ABSTRACT

Master Degree

on:

"Investigation of algorithms with sparse matrices on CUDA"

by

Gorbyk Oleksandr Viktorovich

The actuality

The speed of the algorithms with sparse matrices is critical for many applications, especially for circuit software, working with matrices of large size. The multiplication of sparse matrices on column vector is the basis of many algorithms that are specific to modeling and analysis of electronic circuits. In working with sparse matrices explicit parallelism in the data - processing array elements can occur independently. For effective use of parallelism inherent advisable to use massively-parallel architecture of graphics cards.

CUDA technology provides a powerful and user-friendly interface for general purpose computing on graphics cards. To speed up algorithms must take into account the peculiarities of the format of sparse matrices and of the architecture. In practice, CUDA gives the acceleration of certain algorithms in 10-30 times. Therefore, development and research of effective methods of computing power, in particular, CUDA is a very urgent problem.

The purpose

The aim is to develop algorithms with sparse matrices using a general purpose computing on graphics cards and CUDA study the possibility of using accelerated implementations of algorithms for the application of technical modeling schemes.

Problems that are solved

1. The study features the work of massively-parallel architecture CUDA.
2. Research opportunities to accelerate key algorithms with sparse matrices using CUDA technology.
3. The implementation of the algorithm multiplying sparse matrices by vector on graphics accelerators (matrices formats: DIA, ELL, COO, CSR and ALLTED).
4. Implementation of the method of solving system of linear algebraic equations on graphics accelerator.
5. Feasibility study of implementations of algorithms for applications.

Achieved results

Having solved the problem that put in the work, the author defends:

1. results of the features work on massively-parallel architecture CUDA;
2. results of the acceleration of the basic algorithms of sparse matrices using technology CUDA;
3. implementation of algorithms for multiplication of sparse matrices on vector-side graphics accelerators formats DIA, ELL, COO, CSR and ALLTED;
4. implementation of an iterative method of solving systems of linear tenuous these algebraic equations on the side graphics accelerator;
5. results of possible implementations of algorithms for applications.

Scientific novelty

The scientific novelty of the work is that:

1. study the main features of CUDA technology and demonstration of the possibility of its use in circuit software;

2. the effect of the choice of format for sparse matrices on the speed of the algorithms with them. A software module for the testing of the basic algorithms to work with sparse matrices in the environment of CUDA;

3. a mathematical model interaction between GPU and CPU;

4. developed a software module for testing of the basic algorithms of sparse matrices in the environment of CUDA;

5. developed a kernel function to run on a graphics card that deals with matrix multiplication on vector format that uses PPP ALLTED.

The practical value

The practical value of the work is that:

1. experimentally investigated and demonstrated the effectiveness of environment for CUDA application package schemes of technical modeling;

2. experimentally investigated the implementation of matrix multiplication in a vector format that uses PPP ALLTED.

Conclusions

1. The features work with massively parallel architecture, CUDA. The main problems associated with effective use of graphics hardware.

2. Investigated the acceleration of the basic algorithms of sparse matrices using technology CUDA.

3. Implemented algorithms for multiplication of sparse matrices on vector-side graphics accelerators formats DIA, ELL, COO, CSR and ALLTED. Research in the use of these formats on the side of the graphics card.

4. Implemented an iterative method for solving rarefied system of linear algebraic equations which side graphics accelerator. The analysis of its performance and optimized cost of shipping data.

5. A possibility of using implementations of algorithms for application, particularly in the circuit simulation package ALLTED using exported models Ansys.

The work contains 160 p., 23 figures, 7 tables, 29 sources.

Keywords: THROUGH-OUTPUT ARCHITECTURE, SPARSE MATRIX, CUDA, GENERAL PURPOSE COMPUTATION ON GRAPHICS CARD.