Parallel processing of matrix operations on the Pulsyr configurable architecture

Andriy Lutsyk, Bohdan Rusyn and Oleksiy Lutsyk
Institute of Physics and Mechanics of the Ukrainian National Academy of Sciences
5, Naukova St., 79060 Lviv, Ukraine

The paper describes the application of a pulsating information arrays (Pulsyr) for parallel calculation of matrix operations. The matrix operations form the base of many branches of scientific investigations and belong to computationally cost calculations. The Pulsyr architecture suits well image and video processing and pattern recognition problems, but parallelism inherent to matrix operations makes attractive the use of the Pulsyr for solving tasks in the matrix calculation field.

Presently, there is the following imbalance in computer engineering: contemporary electronics manages to double amount of transistors per chip every couple of years, however sizes of the chips grow slowly. Even in a ten-year period (it is predicted that in 2010 there will be microchips with 1225 square mm area) it does not expect any substantial increases in the area of integrated circuits unless entirely novel ways of processor design are developed. In present paper we offer a solution to this problem to design massively-parallel computer, or homogeneous computing structure, on a whole silicon wafer and implement matrix operations on this structure.

A homogeneous computing structure (HCS) [1, 2] is a regular network, or an array, of locally connected one-bit processors, or cells. A pulsating information array (Pulsyr), presented in the paper, is a subclass of the homogeneous computing structure implemented in silicon, where cells communicate and process information in parallel and multipipeline. It is designed processing arrays, where cells, on-bit processors, are arranged in the whole semiconductor wafer. This allows to substantially increase number of elementary processors in the parallel computer and thus form engineering basis for design of high-performance computing devices. The proposed computing architecture suits well possible implementation of advanced processing technologies, including image processing, neural networks, matrix calculations to name but few. It is developed advanced algorithmic procedures to deal with crystal imperfections of the wafer, the procedure makes the proposed computing system fault-tolerant. Moreover, the modified structure will be capable of self-testing, reconfiguration and restoring their serviceability with faulty of one-bit processors.

The homogeneous computing structure is ready to move towards the wafer and has the advantage over other directions in computer engineering: microprocessor and transputer systems, systolic arrays and field programmable gate arrays [3].

An Architecture of the Pulsyr

The traditional approach of the chip production requires 100% suitability of the chip. Since the Pulsyr is fabricated on the whole wafer, it is obvious, that some parts would be faulty and this can lead to disability of separate computing cells. Architectural features allow to avoid these faulty cells and provide the process of tuning and functioning of the system.

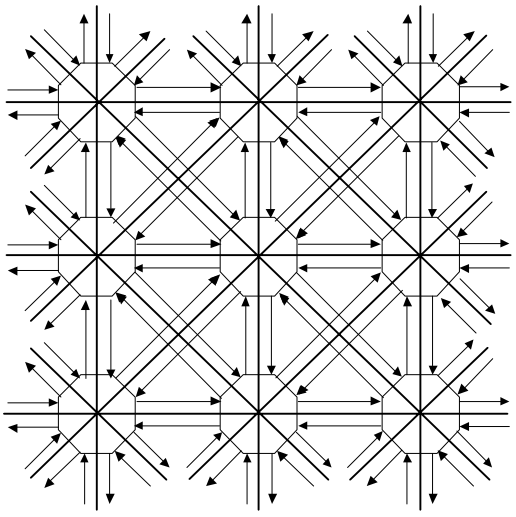

Figure 1. The matrix of the Pulsyr

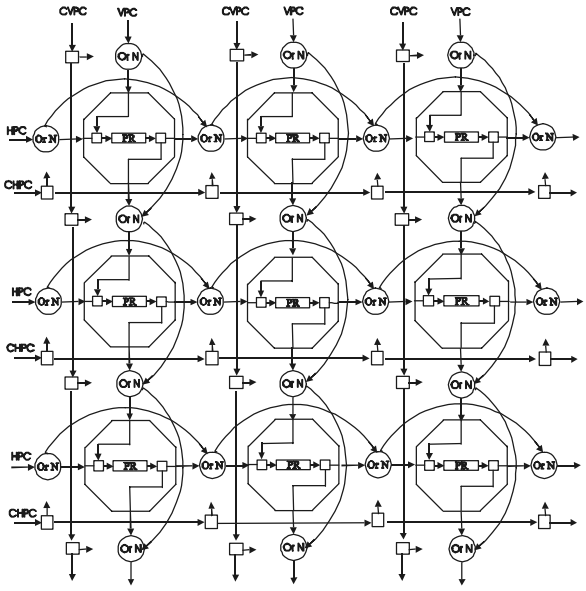

Figure 2. The horizontal and vertical tuning channels with bypassing nodes for instruction registers

Figure 1. shows a Pulsyr matrix. The cell has eight gates for connection with the nearest neighbor cells. Therefore, the Pulsyr cell has enough developed switching possibilities. The Pulsyr cell carries out an appropriate logical operation and as well as functions of switching of information flows. The important feature of the Pulsyr is vertical, horizontal, and two diagonal two-directed buses. Each cell can provide input/output of data on these buses. Figure 1. shows a network of such buses covered on the cell matrix. Including of buses in the Pulsyr field allows to decrease delays and expend switching possibilities. The bus organization provides operational links between cells, when they are not neighbors of each others and also in the case if it is a need to transmit information to remote cells.

A vulnerable place of the Pulsyr is an instruction register of the cell containing a setup of all nodes of the cell to execute required functions. Strengthening of the such important node increases the Pulsyr reliability on the whole. Tuning of the Pulsyr through serially connected cells has essential shortcoming, that is, a defect of the instruction register of one of a cell makes impossible to tune the whole row of computing cells. Therefore, it have been chosen the approach of multichannel access to the instruction register. Such purpose is achieved by organization in the structure of the Pulsyr, horizontal and vertical channels for the access to every of instruction registers.

If the instruction register is disabled, the overriding arc is provided. Fig. 2. shows the Pulsyr scheme which has the overriding arcs and nodes for instruction registers in horizontal and vertical directions.

The cell carries out functions of computation, switching and memory. The cell includes following units: a register channel, a channel of data processing, a channel of memory, a service channel, a transit channel, a bus channel and address channel.

The register channel has the 94 bit instruction register, the content of which influences on work of the whole cell. According to an instruction code, ALU carries out following operations: addition with carry, logical multiplication, logical multiplication with inversion, modulo 2 addition, memorizing of "1", constant generation.

Functioning of the Pulsyr

The process of Pulsyr tuning and the process of computation is shared in time. In the beginning a stream of instructions is formed for the Pulsyr tuning and is written to the instruction registers of every cell, and after that the stream of data is entered. During data execution the instruction codes are kept in the instruction registers of the cells. When computational requirements are changed, the new stream of instruction codes is entered to the structure, swapping hardware configuration of the Pulsyr. Thus, it is possible to execute a series of tasks in rapid succession.

One clock generator synchronizes the work of all cells. Data are processed in the multipipelined mode. The information streams from the input device applied to the information inputs of the Pulsyr are processed in accordance with the instruction codes, moving synchronously with the clock cycles from one cell to another in the matrix. The system has not a control unit, a main memory and a data bus.

The microprogram module (MPM) is a group of computing cells, as a rule, of rectangular shape in the field of the Pulsyr which carries out arithmetical and logical operations. Therefore, a certain region of the field tuned by micro instruction is a specialized processor, programmed on execution of the defined operation of arbitrary but the fixed word length. Synchronization of the large amount of cells using one clock generator allows to perform data streams without intermediate result storage.

When a faulty cell appears inside a microprogram module, the structure of the module can be rearranged to avoid the faulty cell due to some redundancy of the microprogram module or expanding of the size of the module.

Several parallel matrix multiplication algorithms [4] implemented on the Pulsyr with different number of processors ($m \times m$, m, where m is the sizes of the matrix to be processed) are considered. Algorithms of matrix multiplication implemented on the Pulsyr architecture are efficient with linear speed-ups.

References

1. A. Yu. Lutsyk, B. V. Kisil' and O. L. Pelenskyy, Image Processing in Real Time on Configurable Computing Architecture, The Third International Conference on Digital Information Processing and Control in Extreme Situations, Minsk, pp. 124-129. (2002).
2. A. Lutsyk, O. Lutsyk, and O. Pelenskyy, Parallel image processing on configurable architecture, Parallel Numerics'05: Theory and Applications, Ed. M. Vajtersic, R. Trobec, P. Zinterhof, and A. Uhl, pp. 151-163, (2005).
3. A. DeHon, The density advantage of configurable computing, Computer 4, pp. 41-49. (2000).
4. B. Parhami, Introduction to Parallel Processing: Algorithms and Architectures, Kluwer Academic Publishers, (2002).