UNIVERSITI TEKNOLOGI MARA

DESIGN AND DEVELOPMENT OF HIGH PERFORMANCE SWA CELL DESIGN FOR LOCAL DNA SEQUENCE ALIGNMENT

SYED ABDUL MUTALIB AL JUNID

PhD

7th March 2017

Internal Examiner — External Examiner — Supervisor — Chairman — Minute Taker

ABSTRACT

DNA sequence alignment is expected to help in revealing important information related to the human body, disease, genetic and other biological upon discovery of the sequence alignment. Moreover, there have been intensive efforts in improving the performance of the sequence alignment process via hardware-based acceleration using the Field Programmable Gate Array (FPGA). This implementation is becoming popular due to the flexibility of the acceleration design, the ability to reduce the execution cycle, parallel computational solutions, and the ability to increase the performance of alignment at the same time. The performance of the DNA sequence alignment system strongly depends on the algorithm, design architecture and accelerator performance. This study proposed three new DNA sequence alignment accelerator system cell design and architecture based on the Smith-Waterman Algorithm (SWA) named as the new Optimized SWA Linear Gap Penalty (OSL), Optimized SWA Affine Gap Penalty (OSA), and Optimized Recursive Variable Expansion SWA Linear Gap Penalty (ORSL). The OSL, OSA and ORSL proposed a new optimized two bit data representation format for DNA sequence characters, which is previously designed based on the eight bit ASCII 2 characters format. Moreover, the OSL, OSA and ORSL cell designs reduce the complexity of the DNA sequence alignment by rescheduling the process of alignment via a new parallel arrangement of the processing or computational element. On top of that, minimal stages of computation cycle were proposed and introduced in OSL cell design, with a reduction of 25% compared to the previous SWA linear gap penalty design. Next, the new OSL, OSA and ORSL also proposed a new optimized systolic array-based architecture for optimizing and accelerating the performance of the SWA cell design. In this study, the As Soon As Possible (ASAP) optimization has been adopted in the OSA, OSL and ORSL cell designs, which resulted in reducing the stages of the computational cycle. Moreover, the minimal size of the two bit data representation characters format used in the design has resulted in minimizing the temporary memory requirement in the cell design up to 75%. The proposed OSL, OSA and ORSL design code was developed, simulated and implemented using commercial design software, the Altera Quartus II version 12.1 and the NIOS II IDE. The design was targeted to the Altera FPGA platforms, specifically the Altera DE2-115. The performance of the system was characterized by referring to the logic function of the design or known as Logic Element (LE) and Cell Update (CU). The performance of OSL was improved as compared to the previous design with 3.2 GCUPS for single PE and 25.6 GCUPS for a complete architecture. While, the OSA was improved by 9.22% as compared to the previous design with 32.0 GCUPS. The ORSL has shown superior results compared to Recursive Variable Expansion (RVE) and Extended Recursive Variable Expansion (ERVE) techniques with 6.4 GCUPS. Again, the size is slightly bigger since the main target of the proposed design is to improve the design performance. Result attained has proven that, the proposed system produces better performance and design over previous work in accelerating the SWA DNA sequence alignment.