

عنوان مقاله:

A Novel ۲-D BWA-MEM FPGA Accelerator for Short-Read Mapping of the Whole Human Genome

محل انتشار:

مجله پژوهش های کاربردی در مهندسی برق، دوره 1، شماره 2 (سال: 1401)

تعداد صفحات اصل مقاله: 8

نویسندگان:

Mahdi Taheri - *Reliable and Smart Systems Lab (RSS), Shahid Bahonar University of Kerman, Kerman ۷۶۱۶۹۱۳۴۳۹, Iran*

Ali Mahani - *Reliable and Smart Systems Lab (RSS), Shahid Bahonar University of Kerman, Kerman ۷۶۱۶۹۱۳۴۳۹, Iran*

خلاصه مقاله:

The mapping of DNA subsequences to a known reference genome, referred to as “short-read mapping”, is essential for next-generation sequencing. Hundreds of millions of short reads need to be aligned to a tremendously long reference sequence, making short-read mapping very time consuming. Day by day progress in Next-Generation Sequencing (NGS) is enabling the generation of DNA sequence data at ever faster rates and at low cost, which means a dramatic increase in the amounts of data being sequenced; nowadays, sequencing nearly ۲۰ billion reads (short DNA fragments) costs about ۱۰۰۰ dollars per human genome and sequencers can generate ۶ Terabases of data in less than two days. This article considered the seed extension kernel of the Burrows-Wheeler Alignment (BWA) genomic mapping algorithm for accelerating with FPGA devices. We have proposed an FPGA-based accelerated implementation for the seed extension kernel. The Smith-Waterman algorithm is used during the seed extension to find the optimum alignment between two sequences. The state-of-the-art architectures use ۱D-systolic arrays to fill a similarity matrix, based on the best score out of all match combinations, mismatches and gaps are computed. The cells on the same anti-diagonal are calculated in parallel in these architectures. We propose a novel ۲-dimensional architecture. Our new modified algorithm is based on two editing and calculating phases. In each step of calculation, some errors may occur in which all the cells on the same row and the same column are computed in parallel and, thereby, significantly speed up the process. Needless to say, these probable errors will be omitted before the next step of calculation begin. Our simulation results show that the proposed architecture can work with up to ۳۱۲ MHz frequency in Synopsys Design-Compiler for ۱۸۰-nm CMOS technology and be up to ۵۷۰x and ۱.۴x faster than the software execution and the ۱D-systolic arrays, respectively.

کلمات کلیدی:

Bioinformatics, FPGA, Smith-Waterman

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1862726>



