THERMALLY-AWARE LOAD BALANCING USING PREDICTIVE ENERGY CONSUMPTION MODELS IN HIGH-PERFORMANCE SYSTEMS

Power management techniques developed for mobile and desktop computers are not always appropriate for high-performance computing systems because resources on these systems tend be fully utilized.As the server workload increases, so does the thermal stress on the server processors. Modern processors crudely manage this problem through Dynamic Thermal Management (DTM) where the processor monitors the die temperature and applies Dynamic Voltage/Frequency Scaling (DVFS) to slow the processor and reducing energy consumption and thermal impact. However, DVFS methods entail a heavy cost from both a performance and reliability standpoint.

This paper introduces policy-based, thermally-aware load balancing scheduler that utilizes a full-system thermal and energy model to proactively predict and minimize hazardous thermal effects. The dataflow-based systemwide model estimates the energy consumption and thermal impact of an application during its execution. The knowledge of the run-time predictive thermal status of the processor allows the scheduler time to migrate threads and applications to cooler, underutilized cores with minimal performance impact. The model uses key micro-architectural event counters and operating system kernel statistics to keep track of workload related activity within the processor and associated peripherals.

The scheduler uses the die temperature estimates to adjust the load balancing policies to minimize the number of DTM events. The analytical model and the scheduler performance is demonstrated through empirical evaluation to outperform previous approaches. This is shown by extending an existing power-aware scheduler that is part of the OpenSolaris operating system executed on an AMD Opteron processor. A subset of the SPEC CPU2006 benchmark suite is used to simulate application server workloads and evaluate scheduler performance.