



ILLINOIS INSTITUTE OF TECHNOLOGY



Scalable Computing Software Laboratory Technical Report

Department of Computer Science

Illinois Institute of Technology

Reevaluating Memory Stall Time via Concurrent AMAT

Yu-Hang Liu

Department of Computer Science

Illinois Institute of Technology

yliu242@iit.edu

Xian-He Sun

Department of Computer Science

Illinois Institute of Technology

sun@iit.edu

Dec 2, 2013

Technical Report No. IIT/CS-SCS2013-12

<http://www.cs.iit.edu>

10 West 31st Street, Chicago, IL 60616

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of IIT-SCS and will probably be copyrighted if accepted for publication. It has been issued as a Technical Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IIT-SCS prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g. payment of royalties).

ABSTRACT

Memory stall time is the CPU stall time due to memory references. Driven by the infamous memory-wall problem, memory stall time has become a prominent performance bottleneck of high-end computing systems. Intensive studies have been conducted in recent years to reduce memory stall time. However, most of them are developed based on the conventional average memory access time (AMAT) formulation, which has inherent limitations in characterizing concurrency and is not appropriate for modern out-of-order processors. Concurrent-AMAT (C-AMAT) is an extension of AMAT which considers data access concurrency in its formulation. In this research, we use the newly proposed C-AMAT formulation to reevaluate memory stall time. Our contribution is four-fold: first we derive the relationship between memory stall time and C-AMAT by introducing the term of $C\text{-AMAT}_{\text{stall}}$; next we present three analytical results for C-AMAT, and therefore memory stall time, reduction; then simulation results are conducted to verify our theoretical findings; finally a summary is presented to discuss the possible research directions to reduce memory stall time via memory concurrency. In this study, an explicit expression of memory stall time for modern out-of-order processors is given. The influences of memory system concurrency and application memory access concurrency are formally analyzed and explored. Considering the importance of memory stall time, the new memory stall time formulation proposed in this study has the potential to be an effective tool for both future architecture and software design on various modern processors.

Categories and Subject Descriptors

D.3.3 [Computer Systems Organization]: Performance of Systems

General Terms

Performance, Algorithms, Theory, Measurement, Verification

Keywords

Memory stall time; memory wall; memory concurrency; concurrent average memory access time (C-AMAT); average memory access time (AMAT)