University of Victoria | Graduate Studies

Notice of the Final Oral Examination
for the Degree of Doctor of Philosophy

of

ALI ZOLFAGHARI JOOYA

BSc (Azad University, 2003)
MSc (Iran University of Science and Technology, 2009)

"GPGPU Design Space Exploration Using Neural Networks"

Department of Electrical and Computer Engineering

Friday, September 14, 2018
10:30A.M.
Engineering and Computer Science Building
Room 468

Supervisory Committee:

Dr. Nikitas Dimopoulos, Department of Electrical and Computer Engineering, University of Victoria (Co-Supervisor)
Dr. Amirali Baniasadi, Department of Electrical and Computer Engineering, University of Victoria (Co-Supervisor)
Dr. Mihai Sima, Department of Electrical and Computer Engineering, UVic (Member)
Dr. Mohsen Akbari, Department of Mechanical Engineering, UVic (Outside Member)

External Examiner:
Dr. Majid Ahmadi, Department of Electrical Engineering, University of Windsor

Chair of Oral Examination:
Dr. Thomas Saunders, Department of History, UVic

Dr. David Capson, Dean, Faculty of Graduate Studies

Abstract

The enormous computational capability of Graphic Processing Units has motivated the idea of employing them to speed-up the execution of general purpose applications. Although the obtainable performance improvement reported as orders of magnitude for some applications, the challenges of obtaining the optimum performance for non-graphic applications on what was designed to be good at running graphic applications still remains.

To achieve the optimum efficiency, the computation and memory resources of GPU should meet the requirements of the application running on it. For each application, some architectural parameters influence the power and performance of the application more than others. By changing the size of those parameters, one can explore the configuration space of the GPU and if it is done extensively, the optimum configuration can be obtained. However, depending on the characteristics of an application, the search space can be as large as thousands of configurations and exploring such a large search space exhaustively will take a long time which may not be affordable for the user.

We have developed application specific models to facilitate the exploration of GPGPU design space and obtain the power-performance optimum configurations. We have also proposed a method to obtain a configuration that performs well for any given set of applications.

These models can be used by GPGPU programmers to learn about the architectural parameters that most affect the power and performance of an application. This information can be translated into software optimization opportunities. Also, GPU manufacturers can gain insight on the architectural parameters that investing on them would profit GPGPU applications the most in terms of power and performance.