AMSI
AUSTRALIAN MATHEMATICAL
SCIENCES INSTITUTE

National
Collaboration
in the Mathematical
Sciences

Cutting plane approach to parallel sorting on GPUs

Shaowu Liu
Deakin University

This project is in the field of computational mathematics. It concerns the classical problem of sorting. Sorting has many applications in computer science, statistical learning and data analysis, and image processing. Modern applications require fast sorting of very large arrays of data, or order of tens and hundreds of millions of records. High-breakdown robust regression methods is one particular application where multiple sorting of large arrays is needed.

Graphics Processing Units (GPUs) have recently emerged as a promising alternative to traditional high-performance computing. Workstations equipped with high-end graphics cards can be used as desktop supercomputers. Repetitive calculations can be offloaded to GPUs, which can have almost 500 cores and run hundreds of thousands of threads concurrently. Numerous attempts to parallelise sorting for GPUs have been made. The catch is that GPUs execute identical instructions for every block of threads, and hence most of the classical serial sorting methods are not immediately parallelisable, as the instructions deviate.

In this project we developed mathematics and algorithms for an alternative method of parallel sorting on GPUs, based on convex optimization. We aimed at a very efficient parallel bucket sort algorithm, in which the records are distributed into buckets according to pivots—the order statistics of the data. With good pivots the buckets will be of similar size, and their parallel sorting will be efficient.

How do we calculate multiple order statistics without sorting? Already known to Legendre, the median of a tuple (as well as order statistics), is a solution to a univariate convex minimisation problem: minimise the sum of absolute differences between a variable and the data. We applied the cutting plane method to solve such a problem. The challenge of the project is the need to find multiple order statistics (hundreds of thousands) simultaneously, at a cost comparable to that of a single order statistic.

The project involved the following tasks:

1. Formulation and study of the properties of a multi-objective non-smooth optimisation problem with convex objectives.

2. Design of an algorithm for simultaneous and parallel calculation of all the objectives and their subgradients.

3. Design and implementation of parallel cutting plane algorithm for multiple objectives.

4. Proving various properties of the objectives.

5. Implementation of the algorithm on GPUs, and

6. Its testing and benchmarking against other parallel sorting methods.

My contribution to this project was to design the algorithms to the SIMD architecture for parallelisation, evaluate all the objectives simultaneously, and transform the formulas for reduction.

At the time of writing, experiments show that our GPU Bucksort algorithm is competitive with the fastest GPU Radix sort [1] and faster than others [2, 3, 4]. Moreover, our GPU Bucksort algorithm also works with keys larger than double (8-bits), whereas the GPU Radix sort implemented in the Thrust library [5] does not.

In summary, we proposed a new parallel sorting algorithm GPU Bucksort and designed an algorithm for parallel computation of many order statistics based on the minimization problem. We also implemented kelley's cutting plane method to minimize the objectives on GPUs and designed an algorithm for parallel calculation of all the objectives simultaneously.

I enjoyed the project and had a great experience with the Big Day In. The Vacation Research Scholarship has helped broaden my mathematical knowledge and opened my eyes to the wide variety of mathematical research topics that are currently undertaken. I would like to thank my supervisor A/Prof. Gleb Beliakov for teaching and supporting me for this project. I would also like to thank AMSI and CSIRO for their generous funding and for hosting the Big Day In.

Shaowu Liu received a 2011/12 AMSI Vacation Research Scholarship

References

[1] D. Merrill and A. Grimshaw. (2010) *Revisiting sorting for GPGPU stream architectures*. Technical Report CS2010-03, University of Virginia, Department of Computer Science, Charlottesville, VA, USA,.

[2] D. Cederman and P. Tsigas. (2009) ‘GPU-Quicksort: A practical quicksort algorithm for graphics processors’. *ACM Journal of Experimental Algorithmics*, 14:1.4.1-1.4.24..
[3] N. Leischner, V. Osipov, and P. Sanders. (2010). GPU samplesort. In *IEEE International Parallel and Distributed Processing Symposium*, Atlanta, IEEE, DOI: 10.1109/IPDPS. 5470444.
[4] E. Sintorn and U. Assarsson. (2010)‘Fast parallel GPU-sorting using a hybrid algorithm.’ *Journal of Parallel and Distributed Computing*, 68:1381-1388, 2008.
[5] J. Hoberock and N. Bell. Thurs: A parallel template library, http://www.meganewtons.com,.Version 1.3.0.