Programming acceleration the structured way

M. Danelutto

The jungle

- FPGAs, CGRA, SoftCores
- GPUs, GP-GPUs
- TPUs, NPUs
- ASICs
- PIMs
- Neuromorphic

HowTo

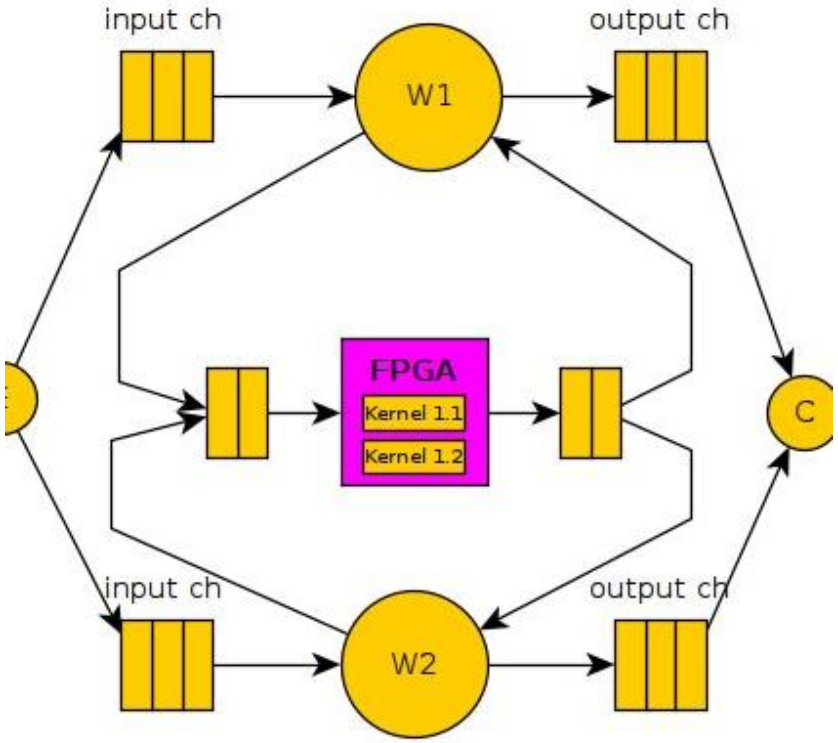
- Host to device
  - Communications
  - Synchronizations
- Device memory management
  - Buffer allocation/deallocation
- Code generation
  - Assembly (e.g. PTX), RTL (e.g. Verilog), unknown (e.g. Google TPU)

Two ways

- Application programmers
  - use libraries, libraries, libraries
- Accelerator developers (system programmers)
  - Provide abstractions
    - (parallel) Patterns
  - Hide mechanisms
    - Garbage collecting device memory space
  - Guarantee performances
    - Exploit low level mechanisms
  - Ensure portability (functional and performance)
    - Really needed ??? GP -> specilization trend -> applicative domain programming tools

FFPGA

- FastFlow
  - Structured parallel programming framework
  - Implements high level streaming and data parallel patterns
- Vitis
  - HLS by Vitis (FPGA)
  - Compiles kernels to FPGA
- Seamless integration of FPGA acceleration in "mainstream" code
  - Efficient management CPU<->FPGA transparent to application programmer
- Second (future) step
  - Generate kernels automatically



Perspective research areas

- Orchestration of accelerated workflows
  - Stream processing with accelerated stages (comm hiding & co.)
- Efficient host-accelerator interface
  - Reducing time-to-production, increasing programmer productivity
- Accelerator-specific optimization techniques
  - Exploitation of non principal mechanisms (e.g. streaming on GPU SM, blas on TPUs
- Reconfigurable hardware prototyping of innovative models

Perspective research areas

- Orchestration of accelerated workflows
  - Stream processing with accelerated stages (comm hiding & co.)
- Efficient host-accelerator interface
  - Reducing time-to-production, increasing programmer productivity
- Accelerator-specific optimization techniques
  - Exploitation of non principal mechanisms (e.g. streaming on GPU SM, blas on TPUs
- Reconfigurable hardware prototyping of innovative models