Critique of Paper presented on November 12, 2008


Presenter – Nicholas Tijoe

The paper describes the hyperthreading technology introduced in the Intel Xeon processor and highlights the performance improvement achieved as a consequence.

Strengths

1. The paper chronicles a commercially successful technology, and discusses in detail the importance of sharing policy, discussing three of them in detail.
2. Simultaneous multithreading represents a marked improvement over time-slice multithreading as it avoids frequent switching.
3. Though conceptually different from the uniprocessor approach of the time, hyperthreading does not add any incompatibility to existing software and hardware.
4. It discusses the merits of replicating versus sharing key resources such as instruction translation look-aside buffer, return stack predictor etc.
5. The boost in performance due to hyperthreading is analyzed for various multithreaded and multitasking workloads.

Weaknesses

1. Limited replication to save on die area significantly increases design complexity. If the goal is to achieve hyperthreading at minimum complexity (and hence cost), then the die area increase would be more than the five percent mark indicated by the authors.
2. Validation issues, as the authors acknowledge, are likely to increase given that logical processors have much more interaction than the physical processors due to sharing of resources and resultant contentions.
3. In figure 6, shared cache is shown as having significant improvements over partitioned cache. However, figure 3 shows shared queue as blocked due to heavy thread. While authors mention that full sharing is better for variable working set sizes and partition better for structures with high utilization, a possible trade-off between the two approaches is not discussed.
4. The authors have not justified why greater performance can be achieved in future implementations given that cost containment is a constant goal. To substantiate this a discussion on impact of technology on hyperthreading could have been included.
5. The analysis on CMP seems a bit dated. Subsequently manufacturing cost of CMP has turned out to be not substantially higher compared to single core chip. Infact, design overhead is much smaller in CMPs due to replication. Therefore, the contention that hyperthreading is the best approach to exploiting TLP may not be always true.
6. While the test system is configuration is described, the manner in which the tests are conducted with the benchmark applications (how performance improvement is quantified and measured) is not described.