No. 87, Feb 2001

THE COMPUTER CORNER

Bits, Bytes, Sectors and Clusters

- by Stan Kaplan, WB9RQR
105 Martin Drive
Port Washington, WI 53074-9654
(262) 284-9346
skaplan@mcw.edu

It is always easiest to learn something new by exploring how it is like something we already know about. That is why analogies are so darn useful. In this issue, we will explore some analogies between the computing world and chemistry, but you won't need to know very much chemistry to make the jump.

The basic building block in the chemical world is the atom or element. The atom is the smallest unit of matter, as we know it. An atom of iron has certain properties that make it iron. If we break an atom of iron down to its constituent parts, it is no longer iron. Calcium, oxygen, lithium, cobalt are all atoms. There are somewhat over 100 different atoms (elements), and everything we know of (including our own bodies) is made of them.

In the computing world, the smallest unit is the bit. Furthermore, there is only one kind of bit, not over 100 as is the case with atoms. The difference is, a bit can be in one of two states. It can be on or off, which is the same as 1 or 0, which is the same as yes or no. Like the toggle switch that turns the lights on or off in your bedroom, it can be in one of two states, and always must be in one or the other. You cannot turn your bedroom light toggle switch half on (forget about dimmer switches for the moment, please!). It clicks to the on position or to the off position. That is it. A computer is nothing more than millions of toggle switches. However, we arrange them in specific, useful ways.

In the chemical world, the next unit up from the element or atom is the compound. A chemical compound is merely a combination of atoms stuck together in some specific proportion, and arranged in some specific way. A compound can be composed of two atoms (carbon monoxide is one atom of carbon and one of oxygen). Or it may be composed of three (rust is one atom of iron and two of oxygen), or four (carbon tetrachloride is one atom of carbon and four of chlorine) or many more. In the computing world, the next step up from the bit is the byte. A byte is simply an adjacent bank of 8 identical bits. They are identical in that all eight are simply toggle switches. What makes them different is their state. Some may be on, some may be off. It is the pattern of on-ness or off-ness among the 8 bits that makes that byte different from other bytes. With a bank of 8 wall switches, there are 256 unique possible patterns of on-off. The table below shows an example of two patterns. The two examples show how the English letter A (upper case) and e (lower case) are represented inside your computer. By the way, please excuse the numbering of the wall switches, shown in the shaded heading of the table. Computer engineers are a weird lot. They start numbering things from the right. Also, no one ever taught them to count properly, so they think the first thing in a series is number zero, the second is number one, and so on. You and I would number those 8 wall switches 1..8, but they number them 0..7, and backwards at that. Weird!

	7	6	5	4	3	2	1	0
A	0	1	0	0	0	0	0	1
e	0	1	1	0	0	1	0	1

See the difference? The only thing not the same between the two letters of the alphabet are bit 2 and 5. Both of those wall switches are off in the letter A and on in the letter e. When you press

the letter e on your keyboard, a chip translates that into 01100101 and sends it to the brain of the computer, the Pentium chip. All the other characters have unique bit patterns, too.

There is another basic unit in the computing world, one step up from the byte, but I really cannot think of any good analogy in the chemical world without getting into polymer chemistry or some other highly specialized subfield. Therefore, we will have to take it as it is. This next-step basic unit deals not with how data is represented in a computer, but rather how much can be handled at a time when writing data to a floppy disk or a hard disk. You see, all software and hardware in the PC world adheres to a single standard, the sector. If you will read the following paragraph, it pretty much explains itself.

> This paragraph is composed of exactly 512 characters and spaces, which emulates one sector in the computer world. A sector is defined as exactly 512 bytes of data, the smallest "chunk" of data that can be written to the surface of a floppy or hard disk at one time. This standard is built in to all hardware devices and software sold today for PCs. There are 18 sectors per track and 80 tracks on each side of a standard 3½-inch, 1.44 Mb floppy disk, so 2,880 of these paragraphs would fit on a single floppy.

When I created this paragraph, the box surrounding it was just about the size of a 3 x 5 inch card, to give you some idea of size. The words fit nicely on the card, with a good margin of white space around them. As noted by the paragraph itself, it represents exactly 512 byes, or one sector. A standard floppy would hold 2,880 copies, each consisting of 512 bytes. If you do the math, 2,880 x 512 = 1,474,560. However, remember that computer geeks are weird. A kilobyte in the computer world is not 1,000 bytes. Rather, a kilobyte is 1,024 bytes, so you must divide the result by 1,024. Dividing 1,474,560 by 1,024 yields 1,440 kilobytes, or 1.440 megabytes, the capacity of a standard 3½-inch floppy. How about that, it works out perfectly!

There is one more level above sector. It is the cluster, also known as the file allocation unit. As its second name hints, it is defined as the smallest amount of data that can be written (or allocated) to a file on a floppy or hard disk. In terms of bytes, it may be as small as one sector or 512 bytes. That is the case with a standard floppy, where one cluster is equal to one sector. On the other hand, some hard drives sport clusters as large as 64 sectors. If I write just WB9RQR to a simple text file on such a hard drive, the data will takes only 6 bytes of space. However, the file containing this data will actually occupy 64 x 512 = 32,768 bytes. None of the unused 32,760 bytes are available for any other purpose until the file is erased. Not very efficient!

Well, there you have it. You have learned about the truly fundamental units in the computing world, bit, byte, sector and cluster. Sometimes, these words are important in other fields, too. This holiday season, there were clusters of grapes and other goodies on the table. Quite often (too often), I bit into a sector of pie. Every byte was good! I hope your holiday season was as good as mine was. Happy computing.