MSc-project proposal
Energy-efficient multigrid solution strategies and their application in dataflow computing

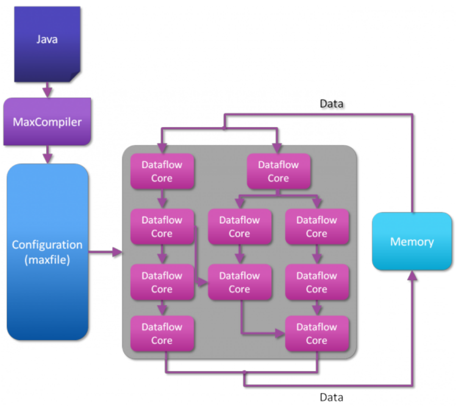
Introduction

Many practical engineering problems require the solution of very large but sparse linear systems of equations stemming from the discretization of partial differential equations (PDEs) by numerical methods such as the Finite Difference Method (FDM), the Finite Volume Method (FVM), or the Finite Element Method (FEM). The direct solution of these equations, e.g., by using Gauss' elimination algorithm, is prohibitively expensive both in terms of computing times and memory required to 'store' the inverse of the matrix.

Iterative solution algorithms compute a sequence of approximations $\{x^n\}_{n=1}^{\infty}$ to the true solution x of the problem $Ax = b$ starting from an initial guess x^0 until the iterated solution has converged to the true one. Basic methods like the Jacobi and Gauss-Seidel method are quite effective in quickly removing the high frequency components of the error in a few iterations but they reduce the low frequencies very slowly. ***Geometric multigrid solvers*** (GMG) overcome this deficiency by employing a hierarchy of nested discretizations so that low frequency components on a fine mesh become higher frequencies on coarser ones and can thus be removed very efficiently.

The main algorithmic components of GMG are the grid transfer operators, the smoothers (e.g. of Jacobi- or Gauss-Seidel-type) and the solver on the coarsest grid. Obviously, the amount of computational work required on fine grids (e.g., with 1 million grid points) is significantly higher than that on coarser grids (e.g. 100 grid points). Thus, large problems clearly benefit from parallel computations, whereas the use of multiple compute units may even increase the computing time if the problem is too small. A recent trend in energy-efficient parallel multigrid solvers [Wl16] is to utilize all available compute units (e.g. cores of a multi-core CPU) on fine grids and send some of them to sleep when operating on coarser grids to reduce the amount of energy consumed and overall runtime.

Commodity hardware like multi-core CPUs only allows for deactivating compute cores to save energy thus leaving a large part of the chip's silicon unused. Field Programmable Gate Arrays (FPGAs) make it possible to realize the solution algorithm directly 'in hardware', so that the available silicon can be split optimally between fine and coarse-grid operations. Maxeler Technologies Ltd. provides a high-performance computing infrastructure based on FPGAs that adopts the dataflow-computing paradigm.



The program source is transformed into a configuration file that describes the operations, layout and connections of computational units. This configuration file is then used to realise the algorithm on the FPGA. In contrast to the traditional computing approach adopted in commodity CPUs, in dataflow computing the data is loaded once from memory and streamed through the entire chain of operations, thereby passing intermediate results from one computational unit to the next without writing them back to off-chip memory.

Problem description

The aim of this project is to develop an energy-efficient geometric multigrid solver for Poisson's equation on Maxeler's dataflow engines [Max]. The equations in two and, possibly, three dimensions are discretized by the Finite Difference Method, which can be easily accomplished by Maxeler's MaxGenFD framework. The main tasks of this project are to realize grid transfer operators and smoothers in an efficient, i.e. most probably matrix-free manner and implement a coarse-grid solver (Gaussian elimination or factorization-based direct solvers) on Maxeler's dataflow engines. Algorithmic components shall be designed with scalability in mind, that is, strategies to split the problem into smaller ones and distribute them over multiple dataflow engines need to be developed.

Challenges

The challenges of this project are formulated in the following research questions:

- ***How to realize the algorithmic components of a geometric multigrid solver efficiently on dataflow engines?*** The use of advanced methods for the (iterative) solution of linear systems of equations on dataflow engines is in its infancy [Bu14, Bu15, Ch14]. The main challenge here is to develop optimal dataflow graphs that maximize the throughput of data streaming from off-chip memory through the compute chain. Matrix-free implementations of grid transfer operators, e.g. by using stencil-based prolongation/restriction have been developed in the context of high-performance CPU- and GPU-computing and need to be extended to dataflow computing.
- ***How to split the available chip silicon between the algorithmic components to maximize the saving of energy?*** The second goal of this project is to implement a geometric multigrid method that directly realizes the concept of 'sending compute cores to sleep on coarser grids' in hardware. This is equivalent to maximizing the problem size that can be handled efficiently by a single dataflow engine without the need to communicate with other dataflow engines.

Time schedule

The following tasks are foreseen:

- Familiarization with Maxeler's data flow computing technology and the development tools (MaxJ programming language, MaxGenFD library)
- Literature study on geometric multigrid methods for Poisson's equation
- Development of a reference CPU code without energy-saving strategies
- Development of the same approach on Maxeler's dataflow engines and analysis of its performance in terms of total computing times with respect to the baseline code
- Thesis writing

Further information

For further information please contact Matthias Möller (m.moller@tudelft.nl) or Georgi Gaydadjiev (g.n.gaydadjiev@tudelft.nl). Access to a dedicated Maxeler server with four MAX3 dataflow engine cards will be granted. Moreover, Maxeler Technologies Ltd. offers the possibility for a short stay at its headquarter in London, UK for hands-on training. Technical support from Maxeler experts will be available during the entire project..

Literature

[Bu14] P. Burovskiy, S. Girdlestone, C. Davies, S. Sherwin, and W. Luk, Dataflow acceleration of Krylov subspace sparse banded problems. In: Proceedings of the 24th Int. Conf. on Field Programmable Logic and Applications (2014) 1-6.

[Bu15] P. Burovskiy, P. Grigoras, S. Sherwin, and W. Luk, Efficient assembly for high order unstructured FEM meshes. In: Proceedings of the 25th Int. Conf. on Field Programmable Logic and Applications (2015) 1-6.

[Ch14] G.C.T. Chow, P. Grigoras, P. Burovskiy, and W. Luk, An efficient sparse conjugate gradient solver using a Benes permutation network. In: Proceedings of the 24th Int. Conf. on Field Programmable Logic and Applications (2014) 1-7.

[Max] Maxeler Technologies Inc. http://maxeler.com

[Wl16] M. Wlotzka, T. Beck, and V. Heuveline, An energy-efficient parallel multigrid method. Presented at: 5th European Seminar on Computing, June 5-10, 2016.