

1. (a) stack machine 利用 stack 做 store.
- (b) 太多的 register 造成 loading 時間很久 (每次切 process 要重新 load), 所以適量即可
- (c) 若切成 caller reg / callee reg 代表切換時, 可同時切成 callee reg \Rightarrow caller 不用額外記錄 / 調用 callee 的 reg 狀況, 適合用在有子程序呼叫之 program
- (d) 使常用指令一般化, 可使 reg 使用效率增加 (因為有多種定址模式)
若只有一種 set reg, 則 reg 要記錄每種 type (使效能較低且繁瑣)

2. (A) $2^{20} + 2^{21} + 2^{22} + 2^{24} + 2^{25} + 2^{26} + 2^{27} + 2^{28} + 2^{29}$ ~~*~~

(B) $0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 1 \ 0 \dots\dots 0 \quad 1.111 \times 2^{-1} \Rightarrow 0.1111 \Rightarrow \frac{15}{16}$ ~~*~~

(C) $0 \ 0 \ 1 \ 1 \ 1 \ 1 \Rightarrow \text{lui}$ ~~*~~

(D) Char byte \Rightarrow 8 bit / 一個, 63, 112, 0, 0 \Rightarrow "? p Nil Nil" \Rightarrow "? p" ~~*~~

(E) 在 x86 上看到的相同, 1B 是 "? p"

(F) 捨入計算誤差. $a = 2 \times 10^{38}$ $b = -2 \times 10^{38}$ $c = 1$

$\left\{ \begin{array}{l} \text{若 } (a+b)+c \Rightarrow \text{result} = 1 \\ \text{若 } a+(b+c) \Rightarrow \text{result} = 0 \end{array} \right.$

3. (A) always taken 跟 1 bit taken prediction

(B) 利用 jr 實現 switch, 或利用 jr 實現 call return \Rightarrow jr \$ra

(C) 因為 indirect branch 沒有確切位置, 要到 run-time 時才能確定跳躍目的位址, 且非常耗能 (stall)

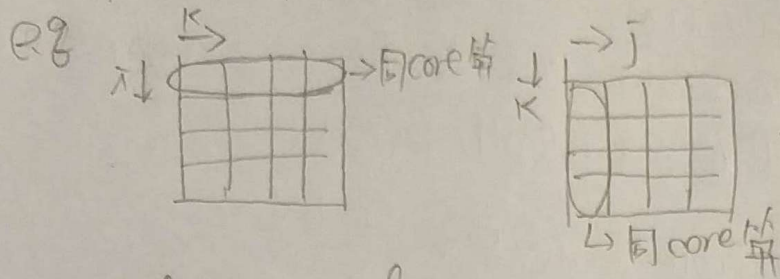
4. A = avg penalty $\begin{matrix} \textcircled{1} & \textcircled{2} \checkmark & \textcircled{3} \end{matrix}$
 $(8+24) \times 0.08 = 2.56 \quad (16+24) \times 0.04 = 1.6 \quad (32+24) \times 0.03 = 1.68$
 block size 16 的 avg penalty 最小

B: optimal word first = 先 fetch 需要的 word, 剩下慢慢填入 cache
 則選 mbs rate 最小的 (需要的第一時間皆可拿到), 提高 hit

5.

(A) Strong-scaling: 可利用加 processor 增加計算效能

(B) 讓每個 core 算同一個 block, 不要平行到同 block 內變數



6. NVM: 有 lower access latency

① 因為一般做 I/O 要花大量時間在等待低速的硬體讀寫, 所以多採用 I/O driven 的方式

② 若增 NVM 方式降低 I/O latency 時間則可考慮用 polling 加快 I/O request 速度

7.

A. 若 time quantum 太小會造成 CPU 使用率低 (因為 switch 的 latency 影響加大)
如此一來多了很多不必要的額外時間, 整體的 turnaround time 變長

B. 使 process 盡可能在同一個 processor 上執行可減少 flush & load 的額外時間
也能使各 cache 不用重新 load \Rightarrow miss rate 下降

8.

```
semaphore chair-num=4;
semaphore mutex=1;
int time[i]=0;
semaphore get-first=1;
```

```
competing(process *p, int i)
{
    time[i]++;
    if (time[i] == 3)
    {
        wait(get-first);
        wait(chair-num);
        wait(mutex);
        c.s & time[i] = 0;
        signal(mutex);
        signal(chair-num);
        signal(get-first);
    }
    else
    {
        if (someone time[i] == 3)
        {
            signal(get-first);
        }
        wait(chair-num);
        wait(mutex);
        c.s & time[i] = 0;
        signal(mutex);
        signal(chair-num);
    }
}
```


9.

- (A) 要使 sys 進入 thrashing 狀態需將可用 memory 全用滿再進行 process 的讀寫 \Rightarrow 造成 memory 不夠需不斷 page fault 且依然不足之狀態
- (B) 若不能搶其他 process 所釋放的 frame 來用, 在原 frame 上有 thrashing 現象之 process 仍會 thrashing \Rightarrow 在該塊 memory 無法被滿足
- (C) 利用 memory 使用局部、集中的特性達到偵測功能 \Rightarrow 計算所需要的 frame 再配置
- (D) 沒辦法完全避免, 假設有 caller function 會利用到額外堆疊呼叫, 則 working set 不一定能偵測得到, 仍有可能發生 thrashing
- (E) 若一個 process 在 VM (A) 產生 thrashing, 其他 VM 不會有任何影響, 因 VM 確保各個 guest OS 上之 simulation Hardware 不會互相影響, 給每個 VM 有固定 Memory

10.

- ① 一般用 read/write 是 linear %
- ② vector % 可在單一個 instruction 中做多個 buffer 空間做同時的 input/output 也可 SIMD 或收集多個 Data 集成單一數據流
- ③ 若 data flow 具有可平行的切割方式 \Rightarrow 能夠做數據處理
- ④ 單一-vector % 可替換成多個 linear %
- ⑤ 可避免執行時單元性問題