Bits and Bytes

Bits and Bytes

Everything in a computer is 0's and 1's. The bit stores just a 0 or 1: it's the smallest building block of storage. Anything with two separate states can store 1 bit.

Group 8 bits together to make 1 byte. One byte = collection of 8 bits.  
For example, 01101011 is stored a byte. One byte can store one character, e.g. 'A' or 't' or '$'

**How many patterns with N bits?**

Every bit can take 2 possibilities. If there is one bit it could be 1 or 0 so two possible options can happen. If there are two bits 00, 01, 10 or 11 could be so there are 4 possible options. When we add one more bit 000, 001, 010, 011, 100, 101, 110, 111 these 8 different patterns can occur. In general, when we add 1 bit, we double the number of patterns.

* 1 bit - 2 patterns
* 2 bits - 4
* 3 bits - 8
* 4 bits - 16
* 5 bits - 32
* 6 bits - 64
* 7 bits - 128
* 8 bits - 256 - one byte

Mathematically: n bits yields 2^n patterns (2 to the nth power)

Q: How many bit combinations are there in a byte?  
A: 256 possible combinations (from 0 to 255) A byte is made of 8 bits. Bits can only be on or off (0 or 1). 00000000 =0 , 00000001 = 1, 00000010 = 2, 00000011 = 3, 00000100 = 4, ... 11111111 = 255.

## Bits and Bytes

### ASCII Code

ASCII (American Standard Code for Information Interchange) is an encoding representing each typed character by a number and each number is stored in one byte (so the number is in 0..255).

| **https://docs.google.com/uc?id=17nj2s3oJyk_5aavVq8agqcbehNLRVf2t** |
| --- |
| Decimal ASCII codes and characters |

**Typing and Bytes**: Each letter is stored in a byte, as below. 1000 typed letters take up 1000 bytes.

| **https://docs.google.com/uc?id=1YlNiBTk9dcKdSXO7C5f-Yeg94VF2LLo6** |
| --- |
| Typing and bytes |

**Unicode** is a universal character encoding standard. It defines the way individual characters are represented in text [files](https://lms.clarusway.com/mod/lesson/view.php?id=1052), web pages, and other types of documents. It is different from ASCII, Unicode was designed to support characters from all languages around the world, mathematical character emojis etc. The standard ASCII character set only supports 128 characters, while Unicode can support roughly 1,000,000 characters. While ASCII only uses one byte to represent each character, Unicode supports up to 4 bytes for each character.

| **https://docs.google.com/uc?id=1jAy8ME9t_O1-y7LGDHCksASEvMROi8sS** |
| --- |
| Unicode of Emojis |

Q: What is ASCII?  
A: Ascii is a character encoding standard adopted by the Institute of Electrical and Electronics Engineers (IEEE) in 1963. ASCII is an abbreviation for American Standard Code for Information Interchange. It is a method of representing text characters in a binary representation recognized by computers, communications equipment, and other technological devices.

### Numbers in computers

Individual characters hold one byte on computer storage. But computers are also good at calculating and playing numbers. We call number holders as **integers**. Integers are typically stored with either 4 or 8 bytes.

4 bytes can store numbers between -2147483648 and 2147483647  
8 bytes can store numbers between -9223372036854775808 and 9223372036854775807

## Kilo/Mega/Gigabytes

### Kilobyte/Megabyte

**Kilobyte** equals to about one thousand bytes (1024 bytes). A page of ordinary Roman alphabetic text takes about 2 kilobytes to store (about one byte per letter). A typical short email would also take up just 1 or 2 kilobytes. While you were sending a message from internet text messages are sent faster than audio or image. So text is compact, requiring few bytes compared to images or sound or video.

**Megabyte** equals to about one million bytes (about 1000 KB). A high-quality digital picture is about 2-5 megabytes and mp3 audio is about 1 megabyte per minute.

### Gigabyte/Terabyte

**Gigabyte** is about a billion bytes (about 1000 MB). Gigabyte is a common metric for [hardware components](https://lms.clarusway.com/mod/lesson/view.php?id=44). For example, while you are looking for computer 4 GB RAM and 256 GB of persistent storage is ordinary computer properties.

**Terabyte** is about 1000 gigabytes or roughly 1 trillion bytes. You can buy 4 TB hard drives today, so we are beginning the time when this term comes into common use.

There is one term too heard a lot. It is **Gigahertz**. Gigahertz is not about bytes. It is about speed. One gigahertz is 1 billion pulses per second (a megahertz is a million pulses per second). Gigahertz is a measure of speed, very roughly the rate that at a CPU can do its simplest operation per second. Gigahertz does not precisely tell you how quickly a CPU gets work done, but it is roughly correlated.

| **https://docs.google.com/uc?id=1i-pmO_Shn4B20CHdAEhdVvdxNPtoCtFl** |
| --- |
| Pyramid of Bytes |

Q: What is the difference between gigabyte and gigahertz?  
A: Gigabyte is storage metric that is equal to 1024 megabyte. But gigahertz is a measure of speed. One gigahertz is 1 billion cycles per second.

### Math Examples

In this section, we will convert [bits and bytes](https://lms.clarusway.com/mod/lesson/view.php?id=886" \o "Bits and Bytes).

Question: 8,000,000 bytes is about how many MB?  
Answer: 8 MB (1,000,000 Bytes = 1 MB )

Question: 25,000 KB is about how many MB?  
Answer: 25 MB (1000KB = 1MB)

Question: 100 KB is about how many MB?  
Answer: 0.1 MB

Question: How many GB is 16,000,000,000 bytes?  
Answer: 16 GB

Question: Say you have many 8 MB JPEG images. How many fit on a 32 GB flash drive?  
(convert everything to MB) Answer: 32 GB is 32,000 MB 32,000 / 8 = 4,000

| **https://docs.google.com/uc?id=1cPBTqwNPuLiPWSU1oFGvp3Vt5-sHTEYM** |
| --- |
| [*Bits and Bytes*](https://lms.clarusway.com/mod/lesson/view.php?id=886) |