Integrating network-attached FPGAs into the cloud using partial reconfiguration

Burkhard Ringlein*†, Francois Abel†, Alexander Ditter*, Christoph Hagleitner† and Dietmar Fey*

*Chair of Computer Architecture
Friedrich-Alexander University
Erlangen-Nürnberg, Germany
{burkhard.ringlein, alexander.ditter, dietmar.fey}@fau.de

†IBM Research – Zurich
Rüschlikon, Switzerland
{ngl, fab, hle}@zurich.ibm.com

***Index Terms*—cloud computing, network- attached FPGA, HPRC, partial reconfiguration, data centers**

Extended Abstract

A. FPGAs for HPC and Cloud

The growth of scalable HPC infrastructures has long been driven by Moore's law and by the use of homogeneous CPU-centric platforms. However, because these traditional systems are suffering from extreme power densities on the chips, modern supercomputers and cloud data centers (DC) are breaking this paradigm by starting to exploit high-performance and low-power heterogeneous devices. Field-Programmable Gate Arrays (FPGA) are such heterogeneous devices. Their energy efficiency and low latency have the potential to improve drastically the power density of HPC compute nodes while also delivering results faster.

On the other hand, the application of FPGAs require high development complexity and profound expertise in the field. Hence, research and industry are focusing on the development of higher-level abstraction concepts and tools to facilitate and accelerate the usage and the deployment of FPGAs.

Meanwhile, large-scale applications ranging from business analytics to scientific simulations have started to scale out over distributed frameworks such as Hadoop, Spark, and Tensorflow. These frameworks will consume an immense amount of FPGAs, which must be provided and managed in an efficient and flexible way.

B. Related Work

This research builds on the work by *Abel et al.* in [1] who proposed a dense FPGA platform that decouples the FPGA from the CPU of the server by connecting the FPGA directly to the DC network. That system turns the FPGA into a disaggregated standalone computing resource and renders the platform particularly cost- and energy-efficient because the number of spread-out FPGAs becomes independent of the number of servers.

Although this approach can deliver large pools of FPGA devices for HPC and cloud environments, the authors of [1] do not address the infrastructure management for serving such network-attached FPGAs as virtual entities or clusters to users.

C. Contributions

We propose an architecture for acquiring network-attached FPGAs and distributing the applications (i.e. bitstream configurations) at a large scale in DC infrastructures.

First, we examine the constraints for building a scalable deployment by investigating the requirements from the DC provider and the key users. This analysis leads to our proposal, which ranges from the global DC-management level all the way down to functional cores inside FPGAs.

Second, the concept of disaggregated FPGAs presumes only one communication path to the FPGA: the network interface. This single interface carries both user and management traffic and therefore the corresponding functionalities must be separated into privileged and non-privileged logic within the FPGA. Otherwise, the integrity of the DC network would be at risk, because the user can take control over the network interface to break out of his designated subnet, start denial-of-service attacks or harm the provider network in other ways.

Hence, we enforce the use of *partial reconfiguration* in our architecture to solve this issue and to physically separate the system functions from the user logic within the FPGA. Partial reconfiguration has the additional advantage of speeding up the deployment as well as protecting the user's intellectual property.

Third, we describe the specific management service that we designed for network-attached FPGA platforms. This service maintains a database of the FPGA resources. It also organizes the setup of the user subnet and the distribution of the partial bitfiles and is controlled via an *RESTfull* application program interface. Therefore our service can be integrated seamlessly into modern modular management software like *OpenStack*. The service is organized in multiple layers that fit the hierarchical structure of the DC into racks, chassis, and FPGAs.

Finally, our partial reconfiguration is realized over a 10 GbE network. To the best of our knowledge, this is the first design that implements the above stack, which enables a scalable and easy-to-use deployment of FPGAs in a Hyperscale DC.

References

[1] F. Abel, J. Weerasinghe, C. Hagleitner, B. Weiss, and S. Paredes, "An fpga platform for hyperscalers," *Proceedings - 2017 IEEE 25th Annual Symposium on High-Performance Interconnects, HOTI 2017*, pp. 29–32, 2017. DOI: 10.1109/HOTI.2017.13.