



Technology Transfer in Computing Systems

D3.16: Individual TTP16 abstract

Project no.: 609491
Funding scheme: Collaborative project
Start date of the project: 1st September 2013
Duration: 36 months
Work programme topic: FP7-ICT-2013-10

Deliverable type: Report
Deliverable reference number: ICT-609491 / D3.16
WP and tasks contributing: WP 3 / all
Due date: 31/10/2015
Actual submission date: 04/02/2016

Responsible Organization: IMPERIAL
Dissemination Level: Public
Revision: 1.0



TETRACOM D3.16: Benchmarking Short Read Mapping Platforms

James Arram, Wayne Luk (Imperial College), Vlad-Mihai Sima, Zaid Al-Ars, Koen Bertels, (Bluebee)

FPGA acceleration of short read alignment

One of the key challenges facing genomics today is how to efficiently analyse the massive amounts of data produced by next-generation sequencing (NGS) platforms. With general-purpose computing systems struggling to meet the performance demands, an increasing amount of interest has been placed in developing customised hardware architectures to accelerate the analysis applications. Reconfigurable hardware, such as the Field-Programmable Gate Array (FPGA), is a promising candidate for accelerating these applications: the multiple levels of exploitable parallelism can provide substantial speed-up, whilst the low operational clock frequencies allow reduced power consumption, and high rack unit densities.

The transfer of technology between Bluebee and Imperial College London focuses on developing an FPGA-based short read alignment design able to achieve substantial speed-up compared to the software tools currently available. This application is the bottleneck of most sequence analysis pipelines, and involves mapping the short sequences of DNA produced by an NGS platform to positions in a reference genome. The project consists of two components: 1) developing benchmarking facilities for measuring the performance of short read alignment on multiple platforms, and 2) developing FPGA-based short read alignment designs and evaluating them using the aforementioned benchmarking facilities.

Hardware components transferred

1. An exact alignment hardware design based on the FM-index. An RTL version of the exact alignment design is created using Vivado HLS. This RTL module is generic enough to be used in most hardware platforms, for example the Convey Computing platform.
2. New hardware optimisations are developed which improve the alignment performance. In particular, the n-step FM-index, and index oversampling are included to reduce the number of off-chip memory accesses required to perform the mapping.
3. Equations are derived to estimate the upper bound design performance of the exact alignment design for different alignment data sets and hardware platforms. These equations can be used to verify the performance of implementations and in benchmarking.
4. Performance benchmarking of the exact alignment design with comparisons to the fastest software alignment tools currently available.

Impact of Hardware transferred

1. Performance estimations indicate that the exact alignment design, run on a Convey HC2-ex system, can be faster than the software tools Soap2, BWA, Bowtie, and Soap3-dp. In particular, the design is up to 91 times faster than Soap2, and up to 34 times faster than Soap3-dp.
2. Our work was presented in the OPENPOWER PERSONALIZED MEDICINE WORKSHOP, which aims to bring together technology providers and consolidate the efforts to create effective genomics computational solutions.