Multiscale Dataflow Processing

Oskar Mencer

Maxeler Technologies

We are at the beginning of the era of Data. While Microprocessors are great for computing with small amounts of data, the microprocessor paradigm does not scale well to Exadata and beyond. In order to scale further, we need to apply the lessons of the industrial revolution and of manufacturing to data processing. It is ridiculous to build hardware which then tries to guess each time it hits a branch if it should go one way or the other and load data from physical locations that get resolved at runtime. You would not start a trip and on the way to the train station decide if you go away for a day or a month, travel to the other side of town, or to another continent. Similarly, when manufacturing billions of articles, the assembly line worker should not be making any decisions but implement simple actions. For dataflow kernels processing large amounts of data, simple arithmetic units organized like a factory can be mathematically locally optimal in achieving any computation, without branches, without load/stores, without virtual memory, and without an instruction stream.

In this talk, I will show tools and results of over a decade of dataflow processing with applications such as climate modelling, Quantum Chromodynamics, seismic imaging, Quantum Espresso, Financial Risk Calculations, CyberSecurity Intrusion Prevention, machine learning, brain simulation, and Etherium mining.