No. 29, March 1996

OVERHANG

Last month we saw that the 512-byte SECTOR is the smallest "chunk" of information that your computer can read from or write to the hard drive at one time. Containing about the amount of data in a paragraph typed on a 3 X 5 inch index card, the size of the sector is a hardware limitations. We also saw that neighboring sectors on the hard drive can be arranged into a software group known as an ALLOCATION UNIT or CLUSTER. If you have a tiny hard drive - below 16 megabytes - a cluster consists of just one 512-byte sector. I also suggested that we create a file named STAN and put just these 6 bytes in it: WB9RQR. When we save the file all 512 bytes will be allocated to STAN even though only 6 bytes are used. That means that 506 bytes of unused hard drive space are tied up permanently in that file, until the file is erased. This unused but unavailable space is called CLUSTER OVERHANG.

Cluster overhang also exists in larger files. Suppose you type a long letter - about 2 pages - and it occupies 5,500 bytes (5.4 kilobytes). On our tiny hard drive, the file will occupy 11 clusters of 512 bytes (11 X 512 = 5,632 bytes). The first 10 will be full, but the last one will contain only 380 bytes. The overhang in this cluster will be 132 bytes.

Cluster overhang tends to waste about half the cluster size for each file on the drive, if there are several hundred files of varying size. Suppose our 16-megabyte drive has 200 files on it. Chances are that the overhang wastes about 51,200 bytes in this case. That is not a big problem with our 16-megabyte drive, but it can be for larger hard drives. The rub is: larger hard drives use clusters that have more than one sector.

If your hard drive is between 16 and 128 megabytes in size, each cluster occupies 2,048 bytes or 4 sectors. The file STAN, containing only 6 bytes, would have a cluster overhang of 2,042 bytes. Even larger hard drives have larger clusters, as shown in the following table.

	SIZE OF EACH CLUSTER IN:		
SIZE OF PARTITION	**bytes**	**kilobytes**	**sectors**
up to 16 megabytes	512	½	1
16 megabytes to 128 Mb	2,048	2	4
128 megabytes to 256 Mb	4,096	4	8
256 megabytes to 512 Mb	8,192	8	16
512 megabytes to 1 Gb	16,384	16	32
1 gigabyte to 2 Gb	32,768	32	64
2 gigabytes to 4 Gb	65,536	64	128

Across the room is a computer that I use for modem work. It has a 100-megabyte hard drive, 59% full. It uses 2,048 byte clusters and has just shy of 2 megabytes of wasted, cluster overhang (4,059 index cards of data). On the other hand, I am writing this on a machine with a 504-megabyte hard drive which is 51% full and which uses 8,192 byte clusters. There is a whopping 16.7 megabytes of space wasted in cluster overhang. Calculations show that if the latter drive was just 9 megabytes larger, cluster size would have gone to 16,384 bytes (32 sectors), and the overhang would have jumped to over 35 megabytes. That is 72,000 index cards full of wasted data space, easily more cards than are used by your local community library!

Now, I have to tell you a secret. That 504-megabyte drive I mentioned above is really a 540-megabyte drive. When I partitioned it (with FDISK), I told the partitioning program to make the biggest drive it could. It came up with the maximum allowed, as explained in last month's article - 504 megabytes. That turns out to be 36 megabytes short of the actual space available on the drive. Had I been more astute, I would have partitioned it into two drives, C: and D:, perhaps each 270 megabytes in size. That way, I would have had use of those 36 megs that I can't use now, which is nearly the equivalent of a 40 megabyte hard drive, something we would have paid several hundred dollars for not many years ago!

On the other hand, if I partition it into two 270 megabyte drives, the cluster size will remain at 16 sectors or 8,192 bytes (see the table) and I would continue to average about 4 kilobytes of wasted space for each file on the drive. If I partitioned it into three drives (C:, D: and E:) of 180 megabytes each, every cluster would contain 8 sectors or 4,096 bytes. The average wasted space for each file would be only about 2 kilobytes. Hmmmm. I am going to repartition that drive as soon as it becomes 2/3 full!

I hope this has clearly illustrated the relationship between hard drive size, cluster size as a function of the number of sectors each cluster contains, and cluster overhang. It should also reveal to you the complexity in making decisions about setting up your machine. A good rule of thumb is to limit the size of your partitions insofar as is possible. for example, don't partition a drive so that it's size is equal to or just over one of the size jumps in the table above. If you must make a single partition, make it 255 megabytes, not 256. Your gain in efficiency due to lower overhang will more than make up for the megabyte of space you loose. Also, even if you have an advanced BIOS or software that will let you go over the usual 504 megabyte drive size limit, keep your drives below 512 megabytes. These rules of thumb come from author and computer guru Jeff Prosise, and he really knows his stuff.

Finally, a word of caution. When you repartition a drive using FDISK (it is really a simple job), you will destroy every byte of data on the drive. Be sure to copy everything to floppies or tape first! And don't forget, when you are done with FDISK, you must first format the drive (using the FORMAT C:/S command) before you can use the drive again. Happy computing!