

A Toolchain for Application Acceleration on Heterogeneous Platforms

Takaaki Miyajima

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

School of Science for Open and Environmental Systems
Graduate School of Science and Technology
Keio University

July 2015

主 論 文 要 旨

No.1

報告番号	甲 乙 第	号	氏 名	宮島 敬明
主 論 文 題 目： A Toolchain for Application Acceleration on Heterogeneous Platforms (異種混在環境におけるアプリケーション高速化のためのツールチェーン)				
(内容の要旨) 近年、CPUに加えてGPU (Graphic Processing Unit) やFPGA (Field Programmable Gate Array) などのアクセラレータを搭載した環境が一般化してきている。このような異種混在型プラットフォーム (Heterogeneous Platforms) は、高い演算能力が必要とされる科学技術計算分野から低消費電力が要求される組み込み分野まで幅広く利用されつつある。設計環境やライブラリなどのプログラミングが整備されてきているが、各プラットフォームでの既存のアプリケーションの高速化は専門知識なしには困難なままである。 これらの背景を受け、本論文では、異種混在型プラットフォームを対象とした既存アプリケーションの高速化のためのツールチェーンの提案を行う。我々が研究中のツールチェーン Courier は、専門知識なしに既存アプリケーション内部の関数群の解析を行い、解析された関数をアクセラレータへオフロードすることができる。Courier は、実行中のアプリケーションのバイナリ解析を行う"Frontend"、解析で得られた処理フローをグラフ化・言語化する"Courier IR"、解析された関数を対応するアクセラレータの処理に入れ替える"Backend"の大きく 3 つの部分から構成される。また、対象とするプラットフォームは、ノード内にCPU-GPUを持つプラットフォーム (CPU-GPU 混在型)、CPU-GPU 混在型ノードを複数利用するプラットフォーム (CPU-GPU 混在クラスタ型)、CPU-FPGAを持つプラットフォーム (CPU-FPGA 混在型)である。CPU-GPU 混在型と CPU-GPU 混在クラスタ型のプラットフォームでは対応する関数とデータ転送処理をコンパイルしたバイナリを、CPU-FPGA 混在プラットフォームではCPUにバス接続されたハードウェアモジュールを利用する。CPU-GPU 混在クラスタ型プラットフォームでは、筑波大学のノード間直接通信機構 PEACH2 (PCI-Express Adaptive Control Hub 2)を利用し、ノード間パイプラインを構築する。また、CPU-FPGA 混在型プラットフォームでは、ソフトウェア処理とハードウェアモジュールのパイプラインを構築する。Courier と各種プラットフォームを用いて画像処理やベクトル演算、フーリエ級数演算などのアプリケーションが高速化できることを各章で示す。				

SUMMARY OF Ph.D. DISSERTATION

School	Student Identification Number	SURNAME, First name
Science and Technology	()	MIYAJIMA Takaaki
Title		
A Toolchain for Application Acceleration on Heterogeneous Platforms		
Abstract <p>Recently, processing platforms which have a host CPU and accelerators such as GPU (Graphic Processing Unit) or FPGA (Field-Programmable Gate Array) have become mainstream. Such heterogeneous platforms are used in a scientific computation domain which takes a lot of processing power and also used in an embedded device domain which requires a power efficient system. Although design tools or useful libraries are provided, it is not easy to accelerate existing applications on such platforms without expertise or special knowledge.</p> <p>In this thesis, a new toolchain for application acceleration called Courier is proposed. It only requires an executable binary of the target application and a corresponding function code for an accelerator. Besides, it doesn't require a source code of the application nor re-compilation of the binary. A work-flow of Courier is a simple and intended for non-expert users. Courier is consisting of Frontend, Courier Intermediate Representation and Backend. Frontend analyzes processing flow of a running application binary, Courier IR makes graph and generates intermediate code of the processing flow, and Backend replaces target functions in the processing flow with accelerated functions which is running on accelerators. Target platforms are a node which has a CPU and a GPU (a single-node mixed CPU-GPU platform), multiple nodes which have a CPU and GPUs (a multi-node mixed CPU-GPU platform), and a node which has a CPU and an FPGA (a single-node mixed CPU-FPGA platform). To accelerate existing applications, Courier uses compiled binaries which include accelerator functions and data transfer functions for a single-node and a multi-node mixed CPU-GPU platforms. It also uses hardware modules including CPU-FPGA bridge bus for a single-node mixed CPU-FPGA platform. In the case of a multi-node mixed CPU-GPU platform, Courier uses a special inter-node communication system developed by Tsukuba University. It is called PEACH2 (PCI-Express Adaptive Control Hub 2) and an inter-node task pipeline is built. In the case of a single-node mixed CPU-FPGA platform, I construct a mixed hardware software pipeline which is composed of software functions on CPU and hardware modules on an FPGA. We conducted case studies so as to show the applicability of Courier. Some applications including image processing, vector processing and Fourier transformation, are accelerated by using Courier on heterogeneous platforms.</p>		