Integration of Hardware Acceleration Techniques in Real-Time Framework using FPGA devices

C. González, M. Ruiz, A. Carpeño, A. Piñas , V. Costa, J. Nieto, E. Barrera

Universidad Politécnica de Madrid

c.gonzalezb@alumnos.upm.es

24th IEEE REAL TIME CONFERENCE
April 22-26 2024
ICISE, Quy Nhon, Vietnam

IEEE NPSS NUCLEAR & PLASMA SCIENCES SOCIETY

INSTRUMENTATION & APPLIED ACOUSTICS RESEARCH GROUP

POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID

Outline

- Motivation
  - ITER Real-Time Framework (RTF)
  - AMD XILINX VITIS acceleration technique
- Tools and methodology
- Results
  - Vector add
  - Matrix Multiplication: GEMM from BLAS
- Conclusion
- Future Work

Motivation

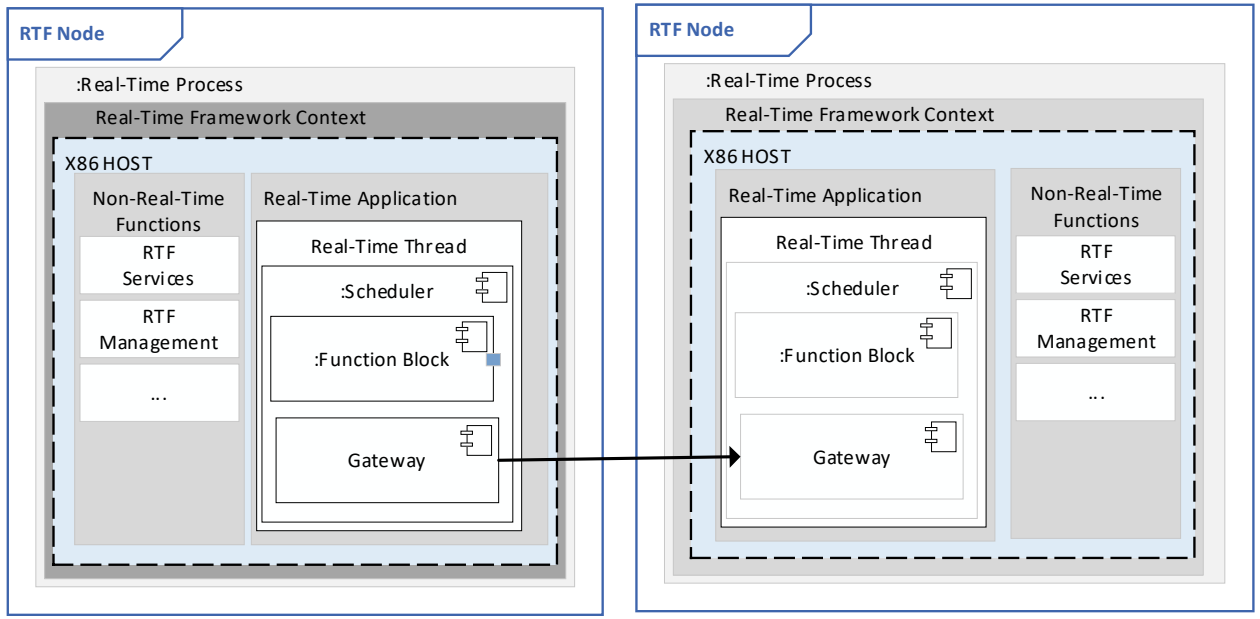
- The use of specific software frameworks to develop and deploy real-time applications is essential in the control systems used in big-science facilities
- Some software frameworks have been implemented in the experimental fusion devices to simplify the development and deployment of real-time applications
  - MARTe
  - ITER Real-Time Framework (RTF)

C. Neto, Sartori,F, Piccolo, F, Vitelli, R, De Tommasi, G, Zabeo, L et al., "**MARTe: A Multiplatform Real-Time Framework**" in IEEE Transactions on Nuclear Science, vol. 57, no. 2, pp. 479-486, April 2010, https://doi.org/10.1109/TNS.2009.2037815

Kadziela M., Jablonski B., Perek P., and Makowski D., "**Evaluation of the ITER Real-Time Framework for Data Acquisition and Processing from Pulsed Gigasample Digitizers**". Journal of Fusion Energy, vol.39, pp. 261–269, November 2020. https://doi.org/10.1007/s10894-020-00264-3

Motivation

- ITER RTF is a collection of software tools that provides common services and capabilities for building real-time applications on a distributed control system integrated into the ***ITER Codac Core System***
- The key element in RTF is the Function Block (FB) that performs from simple operations to complex algorithms
- Reducing the execution time of specific FBs may be crucial for some specific experiments

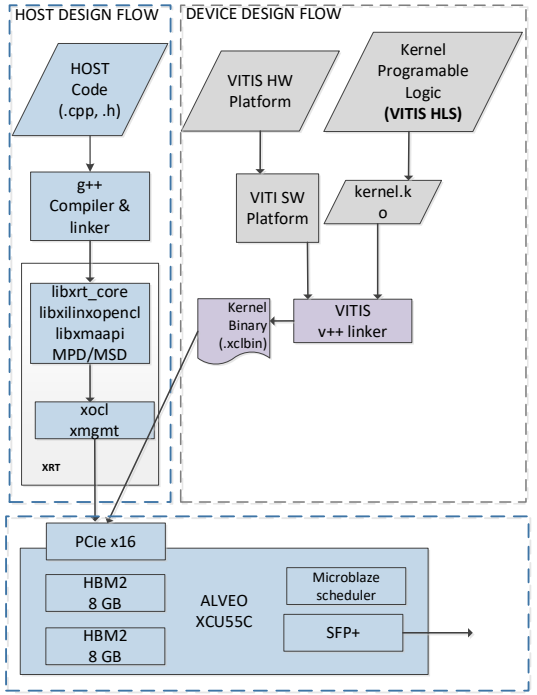


Perek, P., Makowski, D., Kadziela, M., Lee, W. R., Zagar, A., Simrock, et al. “Evaluation of ITER Real-Time Framework in plasma diagnostics applications”, Fusion Engineering and Design, 192. July 2023, https://doi.org/10.1016/j.fusengdes.2023.113623

Motivation

- Implement specific hardware to solve specific problems (Hardware accelerator)
- C-type language to implement the hardware functions
- OpenCL Runtime to manage the hardware accelerator
- PCIe card sharing memory with the host computer

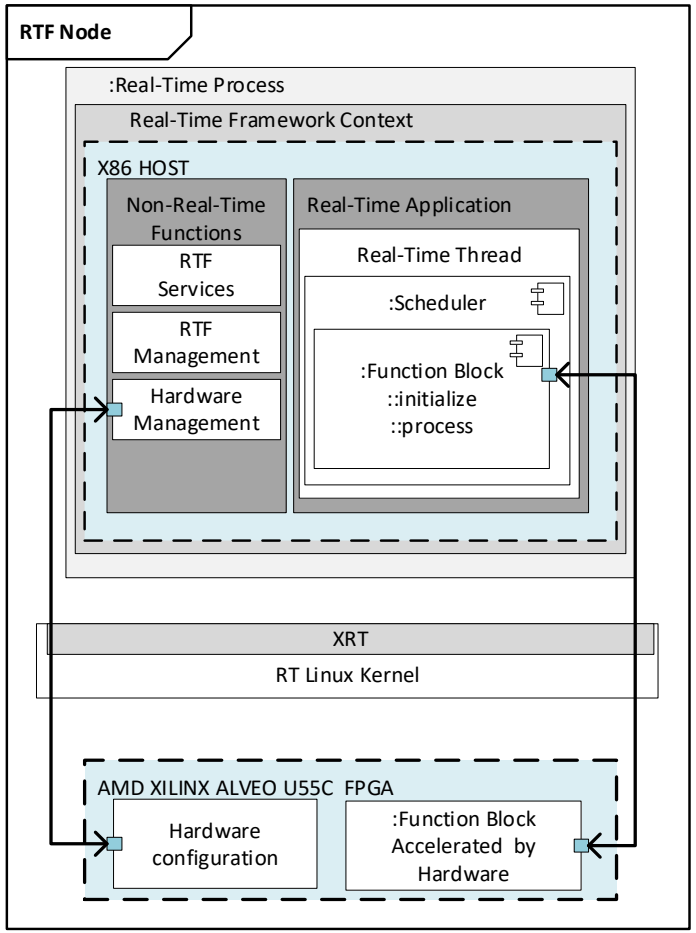
- AMD XILINX hardware acceleration design cycle



Vitis Unified Software Platform Documentation Application Acceleration Development. UG1393 (v2023.2) December 13, 2023.

Tools and Methodology

- Use of ITER CODAC Core System (RHEL 8.5)
- Use of RT Linux Kernel and XILINX XRT
- Use of a XILINX Alveo U55C card
- Use of AMD VITIS for HLS
- Use of C++ 14 for software development in ITER RTF
- Implementation of two applications:
  - Vector add
  - Matrix Multiplications
- Measure the execution time with RTF profiler and AMD Vitis Profiler

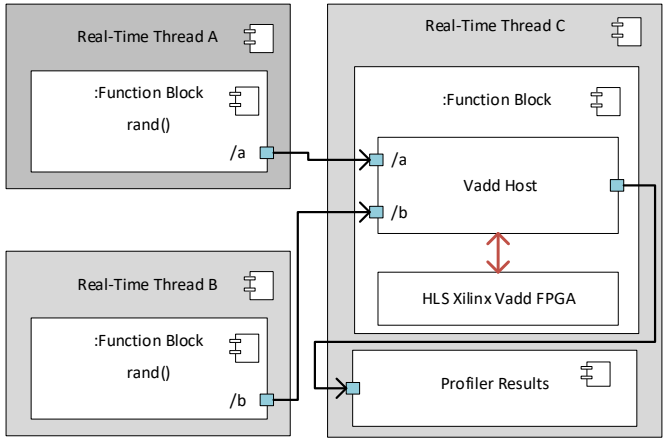


Results: adding vectors N times

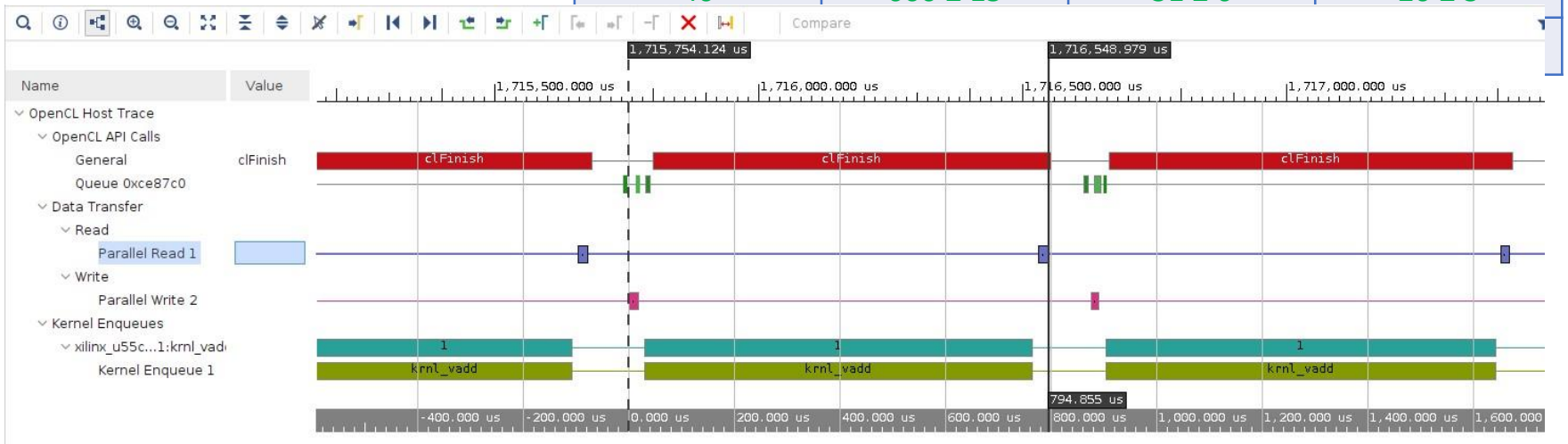
N	CPU (us)	FPGA (us)
	RTF profiler	RTF profiler
1	35.6 ± 1.8	208.0 ± 18.0
10	346.1 ± 2.3	338.5 ± 13.0
20	692.4 ± 13.4	465.0 ± 45.0
30	1037.3 ± 15.1	607.0 ± 24.0
40	1384.8 ± 21.4	744.0 ± 26.0
50	1731.2 ± 27.0	897.0 ± 15.0

Mean value and standard deviation

Vectors of 4096 floats
3000 executions

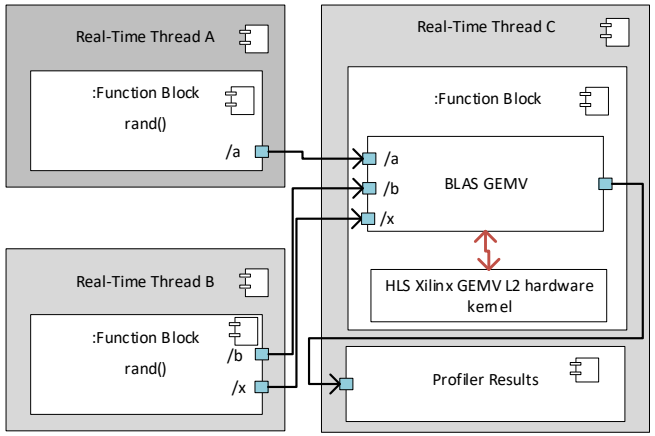


N	Vitis Analyzer (us)		
	Kernel	Write	Read
1	61 ± 6	32 ± 6	22 ± 4
10	183 ± 7	31 ± 7	21 ± 3
20	322 ± 13	31 ± 7	21 ± 7
30	464 ± 6	33 ± 7	20 ± 3
40	600 ± 15	31 ± 6	20 ± 3



Results: Matrix Multiplication

- Use of BLAS library for XILINX VITIS + X



square matrix size	CPU (**ms**)	FPGA (**ms**)
	RTF profiler	RTF profiler
64	0.25 ± 0.002	0.13 ± 0.009
128	2.00 ± 0.007	0.22 ± 0.012
256	22.00 ± 0.038	0.76 ± 0.014
512	222.00 ± 0.112	4.11 ± 0.009
1024	1828.88 ± 1.180	26.418 ± 0.020

3000 executions

Square Matrix Size	Vitis Analyzer (**us**)		
	Kernel	Write	Read
64	40 ± 9	30 ± 3	30 ± 3
128	80 ± 3	60 ± 2	60 ± 2
256	380 ± 5	180 ± 3	180 ± 3
512	2800 ± 5	640 ± 3	640 ± 3
1024	21400 ± 8	2500 ± 4	2500 ± 5

https://docs.amd.com/r/en-US/Vitis_Libraries/blas/overview.html

Conclusions

- Integration of AMD XILINX hardware acceleration technique in ITER RTF framework
- Test with an ITER fast controller configured with an ALVEO U55C and the XILINX XRT using the Red Hat real-time preemptive kernel (4.18.0-348.23.1.rt7.153)
  - No significant issues with the system's latency
  - It requires a specific configuration of the kernel parameters and isolates the specific interruptions of the *xocl* kernel module
- Two applications have been implemented, showing the acceleration obtained using the FPGA
  - The gain obtained with complex matrix multiplication (1024x1024) is 98% of the time used by the CPU

Future Work

- Application to the ITER diagnostics systems implemented with the MTCA platforms, which includes AMC boards with UltraScale MPSoC.
- Investigate the accuracy of the results provided by the profiling tools, and the impact on the latency of the buffer movement implemented by the XILINX XRT.

Project funded by Spanish AEI

- Projects:
  - PID2019-108377RB-C33 MCIN/AEI/10.13039/501100011033
  - PID2022-137680OB-C33 MCIN/AEI/10.13039/501100011033, FEDER, EU

Integration of Hardware Acceleration Techniques in Real-Time Framework using FPGA devices

Thank you for your attention

C. González, M. Ruiz, A. Carpeño, A. Piñas , V. Costa, J. Nieto, E. Barrera

Universidad Politécnica de Madrid

c.gonzalezb@alumnos.upm.es

24th IEEE REAL TIME CONFERENCE
April 22-26 2024
ICISE, Quy Nhon, Vietnam

IEEE NPSS NUCLEAR & PLASMA SCIENCES SOCIETY

INSTRUMENTATION & APPLIED ACOUSTICS RESEARCH GROUP

POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID