The Last Byte

Making the best of those extra transistors

Frank Vahid
University of California, Riverside

MEETING YOUR family's basic needs on a $20,000 annual salary would be a challenge. Increasing that salary to $40,000 would make a big difference, and $80,000 would be even better. But at some point, working for further increases would reach a point of diminishing returns. An increase from $100 million to $200 million per year would be nice, but it probably wouldn't change your life much, and few would even notice the difference between $1 billion and $2 billion.

Meeting your basic embedded-computing needs with a 20,000-transistor silicon budget would also be a challenge. Twenty thousand transistors (roughly the silicon budget two decades ago) are barely enough to implement an 8-bit microprocessor. Increasing the budget to 40,000 transistors would make a big difference, and 80,000 would be even better. But again, at some point, working for further increases would reach a point of diminishing returns. An increase from 100 million to 200 million transistors (modern chip sizes) would be nice, but wouldn't change most designers' systems all that much, and few designers would even notice the difference between 1 billion and 2 billion transistors.

Moore's law says that chip capacity doubles every 18 months. Yet, ASIC vendor data shows that most designs greatly underuse that capacity. Mainstream embedded-system designers are not screaming for higher-capacity chips the way they once were. One reason is that a few hundred million transistors are enough to provide plenty of computing ability. Another reason, known as the productivity gap, is that designer productivity increases haven't kept pace with chip capacity increases, meaning designers often can't create designs that utilize all those transistors.

High-capacity chips will still be made, but more will likely be in the form of prefabricated programmable platform chips, which provide significant time-to-market and cost advantages compared to custom-made chips. Programmability can come in several forms, like general-purpose processors, field-programmable logic, and tunable architecture parameters like reshapable memory hierarchy, segmentable bus structures, optional code and data compression schemes, and variable bit widths. Such programmability uses far more transistors than more customized designs, but those transistors are available anyway, and platform designers will create them in highly replicated and hence easier-to-design patterns. The earlier diminishing-returns argument also applies to the performance overhead of such platforms.

Other uses of those extra transistors are also evolving—extra transistors can dynamically monitor and optimize a chip's execution, or reduce power by executing operations on one of numerous specialized computation units. They could even ease intellectual property distribution—each of a chip's components might be activated only after proper licensing.

A NEW WORLD for transistors may be on the horizon. The days of directly implementing custom functionality in transistors, and hence of designers creating their own chips, may be fading away. Prefabricated programmable platforms, considered grossly inefficient just a decade ago, may prevail, leading to lower costs, faster time to market, lower power, and other benefits perhaps not even foreseeable today. ■

Frank Vahid is an associate professor in the Department of Computer Science and Engineering at the University of California, Riverside. Contact him at vahid@cs.ucr.edu.

Published by the IEEE Computer Society