algorithms to invoke their defragmentation strategy.

Defragmentation software identifies the LCN's it wants to move and the target LCN location. The API integrates these requests with the Windows memory manager, the cache manager, and the file system. The file system then updates the pointers to indicate the files new logical location. Access time to read a file that is logically contiguous versus one that is logically fragmented can improve 30-60%. Performing fewer logical I/O's takes less time, uses less CPU and memory resources, and improves system performance.

Everything discussed so far has nothing to do with the physical placement of clusters on a disk. When a file is created and the file system maps its LCN's, these are then passed to a disk controller. The controller maps the LCN's to physical disk clusters (PCN's). When defragmentation software defrags a file, the new location's LCN will report to the disk controller and the controller software will remap the file on the disk. The disk defragmentation software has no idea where the disk controller will locate the file on the physical disk.

The benefit of disk defragmentation comes not from locating physical clusters on the disk but from reducing file system overhead before any physical disk I/O occurs.

In the case of RAID or SAN, it is indeed likely that a file that defragmentation software makes logically contiguous will be physically fragmented by the disk controller software. This is OK. The user benefits here in two ways. First, the file system is able to access the file in a single logical I/O taking the minimum elapsed time and using scant CPU and memory. Second, the file is in the best possible place for physical access according to the disk controller's intelligence.

Recently, Raxco Software worked with Hewlett-Packard's (HP's) High-End RAID Group to determine what effect defragmentation might have on large RAID configurations. HP tested on a machine using the fastest processors, the maximum amount of memory, and the fastest disk speeds. The idea was to isolate any performance change to disk I/O. The tests were measured using ZD's Server Bench and Intel's IOMeter benchmarking tools. The results showed that defragmentation could improve disk I/O by 50%.

At Microsoft, the Share Point server is a SAN configuration with a 475GB partition with 650GB of compressed data and 20GB changing daily. Microsoft uses Raxco Software's PerfectDisk® to defragment this environment.

The benefit of disk defragmentation comes not from locating physical clusters on the disk but from reducing file system overhead before any physical disk I/O occurs. Once this concept is understood, the need for defragmentation in large RAID or SAN environments becomes obvious.