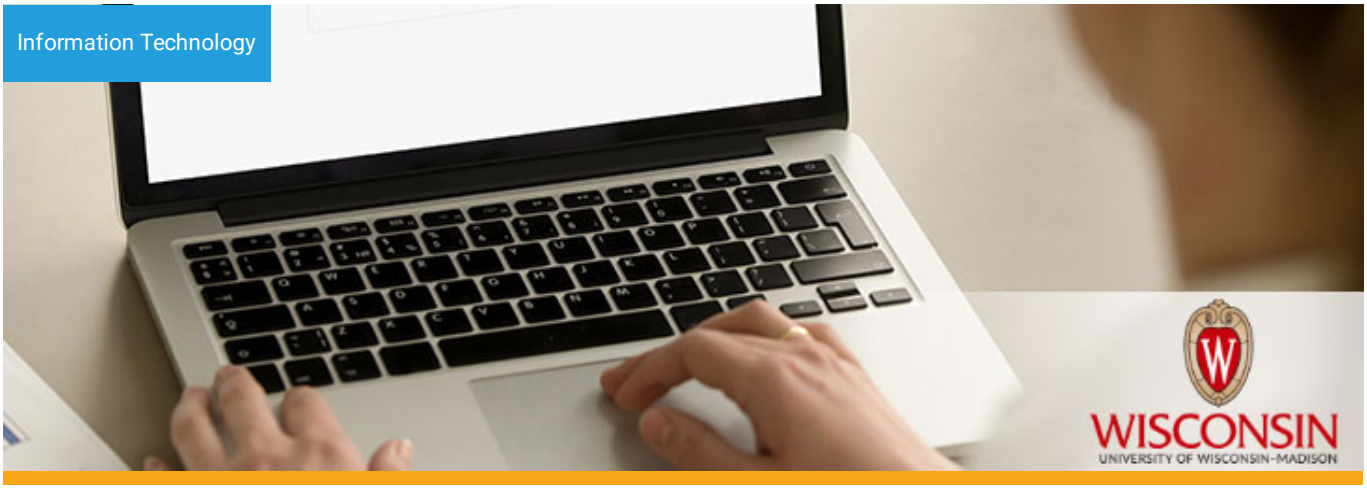
Information Technology



Maximizing Multicore Processor Performance

View U.S. Patent No. 9,606,842 in PDF format.

WARF: P130137US01

Inventors: Nam Sung Kim

The Wisconsin Alumni Research Foundation (WARF) is seeking commercial partners interested in developing a method to improve the performance of power-constrained processors by scaling cores and resources.

Overview

Computer processors typically include a number of central processing unit (CPU) cores for executing instructions. Each core is capable of reading and executing a sequence of instruction code, or 'thread.'

Within each thread, a core also may work in parallel to maximize performance. This is known as instruction level parallelism (ILP). Generally, the more resources (like caches and register files) available to a core, the better it executes applications with high ILP. Additionally, the more cores available to a processor, the better it performs thread level parallelism (TLP), which occurs when multiple cores are executing multiple threads at once.

However, fully operating all cores and all resources can exceed power budgets, leading to failure or damage. In such power-constrained environments, voltage/frequency can be reduced at times to conserve power. Unfortunately, this type of scaling is difficult.

The Invention

A UW–Madison researcher has developed a solution for improving performance while still meeting a maximum power budget for multicore processors operating in a power-constrained environment.

The method provides for joint scaling of both the number of cores in operation as well as the amount of resources per core. Selecting the number of cores and the amount of resources is done by examining the degree of ILP and TLP available for a given application. As such, performance counters (and other characteristics) implemented by a processor are used to determine optimal core/resource configurations.

Performance counters may measure, for example, how many instructions are executed per cycle, length of execution time and cache hit rate. These performance measurements indicate how much ILP and TLP a given application exhibits at a time.

Applications

- Implementation in multicore processors
- Data centers running intensive applications

Key Benefits

- Manages power consumption and performance
- Determines optimal core and resource configuration

- Works within a power constraint
- Works independently of voltage/frequency scaling
- Indicates how efficiently cores and resources are being utilized

Stage of Development

The ability to choose the best resource/core configuration has been shown to improve processor performance by 20 percent on average.

Additional Information

Related Technologies

- WARF reference number P100298US01 describes a method of controlling the voltage at each core of a multicore processor to improve performance and efficiency.
- WARF reference number P110254US01 describes a cache structure that uses transistor cells of variable sizes to maximize energy efficiency in processors.

Tech Fields

- Information Technology : Computing methods, software & machine learning
- Information Technology : Hardware

For current licensing status, please contact Jeanine Burmania at jeanine@warf.org or 608-960-9846