

ARS I AVDITORNE VAJE

1. Imamo računalnik s 5 - stopenjskim cevovodom (kot HIP), ki ga želimo poenostaviti tako, da nima vgrajene logike za ugotavljanje cevovodnih nevarnosti. Programi bodo delovali pravilno le, če programer za reševanje cevovodnih nevarnosti uporabi ukaze NOP.

- a) Popravite program tako, da bo deloval pravilno;
b) Spremenite vrstni red ukazov tako, da se število NOP-ov zmanjša na minimum.

Program:

```
LOOP:  LW    R4, 0(R5)    ; R4 ← M[R5+0]
        LW    R1, 0(R6)    ; R1 ← M[R6+0]
        ADD   R2, R1, R4    ; R2 ← R1+R4
        SW    0(R5), R2    ; M[R5] ← R2
        ADDI  R5, R5, #4    ; R5 ← R5+4
        ADDI  R6, R6, #4    ; R6 ← R6+4
        SUBI  R3, R3, #1    ; R3 ← R3-1
        BNE  R3, LOOP    ; če R3 ≠ 0, potem PC ← LOOP
```

Rešitev:

a)

```
LOOP:  LW    R4, 0(R5)
        LW    R1, 0(R6)
        NOP
        NOP
        NOP
        ADD   R2, R1, R4
        NOP
        NOP
        NOP
        SW    0(R5), R2
        ADDI  R5, R5, #4
        ADDI  R6, R6, #4
        SUBI  R3, R3, #1
        NOP
        NOP
        NOP
        BNE  R3, LOOP
        NOP
        NOP
```

b)

```
LOOP:  LW    R4, 0(R5)
        LW    R1, 0(R6)
        SUBI  R3, R3, #1
        ADDI  R5, R5, #4
        ADDI  R6, R6, #4
        ADD   R2, R1, R4
        NOP
        BNE  R3, LOOP
        NOP
        SW    -4(R5), R2
```

2. Imamo računalnik z dvema nivojema predpomnilnika ($L1$ in $L2$). Dostopni čas do nivoja $L1$ je 1 urina perioda, do nivoja $L2$ pa 10 urinih period.

Kakšna je frekvenca ure f_{CPE} , če je povprečni dostopni čas do pomnilnika 2 ns, dostopni čas do glavnega pomnilnika 50 ns, verjetnost zadetka na prvem nivoju H_{L1} 0,95 in globalna verjetnost zadetka na nivoju $L2$ 0,99.

Rešitev:

$$(1-H_{gl,2})=(1-H_{L1})(1-H_{L2})$$

$$H_{L2}=1-(1-H_{gl,2})/(1-H_{L1})$$

$$H_{L2}=1-0.01/0.05$$

$$H_{L2}=0.8$$

$$t_a=t_{aL1}+(1-H_{L1})*[t_{aL2}+(1-H_{L2})*t_{ag}]$$

$$2ns=t_{aL1}+0.05*[10t_{aL1}+0.2*50ns]$$

$$2ns=1.5t_{aL1}+0.5ns$$

$$t_{aL1}=1ns$$

$$\rightarrow f_{CPE} = 1GHz$$

3. Računalnik z navideznim pomnilnikom z odstranjevanjem ima čas dostopa do glavnega pomnilnika 55 ns. Čas za prenos strani iz navideznega pomnilnika v glavni pomnilnik je 9 ms, verjetnost za napako strani pa je 10^{-6} . Kakšen je povprečni dostopni čas če:
- sta tabeli strani v glavnem pomnilniku,
 - tabeli strani sta v glavnem pomnilniku, imamo pa preslikovalni predpomnilnik z verjetnostjo zadetka 98%.

V obeh primerih je preslikava navideznega naslova v fizičnega dvonivojska (preko dveh tabel strani).

Rešitev:

a)

$$t_a = t_{ag} + t_{ag} + t_{ag} + 1E-6*9ms$$

$$t_a = 55ns + 55ns + 55ns + 9ns$$

$$t_a = 174ns$$

b)

$$t_a = 0.02*[t_{ag} + t_{ag}] + t_{ag} + 1E-6*9ms$$

$$t_a = 0.02*[55ns + 55ns] + 55ns + 9ns$$

$$t_a = 66.2ns$$

4. Analizirajte izvajanje naslednjega programa na hipotetičnem cevovodnem računalniku (HIP):

```
$1F000000      ADDI R2, R0, #1024
$1F000004      ADDI R4, R0, #5
$1F000008      ADD  R3, R0, R0
$1F00000C      LOOP  LW   R1,0(R2)
$1F000010      NOP
$1F000014      ADD  R3, R3, R1
$1F000018      SW   4(R2), R3
$1F00001C      ADDI R2, R2, #8
$1F000020      SUBI R4, R4, #1
$1F000024      BNE  R4, LOOP
$1F000028      NOP
```

Računalnik ima vgrajen ukazni in operandni predpomnilnik. Oba predpomnilnika sta direktna, uporabljata pisanje nazaj in sta velika po 256Kx32 bitov z velikostjo bloka 8x32 bitov. Zgrešitvena kazen je 32 urinih period. Vzemite, da so predpomnilnika in cevovod začetku izvajanja zgornjega programa prazni. Izračunajte:

- Za oba predpomnilnika ugotovite koliko in pri katerih naslovih imamo zgrešitve.
- Koliko urinih period traja izvajanje programa? Predpostavite, da cevovodnih nevarnosti ni (so odpravljene) in da pri zgrešitvah cevovod čaka (nimamo špekulativnega izvajanja ukazov). Izvajanje programa traja do prevzema ukaza NOP na naslovu \$1F000028.
- Kako bi se spremenila rezultata iz točk a) in b), če bi bila predpomnilnika čiste asociativne vrste?

REŠITEV

- Pri ukaznem predpomnilniku imamo 2 obvezni zgrešitvi in sicer na naslovih \$1F000000 in \$1F000020. Pri operandnem predpomnilniku imamo ravno tako 2 obvezni zgrešitvi pri branju in sicer na naslovih 1024 in 1056. Pri pisanju ni zgrešitev, ker so bloki že v predpomnilniku in imamo pisanje nazaj.
- Brez čakanja zaradi zgrešitev traja program $5 \cdot 7$ urinih period v zanki plus 3 za začetne ukaze, skupno torej 38. Temu dodamo še $4 \cdot 32$ urinih period za čakanje pri zgrešitvah, skupno torej $38 + 128 = 166$ urinih period.
- Rezultata bi ostala enaka, ker imamo le obvezne zgrešitve.

5. Imamo računalnik z navideznim pomnilnikom na osnovi ostranjevanja. Čas dostopa do glavnega pomnilnika je 50 ns, čas dostopa do diska 8 ms, hitrost prenosa med diskom in glavnim pomnilnikom je 20 Mbesed/s, velikost strani pa 32 Kbesed.

- Izračunajte verjetnost za napako strani 1 – H (predpostavite, da je v TLB vedno zadetek), če je povprečni čas dostopa do navideznega pomnilnika 1,12 časa dostopa do fizičnega pomnilnika (večji za 12%).
- Zakaj je v primerjavi s predpomnilnikom v tem primeru potrebna tako velika verjetnost zadetka?

REŠITEV

a) $t_a = t_{gp} + (1-H) * t_{disk}$, $t_a = 1.12t_{gp}$ --> $1-H = \frac{0.12t_{gp}}{t_{disk}}$

$$t_{disk} = 8ms + \frac{32Kbese}{20Mbese / s}$$

$$1-H = \frac{0.12 * 50 * 10^{-9} s}{8 * 10^{-3} s + \frac{32 * 10^3}{20 * 10^6} s} = 6.25E-7 \text{ ali } 0.0000625\%$$

b) Ker je kazen za zgrešitev zelo velika.

6. Na cevovodnem računalniku, ki je podoben HIP, se izvaja naslednji program:

```

LOOP1      LW      R3, 0(R2)
           ADD     R8, R3,R3
           SW      0(R2),R8
           BEQ    R8, LOOP2
           ADDI   R2,R2,#4
           J      LOOP1(R0)
LOOP2      SW      4(R2),R2
    
```

Vzemite, da je pri vseh pomnilniških dostopih zadetek v predpomnilniku in izračunajte koliko urinih period traja izvajanje enega obhoda zanke LOOP1 (začetno latenco zanemarite) v naslednjih primerih:

- a) CPE odpravlja cevovodne nevarnosti z vstavljanjem mehurčkov (čakalna stanja).
- b) CPE ima vgrajeno logiko za premoščanje in uporablja zakasnjene skoke. Upoštevajte delovanje zakasnjenih skokov in program spremenite tako, da bo tek el najhitreje. Koliko urinih period traja en obhod zanke LOOP1 sedaj?

REŠITEV:

a) Za LW, ADD imamo 3 čakalna stanja, za BEQ 2, za ADDI ni čakalnih stanj in za J še 2 čakalni stanji. Izvajanje enega obhoda traja 6 u.p. za ukaze plus 10 u.p. za čakalna stanja, skupno torej 16.

```

b) LOOP1      LW      R3, 0(R2)
           ADD     R8, R3,R3
           BNE    R8, LOOP1
           SW      0(R2), R8
           ADDI   R2,R2, #4
LOOP2      SUBI   R2,R2, #4
           SW      4(R2), R2
    
```

Za LW imamo 1 čakalno stanje. Izvajanje enega obhoda traja 5 u.p. za ukaze plus 1 u.p. za čakalna stanja, skupno torej 6 u.p.

Če ne spreminjamo ukazov ampak le vrstni red je najhitrejša rešitev

```

LOOP1      LW      R3, 0(R2)
           ADD     R8, R3,R3
           BEQ    R8, LOOP2
           SW      0(R2),R8
           NOP
           J      LOOP1
           ADDI   R2,R2,#4
           NOP
LOOP2      SW      4(R2),R2
    
```

Za LW imamo 1 čakalno stanje. Izvajanje enega obhoda traja 8 u.p. za ukaze plus 1 u.p. za čakalna stanja, skupno torej 9 u.p.

7. Imamo računalnik z dvema nivojema predpomnilnika (L1 in L2). Dostopni čas do nivoja L1 je 1 urina perioda, do nivoja L2 pa 8 urinih period. Kakšna je frekvenca ure f_{CPE} , če je povprečni dostopni čas do pomnilnika 1,5 ns, dostopni čas do glavnega pomnilnika 60 ns, verjetnost zadetka H_{L1} na nivoju L1 enaka 0,95 in globalna verjetnost zadetka H_{L2} na nivoju L2 enaka 0,995.

REŠITEV:

$$t_a = t_{L1} + (1 - H_{L1}) * t_{L2} + (1 - H_{L2}) t_g$$

$$t_{L1} = \frac{1up}{f_{CPE}} \quad t_{L2} = \frac{8up}{f_{CPE}}$$

$$1.5ns = \frac{1up}{f_{CPE}} + 0.05 * \frac{8up}{f_{CPE}} + 0.005 * 60ns$$

$$f_{CPE} = \frac{1.4up}{1.2ns} \approx 1.17GHz$$

8. Analizirajte izvajanje naslednjega programa na hipotetičnem cevovodnem računalniku (HIP):

```

$40000000      ADDI R2, R0, #$4000
$40000004      ADDI R4, R0, #8
$40000008      ADD  R3, R0, R0
$4000000C      LOOP LW   R1, 0(R2)
$40000010      ADDI R2, R2, #4
$40000014      ADD  R3, R3, R1
$40000018      SUBI R4, R4, #1
$4000001C      BNE  R4, LOOP
$40000020      NOP
$40000024      NOP
$40000028      SW   0(R2), R3

```

Računalnik ima vgrajen skupni ukazni in operandni predpomnilnik. Predpomnilnik je direktni in je velik 16Kx8 bitov z velikostjo bloka 64x8 bitov. Zgrešitvena kazen je 16 urinih period. Vzemite, da sta predpomnilnik in cevovod na začetku izvajanja zgornjega programa prazna.

- Koliko urinih period traja izvajanje programa, če imamo v predpomnilniku pri vseh pomnilniških dostopih zadetek? Predpostavite, da operandne nevarnosti procesor odpravlja z vstavljanjem mehurčkov, pri skokih pa predpostavlja neizpolnjen pogoj. Izvajanje programa traja do prevzema ukaza SW 0(R2),R3 na naslovu \$40000028.
- Upoštevajte delovanje predpomnilnika. Ugotovite pri katerih pomnilniških dostopih pride do zgrešitev. Koliko urinih period traja izvajanje programa sedaj? Ob zgrešitvah cevovod stoji.
- Koliko časa bi trajalo izvajanje programa, če bi imel procesor ločena direktna predpomnilnika za ukaze in operande velikosti 8Kx8 in enako velikost bloka.

REŠITEV:

```

$40000000          ADDI   R2, R0, #$4000
$40000004          ADDI   R4, R0, #8
$40000008          ADD    R3, R0, R0
                    1 čakalno stanje
$4000000C          LOOP   LW    R1, 0(R2)
$40000010          ADDI   R2, R2, #4
                    2 čakalni stanji (tu je tudi prevzem operanda)
$40000014          ADD    R3, R3, R1
$40000018          SUBI   R4, R4, #1
                    3 čakalna stanja
$4000001C          BNE    R4, LOOP
                    tudi pri ponovnem skoku v zanko se izvedeta še oba NOP-a
$40000020          NOP
$40000024          NOP
$40000028          SW     0(R2), R3
    
```

- a) Izvajanje traja $3 + 1\text{čs} + 8 * (5 + 5\text{čs} + 2(\text{nop})) = 100$ ur. period
- b) Zgrešitve so na naslovih
 \$40000000 (ukaz), \$4000 (operand), \$40000018 (ukaz), \$4004 (operand),
 \$40000018 (ukaz), \$4008 (operand), \$40000018 (ukaz), \$400C (operand),
 \$40000018 (ukaz), \$4010 (operand), \$40000018 (ukaz), \$4014 (operand),
 \$40000018 (ukaz), \$4018 (operand), \$40000018 (ukaz), \$401C (operand),
 \$40000018 (ukaz)

Izvajanje se sedaj podaljša za čas kazni pri zgrešitvah ali za $17 * 16 = 272$ ur. period, zmanjšano za $8 * 2$ periodi, ker se 2 čakalni stanji zaradi RAW nevarnosti pri ukazu ADD R3,R3,R1 prekrivata s čakalnimi stanji, ki so posledica zgrešitve v predpomnilniku pri prevzemu ukaza SUBI R4,R4,1 na naslovu 40000018.
 Skupaj torej $272 + 100 - 16 = 356$ ur. period.

- c) Zaradi ločenih predpomnilnikov imamo le dve obvezni zgrešitvi pri dostopih na naslova \$40000000 (ukaz) in \$4000 (operand). Program bi se izvajal $100 + 2 * 16 = 132$ ur. period.

9. Načrtovalec računalniškega sistema ima na voljo RISC procesor z delovnim taktom 250 MHz, ki ima vgrajen L1 predpomnilnik. Proizvajalec procesorja navaja, da lahko procesor v programih, ki uporabljajo le operande v registrih in so vsi ukazi v predpomnilniku, doseže $\text{CPI} = 1$ in da je povprečna verjetnost zadetka v L1 predpomnilniku 0.95. Načrtovalec ima na razpolago tudi dovolj cenen glavni pomnilnik (dinamični RAM) s časom dostopa 72 ns.

- a) Kolikšen CPI doseže računalniški sistem za navedene programe ob upoštevanju zgrešitev v predpomnilniku?
- b) Zaradi zahtevnosti aplikacije v kateri bo na novo zasnovan sistem uporabil, bi načrtovalec rad dosegel vsaj $\text{CPI} = 1.25$. Uporabljeni procesor omogoča enostavno izvedbo zunanjega L2 predpomnilnika. Ali bo načrtovalec sistema L2 predpomnilnik lahko izvedel s cenejšimi SRAM čipi z dostopnim časom 25 ns (poljubne velikosti), ali bo moral uporabiti hitrejše in dražje čipe z dostopnim časom 10ns?

REŠITEV:

Za $\text{CPI} = 1$ imamo en dostop v vsaki periodi, ena perioda pa traja $1/250$ MHz ali 4 ns.

- a) zaradi zgrešitev se CPI podaljša na

$$\text{CPI} = \text{CPI}_{(\text{brez zgrešitev})} + (1 - H) * \text{zgrešitvena kazen} = 1 + 0.05 * (72 \text{ ns} / 4 \text{ ns}) = 1.9$$

- b) predpostavim idealni primer, da se vse zgrešitve v L1 prestrežejo v L2, potem imamo

$$\text{CPI} = \text{CPI}_{(\text{brez zgrešitev})} + (1 - H) * \text{zgrešitvena kazen} = 1 + 0.05 * (28 \text{ ns} / 4 \text{ ns}) = 1.35$$

Dejanski dostopni čas moramo zaokrožiti na prvi večkratnik CPU cikla, ki je večji ali enak dostopnemu času, ki ga navaja proizvajalec. Sistem bi verjetno deloval tudi če bi SRAM vezja "prisilili" na 24 ns dostopni čas, vendar tudi v tem primeru ne dosežemo zahtevanega CPI, ki je 1.25. Z dostopnim časom 24 ns dosežemo namreč le $\text{CPI} = 1.3$. Uporabiti moramo torej hitrejše in dražje SRAM čipe z dostopnim časom 10ns.

10. Na računalniku z navideznim pomnilnikom poženemo program za kopiranje pomnilniškega bloka velikosti 64B od naslova \$0000FD0 dalje na naslove od \$00001FE0 dalje. Pomnilniška beseda je dolga 1B, pri kopiranju se uporabljajo 32-bitni dostopi. Dolžina navideznega in fizičnega naslova je 32 bitov, velikost strani 4KB, preslikovanje naslovov je enonivojsko. Dostopni čas do glavnega pomnilnika je 55ns, čas za prenos strani iz navideznega v glavni pomnilnik znaša 9ms. Tabela strani je v glavnem pomnilniku od naslova \$FFF00000 dalje; del tabele strani in zgradba deskriptorja strani sta prikazani na sliki.

- V katera fizična naslova se preslikata navidezna naslova \$0000FD0 in \$00001FE0?
- Izračunajte povprečni dostopni čas pomnilniškega dostopa do operanda pri zgoraj navedenem kopiranju 64B dolgega bloka (bralni dostopi od naslova \$0000FD0 dalje in pisalni dostopi od \$00001FE0 dalje)? Kolikšen bi bil ta čas v primeru, če pri izvajanju programa ni nobene napake strani.
- V računalnik vgradimo TLB (čisti asociativni) velikosti 32 deskriptorjev strani, ki pred izvajanjem programa ne vsebuje nobenega deskriptorja strani z operandi. Kakšen je odgovor na vprašanji iz točke b) v tem primeru?

Deskriptor in tabela strani:

	številka okvira	parametri	PV
31	1211		1 0

P – bit 0 (stran prisotna): 1-da / 0-ne
V – bit 1 (veljavni bit): 1-veljavno / 0-neveljavno

naslov	vsebina
\$FFF00000	\$00000003
\$FFF00004	\$00007003
\$FFF00008	\$00005001
\$FFF0000C	...
...	...

REŠITEV:

- \$0000FD0 → \$0000FD0, \$00001FE0 → \$00007FE0
- dostopi:

	branje	pisanje
1	\$00000FD0,	\$000001FE0
2	\$00000FD4,	\$000001FE4
3	\$00000FD8,	\$000001FE8
4	\$00000FDC,	\$000001FEC
5	\$00000FE0,	\$000001FF0
6	\$00000FE4,	\$000001FF4
7	\$00000FE8,	\$000001FF8
8	\$00000FEC,	\$000001FFC
9	\$00000FF0,	\$00002000
10	\$00000FF4,	\$00002004
11	\$00000FF8,	\$00002008
12	\$00000FFC,	\$0000200C
13	\$00001000,	\$00002010
14	\$00001004,	\$00002014
15	\$00001008,	\$00002018
16	\$0000100C,	\$0000201C

Pri poudarjeno zapisanem naslovu pride do napake strani (to se zgodi samo enkrat). Torej imamo 31 dostopov, ko napake strani ni, pri enem pa imamo napako strani. Predpostavimo, da ob napaki strani prenos strani iz okvira v glavnem pomnilniku v navidezni pomnilnik ni potreben:

$$t_{acc} = [31 * (55ns + 55ns) + 55ns + 55ns + 9ms] / 32 = 281.36\mu s$$

če napake strani ne bi bilo, imamo 32 dostopov brez napake strani:

$$t_{acc} = [32 * (55ns + 55ns)] / 32 = 110ns$$

- dostopamo do treh strani (v TLB se iz tabele strani prenesejo trije deskriptorji):

$$t_{acc} = [3 * 55ns (tabela strani) + 32 * 55ns (dostop v fiz.pom.) + 1 * 9ms (napaka strani)] / 32 = 281.31\mu s$$

če napake strani ne bi bilo:

$$t_{acc} = [3 * 55ns (tabela strani) + 32 * 55ns (dostop v fiz.pom.)] / 32 = 60.16ns$$