

Advancing the Arizona State University Knowledge Enterprise

Case ID:M13-082P^ Published: 2/26/2020

## Inventors

Ke Bai Aviral Shrivastava

## Contact

Shen Yan shen.yan@skysonginnovations. com

## An infrastructure for memory management on LLM multi-core architectures

Limited Local Memory (LLM) multi-core architectures substitute cache with scratch pad memories (SPM). As a result, SPMs have much lower power consumption compared to other multi-core architectures. However, SPMs lack an automatic memory management system which presents a challenge to programmers as heap data sizes may be variable and data dependent. The heap is a region of computer memory that is not managed automatically and is not as tightly managed by the CPU. Allocating to the heap and removing from the heap require manual calls to specific functions within a computer programming language. Managing heap data of the tasks executing in the cores of an LLM multi-core system has become an important issue.

Researchers at Arizona State University have developed a fully automatic and efficient scheme for heap data management. The scheme comprises two components: (1) an optimized runtime library and (2) a modified compiler. The scheme features code transformation for automation of heap management with support for multi-level pointers as well as improved data structures to more efficiently manage unlimited heap data and unburden the programmer from the task of API function insertion. Experimental results on several benchmarks show an average of 43% performance improvement over previous approaches.

Potential Applications

- Chip Manufacturers reduce programming overhead of implementing scratchpad memory based multi-core processors.
- Application Developers create motivation to efficiently utilize the multitude of multi-core processors available in the market.

Benefits and Advantages

- Automated scheme is fully automated to efficiently manage heap data for LLM multi-core architecture.
- Unlimited Heap Data in Local Memory the compiler can request dynamic memory allocation in the global memory and support unlimited heap data in the local memory.
- Lower Overhead coarser granularity leads to lower management overhead, and therefore leads to less Direct Memory Access (DMA).
- Improved Performance average improvement over all benchmarks by 43%.
- Improved Generality can handle multi-level heap pointers and differentiate with stack pointers.

For more information about the inventor(s) and their research, please see  $\underline{Dr}$ . Aviral Shrivastava 's directory webpage