Silicon Proven Hardware Acceleration of Post-Quantum Cryptography on RISC-V

Jonas Schupp[1*], Patrick Karl[1] and Georg Sigl[1,2]

[1]Chair of Security in Information Technology, TUM School of Computation, Information and Technology, Technical University of Munich
[2]Fraunhofer Institute for Applied and Integrated Security, Germany

Abstract

Hardware/Software co-designs offer a promising way to support different Post-quantum cryptographic algorithms on one platform. Especially, as these algorithms are not yet standardized, the flexibility of such an approach is important to support possible future algorithm changes in the standardization process. To explore the design space of such RISC-V based HW/SW co-designs, we present three different ASICs, designed since 2020, accelerating different subsets of the PQ-algorithms in the NIST competition, one in UMC 65nm and two in Globalfoundries' 22nm. All designs offer significant performance advantages over pure software implementations on the same platform, while largely maintaining the flexibility of a pure software approach.

Introduction

As state-of-the-art asymmetric cryptography can not withstand an attack with a capable quantum computer, new cryptographic primitives need to be found and implemented for secure communications in the future. Current post-quantum algorithms are based on different mathematical problems, among others lattices, isogenies and error-correcting codes, therefore a wide range of mathematical functions needs to be supported in hardware or in software to enable different devices for PQC. While the corresponding algorithms can be run on a general purpose processor core, the resulting performance is insufficient for real world scenarios, especially when talking about embedded devices. To overcome this limitation, we propose different HW/SW co-designs to accelerate different sets of Post-quantum algorithms. Here we focus on algorithms which were promising candidates in the NIST competition at the time of the tapeout. Within the following sections, we present three different designs, all based on the PULPino platform.[1]

Tightly Coupled RISC-V accelerators (2020)

The focus of this tapeout project is to evaluate the possible advantages of tightly coupled hardware accelerators for lattice based cryptography as well as to gain experience in ASIC tapeouts. For this tapeout a rather old technology, i.e. UMC 65nm, was chosen. The algorithms at target are HewHope, Kyber and Saber. For these algorithms, the main performance bottlenecks are polynomial arithmetic and the generation of pseudo random numbers. To accelerate these specific operations, accelerators for NTT and Modular arithmetic, Keccak, Binomial Sampling and Multiply and Accumulate are integrated as shown in Figure 1. This design achieves a speedup factor of up to 11.4 for NewHope, 9.6 for Kyber, and 2.7 for Saber at a maximum frequency of 45.47 MHz [1]. An instruction set extension to use these accelerators is proposed in this work as well.

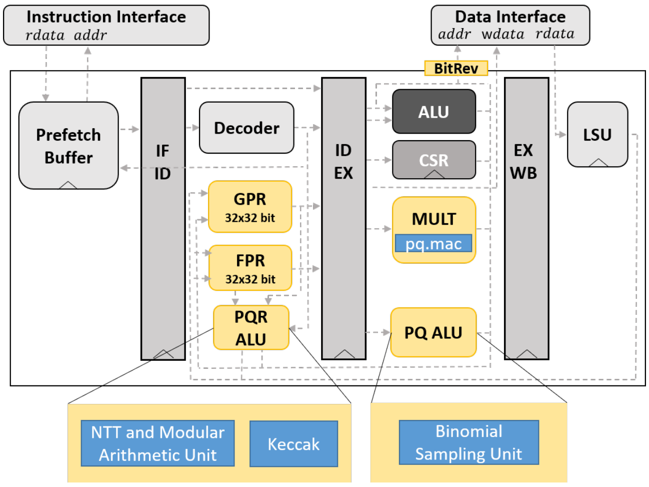

Figure 1: *Integration of tightly coupled accelerators in the RISC-V pipeline*

Tightly and Loosely Coupled (Masked-)Accelerators (2021)

In difference to the afore mentioned tapeout, which mainly focuses on lattice-based cryptography, one goal for this tapeout is to support Key Encapsulation Mech-

*Corresponding author: Jonas.Schupp@tum.de

[1] https://github.com/pulp-platform/pulpino

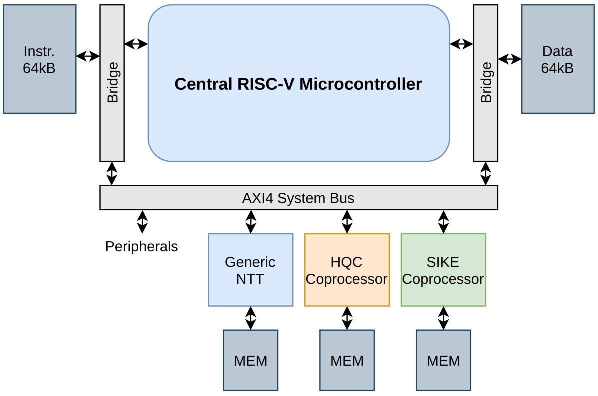

Figure 2: *Overview of the accelerators in the tapeout with tightly and loosely coupled accelerators*

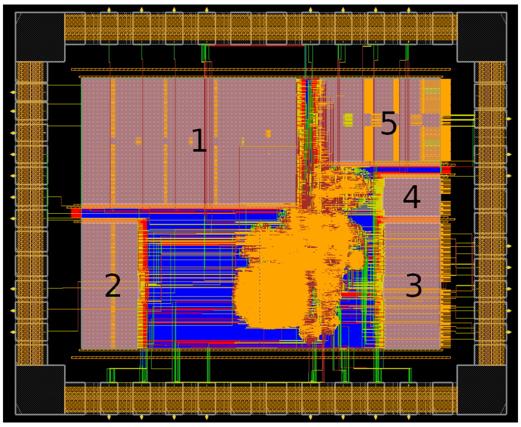

Figure 3: *Floorplan of the ASIC with accelerators for post-quantum signatures*

anisms (KEMs) from three different families of post-quantum cryptography. This includes schemes which are lattice-based, i.e. Kyber and Saber, SIKE (isogenies) and HQC (code-based). While the accelerators for lattice based crypto are partially integrated into the RISC-V core and are partially loosely coupled to the bus of the PULPino platform, mainly due to their larger flexibility if not being tightly coupled, the accelerators for SIKE and HQC are standalone accelerators which are connected via the bus of the PULPino platform. An overview of the platform is visible in Figure 2. Another goal for this work is to protect parts of the platform and the accelerators against side-channel attacks, as discussed in [2]. This includes hardware extensions to convert boolean to arithmetic masking and vice versa. This tapeout furthermore focuses on high performance applications, which also caused the need for a smaller node, leading to the choice of Globalfoundries‘ 22nm. This design achieves a reduction factor in the cycle count between 10.05 (Decapsulation) and 12.69 (Key Generation) for unprotected Kyber-1024 running on the same platform, while running at an unchanged frequency of 500 MHz.

Accelerators for Post-Quantum Signatures (2022)

As the two previous tapeouts are focused on KEMs, the goal in this project is to evaluate how well accelerators built to accelerate KEMs can be used to accelerate signature algorithms on the same hardware. The two algorithms of interest here are the two lattice-based signature algorithms which NIST selected for standardization in July 2022, i.e. CRYSTALS-Dilithium and Falcon. As both parts of Falcon's key generation as well as signature generation require double precision floating point arithmetic, they are challenging to implement on the PULPino platform. We therefore focused on all three operations of Dilithium and on the Falcon Signature verification only [3]. The results are rather promising though a limiting factor might be memory size as both signature algorithms have rather large storage requirements compared to KEMs on the same platform. The floorplan of the design is visible in Figure 3, as one can see, the area is dominated by the memory sizes, while sufficient area for additional logic would still be available. With a target frequency of 800 MHz, this design is able run key generation, signature generation and signature verification of Dilithium-II and signature verification of Falcon faster than state-of-the-art RSA and ECDSA implementations on the Cortex M4 (compared at the same frequency of 180 MHz), proving its real world applicability.

Contribution to the Summit

Our goal is to explain how the accelerators are implemented and which challenges arose during implementation. Finally, we want to discuss the performance and energy gain in relation to the area cost.

References

[1] Tim Fritzmann, Georg Sigl, and Johanna Sepúlveda. "RISQ-V: Tightly Coupled RISC-V Accelerators for Post-Quantum Cryptography". In: *IACR Transactions on Cryptographic Hardware and Embedded Systems* 2020.4 (Aug. 2020), pp. 239–280. DOI: 10.13154/tches.v2020.i4.239-280. URL: https://tches.iacr.org/index.php/TCHES/article/view/8683.

[2] Tim Fritzmann et al. "Masked Accelerators and Instruction Set Extensions for Post-Quantum Cryptography". en. In: *IACR Transactions on Cryptographic Hardware and Embedded Systems* 2022.1 (Nov. 2021), pp. 414–460. DOI: https://doi.org/10.46586/tches.v2022.i1.414-460.

[3] Patrick Karl et al. "Post-Quantum Signatures on RISC-V with Hardware Acceleration". In: *ACM Trans. Embed. Comput. Syst.* (2023). DOI: 10.1145/3579092.