Open Hardware Custom Architectures for Near-Sensor Signal Processing and Machine Learning

Joseph Wagane FAYE, email : joseph-wagane.faye@insa-rennes.fr[1], Maxime PELCAT[1], Jean-François NEZAN[1], Kevin MARTIN[2], and Shuvra S. BHATTACHARYYA[3,1]

[1]IETR - UMR CNRS 6164, INSA Rennes, FRANCE
[2]Lab-STICC - UMR CNRS 6285, Lorient, FRANCE
[3]University of Maryland, College Park, USA

April 1, 2022

1 Abstract

In order to improve the energy efficiency of distributed systems, signal processing and machine learning (SPML) workloads tend to be processed closer to the network edge. These workloads make heterogeneous and specialized architectures necessary in high performance embedded systems so as to comply with systems latency, throughput and energy constraints. Exploiting the potential of these specialized architectures remains difficult and costly. The stream processing nature of SPML workloads makes it possible to exploit data, task and pipeline parallelism to optimize the usage of memory and computation hardware resources. The Dataflow Models of Computation (MoCs) have proven useful to formalize the aforementioned applicative properties and feed automated optimization processes. From these abstractions, a datapath analysis can be employed to determine the longest latency path of the system and identify its latency limiting operations [3]. Additionally, taking advantage of the observer pattern, notifying memories can reduce the amount of memory accesses of SPML workloads [2]. A generalized trend is the use of caches in memory systems so as to mask the latency of external memories. However, SPML workloads tend not to benefit from these cache hierarchies [1].

These elements motivate for proposing methods for specializing hardware architectures for SPML near sensor systems, building on the dynamism of the Open Hardware and RISC-V community. This poster will discuss PathTracer and Notifying Memories, as well as research objectives on exploiting a high-level knowledge of the application datapath to reduce the memory access costs of SPML systems.

References

[1] Alemeh Ghasemi, Rodrigo Cataldo, Jean-Philippe Diguet, and Kevin J. M. Martin. On cache limits for dataflow applications and related efficient memory management strategies. In Tomasz Kryjak and Andrea Pinna, editors, *DASIP '21: Workshop on Design and Architectures for Signal and Image Processing (14th edition) - in conjunction with HiPEAC 2021, Budapest, Hungary, January 18-20, 2021*, pages 68–76. ACM, 2021.

[2] Kevin J. M. Martin, Mostafa Rizk, Martha Johanna Sepúlveda, and Jean-Philippe Diguet. Notifying memories: a case-study on data-flow applications with noc interfaces implementation. In *Proceedings of the 53rd Annual Design Automation Conference, DAC 2016, Austin, TX, USA, June 5-9, 2016*, pages 35:1–35:6. ACM, 2016.

[3] Claudio Rubattu, Francesca Palumbo, Shuvra Bhattacharyya, and Maxime Pelcat. Pathtracer: Understanding response time of signal processing applications on heterogeneous mpsocs. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, 02 2022.