Cluster 2017 – Keynote Speech

Specializing Data Centers using Reconfigurable Logic

Dr. Derek Chiou, Microsoft and The University of Texas at Austin

Abstract:
Introducing reconfigurable logic into data center servers provides both the benefits of specialized hardware and the convenience of homogeneous hardware. Placing an FPGA in the network path as well as attached to the server via PCIe enables an FPGA-centric computational model, in contrast to the CPU-centric computational model that pervades computing today. In an FPGA-centric model, the FPGA is the first to process each packet and only passes the packets it cannot handle to the CPU that acts as a complexity offload engine. Microsoft has deployed such an architecture throughout its cloud and implements a wide range of capabilities, including deep neural networks and software defined networking acceleration, on it. I will describe Microsoft’s Configurable Cloud, some cases of how it is used, and the resulting performance.

Biography:
Derek Chiou is a Partner Architect at Microsoft where he leads the Azure Cloud Silicon team working on FPGAs and ASICs for data center applications and infrastructure and a researcher in the Electrical and Computer Engineering Department at The University of Texas at Austin. His research areas are FPGA acceleration, high-performance computer simulation, rapid system design, computer architecture, parallel computing, Internet router architecture, and network processors. Before going to UT, Dr. Chiou was a system architect and lead the performance modeling team at Avici Systems, a manufacturer of terabit core routers. Dr. Chiou received his Ph.D., S.M. and S.B. degrees in Electrical Engineering and Computer Science from MIT.
Cluster 2017 – Keynote Speech

What is a Proxy App and Why Should I Care?

Dr. David Richards, Lawrence Livermore National Laboratory

Abstract:
Proxy apps are small, nimble codes that can be used to represent larger applications in situations where it would be difficult or impossible to use the real application. Proxies are used to design and test new computer architectures, software technologies, and programming techniques, and will play a role in selection process of the exascale systems that will be purchased by the Department of Energy (DOE). This talk will discuss the role of proxy apps in DOE’s Exascale Computing Program (ECP) and provide examples of how proxies can be used (and misused) in co-design with vendors, research, and development. We will also describe the ECP Proxy App Project and show how that project can benefit anyone who wants to create or use proxy apps.
Cluster 2017 – Keynote Speech

At the Crossroads of HPC and Big Data

Dr. Kate Keahey
Mathematics & CS Division, Argonne National Laboratory
Computation Institute, University of Chicago

Abstract:
Experiments, data, and computation have always been inextricably linked and are even more so today. Large experimental instruments, equipped with millions of sensors, and producing hundreds of terabytes of data per experiment will be used more efficiently if extended with a computational facility providing the scientist with ongoing insight into data. This relationship is becoming stronger as recently these sensors have left the lab and started multiplying at large: inexpensive and increasingly sophisticated sensor devices now allow scientists to instrument forests, oceans or cities turning our planet into an “instrument at large” and providing unprecedented opportunities in geophysical, environmental, and social sciences. All this is creating demand to process more data, faster, and produce results in a more timely fashion.

This presentation will describe how emergent technology is creating potential for new avenues of exploration and how this potential is translated into new scientific applications – but also new infrastructure requirements and new ideas on how computing can support science. I will give examples of different approaches explored by various scientific application groups and discuss ideas on what we can do to catalyze change in tools and infrastructure – from specific solutions to changes in experimental approach -- to support new modes of usage.