Review of "Understanding Sources of Inefficiency in General-Purpose Chips" [1]

Riley Wood

April 23, 2016

Using three late days!

Key Ideas

The paper seeks to explore how specialization can benefit a chip multiprocessor (CMP) in terms of performance and power consumption. It is generally known that a specially-designed ASIC will outperform a CMP in these areas, but the paper aims to point out exactly why, in an effort to appropriate the most beneficial characteristics for new CMP designs. They use H.264 video encoding as their primary application for demonstration. H.264 video encoding is something which is typically carried out by an ASIC rather than a general-purpose processor because the ASIC ends up being much better in all respects (power, area, and performance). They begin with an implementation of H.264 on a general processor that lags behind the ASIC drastically. They gradually apply improvements to the CMP which incorporate aspects of the ASIC to improve performance and resource consumption of the CMP. They conclude that reducing the energy cost of CMPs is difficult, as energy overhead will tend to dominate.

Review

I believe they selected an appropriate application for testing, since H.264 video encoding is today widely implemented using hardware acceleration. However, the fact that industry has chosen hardware acceleration over specialized processors means that industry trends have not aligned with the paper's conclusion: that specialization of the processor itself is the best way to maximize performance given energy constraints. I don't think the paper addresses the hardware acceleration school of thought well enough. In the introduction, the authors mention that work is being done to make generating customized hardware easier. They never weigh this option against augmenting the processor. In fact, in their conclusion, they admit that extending the processor is difficult and additional tools will be needed if this is to be made easy for designers. So this approach is no easier than custom designing hardware.

Conclusion

This paper offers a lot of insight into the ways in which ASICs outperform processors in ways that are very relevant to today's trend toward power efficiency. Unfortunately, I disagree with their claim that extending the processor with specialized functionality is the best approach to averting the utilization wall. I am not convinced that this beats integrating what are effectively ASICs into the chip die as accelerators.

References

[1] Rehan Hameed, Wajahat Qadeer, Megan Wachs, Omid Azizi, Alex Solomatnikov, Benjamin C. Lee, Stephen Richardson, Christos Kozyrakis, and Mark Horowitz. Understanding sources of inefficiency in general-purpose chips. In *Proceedings of the 37th Annual International Symposium on Computer Architecture*, ISCA '10, pages 37–47, New York, NY, USA, 2010. ACM.