The heat death of information processing and why interconnects matter more than logic

David A. B. Miller
Stanford University, USA

Today essentially all information communication is Internet traffic and that traffic is growing by approximately a factor of 100 every 10 to 12 years. The installed base of general-purpose information processing logic is growing at a similar rate. The capacity of connections in wires at shorter distances does not, however, scale to keep up with the logic. If we propose new logic for which communications are more difficult than the present charge-based signaling on wires (and many novel logic systems do exactly this), then we make the problem worse still. Therefore there is a growing problem in the density of communications. Information processing machines will choke on their own wiring if we do not find another interconnect solution.

Information processing and communications today take a few percent of our energy consumption and carbon footprint. If any reasonable fraction of this energy cost were to scale in proportion to Internet traffic and if that traffic were to grow at current rates, then in a few decades all energy generation would be for handling information. Today, large data centers are built beside rivers for cooling and for cheap hydroelectric power. The clock rate of current chips is set not by the speed of the transistors but by the necessity to minimize power dissipation, both to avoid heating and to reduce power costs. The performance and scale of information processing is also therefore significantly limited by power generation and dissipation.

The energy required to communicate information from one side of a logic gate to the other is approximately equal to the energy required to switch the transistors in the gate. Essentially all the rest of the energy consumption in chips and information processing generally is in sending information or clock signals over longer distances. Even in 2004, more than half the power dissipation on a chip was estimated to be in interconnects and clock distribution. That fraction is expected to grow, perhaps to 80%. The energy cost of running servers already is comparable to or exceeds their purchase cost. The energy dissipated in interconnects in servers already exceeds the total of all solar power generation. Therefore we need technologies that can reduce interconnect energy per bit.

Arguably, therefore, the issues of interconnect density and energy may come to dominate the advance of information processing technology. These issues may become more important than the density and speed of logic devices. The use of optics at increasingly shorter distances is one approach that, at least on the basis of the physics, could address both density and energy of interconnects, though the technology challenges are substantial. Whether or how soon optics solves these problems is debatable. If not, the issues of interconnect density and energy dissipation will come to dominate information processing hardware and systems in the coming decades and may end the remarkable decades-long exponential growth in information processing and communications.