

Analisis Performa Algoritma Needleman-Wunsch (NW) Sekuensial Pada Raspberry Pi

Ramdan Sotra¹, Muhammad Fuad², Hedi Hestriyandi³, Mohamad Iqbal⁴, Muhammad Iqbal⁵
Universitas Muslim Indonesia¹, Universitas Seiwijaya², Politeknik Negeri Sambas³, Universitas Pakuan⁴, Teknorat⁵
Indonesia

ramdanstr@gmail.com¹, fuadphysics@gmail.com², HediHedi@gmail.com³, mohamad.iqbal@unpak.ac.id⁴, m.iqbal@teknokrat.ac.id⁵

Abstrak - Algoritma Needleman-Wunsch (NW) merupakan metode untuk menjajarkan DNA. Penjajaran DNA merupakan hal yang sangat penting untuk penelitian di bidang Biologi, diantara penamfaatan penjajaran DNA adalah untuk melihat kemiripan sifat-sifat sepasang makhluk hidup. Algoritma NW menggunakan pendekatan Dynamic Programming yang memiliki kompleksitas waktu komputasi (n^2). Pengujian performa komputasi algoritma ini telah banyak dilakukan pada mesin PC (Personal Computer) dan belum ada yang melakukan untuk mesin Raspberry Pi. Penelitian ini memperlihatkan kemampuan Raspberry Pi dengan spesifikasi prosesor ARM dan memory RAM 512 MB telah berhasil memproses penjajaran dua pasang DNA dengan panjang 11500 bp.

Keywords: *Needleman-Wunsch, Dynamic Programming, komputasi, DNA.*

I. PENDAHULUAN

Bioinformatika adalah penerapan teknologi komputer untuk manajemen informasi biologi. Komputer digunakan untuk mengumpulkan, menyimpan, menganalisis dan mengintegrasikan informasi biologi. Pada bioinformatika ada masalah yang sangat mendasar dan penting yaitu masalah penjajaran sekuen (*sequence alignment*). *Sequence alignment* dapat didefinisikan sebagai masalah untuk menemukan bagian mana dari *sequence* yang sama dan bagian mana yang berbeda. Manfaat *sequence alignment* diantaranya adalah menemukan kemiripan sifat sifat genetik dari organisme. Ada dua metode yang digunakan untuk *sequence alignment* yaitu *global alignment* menggunakan algoritma Needleman-Wunsch dan *local alignment*

menggunakan algoritma Smith Waterman.

Penerapan algoritma pada Needleman-Wunsch telah banyak diterapkan pada Personal Computer (PC), namun algoritma ini belum diterapkan pada Raspberry Pi. Raspberry Pi merupakan sebuah komputer papan tunggal (*single board computer*) yang dikembangkan pertama kali di Inggris pada tanggal 29 Februari 2012 oleh yayasan raspberry pi dengan prosesor ARM1176JZF-S, dilengkapi dengan VideoCore IV GPU dan memory 512 MB. Raspberry Pi memiliki ukuran fisik yang kecil mendukung system operasi linux dan dilengkapi SDCard sebagai tempat penyimpanan.

Penelitian mengenai Raspberry Pi masih sangat jarang, oleh sebab itu pada penelitian ini akan dilakukan pengujian performa dengan melihat komputasi Algoritma Needleman-Wunsch pada Raspberry Pi, kemudian dibandingkan dengan komputasi algoritma Needleman-Wunsch pada PC.

II. LANDASAN TEORI

Algoritme Needleman-Wunsch

Algoritma Needleman-Wunsch ini pertama kali diusulkan oleh Saul B. Needleman dan Christian D. Wunsch pada tahun 1970, Algoritma Needleman-Wunsch melakukan *sequence alignment* global, yaitu melihat total global *sequence alignment* yang memiliki tingkat kemiripan yang tinggi dari global penjajaran sekuen.

Tahapan algoritma Needleman-Wunsch adalah

1. Inisialisasi matriks

Pada tahap ini *pairwise sequence* dibuat

menjadi matriks $M \times N$, dimana M adalah sekuens pertama dan N adalah sekuens kedua, sebagai contoh, ada sepasang sekuens, dimana **ATGCGC** adalah *sequence 1* atau M dan **AGACT** adalah *sequence 2* atau N , maka kedua sekuens akan menjadi matriks $[m,n]$ atau kita namakan matriks $A_{m,n}$, seperti Gambar 1.

	A	T	G	C	G	C
A	0	0	0	0	0	0
G	0					
A	0					
C	0					
T	0					

Gambar 1 Matriks $A_{m,n}$

Variabel yang digunakan :

- m menjelaskan baris dan n menjelaskan kolom.
- A nilai matriks untuk setiap *cell* dinyatakan sebagai $(A_{m,n})$
- S adalah skor setiap *cell* dinyatakan sebagai $(S_{m,n})$
- W sebagai *gap alignment*

2. Pengisian nilai tiap *cell* matriks (*matrix filling*)

Pengisian *cell* matriks untuk matriks $A_{m,n}$, didapatkan dengan menggunakan rumus dibawah ini :

$$A_{m,n} = \text{Maksimum} [A_{m-1,n-1} + S_{m,n}, A_{m,n} + W, A_{m-1,n} + W + 0]$$

Untuk penentuan skor tidak ada ketentuan khusus, akan tetapi bebas memberikan nilai, misalnya dalam kasus ini, untuk elemen sekuens yang sesuai (*matching*) diberikan nilai $+5$, tidak sesuai (*mismatch*) diberikan nilai -3 dan *gap* dikenakan penalti -4 . Untuk mengisi *cell* matriks $A_{1,1}$, digunakan rumus dibawah ini :

$$\begin{aligned} A_{1,1} &= \text{Maksimum} [A_{0,0} + S_{1,1}, A_{1,1} + W, A_{0,1} + W + 0] \\ &= \text{Maksimum} [0 + (5), 0 + (-3), 0 + (-3) + 0] \\ &= \text{Maksimum} [5, -3, -3, 0] \end{aligned}$$

$$A_{1,1} = 5$$

Dan apabila diteruskan, maka akan terisi semua *cell* matriks $A_{m,n}$, seperti pada Gambar 2.

	A	T	G	C	G	C
A	0	5	2	-1	-3	-3
G	0	2	2	7	4	2
A	0	5	2	4	4	1
C	0	2	2	1	8	8
T	0	-1	7	4	6	3

Gambar 2 Pengisian *cell* matriks $A_{m,n}$

3. Pelacakan kembali (*trace back*)

Pelacakan kembali (*trace back*) merupakan tahap terakhir dari algoritma Needleman-Wunsch, pada tahap ini mencari semua kemungkinan jalur dalam global penjajaran sekuens dapat dilihat pada Gambar 3.

	A	T	G	C	G	C
A	0	5	2	-1	-3	-3
G	0	2	2	7	4	2
A	0	5	2	4	4	1
C	0	2	2	1	8	8
T	0	-1	7	4	6	3

Gambar 3 Jalur global yang mungkin pada matriks $A_{m,n}$

Setelah mendapatkan beberapa jalur global maka dipilih maka dipilih jalur global yang paling maksimal nilainya, untuk kasus ini didapatkan dua jalur global yang maksimal nilainya seperti ditunjukkan pada Gambar 4.

	A	T	G	C	G	C
A	0	5	2	-1	-3	-3
G	0	2	2	7	4	2
A	0	5	2	4	4	1
C	0	2	2	1	8	8
T	0	-1	7	4	6	3

Gambar 4 Jalur global *sequence alignment* yang optimal

Maka didapatkan hasil global *sequence alignment* pada Gambar 5.

A	T	G	-	C	G	C	A	T	-	C	G	C
A	-	G	A	C	-	T	A	-	G	A	C	T
5	-3	5	-3	5	-3	-3	5	-3	5	-3	5	-3

Gambar 5 Hasil global *sequence alignment*

Raspberry Pi

Raspberry Pi adalah sebuah komputer papan tunggal (*single board computer*) yang

dikembangkan pertama kali di Inggris oleh yayasan Raspberry Pi dengan tujuan mempromosikan pengajaran ilmu komputer dasar. Raspberry Pi dilengkapi dengan chip Broadcom BCM2835, mencakup prosesor ARM1176JZF-S 700 Mhz, untuk pengolahan grafis didalamnya terdapat VideoCore IV GPU serta dilengkapi memory 512 MB. Built-in hard disk atau *solid-state drive* bersifat optional tetapi dapat menggunakan memory card untuk *booting* dan penyimpanan data. Raspberry pi sepenuhnya dibangun dengan lisensi *Open Source* dan menggunakan sistem operasi linux. Perkembangan komputer mini *raspberry pi* telah membuka peluang besar pada sistem komputasi untuk diterapkan pada sejumlah area penelitian. Teknologi raspberry berpotensi dalam memberikan solusi komputasi di negara berkembang.



Gambar 6 Struktur Raspberry Pi

METODE PENELITIAN

Penelitian secara umum dilakukan dengan tiga tahapan, yaitu 1) Desain aplikasi sekuensial *sequence alignment* dengan algoritma Needleman-Wunsch menggunakan bahasa pemrograman C, 2) Menjalankan aplikasi sekuensial dan melakukan 5 kali percobaan untuk mengetahui rata-rata waktu komputasi aplikasi sekuensial pada Raspberry Pi dan PC, dan 3) menganalisis performa aplikasi sekuensial. Waktu sebagai parameter untuk mengukur performa algoritma Needleman-Wunsch sekuensial pada Raspberry Pi dan PC.

HASIL PENELITIAN

Penelitian ini melakukan peninjauan sekuen (*sequence alignment*) dengan panjang

sekuen DNA berbeda-beda. Panjang sekuennya yaitu 1000, 2000, 4000, 8000, 10000, 11000, 11500 dan 12000 karakter sekuen DNA. Kami melakukan 5 kali percobaan untuk masing-masing panjang sekuen DNA pada Raspberry Pi dan PC. Hasil komputasi yang didapatkan untuk *sequence alignment* DNA dengan algoritma Needleman-Wunsch menggunakan pendekatan Dynamic Programming ditampilkan pada Tabel 1, Tabel 2, Tabel 3 dan Gambar 7.

Tabel 1 Hasil komputasi *sequence alignment* DNA pada Raspberry Pi

Panjang Sekuen DNA	Percobaan dan Waktu (second)				
	1	2	3	4	5
1000	0,27	0,324	0,324	0,335	0,320
2000	1,08	1,096	1,106	1,083	1,098
4000	4,35	4,348	4,345	4,354	4,353
8000	17,6	17,682	17,724	17,698	17,67
10000	27,7	27,723	27,708	27,695	27,70
11000	38,3	37,490	34,373	36,150	34,01
11500	44,4	42,416	40,206	40,580	40,46
12000	Failed	Failed	Failed	Failed	Failed

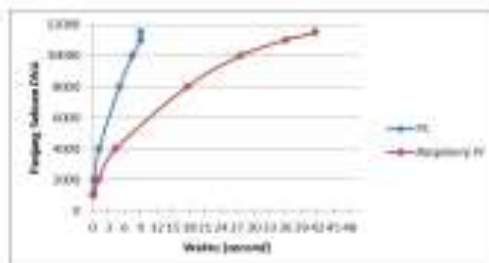
Tabel 2 Hasil komputasi *sequence alignment* DNA pada PC

Panjang Sekuen DNA	Percobaan dan Waktu (second)				
	1	2	3	4	5
1000	0,07	0,081	0,067	0,079	0,085
2000	0,29	0,296	0,309	0,312	0,298
4000	1,21	1,195	1,27	1,171	1,216
8000	5,4	4,717	5,52	4,824	5,118

	11				
	3				
	7,				
	37				
10000	1	7,25	7,509	7,535	7,932
	9,				
	64				
11000	7	9,098	8,805	8,836	8,818
	9,				
	16				
11500	6	8,92	8,877	9,19	9,193
	10				
	,5				
12000	7	10,81	10,43	10,8	11,47

Tabel 3 Rata-rata komputasi *sequence alignment* DNA pada Raspberry Pi dan PC

Panjang Sekuen DNA	Rata-rata komputasi dalam <i>second</i>		Perbandingan (Rasio)
	Raspberry Pi	PC	
1000	0,315	0,076	4,120
2000	1,093	0,303	3,613
4000	4,350	1,213	3,586
8000	17,693	5,058	3,498
10000	27,711	7,519	3,685
11000	36,078	9,041	3,990
11500	41,612	9,069	4,588
12000	Failed	10,81	-
Rata-rata perbandingan kecepatan			3,869



Gambar 7 Grafik perbandingan komputasi *sequence alignment* pada Raspberry Pi dan PC

Hasil percobaan diatas menunjukkan bahwa

waktu komputasi *sequence alignment* DNA pada PC lebih cepat dibandingkan dengan Raspberry Pi. Dimana rata-rata perbandingan kecepatan antara PC terhadap Raspberry Pi adalah $3,86 \approx 4$ kali lebih cepat PC terhadap Raspberry Pi. Jika melihat perbandingan spesifikasi maka untuk komputasi *sequence alignment* DNA pada Raspberry Pi masih dapat dikatakan baik. Raspberry Pi dapat melakukan penjajaran sekuen DNA (*sequence alignment* DNA) dengan panjang sekuen DNA hingga 11500 karakter. Perhatikan Tabel 4 dengan perbandingan yang dilihat secara umum yaitu 8 : 1 sementara hasil proses komputasi yang dilakukan 4 : 1. Maka Raspberry Pi dapat dikatakan baik melakukan komputasi *sequence alignment* DNA untuk jumlah sekuen maksimal 11500 karakter.

Tabel 4 Rincian perbandingan secara umum Raspberry Pi dan PC

Spesifikasi	Raspberry Pi	PC	Perbandingan
Kecepatan processor	700 MHz x 1 core	1,6 GHz x 2 core	8 : 1
Ukuran RAM	512 MB	4 GB	8 : 1
Harga	Rp 500.000	Rp 4.000.000	8 : 1
Rata-rata perbandingan			8 : 1

IV. KESIMPULAN

Berdasarkan hasil dari perbandingan komputasi dan spesifikasi secara umum antara Raspberry Pi dan PC, maka dapat disimpulkan :

1. Rata-rata perbandingan kecepatan antara Raspberry Pi terhadap PC adalah $3,86 \approx 4$ kali lebih cepat PC terhadap Raspberry Pi.
2. Perbandingan spesifikasi umum PC terhadap Raspberry Pi adalah PC 8 kali lebih baik dibanding Raspberry Pi (8:1). Namun perbandingan komputasi *sequence alignment* DNA pada PC terhadap Raspberry Pi adalah PC 4 kali lebih cepat dibanding Raspberry Pi (4:1). Komputasi *sequence alignment* DNA lebih cepat 4 kali dari yang seharusnya (8:1) apabila dilihat dari spesifikasi umum

Raspberry Pi dengan PC. Oleh sebab itu Raspberry Pi dapat dikatakan baik performanya untuk komputasi *sequence alignment* DNA.

3. Waktu komputasi *sequence alignment* DNA menggunakan Raspberry Pi tidak begitu berbeda signifikan dibandingkan PC untuk panjang sekuen DNA antara 1000 – 11500 karakter.
4. Perbandingan biaya untuk melakukan *sequence alignment* DNA lebih murah menggunakan Raspberry Pi dibandingkan dengan PC. Karena biaya perangkat keras Raspberry Pi lebih murah dibandingkan PC.

DAFTAR PUSTAKA

- [1] Manoharan A, Kanagavel B, Muthuchidambaram A, Kumaravel JPS. 2011. Bioinformatics Research – an Informetric View. *International Conference on Information Communication and Management IPCSIT.16* (2011). IACSIT Press, Singapore.
- [2] Junior SAC. 2003. Sequence Alignment Algorithms. *Department of Computer Science School of Physical Sciences & Engineering King's College London*.
- [3] Huang KF, Yang CB, Tseng KT. 2002. An Efficient Algorithm For Multiple Sequence Alignment. *Proc. of the 19th Workshop on Combinatorial Mathematics and Computation Theory*.
- [4] Global alignment of two sequences- Needleman-Wunsch Algorithm 2013, [internet]. <http://jdiunihub2013.oktober12j.tersedia.pada> <http://amrita.vlab.co.in/?sub=3&brch=274&sim=1431&cnt=1>.
- [5] Charter K, Schaeer J, Szafron D. 2000. Sequence Alignment using FastLSA. *The Department of Computing Science at the University of Alberta Canada*.
- [6] Khajeh-Saeed A, Poole S, Perot JB. 2010. Acceleration of the Smith-Waterman algorithm using single and multiple graphics processors. *Journal of Computational Physics*. 229 (2010) 4247–4258. doi:10.1016/j.jcp.2010.02.009.
- [7] Siriwardena TRP, Ranasinghe DN. 2010. Global Sequence Alignment using CUDA compatible multi-core GPU. *IEEE Journal*.

Page 201 – 206. ISBN 978-1-4244-8549-9. doi : 10.1109/ICIAFS.2010.5715660.

- [8] <http://www.raspberrypi.org/> akses 9 Desember 2013.
- [9] Ali, Murat. *et al.* 2013. Technical Development and Socioeconomic implications of the Raspberry Pi as a Learning Tool in Developing countries. *5th Computer Science and Electronic Engineering Conference (CEECE)*. Hal 103-108. *Proceedings Of The IEEE*, 2013.