Moore’s Law, which is actually just an observation, states that the transistor density would double every two years.

Processors are just a packet of transistors which are used to do computation. If a transistor gets smaller, it switches faster. This means that computations can be done more quickly. So if the amount of transistors on a processor gets doubled, it means that the transistors get smaller, which leads to an increase in speed. The problem however is that each transistor consumes power, so if you have a processor with a lot of transistors, you need a lot of power. All this power results in an increase of temperature (high power means high temperature) and without a cooling mechanism it will melt the processor. In most computers the processor is cooled using air, and air can only cool down the processor for this much.

A way to tackle this problem, to save power (which lowers temperature), is to scale down the amount of voltage of a transistor. However there are physical restrictions which constrain you from ever lowering the voltage. One, the provided voltage must be higher than the threshold voltage of a transistor. Two, because noise problems occur, having a low voltage swing makes it hard / impossible to tell whether a transistor should be turned on or not. You become less noise tolerant.