Improving Memory Access in Asymmetric Memories

[View U.S. Patent No. 9,959,205 in PDF format.](#)

WARF: P150172US01

Inventors: Nam Sung Kim, Hao Wang

The Wisconsin Alumni Research Foundation (WARF) is seeking commercial partners interested in developing memory architecture that boosts execution speed and makes page migration more efficient.

Overview

In computer memory there is a tradeoff between capacity (how much data can be stored) and latency (how quickly it can be accessed). To achieve a balance, typical memory systems employ a hierarchy of different memory structures with different capacities and latencies.

The same goal can be achieved using a single asymmetric memory. In this type of memory, data is allocated between a fast (low capacity) bank and a slow (high capacity) bank. Any given data element is held in only one of the banks according to its access frequency.

However, there are several drawbacks that need to be addressed with this type of memory. For example, it takes a lot of time and processor resources to move data between the two banks. Such movement entails reading an entire page of data and then writing it back to a new location. Problems are exacerbated by changing workload patterns.

The Invention

UW–Madison researchers have developed a shared row buffer system for asymmetric memory. The system better accommodates changing patterns of data by sharing row buffers between fast and slow memory banks.

In general, the row buffers provide a vehicle for swapping data. They can be loaded from one memory bank and then reassigned to the other memory bank. The system provides a lightweight method of moving data between memory banks without incapacitating the memory channel or involving the processor.

Applications

- Memory architecture
- Applies to resistive memory technology such as PCM and MRAM; possibly DRAM

Key Benefits

- Accommodates dynamic workload conditions
- Flexible and scalable
- Improves speed
- Permits lightweight memory swap
- Low processor/memory channel overheads

Stage of Development

Performance evaluation benchmarks.

The new hardware can improve the performance of a system running multiprogram workloads by up to 10 percent, and can compete with memory architecture comprised of only fast banks.

Additional Information

Related Technologies

- WARF reference number P120224US01 describes a method for improving GPU performance using memory-link compression.

Tech Fields

- Information Technology : Hardware

For current licensing status, please contact Jeanine Burmania at jeanine@warf.org or 608-960-9846

WARF | info@warf.org | 608.960.9850