**Introduction**

Background

Cache memories have a significant impact on performance of modern microprocessors.

- Multi-core processors can accelerate an application by using multiple cores.
- In general, LLC is shared by multiple cores on processors.
- Accesses from cores concentrate on LLC when multi-threaded applications are executed.
- Data with various characteristics are stored.

In order to improve performance, it is necessary to use LLC efficiently based on data characteristics.

Shared data

Shared data are stored in LLC when multi-threaded application are run.

- Private Data - Cache line accessed by only one core
- Shared Data - Cache line accessed by more than two cores

Shared data are frequently reused since they are accessed by multiple threads.

Locality of shared data

- Experimental Setup
  - Processor Cores - 2 GHz, 8 cores
  - L1 DCache, L1 ICache - 32 kB, 2-way, 64 B block
  - L2 Cache - 256 kB, 8-way, 64 B block
  - LLC - 8 MB, 16-way, 64 B block

More than 98% of lines are accessed by multiple threads.

**Shared Data-Aware Partitioning**

Problems

Private data kick out shared data that have a high reusability on LLC.

In order to improve performance, it is necessary to keep shared data on LLC as possible.

LLC management mechanism to protect shared data

- Shared data-aware partitioning
  - An protected area that prevents shared data from being kicked out by private data
  - Shared data are stored in the area
  - Private data are stored in another area
  - Private data do not kick out shared data

- The way-based partitioning mechanism
  - The protected area is identified by a way
    - For example, the area from the way 0 to 4 becomes the protected area
    - Each way has a bit to identify whether the way is protected area or not
  - If data are shared ones, they are stored in the protected area on LLC
  - Otherwise private data are stored in the other area

- LLC management policy
  - We propose two LLC management policies
    1. Shared data accessed by more than two threads are stored in the protected area
    2. Shared data accessed by all threads are stored in the protected area

This approach prevents shared data from being kicked out by private data on LLC.

**Evaluations**

Experimental Setup

- Processor Cores - 2 GHz, 8 cores
- L1 DCache, L1 ICache - 32 kB, 2-way, 64 B block, 2-cycle latency
- L2 Cache - 256 kB, 8-way, 64 B block, 10-cycle latency
- LLC - 8 MB, 16-way, 64 B block, 20-cycle latency
- Main Memory - 200-cycle latency
- Benchmark - PARSEC suite
- # of Way for Shared Data - 1, 2, 15-way

- Benchmarks are run for the entire ROI
- Based on the profiling evaluation, the shared data is identified in advance
- The evaluations are conducted by using two LLC management policies
- For each benchmarks, all the numbers of ways for the protected data are examined, and the number of ways of the best performance is selected

**Performance Evaluation**

Policy 1: Shared data accessed by more than two threads are stored in the protected area

Improving performance by up to 9.1%, and 1.6% on average

Policy 2: Shared data accessed by all threads are stored in the protected area

Improving performance by up to 12.8%, and 3.9% on average

**Conclusions**

- Shared data have a higher reusability than that of private data.
- The shared data-aware LLC partitioning prevents shared data from being kicked out by private data.
  - Shared data are stored in the protected area, on the other hand, private data are stored in the other area.
- The evaluation results show that the proposed mechanism can improve performance for multi-threaded applications.

SC14 New Orleans, Louisiana

(URL) http://www.sc.isc.tohoku.ac.jp/
(E-mail) nishimu@sc.isc.tohoku.ac.jp