

RoMoL: Riding on Moore's Law

Description

The most common interpretation of Moore's Law is that the number of components on a chip and accordingly the computer performance doubles every two years. This experimental law has been holding from its first statement in 1965 until today. At the end of the 20th century, when clock frequencies stagnated at ~3 GHz, and instruction level parallelism reached the phase of diminishing returns, industry turned towards multiprocessors, and thread level parallelism. However, too much of the technological complexity of multicore architectures is exposed to the programmers, leading to a software development nightmare that horrifies the entire computing industry.

We propose a radically new conception of parallel architectures, built using a higher level of abstraction. Instead of expressing algorithms as a sequence of instructions, we will group instructions into higher-level tasks that will be automatically managed by the architecture, much in the same way superscalar processors managed instruction level parallelism. We envision a holistic approach where the parallel architecture is partially implemented as a software runtime management layer, and the remainder in hardware. The hardware gains the freedom to deliver performance at the expense of additional complexity, as long as it provides the required support primitives for the runtime software to hide complexity from the programmer. This regained freedom in hardware design enables us to revisit a number of previously proposed architectural concepts in light of the new technology context for unforeseen impact in performance an energy efficiency. This now becomes possible thanks to the joint development of the runtime layer with the parallel architecture.

We will focus our architecture research on a most efficient form of multicore architecture coupled with vector accelerators for exploiting both thread and data level parallelism. Our architecture will also be extended with hardware components and primitives that enable the runtime to handle the complexity and novel proposals to optimize the data movement, e.g., user-defined on-the-fly data transformations, moving intelligence to the memory instead of moving data back and forth to the processors. Furthermore, we will integrate our vector proposal with the energy efficient, made in Europe, ARM architecture that is established world-wide leader in the mobile domain.

This will allow the European industry to strengthen its leading position in the embedded market and gain significant advantage for the next phase of the high-performance computing race. The holistic approach towards parallel architectures offers a single solution that could solve most of the problems we encounter in the current approaches: handling parallelism, the memory wall, the power wall, and the upcoming reliability wall in a wide range of application domains from mobile up to supercomputers. Altogether, this novel approach toward future parallel architectures is the way to ensure continued performance improvements, getting us out of the technological mess that computers have turned into, once more riding on Moore's Law.

Source URL (retrieved on 15 jul 2024 - 21:19): <https://www.bsc.es/ca/research-and-development/projects/romol-riding-moores-law>