



# OpenACC to FPGA: A Directive-Based High-Level Programming Framework for High-Performance Reconfigurable Computing

Seyong Lee<sup>1</sup>, Jacob Lambert<sup>2</sup>, Jungwon Kim<sup>1</sup>,  
Jeffrey S. Vetter<sup>1</sup>, and Allen D. Malony<sup>2</sup>

<sup>1</sup>Oak Ridge National Laboratory <sup>2</sup>University of Oregon  
<http://ft.ornl.gov/research/openarc>

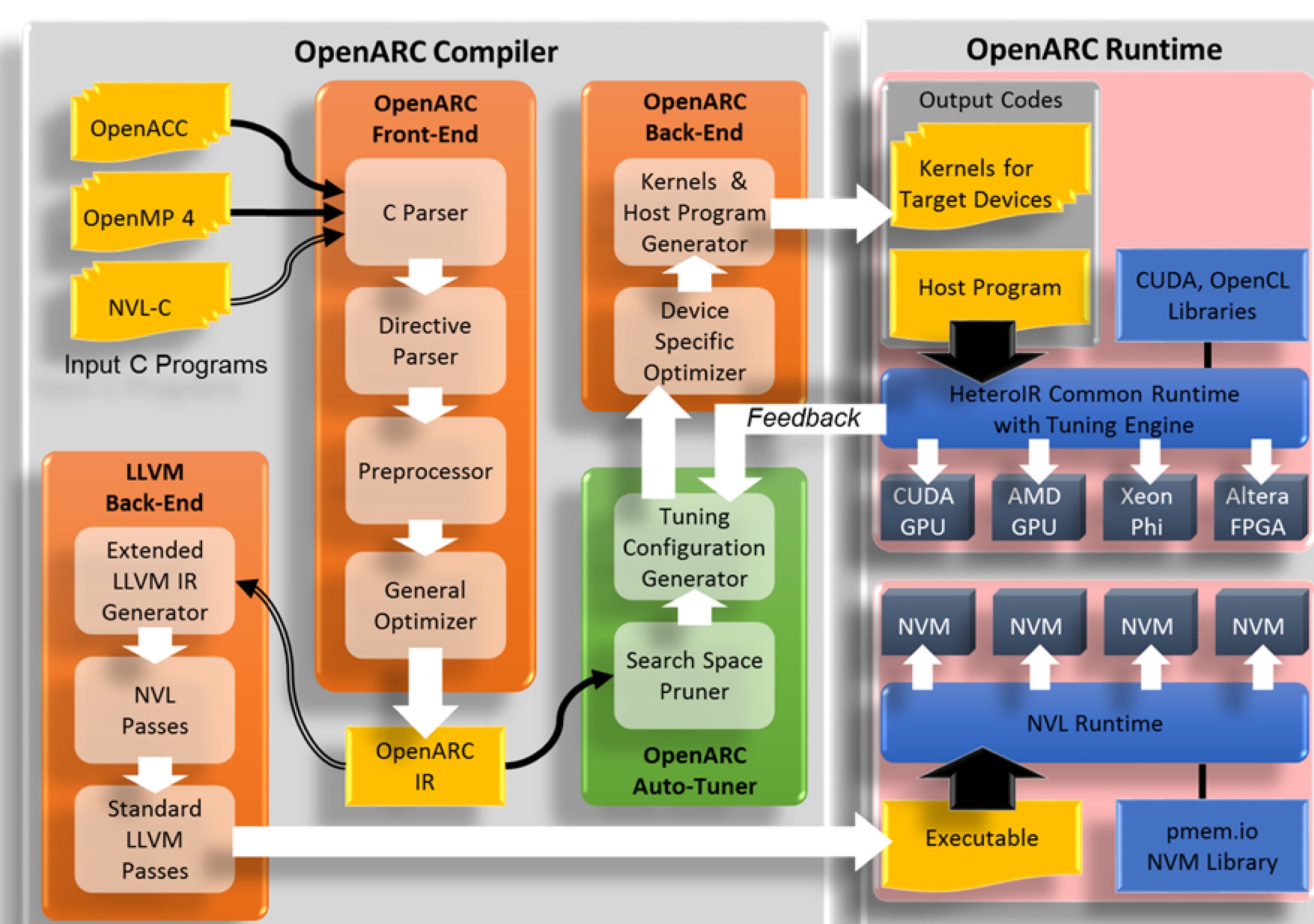


## Motivation

- Directive-based accelerator programming models provide abstraction over architectural details and low-level programming complexities.
- FPGAs provide high-performance, low-power execution of many applications, but often require low-level programming.
- Proposed OpenACC-to-FPGA framework, implemented in OpenARC, offers a high-level, directive-based alternative for FPGA programming.

## OpenARC: Open Accelerator Research Compiler

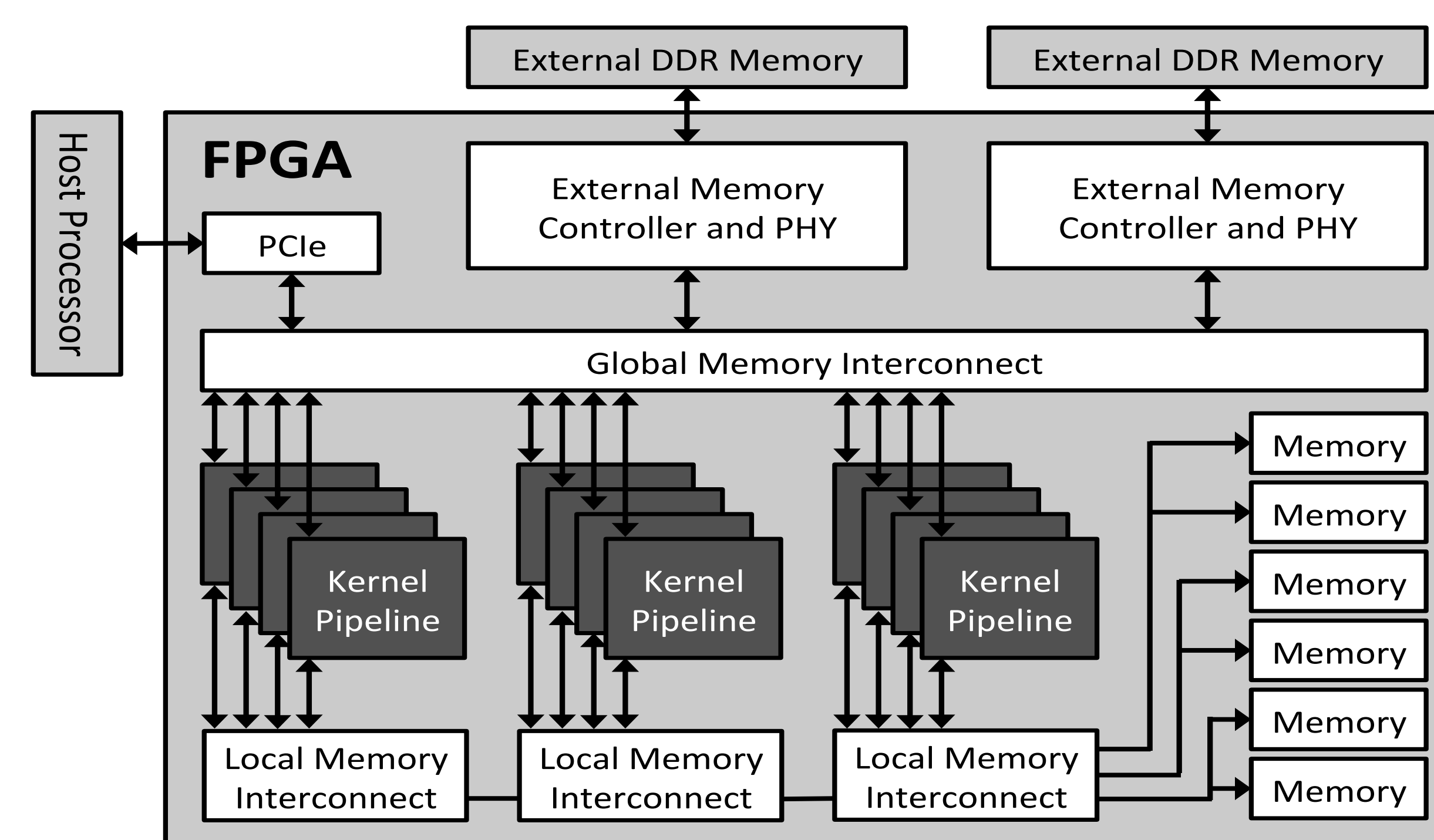
- Open-Sourced, High-level Intermediate Representation (HIR)-based, extensible compiler framework.
  - Performs source-to-source translation from OpenACC/OpenMP C to CUDA/OpenCL C/C++ customized for target architectures.
  - For targeting FPGA architectures, OpenARC supports the Altera SDK for OpenCL (AOCL) and Altera Offline Compiler (AOC) as its backend.



OpenARC System Architecture

## Field Programmable Gate Array (FPGA)

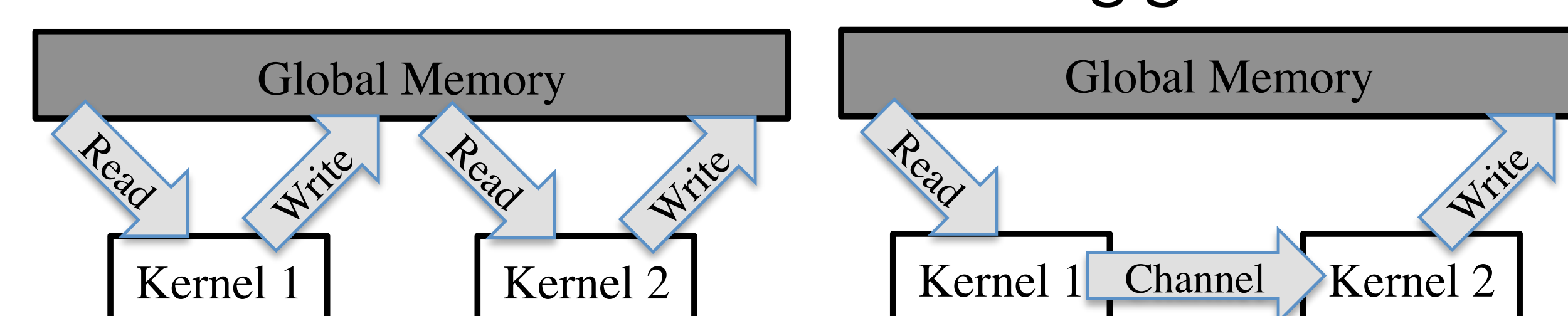
- High-performance, flexible, and power-efficient devices reconfigured at the hardware level to execute specific programs.
- Required knowledge about low-level HDLs like VHDL and Verilog, or more recently, Altera OpenCL.
- OpenARC offers those familiar with OpenACC/OpenMP a high-level alternative to HDLs and OpenCL for FPGA programming.



FPGA OpenCL Architecture

## FPGA-Specific Optimizations

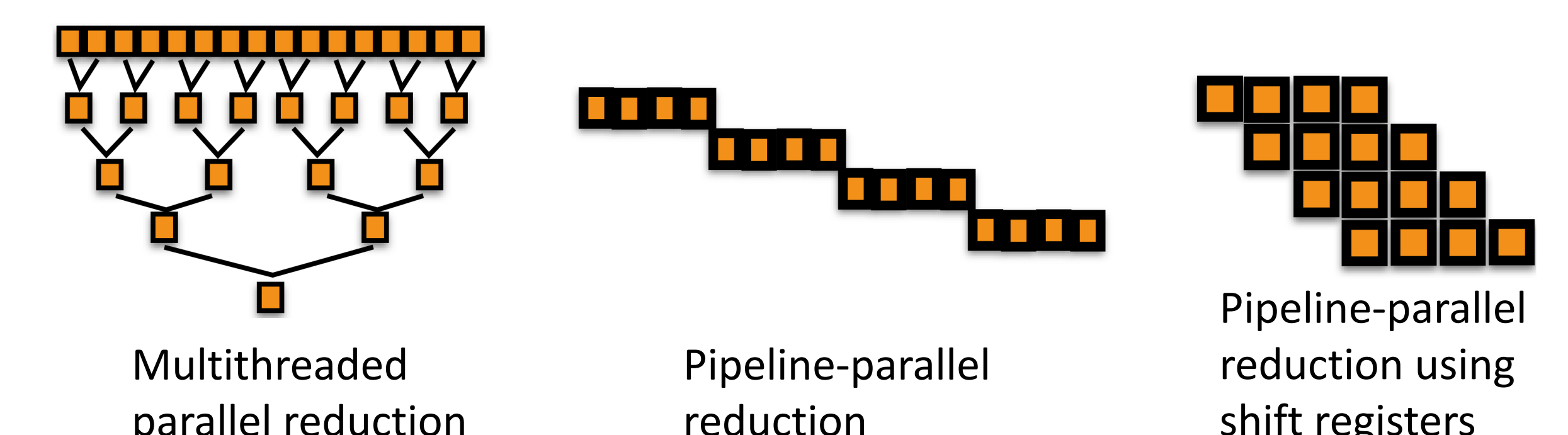
- Dynamic Memory-Transfer Alignment
  - AOCL uses Direct Memory Access (DMA).
  - Transfers not 64B-aligned degrade performance.
  - OpenARC's optimization automatically aligns memory transfers to exploit DMA.
- Kernel Pipelining Transformation
  - Proposed directive extensions and compiler transformations exploiting Altera OpenCL channels to allow direct communication between kernels without using global memory.



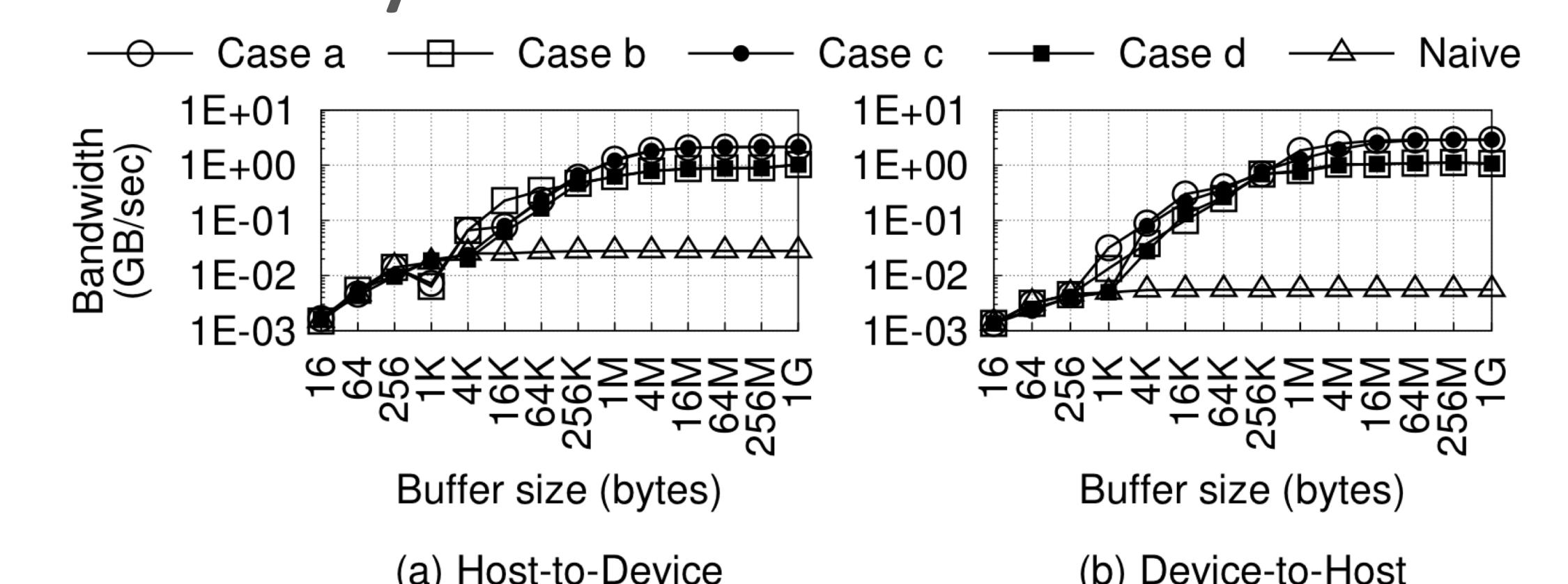
(a) Global Memory Access Without Channels (b) Global Memory Access With Channels

## Optimizations to Exploit Shift Registers

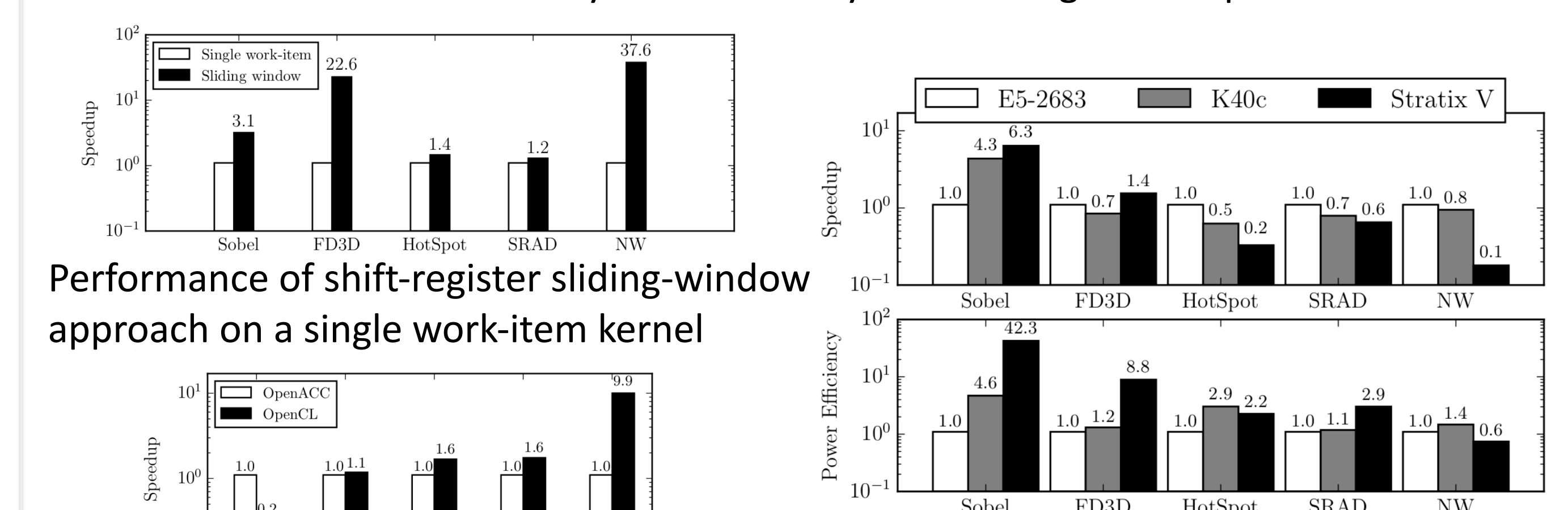
- Sliding Window Transformation
  - FPGA executions benefit from deep pipelines.
  - Shift registers perform well on FPGAs.
  - AOC infers a shift register by recognizing specific programming patterns.
  - Proposed a new directive and transformations to automatically generate the programming pattern needed by AOC to infer shift registers.
- Reduction Optimization
  - Developed a compiler transformation to generate pipeline parallel reduction, which uses shift registers to increase dependence distance, improving pipeline efficiency.



## Preliminary Evaluation



(a) Host-to-Device (b) Device-to-Host  
Performance of dynamic memory-transfer alignment optimization



Performance of shift-register sliding-window approach on a single work-item kernel  
Comparison of OpenACC code with FPGA-specific directives and hand-optimized OpenCL

Comparison of directive-based FPGA approach with directive-based CPU and GPU Approaches