

No Need to Constrain Many-Core Parallel Programming

Seminar by Prof. Uzi Vishkin, U. of MD

Th Nov.12,2009, 4:30pm, Eng. Rm.4457

The transition in mainstream computer science from serial to parallel programming for many-core on-chip computing offers parallel computing research the wonderful impact opportunity it had sought all along. However, such transition is a potential trauma for programmers who need to change the basic ways in which they conduct their daily work. The long experience with multi-chip parallel machines only adds to the apprehension of today's programmers. Many people who tried (or deliberated trying) to program these multi-chip machines consider their programming "as intimidating and time consuming as programming in assembly language" (--2003 NSF Cyberinfrastructure Blue Ribbon Committee), and have literally walked away. Programmers simply did not want to deal with the constraint of programming for locality in order to extract the performance that these machines promise. Consequently, their use fell far short of historical expectations. Now, with the emerging many-core computers, the foremost challenge is ensuring that mainstream computing is not railroaded into another major disappointment. Limiting many-core parallel programming to more or less the same programming approaches that dominated parallel machines could again: (i) repel programmers; (ii) reduce productivity of those programmers who hold on: getting the performance promise requires high development-time and leads to more error-prone code; (iii) raise by too much the minimal professional development stage for introducing programmers to parallel programming, reducing further the pool of potential programmers; and overall (iv) fail to meet expectations regarding the use of parallel computing; only this time for many-cores.

The talk will overview a hardware-based PRAM-On-Chip vision that seeks to rebuild parallel computing from the ground up. Grounded in the richest and easiest known theory of parallel algorithms, known as PRAM, where the programmer only needs to identify at each step operations that can be executed concurrently, an on-chip architecture that scales to thousands of processors on chip called XMT (for explicit multi-threading) was introduced. Significant hardware and software prototyping of XMT will be reported, including a 64-processor FPGA-based machine and two ASIC chips fabricated using 90nm CMOS technology, as well as strong speedups on applications. By having XMT programming taught at various levels from rising 6th graders to graduate students, we developed evidence that the stage at which parallel programming can be taught is earlier than demonstrated by other approaches. For example, students in a freshman class were able to program 3 parallel sorting algorithms. Software release of the XMT environment can be downloaded to any standard PC platform along with extensive teaching materials, such as video-recorded lectures of a one-day tutorial to high school students and a full-semester graduate class, class notes and programming assignments. Preliminary thoughts on encapsulating XMT into a hardware-enhanced programmer's workflow will also be presented and the prospects for incorporating it as an add-on into some other many-core designs be discussed.

Short biography

Uzi Vishkin got his BSc and MSc degrees in Mathematics from the Hebrew University, Israel, and his DSc degree in CS from the Technion, Israel in 1981. He then worked at IBM T.J. Watson and New York University. He was affiliated with Tel Aviv University between 1984 and 1997, and was Chair of CS there in 1987-8. He has been Professor of Electrical and Computer Engineering at the University of Maryland Institute for Advanced Computer Studies (UMIACS) since 1988. The Shiloach-Vishkin work-depth methodology for presenting parallel algorithms provided the presentation framework in several parallel algorithm texts that also include quite a few parallel algorithms he co-authored. He is the inventor of the PRAM-On-Chip desktop supercomputer framework under development since 1997 at UMD. He was elected ACM Fellow in 1996 for, among other things, having "played a leading role in forming and shaping what thinking in parallel has come to mean in the fundamental theory of Computer Science", is an ISI-Thompson Highly Cited Researcher, and was named a Maryland Innovator of the Year for his PRAM-On-Chip venture.