Heterogeneous bioinformatic data encryption on portable devices

Hao Chen, Lianhui Lin, Ziqiang Liao, Xiaotong Wang, Liyan Chen, Xiaogang Cui, Jie Cheng, Xing Gao, Juncong Lin

***Abstract*—With the popularity of mobile health monitoring and genome sequencing techniques, the scale of biomedical and genomic data grow rapidly, with their privacy receiving more and more concerns. Encryption technique plays an important role in many aspects of security guarantee for these data. For mobile devices, encryption is even more important yet more challenging, as these devices are usually used in environments that may not be well protected and have rather limited computing resources. With heterogeneous multi-core processors becoming popular on mobile devices to satisfy different needs of applications, designing heterogeneous algorithms to harness all the available resources are tricky but have the potential to deliver high performance. In this research, we study how to design the heterogeneous version for AES algorithm, a representative encryption algorithm, on such processor to improve throughput and energy efficiency. To alleviate the overhead, we proposed a hybrid strategy to firstly find optimal workload allocations for cores of the processor in the offline stage and then dynamically adjust the balance in the online stage to match the running environment. We do a series of experiments on common genome data, with results showing $25\% - 400\%$ improvements in throughput, and $5.5\% - 2800\%$ improvements in energy efficiency.**

***Index Terms*—Genome data encryption, MPSoC, Heterogeneous algorithms.**

I. Introduction

THE rapid development of sequencing techniques not only drastically decreases the cost, but also makes sequencing become more and more convenient, especially with portal devices such as the MinION device [16]. As a consequence, the amount of genome data start to sprint. Meanwhile, with the advances in mobile health technologies (including remote monitoring, wearable devices, and embedded sensors), long-term and continuous health monitoring is enabled, leading to high volume of biomedical data. With these data used more and more in scientific research, healthcare, legal and forensic, and even direct-to-consumer, privacy concerns are raised naturally on the improper management and abuse of them. Although segregating the data, either centralized or de-centralized, to limit access is a commonly adopted and effective strategy, it also deteriorates value of the data. Many efforts [8], [10], [33] have been devoted to the retrieval of sensitive data while maintaining individual privacy. No matter how, encryption algorithm plays an important role in the storage, transmission and retrieval of such data.

While mobile devices have enabled data collection and analysis to be performed in the field or the clinic under various contexts, they are also easier to be attacked as the operating environment are usually not so securely protected. Thus, efficient encryption on these devices is even more critical either for temporary storage, transmission and even retrieval. However, encryption of data on these devices in time is not so easy as they usually have limited computing resources. Some acceleration algorithms for Field-Programmable Gate Arrays (FPGAs) have been proposed and achieved impressible progress [24], [31]. However, FPGAs are currently not yet suitable to be used directly by an audience untrained in circuit design though many efforts have been made to improve the ease of programming. Besides, the development cycles of FPGAs are much longer and require logic design expertise, when compared with CPUs and GPUs. The benefits are further reduced considering the long time cost to synthesize a design for FPGAs.

In the field of mobile/portable devices, ARM processors are commonly used today. Due to issues in thermal/power constraints, reliability issues and design complexity etc, multi-core architecture has become an irreversible trend in the development of MPSoC. Especially, heterogeneous multi-core architecture, which consists of different core types, can offer significant advantages in performance, power, area, and delay, compared with homogeneous counterpart (consisting a set of identical cores). Besides, heterogeneous multi-core are perfect fit for the dark silicon regime (the increasing power density on chip prevents all the cores to be switched on at the same time) as only the cores suited for an application need to be switched on. The heterogeneity of MPSoC can be either in performance or in function. Taking the Samsung Exynos 7420 MPSoC as an example, which is designed with the ARM big.LITTLE architecture (a classical heterogeneous architecture in performance), the MPSoC integrates four high-performance out-of-order ARM Cortex-A57 cores (big cores) and four low-power in-order ARM Cortex-A53 cores (little cores). It also has a GPU, Mali-T760 MP8, which is functionally different from the CPUs (see Figure 1 for details on the architecture of MPSoC).

These mobile processors have the potential to deliver high-performance if all the available resources can be harnessed by the software, in a way to use the cores that are most power efficient for the current computing need without negatively impacting the performance. However, designing algorithms to run efficiently on such heterogeneous MPSoC is challenging: First of all, we need to select appropriate core type according to the affinity between cores and tasks; Besides, data sharing overhead need to be alleviate as much as possible, otherwise it will easily compensate the gaining from collaboration of different tasks. Different cores prefer

0000–0000/00$00.00 © 2021 IEEE