Research on cache block compression algorithm for reducing power consumption

Takahiro Kawahara (510026)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 8, 2007

Keywords: energy reduction, cache memory, data compression.

1 Introduction

In recent years, a battery life time of mobile computer becomes short and radiation of high performance processor increases rapidly because of the increase of power consumption of the processors. Especially, leakage current is increasing as the process rule of transistors becomes deeper. Leakage current is the phenomenon in which current flows continuously even when the transistor is off. As for high performance processors of which cache memory tends to increse, the increase of the leakage current leads to a problem. Therfore, it is important to reduce the power consumption of a processor and various researches of energy reduction of a processor have been done. This research aims improvement in the energy reduction method by Gated-Vdd using a data compression[1] which is one of the power reducing methods of L2 cache memory. Therefore, I explore an effective algorithm to compress data that is stored into a cache block. I investigate the granularity of power supply control based on the compression size of a block. By the fine-grain control, the compression of many blocks is successfully made and power reduction is performed efficiently.

2 Related works

2.1 Gated-Vdd[2]

There is the Gated-Vdd as a technique for controlling the power supply of cache blocks. A Gated-Vdd transistor with a high threshold is placed between the SRAM cells which are used for cache memory and GND. Leakage current is reduced by turning off this transistor and turning off power supply.

2.2 Power redection by Gated-Vdd and used data compression[1]

The system of this research tries compression for the data stored in the L2 cache memory. The data is stored into the L2 cache memory in the compressed form if the data are compressed into smaller than half of the block size. Power consumption is reduced by turning off power supply using Gated-Vdd to the vacated space by compression. Data loss by turning off power supply of the cache block is avoided by storing the data in the compressed form.

3 Energy-saving cache memory that uses data compression

This research aims at improvement in the rate of power reduction by reconsidering compression algorithms in the power reduction method by Gated-Vdd and data compression. First of all, I compress data of cache blocks using various compression algorithms. Then, I find the optimal compression algorithm for compression of the data of cache blocks. Here, I explain compression algorithms that are candidates for evaluation in this research. The Frequent Pattern compression[3] used by the conventional research compresses a data by using the rules that are based on bit patterns of data. The Frequent Value compression[4] compresses a data by using frequent values. The X-Match algorithm[5] compresses the data by using difference with data referred to in the past. The X-RL algorithm is an algorithm to take the concept of run length to the X-Match algorithm.
I consider granularity of compression size in addition to the compression

algorithms. Here, the distribution of compression size by the distribution, Frequent pattern compression is shown. With the distribution, it is effective to set the granularity of data compression to 1/4 from conventional 1/2. I propose hardware organization required for the power control by 1/4 measure.

4 Evaluation

I evaluate the results of simulations. SPECint95 benchmark[6] was used as the programs for evaluation. The rate of power reduction, the number of blocks per compression size and execution clock cycles were shown for each compression algorithms for every benchmark programs. As a result of the experiment, the X-RL algorithm reduced the power by about 30% on average compared with the non-compressing execution. Since the increase in execution clock cycles is suppressed within 2%, there is little influence on execution speed.

5 Conclusion

To improve the performance of the power reduction method by Gated-Vdd and data compression, the compression rate of cache blocks should be improved. First, I focus attention on the compressed data size in the L2 cache memory by Frequent pattern compression. Then, I decided to employ the three-level power control of 1/4, 1/2 and 3/4, rather than only 1/2 control. And I proposed the hardware organization that achieves the three-level control. By simulating SPECint95 programs, I evaluate several compression algorithms in terms of the rate of power redection.

References

[1] A.Matsuda. 'Research on power consumption reduction that uses data compression', Japan Advanced Institute of Science and Technology Master thesis 2006.

[2] M.Powell, S.Yang, B.Falsafi, K.Roy, T.N.Vijaykumar. 'Gated-Vdd:A Circuit Technique to Reduce Liakage in Deep-Submicron Cache Memories', Proc. of ISLPED, pp. 90–95, 2000.

[3] A.T.Alameldeen and D.A.Wood. 'Frequent Pattern Compression:A Significance-Based Compression Scheme for L2 Caches', Technical Report 1500, Computer Sciences Dept., UW-Madison, April 2004.

[4] J.Yang, R.Gupta. 'Energy Efficient Frequent Value Data Cache Design', Proc. of MICRO-35, pp. 197–207, 2002.

[5] M.kjelso, M.Gooch, S.Jones. 'Design and Performance of a Main Memory Hardware Data Compressor', Rroc. of EuroMicro, PP.423–430, 1996.

[6] Standard Performance Evaluation Corporation. 'SPEC CINT95 Benchmarks', http://www.spec.org/cpu95/CINT95/.