Moore's Law

Moore's law is the observation that the number of transistors in a dense integrated circuit doubles about every two years. Even though it's called a law, it's really an observation; one that is no longer accurate since the rate of doubling has slowed down significantly.

Temperature

The main concept that explains the rate's slowing is that of power density; which is a unit of power per unit of area. Let's say we double the number of transistors on a chip. This means that the power required to run the chip will also double. If the area of the chip doesn't increase, then the power density will also double. This is a problem because the chip can only dissipate so much heat before it starts to malfunction.

If you think that we can reduce the heat generation to nothing, then you'd be mistaken. Landauer's principle states that there is a minimum amount of energy required to perform a computation. This means that there is a minimum amount of heat that will be generated when a computation is performed. We are nowhere near this limit, but it's a limit nonetheless.

Dynamic Voltage Scaling

Another issue is that voltage scaling reduces the power consumption of a chip. However, the voltage cannot be reduced indefinitely. This is because the voltage has to be higher than the threshold voltage of the transistor. If the voltage is too low, then the transistor will not be able to switch on and off. This is a problem because the threshold voltage has not decreased at the same rate as the supply voltage.

Furthermore, we cannot prevent the leakage of current. This is because the transistor is never truly off. There is always some current that flows through the transistor. This is a problem because the leakage current increases as the transistor size decreases.

Finally, the amount of voltage scaling is limited to the ration of the threshold voltage to the noise margin. If the voltage is scaled too low, then the noise margin will be too small. This is a problem because the noise margin is the amount of noise that can be tolerated before the signal is misinterpreted.