|  |  |  |
| --- | --- | --- |
| **BAR-ILAN UNIVERSITY** Engineering Faculty |  | אוניברסיטת בר-אילן הפקולטה להנדסה |

ארכיטקטורת מחשבים

# תשע"ד סמס' ב' מועד ג'

**83-248**

* מרצה : פרופ' שמואל וימר
* מתרגל : מר יוסף לונדון
* **חומר עזר מותר:** מחברות/ דפים מהרצאות, תרגולים ותרגילי בית, ספר computer organization and design, מילון עברי צרפתי.
* משך המבחן שלוש שעות
* סך כל הנקודות הוא 110. הציון המרבי הינו 100.
* משקל השאלות השונות נתון בגוף השאלות.
* **יש לנמק את כל תשובותיכם**. אין צורך לפתח מחדש תוצאות שהוכחו בכיתה, אלא אם כן נאמר מפורשות לעשות כן.
* יש לשרטט דיאגראמות באופן ברור !
* הכתיבה בעט בלבד. כתיבה בעפרון לא תיבדק.

**בהצלחה!**

**שאלה 1 (40 נקודות)**

שאלת מטמון:

נתונה חלק מתוכנית בC:

char a[STOP];

int i, j, temp;

for (i=0 ; i<10,000; i=i+1)

 for (j=0 ; j<STOP; j=j+STEP)

 sum=sum+a[j];

הניחו גודל של CHAR הוא byte יחיד. הניחו שבזכרון המרכזי שמור רק ערך המערך a.

הרעיון של השאלה הוא ניתוח מבנה המטמון (CACHE) מתוך זמן ההרצה של התוכנית הנ"ל.

התוכנית רצה כתוצאה מהלולאה החיצונית 10,000 פעמים. הלולאה הפנימית קוראת מידע ממערך של bytes כתלות בגודל של STEP. כל המשתנים נשמרים ברגיסטרים כך שזה מהווה "אפס זמן" גישה וכן שחישוב סכום וכדו' מהווה "אפס זמן".

מריצים את התוכנית על מעבד עם ערכים שונים לצורך בחינה של מבנה הCACHE ומודדים את הזמן הממוצע (!) לביצוע שורת הקוד: sum=sum+a[j];

|  |
| --- |
| **STEP** |
| **128** | **64** | **32** | **16** | **8** | **4** | **2** | **1** |  | **STOP** |
| 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | **8** |
| 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | **16** |
| 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | 10ns | **32** |
| 10ns | 10ns | 10ns | 10ns | 50ns | 50ns | 30ns | 20ns | **64** |
| 10ns | 10ns | 10ns | 50ns | 50ns | 50ns | 30ns | 20ns | **128** |
| 10ns | 10ns | 50ns | 50ns | 50ns | 50ns | 30ns | 20ns | **256** |

רמזים לפתרון השאלה:

* שימו לב לגודל של STEP ונסו להדגים לעצמכם מה יקרה עבור גדלים שונים. השוו עם תוצאות הטבלה.
* $AMAT=Hit time+Miss Rate\*Misspenalty$
* זכרו, הזמנים הם זמנים ממוצעים. כלומר איטרציות יחידאיות מבחינת זמן, נהיים זניחים
1. מהו גודל הCACHE מבחינת DATA? כלומר, מהו כמות המידע שמכיל הCACHE?
ניתן לזהות די מיידית כי הלולאה הפנימית מבצעת מספר גישות לזכרון בהתאם לחסם שבחוץ. כאשר החסם (STOP) באורך של 32 ויחידת הקפיצה (STEP) באורך של 1, ניגשים צעד אחר צעד בזכרון. כפי שניתן לראות, עד חסם של 64 אין שינוי בזמני הגישה ולכן כנראה עד לשם יש hit. מה שאומר, שגודל המטמון הוא כנראה 32 byte שכן עבור צעד של 1 byte תמיד אנחנו בHIT וברגע שנחרוג ל64 אנחנו מתחילים לפספס.

נדגיש שוב – הלולאה רצה 10,000 פעם כך שגם אם בפעם הראשונה תמיד פספסנו, בפעם השניה (וכן הלאה עד 10000) המידע כבר שמור במטמון.

כבר בפעם הראשונה שמבצעים את הלולאה הcache מתמלא במידע ובזה נגמר הסיפור. לא משנה מהו אורך הבלוק או האסוציאטיביות של הCACHE , עבור CACHE בגודל הנ"ל תמיד נוכל להיות עם כל המידע של המערך מההתחלה ועד לbyte ה32.

1. מהו גודל בלוק יחידי בCACHE?

נסתכל על החסם של 64. אנחנו רואים כי זמני הגישה גדלים עם גודל הקפיצה.

$$AMAT=Hit time+Miss Rate\*Misspenalty$$

כלומר – התחלנו לפספס, ו**זמן הפספוס** לכל ניסיון גישה לא משתנה – לכן רק **אחוזי\קצב הפספוס** משתנה. תחת ההנחה הסבירה שזמן גישה הוא 10ns, (כך הבנו מסעיף א) אנו רואים שניתן לכתוב:

$$20=10+Miss Rate\_{1}\*Misspenalty ; 30=10+Miss Rate\_{2}\*Misspenalty $$

$$10ns=MR\_{1}\*T ;20ns=MR\_{2}\*T$$

כלומר, קצב הפספוס גדל פי 2 בין צעד יחיד לצעד כפול.

כמו כן הוא גדל פי 4 לצעד של 4 ו**לא** משתנה לצעד של 8. מכאן ניתן להסיק גודל בלוק של 4.
עבור צעד יחיד – יש פספוס כל byte ראשון ואז hit במשך 3 byte הבאים. המידע גם מתחלף במשך כל לולאה. על כן אחוזי הפספוס הינם 25% במקרה הנ"ל.

עבור צעד כפול – פספוס כל byte שני. לכן 50% פספוס.

עבור צעד של 4 – כל byte מבוקש מביא איתו את ה3 byte הצמודים לו אך הם אינם מענייננו שכן אנו קופצים בארבעה בתים ולכן יש 100% פספוס. וכנ"ל לצעד של 8.

הצעדים שלאחר מכן יתבהרו בהמשך.

1. מהו תצורת הCACHE directly mapped\fully associative\n-way associative (אם n אז למה שווה n)?

לפי מה שאמרנו יש לCACHE 32 byte ובבלוקים של 4 ולכן סה"כ שמונה בלוקים. המעבר בין זמן גבוה עבור חסם של 64 וצעד של 8 לזמן מאוד נמוך עבור אותו חסם בצעד של 16 נותן אינדיקציה טובה להבנה.

צעד של 8 ניגש למקומות ה -0,8,16,24,32,40,48,56. גם עבור directly mapped לכל אינדקס כזה בcache של 8 בלוקים נקבל אותו אינדקס בCACHE (כולם mod8 =0) (וגם נדרוס את המידע בתוך כל לולאה) אבל עבור צעד של 16 זה לא אמור לרדת פתאום. אלא, בצעד של 16 כל יחידת מידע אומנם ממופה לאותו אינדקס אך יש 4 סטים ולכן הם לא דורסים האחד את השני.

1. מהם זמני Hit\Miss?
מתוך מה שאמרנו זמני הפגיעה והפספוס ברורים והינם 10ns ו40ns בהתאמה.
מבנה הCACHE:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SET 4 | SET 3 | SET 2 | SET 1 |  |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

 | MOD2=0 |
|

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

 | MOD2=1 |

שאלה 2 **(40 נקודות)**

**עבור מעבד ה** MIPS**, נתונה טבלת ההשהיות הבאה של מרכיביו:**

|  |  |  |
| --- | --- | --- |
| Delay (pSec) | Parameter | Element |
| 30 | *t*pcq | register clock-to-Q |
| 20 | *t*setup | register setup |
| 25 | *t*mux | multiplexer |
| 40 | *t*eq | equality comparator |
| 15 | *t*AND | AND gate |
| 200 | *t*ALU | ALU |
| 250 | *t*MEMread | memory read |
| 220 | *t*MEMwrite | memory write |
| 150 | *t*RFread | register-file read |
| 100 | *t*RFwrite | register-file write |
| 20 | *t*RFsetup | register-file setup |

1. **חשבו מהו תדר השעון המרבי בו ניתן להפעיל את המעבד.**
* **תארו בפירוט מלא את חישוביכם.**
* **רשמו ביטוי מפורש לזמן מחזור השעון במונחי הפרמטרים המופיעים בטבלה.**
* **לאחר מכן השתמשו בערכיהם לחישובים מספריים.**

**הנחיות:**

1. **היעזרו בתרשימים המצורפים וסמנו לכל דרגה ודרגה את מסלול החישוב הארוך ביותר. שימו לב שזהו תרשים של** MIPS **שנלמד בקורס, בתוספת** MUX **במוצא של ה** REGISTERFILE**, המשמשים לניתוב נתונים מקודמים (**FORWARD**) במקרה של** HAZARD**.**
2. **שימו לב שלחלק מהרגיסטרים שעון שונה מלאחרים. חלקם מגיבים לעליית שעון וחלקם לירידת שעון.**

The clock cycle is dictated by the longest delay path of the processor.

Consider first the critical path at every pipeline stage. Such paths always start at the clock of a register and terminate at the data input of a register.

Notice that we need to be able to complete each task in half a cycle.

**Instruction fetch stage:**

*Tcyc min* IF =2 x ( *t*pcq + *t*MEMread + *t*setup ) = 2 x (30 + 250 + 20) = 600 pSec

**Decode stage:**

*Tcyc min* ID = 2 x (*t*RFread + *t*mux + *t*eq + *t*AND + *t*mux + *t*setup) =

 2 x (150 + 25 + 40 + 15 + 25 + 20) = 550 pSec

**Execution stage**:

*Tcyc min* EXE = 2 x (*t*pcq + *t*mux *+ t*mux *+ t*ALU + *t*setup) =

 2 x ( 30 + 25 + 25 + 200 + 20) = 600 pSec

**Memory stage:**

*Tcyc min* MEM = 2 x ( *t*pcq + max{ *t*MEMread , *t*MEMwrite } + *t*setup) =

2 x (30 + 250 + 20 )= 600 pSec

**Writeback stage:**

*Tcyc min* WB = 2 x (*t*pcq + *t*mux + *t*RFwrite) = 2 x (30 + 25 + 100) = 310 pSec

*T*CLK = max *Tcyc min* {IF , ID , EXE , MEM , WB} = 600 pSec.

1. **כמה זמן תימשך ריצת אוסף תכניות המכילות 100 ביליון פקודות מכונה. הניחו שמספר המחזורים הממוצע לפקודה הוא CPI = 1.15?**

*T*total = (100 billion instructions) x (1.15 cycles/instruction) x (600 pSec) = 1011 x 1.15 x 600 x 10-12 = 69 Sec

1. **מהו הTHROUPUT של המערכת? מהו הLATENCY?**כאן ניתנו נקודות על הסברים שונים כל עוד הם הוסברו.
הLATENCY הוא זמן שלוקח לפקודה לעבור – 5 מחזורים.

הTHROUGHPUT לפקודה יחידה היא מחזור יחיד. אבל בממוצע הTHOUGHPUT היה יותר ממחזור יחיד שכן CPI היה גדול מאחד.

1. **נמדוד את ביצועי המעבד ע"י מספר הפקודות לשנייה בתדר שעון מרבי. בניסיון לשפר ביצועים, ראובן, מהנדס מתחיל בחברה (בוגר בר-אילן) מציע להעביר את בדיקת תנאי השוויון של קפיצה מותנית משלב ה** ID **לשלב ה** EXE**, ואילו שמעון, מהנדס וותיק בחברה (בוגר אוניברסיטה אחרת) מציע להעביר את בדיקת תנאי השוויון לשלב ה** MEM**.**
* **האם כל אחת מההצעות הנ"ל תשפר, תקלקל או לא תשפיע על הביצועים? הסבר בפרוטרוט את תשובתך.**
* **במידה והביצועים משתנים, חשב באופן מספרי פי כמה (Tcycle, THROUGHPUT, LATENCY , cpi).**

**נתון**BENCHMARK **כדלהלן:**

**25% פקודות קריאה (**LOADS**), ב 40% מהן מופיעה מיד פקודה המשתמשת בנתון שנקרא.**

**10% פקודות כתיבה (**STORES**).**

**11% הסתעפויות (**BRANCHES**), ב 25% מתוכן ניחוש הקפיצה היה שגוי.**

**2% פקודות קפיצה. הנח שהפקודה שבאה מיד אחרי קפיצה מרוקנת (**FLUSH**) מהצינור.**

**52% פקודות אריתמטיות (**R-TYPE**).**

The critical delay paths of the ID stage gets shorter by *t*eq + *t*AND + *t*mux = 80pSec.

*Tcyc min* ID = 2 x (*t*RFread + *t*mux + *t*setup) = 390 pSec

The critical path in the EXE stage gets longer by 80pSec at most in the first case and the critical paths in the MEM gets longer by 80pSec at most in the second case.

**Execution stage**:

*Tcyc min* EXE = 2 x (*t*pcq + *t*mux *+ t*mux *+ t*ALU + *t*setup) + 160 =

 2 x ( 30 + 25 + 25 + 200 + 20) + 160 = 760 pSec

**Memory stage:**

*Tcyc min* MEM = 2 x ( *t*pcq + max{ *t*MEMread , *t*MEMwrite } + *t*setup) + 160=

2 x (30 + 250 + 20 ) + 160 = 760 pSec

So it makes no difference which of the two – they will both dictate a slower clock rate!

Perfbranch\_ID original = (1/(600 x 10-12))/1.1475 = 1.4524 x 109 instructions/Sec

Ignoring the branch prediction differences:

Perfbranch\_ID =(1/(760 x 10-12))/1.1475 = 1.1467 x 109 instructions/Sec

Unfortunately we can't therefor:

If equality check takes place at EXE stage two instructions are flushed in miss:

CPI\_BRANCH = 0.75 x 1 + 0.25 x 3 = 1.5.

CPI = 0.25 x 1.4 + 0.1 x 1 + 0.11 x 1.5 + 0.02 x 2 + 0.52 x 1 = 1.175.

Perfbranch\_EXE = (1/(760 x 10-12))/1.175 = 1.1198 x 109 instructions/Sec

If equality check takes place at MEM stage three instructions are flushed in miss:

CPI\_BRANCH = 0.75 x 1 + 0.25 x 4 = 1.75.

CPI = 0.25 x 1.4 + 0.1 x 1 + 0.11 x 1.75 + 0.02 x 2 + 0.52 x 1 = 1.2025.

Perfbranch\_MEM = (1/(760 x 10-12))/1.2025 = 1.0942 x 109 instructions/Sec



שאלה 30 נק':

לרשותכם מערכת MIPS. רכיב הALU במערכת שלכם יכולה לבצע פקודות ADD, SUB, OR, AND.

נתון כי במקום מסוים בזכרון המכונה LABEL1 שמור ערך בזכרון – X. כמו כן במקום אחר המכונה LABEL2 שמור הערך – Y.

הכתובות בזכרון של X,Y שמורים ברגיסטרים $S0, $S1 בהתאמה.

1. כתבו תוכנית בשפת ASSEMBLY המבצעת את הפעולה X(XOR)Y.

דגשים:

* שימו לב לSYNTAX נכון ושימוש נכון ברגיסטרים הנכונים.
* זכרו ששיטת הקידוד בMIPS הוא במשלים ל-2.
1. בכמה מחזורים יתבצע התוכנית בMIPS מסוג SINGLE CYCLE\MULTI CYCLE\PIPELINE?

מבחינת פעולות לביצוע NOT על מספר המיוצג בשיטת המשלים ל-2:

$$X+X^{'}= -1==>X^{'}= - X-1$$

מכיוון שהם לא למדו את הפעולה la שעושה load של label לתוך רגיסטר אני אשנה את השאלה כך שרגיסטרים $S0 , $S1 יכילו אותם בהתאמה.

la $S0 , LABEL1

la $S1 , LABEL2

lw $S0 , 0($S0) // s0 contains X

lw $S1 , 0($S1) // s1 contains Y

sub $t0, $ZERO, $S0

sub $t1, $ZERO, $S1

addi $t0, $t0, -1 // t0 contains X'

addi $t1, $t1, -1 // t1 contains Y'

and $t0, $t0,$S1 // to contains X'Y

and $t1 , $t1, $S0

or $t0, $t0 , $t1

את החישובים של מספר מחזורים לכל פקודה וכדו – לפי ההתפלגות הידועה. (בצינור זה מחזור לפקודה וכן וsingle cycle ובmulti לפי הנלמד בהרצאה).

במבחן אמרתי שלא חובה להתייחס כאן מHAZARDS. אבל , מי שיתחשב צריך להתחשב לו יותר בציון.