McPAT 1.0: An Integrated Power, Area, and Timing Modeling Framework for Multicore Architectures

Authors - Sheng Li, Jung Ho Ahn, Jay B. Brockman , Norman P. Jouppi
University of Notre Dame, Hewlett-Packard Labs, Seoul National University
Paper Summary by Dhruv Jain

Multicore/manycore and multithreaded processors is a hot research topic these days which are overall powerful and efficient. In this paper I am going to summarize about McPAT framework. Wattch was the first tool that studied power related architectures and estimate power. However Wattch couldn't model timing and area with only calculating power for dynamic consumption. Orion is another tool that models power in networks-on-chip(NoC). Orion included power models for area, dynamic power, and gate leakage but not for short-circuit power or timing. Heat and power Dissipation has become most critical design constraint in today's world.
McPAT or (Multicore Power, Area, and Timing) is an integrated power, area, and timing modeling framework for multithreaded, multicore, and manycore architectures. It has models for the fundamental components of a chip multiprocessor, including in-order and out-of-order processor cores, networks-on-chip, shared caches, integrated memory controllers, and multiple-domain clocking. It models power, area, and timing simultaneously and consistently and supports comprehensive early stage design space exploration for multicore and manycore processor configurations ranging from 90nm to 22nm and beyond. The latest release of McPAT includes added support for SRAM and DRAM cache memory. McPAT helps in energy-delay-area-product, which quantifies the cost of new architectural ideas. It ignores many of low-level details of the components being modeled. McPAT also models not just dynamic power but also static and short-circuit power which gives a complete view of the power envelope of multicore processors. McPAT models the power of important components of multicore processors. McPAT handles newer technologies that are no longer modeled by linear scaling assumptions. McPAT also supports detailed and realistic models based on existing out-of-order processors (OOO). Using McPAT it is easy to find out cost of new architectural ideas using Energy-Delay-Area-Product (EDAP).

McPAT uses XML-based interface with the performance simulator. McPAT communicates with performance simulator which make it flexible and easy to port to other performance simulators. These are McPAT components - the hierarchical power, area and timing models, optimizer for determining circuit level implementations and the internal chip. McPAT is divided into two phases- the initialization phase and the computation phase. During the initialization phase a user will have to tell the number of cores, the number of routers, shared level cache parameters, core issue width, number of threads, etc. The McPAT combines with performance simulator to calculate power after the XML interface file for McPAT is generated. McPAT uses certain optimization techniques to optimize the components and the communication among them. In this phase, internal chip

representation is generated. The computation phase is initiated by performance simulator to calculate power. Performance simulator, which run separately from McPAT, calculates a statistic called activity factor of different components to McPAT via the XML interface file. Since the McPAT models power, area and timing simultaneously, the results from McPAT are steady electronically. McPAT consists of three levels - architectural, circuit and technology which allows for user flexibility in terms of configuration.
McPAT power modeling allows a variety of functions including dynamic, subthreshold leakage, gate leakage and short-circuit power. McPAT is the first tool to support clock gating and power management schemes with multiple power-saving states. McPAT includes the ability to determine power and area optimal configurations for array structures and interconnects, given specified targets for the timing value and optimization function. The parameters - cache capacity and core issue width, in the internal chip representation are directly set by the input parameters. These optimizations decreases the burden on the architect to find details, and hence lowers the learning curve to use the tool. The McPAT users have the flexibility to turn off these features and set the circuit-level implementation parameters.

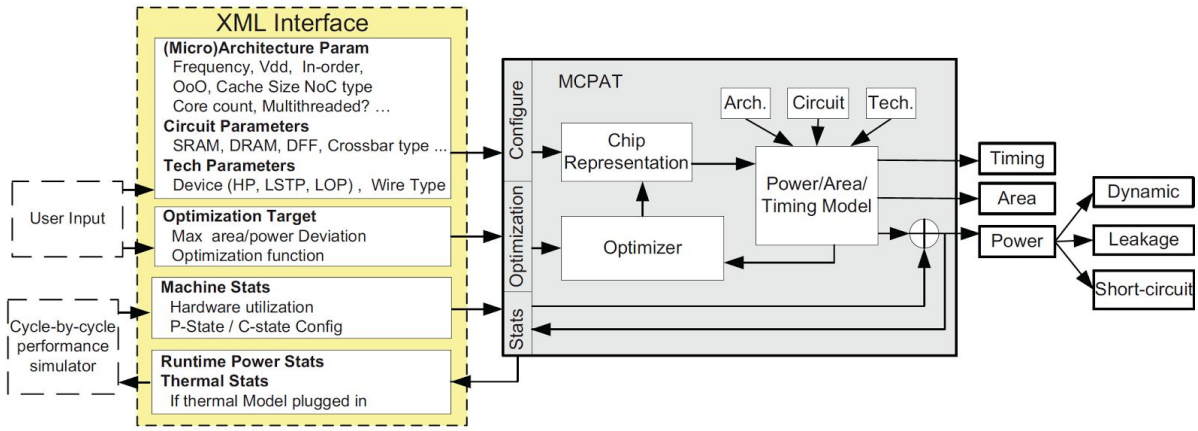

Figure 1: Block diagram of the McPAT framework.

The paper is very detailed and a good starting point for someone starting to study McPAT. I learned a great deal from reading the paper about the manycore/multicore architecture and the power consumption. McPAT is one of the first tools of its kind which integrates power, area and timing into a processor which also includes components needed to model multithreaded and multicore/manycore systems. The previous tools before McPAT only integrated with a specific performance simulator, but McPAT uses XML interface that differentiate and distinguish between power, area and timing analysis which is useful because it allows it to be used with a variety of performance simulators. McPAT and performance simulator results explored the interconnect options of future manycore processors by varying the degree of core clustering over several generations of process technologies.
References - http://www.hpl.hp.com/research/McPAT/McPATAlpha_TechRep.pdf