

# Toward Realizing a PRAM-on-a-Chip Vision

Uzi Vishkin

University of Maryland

University of Maryland Institute for Advanced Computer Studies (UMIACS)  
College Park, Maryland, USA

## Abstract

Serial computing has become largely irrelevant for growth in computing performance at around 2003. Having already concluded that to maintain past performance growth rates, general-purpose computing must be overhauled to incorporate parallel computing at all levels of a computer system – including the programming mode – all processor vendors put forward many-core roadmaps. They all expect exponential increase in the number of cores over at least a decade. This welcome development is also a cause for apprehension. The whole world of computing is now facing the same general-purpose parallel computing challenge that eluded computer science for so many years and the clock is ticking. It is becoming common knowledge that if you want your program to run faster you will have to program for parallelism, but the vendors who set up the rules have not yet provided clear and effective means (e.g., programming models and languages) for doing that. How can application software vendors be expected to make a large investment in new software developments, when they know that in a few years they are likely to have a whole new set of options for getting much better performance?! Namely, we are already in a problematic transition stage that slows down performance growth, and may cause a recession if it lasts too long. Unfortunately, some industry leaders are already predicting that the transition period can last a full decade.

The PRAM-On-Chip project started at UMD in 1997 foreseeing this challenge and opportunity. Building on PRAM – a parallel algorithmic approach that has never been seriously challenged on ease of thinking, or wealth of its knowledge-base – a comprehensive and coherent platform for on-chip general-purpose parallel computing has been developed and prototyped. Optimizing single-task completion time, the platform accounts for application programming (VHDL/Verilog, OpenGL, MATLAB, etc), parallel algorithms, parallel programming, compiling, architecture and deep-submicron implementation, as well as backward compatibility on serial code. The approach goes after any type of application parallelism regardless of its amount, regularity, or grain size. Some prototyping highlights include: an eXplicit Multi-Threaded (XMT) architecture, a new 64-processor, 75MHz XMT (FPGA-based) computer, 90nm ASIC tape-out of the key interconnection network component, a basic compiler, class tested programming methodology where students are taught only parallel algorithms and pick the rest on their own, and up to 100X speedups on applications.

The talk will overview some future plans and will argue that the PRAM-On-Chip approach is a promising candidate for providing the processor-of-the-future. It will also posit that focusing on a small number of promising approaches, such as PRAM-On-Chip, and accelerate their incubation and testing stage, would be most beneficial both: (i) for the field as a whole, and (ii) for an individual researcher who is seeking improved impact.

URL: <http://www.umiacs.umd.edu/~vishkin/XMT>