

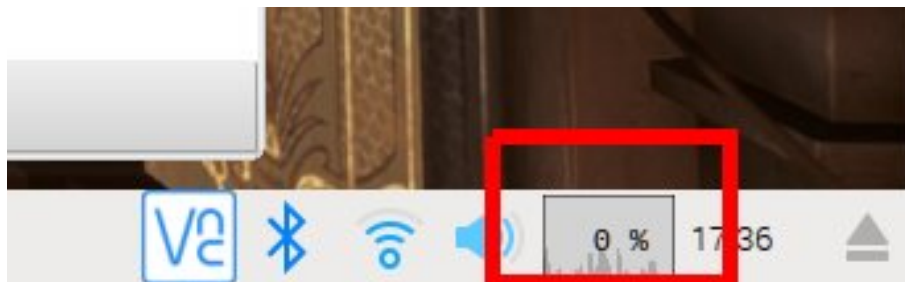
ภาคผนวก M

การทดลองที่ 13 การพัฒนาอัลกอริธึมแบบขนาน ด้วยไลบรารี OpenMP

การพัฒนาอัลกอริธึมแบบขนานบนซีพียูชนิดมัลติคอร์ในปัจจุบันจำเป็นต้องอาศัยภาษาคอมพิวเตอร์ระดับสูง เช่น ภาษา C/C++ ภาษา Java เป็นต้น เพื่อช่วยลดเวลารัน (Run Time) ซึ่งเท่ากับเร่งความเร็ว (Speedup) ให้ อัลกอริธึมหรือโปรแกรม โดยการสร้างเธรดผู้ช่วย (Worker Thread) และมอบหมายงานให้ไปรันบนแกนประมวลผลที่ยังว่างอยู่ ผู้อ่านสามารถประยุกต์ใช้หลักการนี้บนเครื่องคอมพิวเตอร์ทั่วไปจนถึงเครื่องซูเปอร์คอมพิวเตอร์ ตามเนื้อหาในบทที่ 8 ดังนั้น การทดลองมีวัตถุประสงค์ดังนี้

- เพื่อพัฒนาโปรแกรมภาษา C ด้วยไลบรารี OpenMP ให้สามารถทำงานแบบมัลติเธรดและใช้งานซีพียูชนิดมัลติคอร์ได้เต็มที่
- เพื่อเรียนรู้การวัด CPU Utilization (%CPU) เวลาจริง (T_{real}) เวลาผู้ใช้ (T_{user}) และเวลาระบบ (T_{sys}) ในซีพียูชนิดมัลติคอร์
- เพื่อทำความเข้าใจการวัดประสิทธิภาพของอัลกอริธึมแบบขนานจากการประเมินความซับซ้อนเชิงเวลา ด้วยพีชคณิต BigO ที่มา: Rosen (2002) และตัวชี้วัด Speedup ที่มา: Patterson and Hennessy (2016) จากเวลาที่วัดได้

M.1 การวัด CPU Utilization



รูปที่ M.1: กราฟแสดงการใช้งานซีพียู (CPU Usage Monitor) ย้อนหลังและค่าสรุป ณ เวลาปัจจุบัน ที่มา: abload.de

ผู้อ่านสามารถติดตั้งเครื่องมือและกราฟในรูปที่ M.1 แสดงการใช้งานซีพียู (CPU Usage Monitor) ย้อนหลังและค่าสรุป ณ เวลาปัจจุบันของบอร์ด Pi ประกอบการทดลองที่ 13 ตามขั้นตอนเหล่านี้

1. เลื่อนเมาส์ไปบนตำแหน่งว่างของ Task Bar
2. คลิกขวา เพื่อให้เมนูต่อไปนี้ปรากฏขึ้นแล้วคลิกซ้ายเลือก Add/Remove Panel Items
3. คลิกที่แท็บ Panel Applets
4. เลื่อนรายการขึ้นลงเพื่อหารายการชื่อ CPU Usage Monitor แล้วคลิก Add
5. กดปุ่ม Up และ Down เพื่อวางตำแหน่งของ CPU Usage Monitor ในตำแหน่งที่ต้องการ โปรดสังเกตรายชื่อ เมื่อได้ตำแหน่งที่ต้องการแล้วกด Close หมายเหตุ **Spacer** หมายถึง ช่องว่างที่คั่นระหว่าง Applet ที่อยู่บน Task Bar
6. สังเกตด้านขวาของ Task Bar จะมีจอสีเทาขนาดเล็กแสดงเป็นกราฟแท่ง โดยแท่งขวาสุดคือ วินาทีล่าสุด
7. เลื่อนเมาส์ไปบนกราฟแล้วคลิกขวาเพื่อเพิ่มการแสดงผลเป็นตัวเลขหน่วยเป็นเปอร์เซ็นต์ (%)
8. ทดสอบการทำงานโดยการเปิดคลิปเดียวกันบน [YouTube.com](https://www.youtube.com) ที่ความละเอียดแตกต่างกัน เช่น 240p, 480p และ 720p ที่ละค่าเพื่อให้เห็นค่า $\%CPU_{max}$ ที่แตกต่าง

M.2 การคูณเมทริกซ์แบบขนาน

$$C = A \times B \quad (M.1)$$

การคูณเมทริกซ์เป็นพื้นฐานของการคำนวณพื้นฐานทางวิทยาศาสตร์ และวิศวกรรมศาสตร์ กำหนดให้เมทริกซ์จัตุรัส A ขนาด $N \times N$ สามารถเขียนในรูปแบบของอาร์เรย์ 2 มิติในภาษา C/C++ ได้ดังนี้

$$A = (A[i][j])$$

โดยดัชนีตัวแรก i คือ หมายเลขแถว มีค่าตั้งแต่ 0 ถึง $N-1$ ดัชนีตัวที่สอง j คือ หมายเลขคอลัมน์ มีค่าตั้งแต่ 0 ถึง $N-1$ ดังนั้น

$$A = \begin{pmatrix} A[0][0] & A[0][1] & \dots & A[0][N-1] \\ A[1][0] & A[1][1] & \dots & A[1][N-1] \\ \vdots & \vdots & \ddots & \vdots \\ A[N-1][0] & A[N-1][1] & \dots & A[N-1][N-1] \end{pmatrix}$$

เมื่อทำความเข้าใจพื้นฐานของเมทริกซ์ในรูปแบบของอาร์เรย์ 2 มิติแล้ว ผู้อ่านสามารถทำการทดลองตามขั้นตอนต่อไปนี้

1. สร้างไดเรกทอรี `/home/pi/asm/Lab13` บนโปรแกรม Terminal ด้วยคำสั่งต่อไปนี้ตามลำดับ

```
$ cd /home/pi/asm
$ mkdir Lab13
```

```
$ cd Lab13
$ nano multMatrix.c
```

2. กรอกโปรแกรมต่อไปนี้ด้วยโปรแกรม nano และบันทึกในไฟล์ชื่อ **multMatrix.c** ในไดเรกทอรีที่สร้างไว้

```
#include <stdio.h>
#include <stdlib.h>
#include <time.h>
#include <omp.h>

#define N 200
float A[N][N], B[N][N], C[N][N]; // matrices of NxN elements

int main () {
/* DECLARING VARIABLES */
int i, j, k; // indices for matrix multiplication
double t_mul; // Multiply time
double start, stop; // start time and stop time

/* FILLING MATRICES WITH RANDOM NUMBERS */
srand ( time(NULL) );
for(i=0;i<N;i++) {
    for(j=0;j<N;j++) {
        A[i][j]= rand();
        B[i][j]= rand();
    }
}

/* MATRIX MULTIPLICATION */
printf("Max number of threads: %i \n",omp_get_max_threads());
#pragma omp parallel
    printf("Number of threads: %i \n",omp_get_num_threads());
    start=omp_get_wtime(); // time measure: start time
    #pragma omp parallel for private(k, j)
        for(i=0;i<N;i++) {
            for(j=0;j<N;j++) {
                C[i][j]=0; // set initial value of resulting matrix C = 0
                for(k=0;k<N;k++) {
                    C[i][j]=C[i][j]+A[i][k]*B[k][j];
                }
            }
        }
}
```

```

    stop=omp_get_wtime(); // time measure: stop time
    t_mul = stop-start; // Multiply time → เวลาที่แท้จริงจริง
    printf("Mutiply Time: %2.4f \n",t_mul);

    /* TERMINATE PROGRAM */
    return 0;
}

```

3. exit ออกจากโปรแกรม nano เพื่อคอมไพล์โปรแกรมด้วยคำสั่งต่อไปนี้

```
$ gcc -fopenmp multMatrix.c -o mulMatrix
```

แก้ไขหากมีข้อผิดพลาดจนกว่าจะคอมไพล์โปรแกรมสำเร็จและมีไฟล์ชื่อ mulMatrix

4. ตั้งค่าจำนวนเธรด $n=1$ ของโปรแกรมด้วยคำสั่งต่อไปนี้

```
$ export OMP_NUM_THREADS=1
```

5. รันโปรแกรมจับเวลาด้วยคำสั่ง time ดังนี้จำนวน 5 ครั้งเพื่อหาค่าเฉลี่ย ขณะทำการทดลองขอให้ผู้อ่านใช้คลิกข้อมือจับเวลาไปพร้อมๆ กัน เพื่อเปรียบเทียบกับค่าของ $T_{mul,n}$ และ T_{real}

```
$ time ./mulMatrix
```

ซึ่งจะรายงานผลการจับเวลาการทำงานของทั้งโปรแกรมในแง่มุมต่างๆ

6. จดบันทึกค่า CPU Utilization สูงสุดหรือ $\%CPU_{max}$ ที่สังเกตได้ หาค่าเฉลี่ยของ $T_{mul,n}$ T_{real} T_{user} และ T_{sys} ที่ได้เป็นวินาทีลงในตารางที่ [M.1](#)

ตารางที่ M.1: เวลาและ $\%CPU_{max}$ ของการคูณเมทริกซ์ที่ขนาด N และจำนวนเธรดเท่ากับ 1, 2, 4, 8 เธรด

เวลาเฉลี่ย (วินาที)	$N=200$ (วินาที)	$N=400$ (วินาที)	$N=800$ (วินาที)	$N=1000$ (วินาที)
$n=1$ เธรด				
$T_{mul,1}$	0.0124	0.0862	0.7919	1.5649
T_{real}	0.014	0.091	0.804	1.593
T_{user}	0.090	0.644	6.034	12.261
T_{sys}	0.000	0.000	0.000	0.020
$\%CPU_{max}$	0%	9%	5%	9%
$n=2$ เธรด				
$T_{mul,2}$	0.0095	0.0893	0.7534	1.6506
T_{real}	0.013	0.094	0.767	1.668
T_{user}	0.041	0.682	5.934	12.948
T_{sys}	0.000	0.000	0.000	0.040
$\%CPU_{max}$	17%	9%	10%	11%
$n=4$ เธรด				
$T_{mul,4}$	0.0139	0.0956	0.8194	1.6904
T_{real}	0.019	0.101	0.831	1.709
T_{user}	0.110	0.748	6.441	13.121
T_{sys}	0.000	0.000	0.010	0.010
$\%CPU_{max}$	10%	9%	10%	11%
$n=8$ เธรด				
$T_{mul,8}$	0.0109	0.0880	0.7849	1.549
T_{real}	0.014	0.092	0.796	1.574
T_{user}	0.060	0.689	6.126	12.339
T_{sys}	0.000	0.000	0.000	0.010
$\%CPU_{max}$	10%	9%	11%	12%

7. เปลี่ยนจำนวนเธรดเท่ากับ $n=2$ เธรด ด้วยคำสั่งต่อไปนี้

```
$ export OMP_NUM_THREADS=2
```

แล้ววนกลับไปทำข้อ 5 เพื่อกำหนดค่าเฉลี่ยเวลาในตารางที่ M.1 จนครบ แล้วจึงเปลี่ยนจำนวนเธรด $n=4$ และ 8 เธรด

8. เปลี่ยนขนาดข้อมูล $N=400$ แล้วกลับไปเริ่มทำข้อ 3 จนถึงข้อ 8 จนครบ $N=800$ และ 1000

จากตารางที่ M.1 ผู้อ่านสามารถใช้ประกอบการคำนวณประสิทธิภาพการคำนวณแบบขนานในหัวข้อถัดไป

M.3 ความซับซ้อน (Complexity) ของการคำนวณ

ความซับซ้อนเชิงเวลา (Run Time Complexity) ของการคูณเมทริกซ์เท่ากับ $O(N^3)$ ในทางทฤษฎี ผู้อ่านสามารถประยุกต์ใช้อัตราส่วนระหว่างระหว่าง $O(N_2^3)$ และ $O(N_1^3)$ ที่ภาระงานขนาด $N_2:N_1$ และจำนวน n เรดเหมือนกัน เพื่อวัดความซับซ้อนของอัลกอริธึมได้ดังสมการต่อไปนี้

$$\frac{O(N_2^3)}{O(N_1^3)} = \frac{T_{mul,N_2}}{T_{mul,N_1}} \quad (M.2)$$

สำหรับการคูณเมทริกซ์ $T_{mul,N}$ คือ เป็นระยะเวลาเฉลี่ยของการคูณเมทริกซ์ขนาด $N \times N$ ด้วยจำนวน n เรด จากหัวข้อที่ผ่านมา ผู้อ่านสามารถคำนวณค่าอัตราส่วนของเวลาในตารางที่ M.2 เพื่อใช้วิเคราะห์ต่อไป

ตารางที่ M.2: อัตราส่วนเวลาการคูณเมทริกซ์ที่ขนาด N และเวลาที่ขนาด 200 ที่จำนวนเรดเท่ากับ 1, 2, 4, 8 เรด จากสมการที่ (M.2)

	$N=200$	$N=400$	$N=800$	$N=1000$
$n=1$ เรด				
$T_{N,1}/T_{200,1}$	1.00	6.9516	13.8629	126.1855
$\sqrt[2]{T_{N,1}/T_{200,1}}$	1.00	2.6366	7.9914	11.2332
$\sqrt[3]{T_{N,1}/T_{200,1}}$	1.00	1.9085	3.9971	5.0156
$n=2$ เรด				
$T_{mul,N}/T_{mul,200}$	1.00	9.2947	74.3093	173.7474
$\sqrt[2]{T_{mul,N}/T_{mul,200}}$	1.00	3.0487	8.6253	13.1813
$\sqrt[3]{T_{mul,N}/T_{mul,200}}$	1.00	2.1025	4.2964	5.5800
$n=4$ เรด				
$T_{mul,N}/T_{mul,200}$	1.00	6.9275	59.9766	122.5145
$\sqrt[2]{T_{mul,N}/T_{mul,200}}$	1.00	2.6320	7.7566	11.0686
$\sqrt[3]{T_{mul,N}/T_{mul,200}}$	1.00	1.9067	3.9013	4.9666
$n=8$ เรด				
$T_{mul,N}/T_{mul,200}$	1.00	7.7982	72.0092	142.1101
$\sqrt[2]{T_{mul,N}/T_{mul,200}}$	1.00	2.7925	8.4898	11.9210
$\sqrt[3]{T_{mul,N}/T_{mul,200}}$	1.00	1.9830	4.1609	5.2185

M.4 ประสิทธิภาพ (Performance) ของการคำนวณแบบขนาน

ผู้อ่านสามารถวัดประสิทธิภาพ (Performance) ของอัลกอริธึมใดๆ ได้จากอัตราส่วนของเวลาเดิม (T_{old}) และเวลาใหม่ (T_{new}) ที่ได้ทำการปรับปรุงอัลกอริธึมนั้นๆ ที่มา: [Patterson and Hennessy \(2016\)](#)

$$\frac{Perf_{new}}{Perf_{old}} = \frac{T_{old}}{T_{new}} \quad (M.3)$$

ดังนั้น ประสิทธิภาพของการคำนวณแบบขนานสามารถวัดได้จากอัตราส่วนระหว่างระยะเวลา $T_{alg,1}$ ของ 1 เธรดและ $T_{alg,n}$ ของ n เธรด และตั้งชื่อเรียกว่า $Speedup(n)$ ด้วยสมการต่อไปนี้

$$Speedup(n) = \frac{T_{alg,1}}{T_{alg,n}} \rightarrow \begin{matrix} \text{เวลาที่มากที่สุด} \\ \text{เวลาที่น้อยกว่า} \end{matrix} \quad (M.4)$$

โดย $T_{alg,n}$ คือ ช่วงการรันโปรแกรมอัลกอริธึมด้วยจำนวน n เธรด โดยไม่รวมช่วงเวลาอื่นๆ ซึ่งไม่ได้เกี่ยวข้องกับ การอัลกอริธึมแบบขนาน ผู้อ่านสามารถประยุกต์ตัวชี้วัดนี้กับอัลกอริธึมการคูณเมทริกซ์ ดังนี้

$$Speedup(n) = \frac{T_{mul,1}}{T_{mul,n}} \quad (M.5)$$

โดย $T_{mul,n}$ คือ ช่วงการรันโปรแกรมคำนวณเมทริกซ์จริงๆ ด้วยจำนวน n เธรด ที่ขนาด N เท่ากันโดยไม่รวมช่วง เวลาสุ่มค่าตั้งต้น และการแสดงผลอื่นๆ ผู้อ่านคำนวณค่า $Speedup(n)$ และกรอกในตารางที่ M.3 เพื่อวิเคราะห์ ผลการคำนวณที่ได้โดยตอบคำถามในกิจกรรมท้ายการทดลอง

ตารางที่ M.3: ผลการคำนวณ $Speedup(n)$ ของการคูณเมทริกซ์ที่ขนาด N และจำนวนเธรดเท่ากับ 2, 4, 8 เธรด เทียบกับ 1 เธรดด้วยสมการที่ (M.5)

Speedup	$N=200$	$N=400$	$N=800$	$N=1000$
$n=2$ เธรด $Speedup(2) = T_{mul,1}/T_{mul,2}$	1.3053	0.9762	1.0511	0.9480
$n=4$ เธรด $Speedup(4) = T_{mul,1}/T_{mul,4}$	0.8986	0.9017	0.9664	0.9255
$n=8$ เธรด $Speedup(8) = T_{mul,1}/T_{mul,8}$	1.1576	1.0141	1.0089	1.0101

M.5 กิจกรรมท้ายการทดลอง

1. เหตุใดการทดลองจึงต้องใช้การหาค่าเฉลี่ยเวลาต่างๆ

2. T_{sys} หมายถึง เวลาซีพียูทำงานประเภทไหน

3. T_{user} หมายถึง เวลาซีพียูทำงานประเภทไหน

4. T_{real} มีความสัมพันธ์กับ T_{mul} อย่างไร

5. T_{user} มีความสัมพันธ์กับ T_{mul} และจำนวนเธรด n อย่างไร

6. เมื่อ $N_1=200$ จงเปรียบเทียบค่าผลการคำนวณของ $\sqrt[3]{T_{mul,N_2}/T_{mul,200}}$ และ $\sqrt[3]{T_{mul,N_2}/T_{mul,200}}$ ที่ได้ในตารางที่ M.2 เมื่อ $N_2= 400, 800$ และ 1000 และ $n= 1, 2, 4$ และ 8 เธรดตามลำดับ ว่ามีค่าใกล้เคียงกับ $N_2/200= 2, 4, 5$ ตามลำดับอย่างไร เพราะเหตุใด

7. จำนวนเธรด และ จำนวนแกนประมวลผล มีผลต่อค่า Speedup อย่างไร วิเคราะห์ทั้งหมด 3 กรณีดังนี้

- จำนวนเธรด < จำนวนแกนประมวลผล

Speed up น้อย , มีค่า speed up < 1

- จำนวนเธรด = จำนวนแกนประมวลผล

Speed up ทั่ว , มีค่า speed up ใกล้เคียงหรือเท่ากับ 1

- จำนวนเธรด > จำนวนแกนประมวลผล

Speed up มาก , มีค่า speed up > 1

8. เหตุใดค่าเฉลี่ยเวลา T_{user} จึงไม่แตกต่างกัน ที่ N คงที่

9. เวลาส่วนใหญ่ของการรัน T_{real} T_{user} และ T_{sys} สัมพันธ์กันอย่างไร จงสร้างสมการ

10. จำนวนเธรดที่เพิ่มขึ้นทำให้การคำนวณเร็วขึ้นอย่างไร มีข้อจำกัดหรือไม่

11. ที่ขนาดข้อมูล $N=1000$ จำนวนเธรดที่เพิ่มขึ้นทำให้ T_{user} เปลี่ยนแปลงอย่างไร มีข้อจำกัดหรือไม่

12. ที่ขนาดข้อมูล N ต่างๆ ค่า $\%CPU_{max}$ มีการเปลี่ยนแปลงและมีความสัมพันธ์กับจำนวนเธรด n อย่างไร

13. ขนาดข้อมูล N ที่เพิ่มขึ้นมีผลต่อ $Speedup(n)$ ที่ $n=1, 2, 4$ และ 8 หรือไม่ อย่างไร

หมายเลข CPU : Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz

$\sqrt[3]{T_{mul,N_2}/T_{mul,200}}$ สอดคล้องใกล้เคียงกับ $\frac{N_2}{200}$
ค่าแค่ 2,4,5
เนื่องจาก N_2 เป็นค่าทวีคูณของ N_1