

1.

	Miss Rate		
	sum1	sum2	sum3
N = 64	0.25	1	0.125
N = 60	0.25	1	0.125

พิจารณา N = 64, N=60

sum1 - ทำงานแบบ row major ซึ่งจะทำให้ miss แค่ครั้งแรกรั้งเดียว ดังนั้น miss rate =  $\frac{1}{4}$

sum2 - ทำงานแบบ column major จึง miss ทุกครั้งที่เรียก ดังนั้น miss rate = 1

sum3 - ทำงานแบบ row major แต่เพราะทำทีละสองแถว นั่นคือเรียกทีละ 4 ช่องจากเดิม 1 ช่อง และจะ miss แค่ตัวแรกสุดของทุกแถว miss rate =  $\frac{1}{8}$

\* block มีขนาด 16 byte ดังนั้นเก็บ int ได้ 4 ตัวต่อ block

2.

### mm

- ลำดับรูปแบบใดที่มีค่า cycles / loop iteration สูงที่สุด
  - jki
- ลำดับรูปแบบใดที่มีสมรรถนะสูงที่สุด
  - ikj
- อธิบายเปรียบเทียบลำดับที่มีสมรรถนะสูงที่สุดกับลำดับที่มีสมรรถนะต่ำสุดว่ามีความแตกต่างกันอย่างไร
  - Loop jki fixed b[k][j] แล้วจากนั้น \* a[i][k] แล้วไล่บวกตาม column i ซึ่งการไล่ตาม column ทำให้ miss สูง
  - Loop ikj จะ fixed ตำแหน่ง a[i][k] ก่อน จากนั้น \* b[k][j] โดยบวกไล่ตามแถว j จะได้ผลลัพธ์แต่ละจุดของ c[i][j] ออกมา ด้วยความที่ไล่ตาม row ทำให้ miss ต่ำ
- ทำไมเมื่อ n มีค่ามากขึ้น ค่า cycles / loop iteration จึงมีค่าสูงขึ้น
  - เพราะค่า n ทำให้ size ของ matrix ใหญ่ขึ้น จำนวน cycle/loop จึงมากขึ้นตาม

### bmm

- ทำไมเมื่อใช้ blocking โดยให้ block size มีขนาดคงที่ (50) ค่า cycles / loop iteration สำหรับลำดับ ijk จึงไม่แปรปรวนและเพิ่มขึ้นเมื่อ n เพิ่มขึ้นเหมือนในกรณีที่ไม่ใช้ blocking
  - การใช้ blocking เป็นการคำนวณ matrix ย่อยๆ ขนาดเล็กๆ ที่ก่อนทำให้ค่าที่ได้ไม่แปรปรวน เพราะคำนวณในรูปแบบเดียวกันเสมอ ไม่ว่า n จะเปลี่ยนไปอย่างไรก็ตาม
- จากที่เราได้เรียนรู้มาว่าจำนวน miss เมื่อทำ blocking จะมีค่าประมาณ  $\frac{1}{(4B)} * n^3$  นั่นคือเมื่อขนาดของ block B มีค่ามากขึ้น จำนวน miss น่าจะมีค่าลดลง แต่ทำไมผลจากการรันจึงดูเหมือนไม่เป็นไปตามสูตรนี้
  - เนื่องจากกำหนดขนาดของ block ใกล้เคียงกับ ขนาด matrix(array) ดังนั้นเวลารันจึงเหมือนกับการรันแบบไม่ทำ block ทำให้จำนวน miss มีค่าใกล้เคียงกับเดิม

3.

	100	200	500	1000	2000	5000
Noblocking	0.005	0.018	0.257	1.285	15.833	139.351
Block = 20	0.007	0.039	0.277	1.245	4.618	32.409

สรุป การทำ blocking และไม่ทำ blocking มีผลแตกต่างกันมากซึ่งจะเพิ่มขึ้นตาม size matrix (big-o =  $O(n^2)$ )  
ขณะเดียวกันเมื่อทำ blocking ก็จำเป็นต้องกำหนดขนาดของ block ให้มีความพอดีเพื่อให้ไม่จำเป็นต้องวนลูปอย่างเสียเปล่า และคำนึงขนาดของ size matrix ให้เหมาะสมด้วย