Why make transistors smaller, and why is Moore’s law the case?
Daniel Wilcox
{} – represents slide change
[] – represents a cited source

{Slide 1} Gordon Moore, the co-founder of Intel, predicted that every two years, the number of transistors that can fit into a given space doubles. This prediction is known as Moore’s law, and it has been consistent since 1975. Why Moore’s law has been so relevant in the past 50 years, however, is because of the actual relationship between transistors and computing speed. Because of the superscalar designs of CPUs, the number of transistors has a direct effect on clock speeds and processing power. Furthermore, semiconductor companies such as Intel or Samsung have been following Moore’s law because it functions as an economic goal for the market. Moore’s law has been the case for fifty years, not only because of the direct relationship between number of transistors and clock speed, but also because of the competitive semiconductor market Intel has created.

{Slide 2} The reason why computer chip companies have been making transistors smaller is because of the direct relationship between transistors and computing power. A primary example of this is the CPU of a computer. The CPU is the Central-Processing-Unit, and it runs programmes, determines performance and computing power, and is the only thing that really makes a computer a computer. Modern CPUs have *superscalar* designs, meaning they can have multiple “cores,” and multiple *execution units*. Modern CPUs can have as many as 30 execution units, which are the individual units that process chunks of information [1][2]. When a CPU receives programmes, or, “instructions,” to run, it starts the *pipelining* process: A superscalar CPU will break the instructions down into segments that can be run in *parallel*, and then distributes them into as many execution units as possible, before combining the results together at the end. This concept of “parallelism” is the focal point for CPU processing speed. The more a programme can be run in parallel, the faster the execution time. In fact, advanced software designers will design their programmes to run in parallel as much as possible in order to increase execution speed [3]. But ultimately, the speed that a programme can execute at depends on the number of execution units in the CPU, and its superscalar architecture and design. This is where the number of transistors starts to have a direct impact on processing speed.

{Slide 3} Execution units are made of transistors, and the number of transistors that can fit inside a CPU effects the number of execution units: more transistors means more units. One or two extra units, however, doesn’t make much of a difference. This is why modern CPUs have been specifically designed with parallelism in mind [4]. Modern CPUs have *distributors*, whose job is to separate programme instructions into segments that can be run in parallel. The evolution of scalar to superscalar CPUs has resulted in distributors that are specifically designed to work with multiple execution units. This is proven with a study from 1997, that tested the clock speeds of superscalar CPUs. Instructions were sent, pipelined through the distributor, and into the execution units. The results demonstrated that performance is improved because individual blocks of instructions could be selected and dealt with quicker [1]. Execution units, however, are just one example of how increasing the number of transistors in a given space directly improves computing power. Superscalar designs were invented and popularised as a result of the innovation Moore’s law created [10]. We, as a society, make transistors smaller because there is a positive, proportionate relationship between transistors and clock speed.

{Slide 4} But the relationship between transistors and performance doesn’t explain the constant rate of Moore’s law: doubling every other year, for the past fifty years. This is no coincidence, however, nor is it an exact ‘law’ per se. Moore’s law is, more accurately, a prediction that turned into a procedure. Moore's law started out as a prediction that the number of transistors that would fit into a given space would double every year, and then it was revised to every two years [5]. Gordan Moore made his prediction based on the approaches used by technology manufacturers, including the micro-assembly of individual components and semiconductor integrated circuits [6]. Around the 70's, however, when Moore's law was more widely known, it became a "goal" of sorts for the semiconductor industry to follow. That's how Moore's law became a procedure: since then, the Semiconductor Industry Association has been using a roadmap as a guide [12]. The roadmap, made in 1993, demonstrates landmarks in the technology every 3 years. The four most important characteristics of this roadmap are *feature siz*e (size of individual chips), *gates per chip*, *wafer processing cost*, and *wafer diameter*. The wafer diameter remains fairly constant throughout all 15 years, ranging between 200-400 mm. Furthermore, while the actual feature size decreases, the number of gates, or, "transistors / MOS-FETs" in each feature increases. The increments between gates per chip also correspond with Moore's law: by multiplying every value by $2^{1.5}$ (because the interval is every 3 years), we get a very close result to the next value. The slight difference is likely due to rounding down, not wanting to make predictions too specific, or some other decision by the SIA. Regardless, this information demonstrates the increasing number of chips per wafer, and transistors per chip, while still following Moore's law.

Finally, the fourth most important characteristic in this roadmap is the Wafer processing cost. Why is this one of the most important aspects? Because it brings attention to the driving force behind Moore's law for the past fifty years: economics. {Slide 5} When Gordan Moore announced his vision for the future of semiconductor manufacturing, Intel strived to follow it and back it up fully. Considering Intel was one of the inventors of the microprocessor, as well as was co-founded by Gordan Moore himself, focusing all of its energies on creating CPUs and enforcing Moore's law became a top priority [8]. This focus has made Intel a key competitor in the semiconductor marketplace since 1980, and other companies such as Samsung, AMD, and TSMC have been competing. Intel's lead in microprocessor manufacturing has enforced competition in the market, and thus enforced Moore's law. The reason transistors become smaller at a predictable rate, and why computers become faster at a predictable rate, is ultimately due to the economic competition Intel and Gordan Moore have created.

Cited sources, via Google Scholar:

1. Palcharla, Subbarao, et al. “Complexity-Effective Superscalar Processors.” *Complexity-Effective Superscalar Processors | Proceedings of the 24th Annual International Symposium on Computer Architecture*, ISCA '97: Proceedings of the 24th Annual International Symposium on Computer Architecture, 1 May 1997, dl.acm.org/doi/abs/10.1145/264107.264201.
2. Fite, Elaine H. “US5067069A - Control of Multiple Functional Units with Parallel Operation in a Microcoded Execution Unit.” *Google Patents*, Google, 3 Feb. 1989, patents.google.com/patent/US5067069A/en.
3. Kanter, David. “Intel's Haswell CPU Microarchitecture.” *Real World Tech*, 13 Nov. 2012, www.realworldtech.com/haswell-cpu/4/.
4. “Reading: The Central Processing Unit.” *Lumen*, courses.lumenlearning.com/zeliite115/chapter/reading-the-central-processing-unit/.
5. “22nm Technology Backgrounder.” *OVER 6 DECADES OF CONTINUED TRANSISTOR SHRINKAGE, INNOVATION*, May 2011.
6. Moore, Gordon E. “Cramming More Components Onto Integrated Circuits.” *Proceedings of the IEEE*, vol. 86, no. 1, 19 Apr. 1965, pp. 82–85., doi:10.1109/jproc.1998.658762.
7. Guenthner, Russell W. “EP0106670A2 - CPU with Multiple Execution Units.” *Google Patents*, Google, 1984, patents.google.com/patent/EP0106670A2/en.
8. Gulett, Mike. “Moore's Law Is Not Really A Law But It Is Still Obeyed.” *Medium*, Data Driven Investor, 29 Apr. 2019, medium.com/datadriveninvestor/moores-law-is-not-really-a-law-but-it-is-still-obeyed-2e0c2b596525.
9. Maliniak, Lisa. “Ten Notable Flops: Learning From Mistakes.” *ElectronicDesign*, ElectronicDesign, 21 Oct. 2002, www.electronicdesign.com/technologies/boards/article/21772296/ten-notable-flops-learning-from-mistakes.
10. Johnson, Mike. *Superscalar Microprocessor Design*. Prentice Hall, 1991.
11. *“Model of a Subscalar CPU.” Reading: The Central Processing Unit, Lumen Learning, 2010, courses.lumenlearning.com/zeliite115/chapter/reading-the-central-processing-unit/.*
12. Semiconductor Industry Association, 1993. The First Semiconductor Roadmap, Released In 1993 By The SIA. [Image].
13. “2014F Top 20 Semiconductor Sales Leaders.” *IC Insight's Strategic Reviews Database*, IC Insights, 2014, www.icinsights.com/.
14. “A Simple Superscalar Pipeline.” Reading: The Central Processing Unit, Lumen Learning, 2010, courses.lumenlearning.com/zeliite115/chapter/reading-the-central-processing-unit/.