FPGA-based Accelerated Cloud Computing with AWS EC2 F1 and SDAccel

Conference: International Conference on Computer Architecture – ISCA 2018

Date and Venue: June 2-6, 2018; InterContinental Los Angeles Downtown, Los Angeles, CA

Tutorial Date and Time: June 3, 2018; 8:30 AM to 12:30 PM

Presenter: Parimal Patel, XUP Senior Systems Engineer

Abstract: The increasing computational requirements of next-generation Cloud and High-Performance Computing (HPC) applications are pushing the adoption of accelerated computing based on heterogeneous architectures into mainstream, as traditional CPU technology is unable to keep pace. FPGA accelerators complement CPU-based architectures and deliver significant performance and power efficiency improvements.

In this regard, Xilinx FPGAs are now available on the Amazon Elastic Compute Cloud (EC2) F1 instances, which are designed to accelerate data center workloads, including machine learning inference, data analytics, video processing, and genomics. These are available in two different sizes that include up to eight Virtex® UltraScale+ VU9P FPGAs with a combined peak compute capability of over 170 TOP/sec (INT8). Furthermore, Amazon Web Services offers the SDAccel™ Development Environment for cloud acceleration, enabling the user to easily and productively develop accelerated algorithms and then efficiently implement and deploy them onto the heterogeneous CPU-FPGA system.

SDAccel completely automates the step of the hardware design flow, offering an easy to use environment for FPGA application design. It offers the possibility to specify a compute kernel using C and C++ for higher-level algorithmic implementation, or using hardware description languages for RTL designs, while using OpenCL APIs to control run-time behavior. The high performance and high-level of scalability offered by F1 instances, paired with the power and ease of use of Xilinx SDAccel, is very appealing for the development of high high-performance FPGA-based accelerated solutions, and will be the focus of this workshop.

Topics to be covered:

Introduction to FPGA-based acceleration, development framework, platform, and use cases

Demonstration and hands-on-experience

- How to deploy an AWS EC2 F1 instance
- How to design using SDAccel with the Makefile flow
- How to design using SDAccel with the GUI flow
- Developing, profiling and optimizing F1 applications with SDAccel

Attendees will use their laptops to connect to the WiFi network and use Amazon AWS