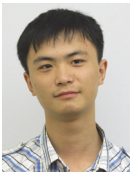
A Simulation Infrastructure for Exploring and Optimizing Chip-Multiprocessors


Da-Cheng Juan


Pinjie Huang


Xianglong Fu


Vikram Rajkumar


Arnold Pereira


Diana Marculescu

Moore's law has pushed more and more cores to be integrated within one chip, enabling high-performance multi-cores instead of a classic, monolithic single-core. The use of these chip-multiprocessors (CMPs) can incur high power density or high thermal profiles due to the spatial and temporal variability arisen from the usage of on-chip resources by workloads. Therefore, modern processors usually support various control policies for dynamic power management by using on-chip sensors or performance counters. In general, the development and evaluation of these policies cannot be performed on real hardware due to the lack of observable metrics and implementation flexibility. In this context, a full-system simulation infrastructure that can reflect real-time processor states is highly required for both development and evaluation of optimal control policies.

A full-system simulator needs to be capable of accurately modeling the entire system evolution, including data flow, coherency conflicts and complex memory latency, while emulating the same performance counters and on-chip sensors of the real processor. In this work, we present a complete simulation infrastructure with the following features: (1) a high-level functional model that integrates timing, power and thermal models for fast and accurate simulations, and (2) dynamic voltage and frequency scaling (DVFS) on both processing cores and on-chip communication fabrics, which allows us to tune the respective voltage and frequency of each on-chip resource for performance, power or thermal optimization. Figure 1 illustrates the power consumption of a 16-core CMP under four different control schemes, and Figure 2 shows the corresponding network latency of the on-chip communication fabric. The developed simulation infrastructure can be readily applied to efficiently evaluating and optimizing performance and power, while also providing support for power and thermal management evaluation for advanced CMPs.

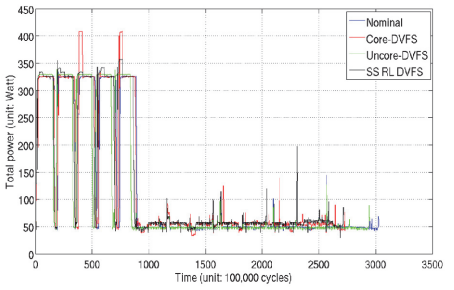


Fig. 1: Power consumption analysis

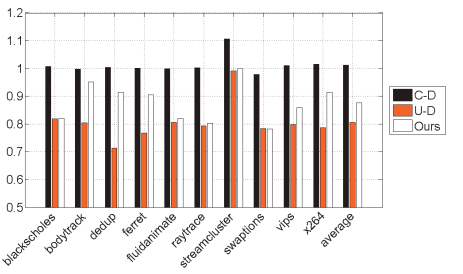


Fig 2: End-to-end network latency