



September 23, 2024

An Open-Source RISC-V-based GPGPU Accelerator for Machine Learning-based Edge Computing Applications



EPFL - Embedded Systems Laboratory (ESL)

simone.machetti@epfl.ch

Motivation

Analyzing the state-of-the art, we realized the need for a GPU...

Motivation

Analyzing the state-of-the art, we realized the need for a GPU...

Open-source

Motivation

Analyzing the state-of-the art, we realized the need for a GPU...

Open-source

Natively Configurable

Motivation

Analyzing the state-of-the art, we realized the need for a GPU...

Open-source

Natively Configurable

RISC-V-based

Motivation

Analyzing the state-of-the art, we realized the need for a GPU...

Open-source

Natively Configurable

RISC-V-based

Fully Synthesizable

Motivation

Analyzing the state-of-the art, we realized the need for a GPU...

Open-source

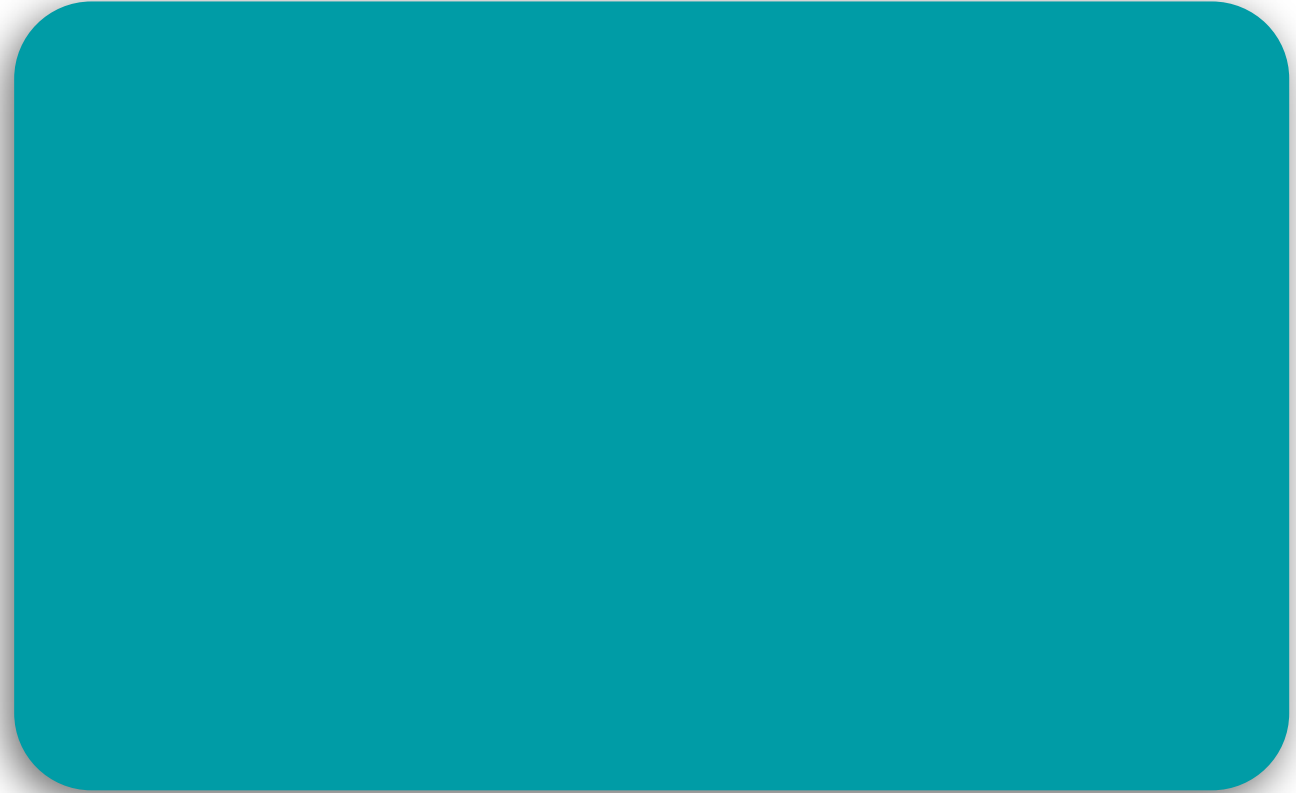
Natively Configurable

RISC-V-based

Fully Synthesizable

OpenCL Support

GPGPU Accelerator



GPGPU Accelerator



Streaming Multiprocessor

GPGPU Accelerator



Streaming Multiprocessor

Memory Hierarchy



GPGPU Accelerator

Configurable



Streaming Multiprocessor

Memory Hierarchy

GPGPU Accelerator

Configurable


-  Number of threads

Streaming Multiprocessor

Memory Hierarchy

GPGPU Accelerator

Configurable


-  Number of threads
- Number of warps

Streaming Multiprocessor

Memory Hierarchy

GPGPU Accelerator

Configurable


-  Number of threads
- Number of warps
- Floating-point unit

Streaming Multiprocessor

Memory Hierarchy

GPGPU Accelerator

Configurable

-  Number of threads
- Number of warps
- Floating-point unit
- Memory hierarchy




The diagram shows a teal rounded rectangle containing two stacked rounded rectangles. The top one is red and labeled 'Streaming Multiprocessor'. The bottom one is gray and labeled 'Memory Hierarchy'.

Streaming Multiprocessor

Memory Hierarchy

GPGPU Accelerator

Configurable


-  Number of threads
- Number of warps
- Floating-point unit
- Memory hierarchy
 - Scratchpad-based

Streaming Multiprocessor

Memory Hierarchy

GPGPU Accelerator

Configurable

-  Number of threads
- Number of warps
- Floating-point unit
- Memory hierarchy
 - Scratchpad-based
 - Cache-based

Streaming Multiprocessor

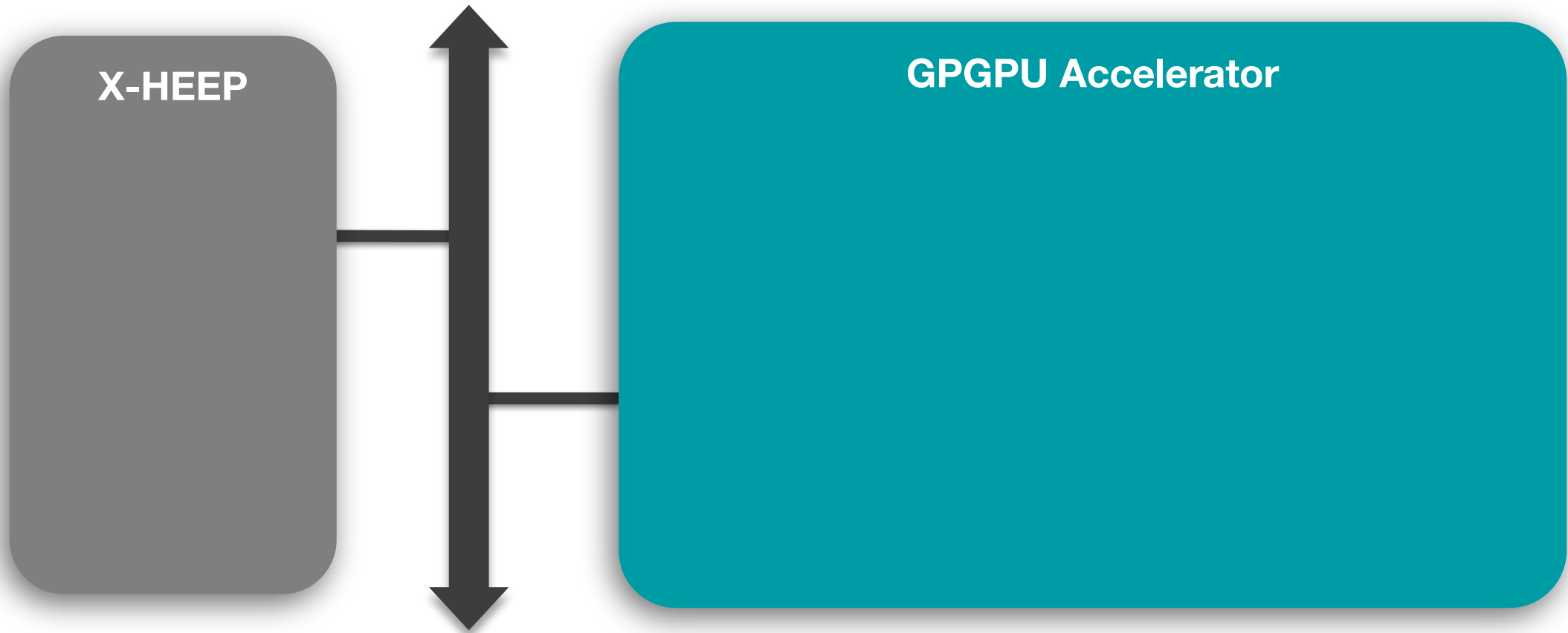
Memory Hierarchy

Accelerated Processing Unit (APU)

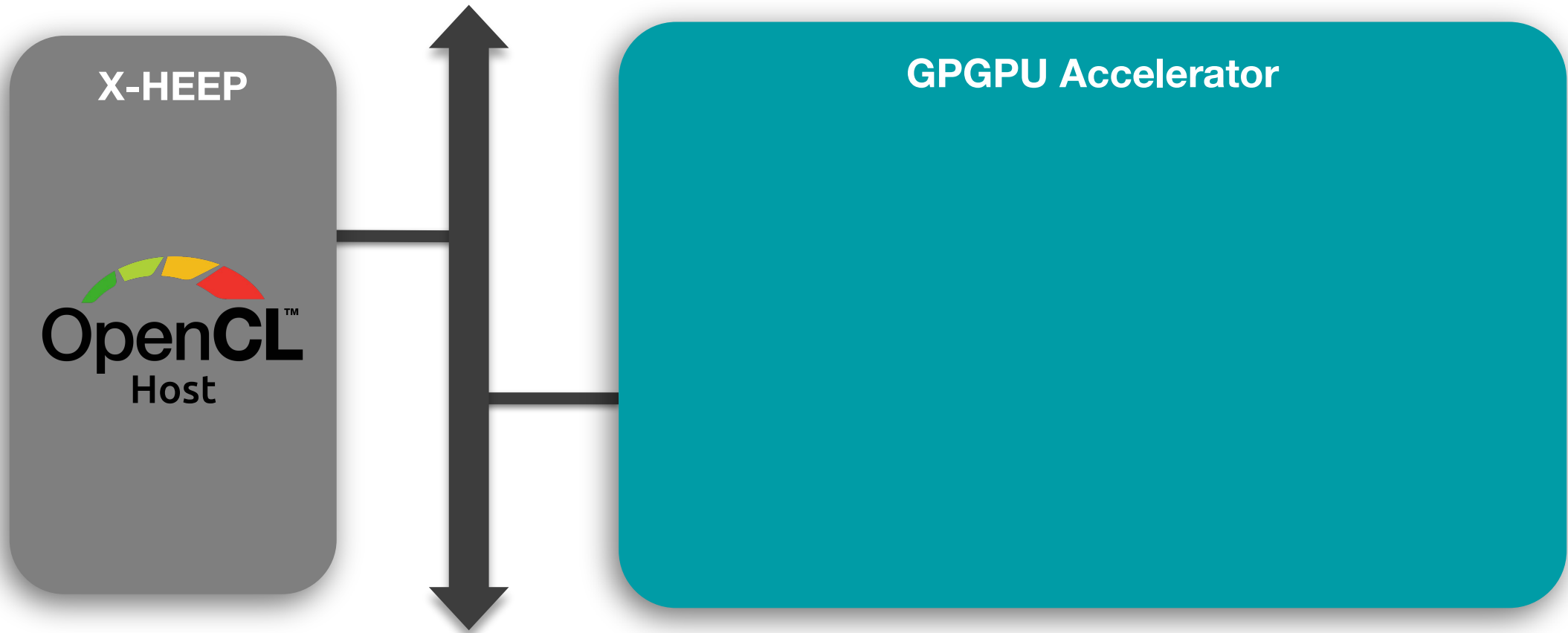


GPGPU Accelerator

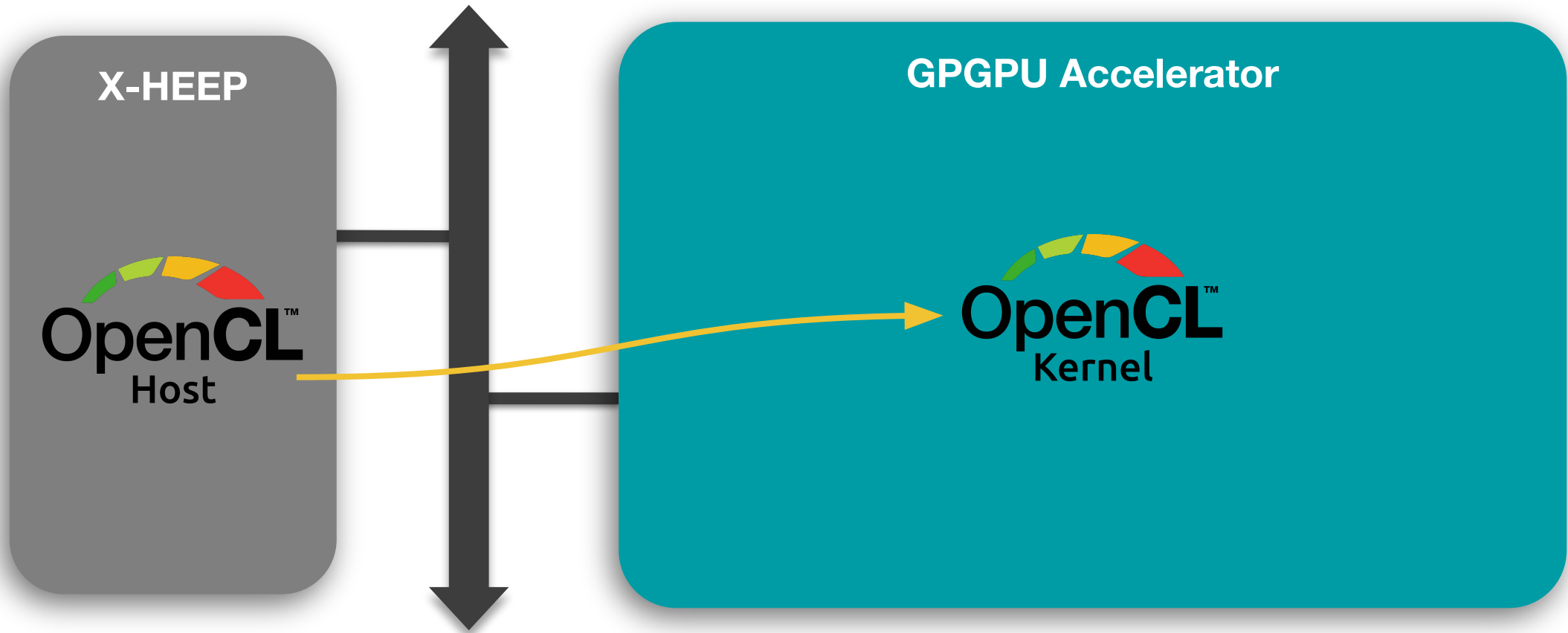
Accelerated Processing Unit (APU)



Accelerated Processing Unit (APU)



Accelerated Processing Unit (APU)



Conclusion

The APU code and documentation will be 100% open-source and the first version will be released very soon...

Conclusion

The APU code and documentation will be 100% open-source and the first version will be released very soon...



Conclusion

The APU code and documentation will be 100% open-source and the first version will be released very soon...



Stay Tuned!



Thank you for your attention!



EPFL - Embedded Systems Laboratory (ESL)

simone.machetti@epfl.ch