|  |  |  |
| --- | --- | --- |
| **BAR-ILAN UNIVERSITY (RA)**  Faculty of Engineering  Ramat-Gan 52900, Israel | **Tel: 03-5317722**  **engbi@mail.biu.ac.il** | אוניברסיטת בר-אילן (ע"ר) הפקולטה להנדסה  רמת-גן 52900 |

מבנה מחשבים ספרתיים

# תשע"ט סמסטר ב' מועד ג'

**83-301**

**מרצה:** פרופ' שמואל וימר

**מתרגל:** מר בנימין פרנקל

* **יש לקרוא היטב את ההוראות.**
* **חובה לענות על כל השאלות.**
* **ציון מקסימלי בבחינה: 100 נקודות.**
* **יש לנמק הייטב את כל תשובותיכם, לפרט חישובים ולהסבירם**.
* יש להקפיד על כתב יד קריא!
* יש לרשום תשובות בתוך הטבלאות המצורפות במקום שנדרש.
* **חומר עזר מותר בשימוש**: מחשבון, ספר הקורס ושקפי ההרצאות בלבד! כל חומר עזר אחר אסור בהחלט!
* **משך הבחינה:** שלוש שעות.
* **יש לצרף את שאלוני הבחינה למחברת!**

**בהצלחה!**

**שאלה א' 55 נק' (זכרון מטמון)**

1. נתונים שני מעבדים הפועלים באותו תדר ומריצים אותה תכנית בדיוק.

במעבד הראשון miss penalty הינו , ואילו cache hit אורך . ידוע ש hit קורה ב 90% מהזמן.

במעבד השני miss penalty הינו , ואילו cache hit אורך . ידוע ש hit קורה ב 95% מהזמן.

1. (5 נק') חשב בצורה מלאה את ביצועי הזכרון של שני המעבדים. לאיזה מהם ביצועים טובים יותר?
2. (5 נק') עבור איזה cache hit ביצועי המעבדים משתווים?

**תשובה**:

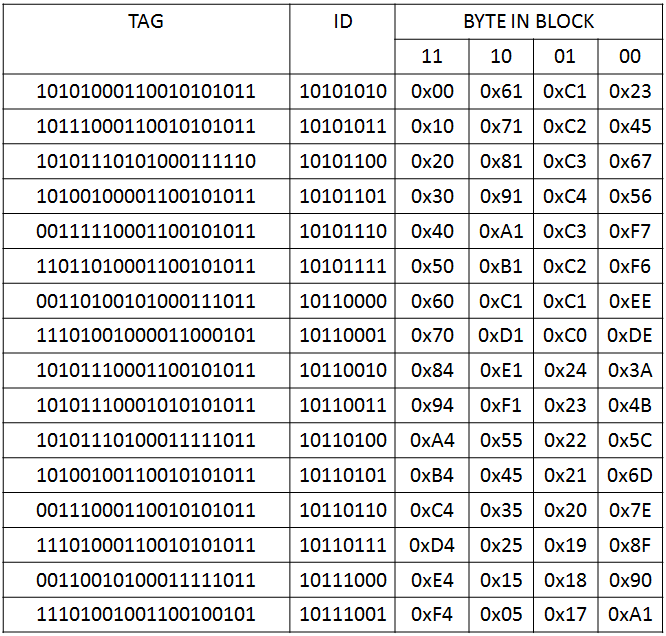
In the 1st processor 90% of the time the access will be 2 ns while the rest of the time (10%) the access time will be 200 ns. The average is

(0.9 × 2ns) + (0.1 × 200ns) = 1.8ns + 20ns = 21.8ns

In the 2nd processor 95% of the time the access will be 1ns while the rest of the time (5%) the access time will be 500 ns. The average is

(0.95 × 1ns) + (0.05 × 500ns) = 0.95ns + 25ns = 25.95 ns

1. לפניך חלק מתוך CACHE של מעבד כלשהו. בכל אחד מהסעיפים שלהלן צרף לתשובתך **הסבר מלא וחישוב מדוייק**.



1. (5 נק')

מאיזה סוג הCACHE הנ"ל?

מהו גודל מרחב הכתובות של המעבד?

**תשובה**: Directly mapped וישנן 30 ביטים בכתובת סה"כ (20 + 8 +2) לכןBytes

1. (5 נק')

כמה בלוקים יש במרחב הכתובות?

מהו גדלו של ה CACHE בבתים (Bytes) ?

**תשובה**: בכל בלוק ישנם 4 בתים,

1. (10 נק') נדרש לקרוא ממרחב הזיכרון את הבית Byte)) הנמצא בכתובת 0x01BA3ED1 האם ניתן לדעת מה ערכו של הנתון בכתובת הנ"ל? ובמידה וכן אז מהו?

**תשובה**:

0x01BA3ED1=101011101000111110111011010001 ולכן:

Byte offset = 01

Index = 10110100

Tag=10101110100011111011

ולכן הכתובת מופיעה ב cache והערך הנמצא בה הוא 0x22.

1. (10 נק') בקוד שרץ על המעבד הנ"ל ישנה לולאה בת חזרות ובתוכה רצף גישות למלים הנמצאות בכתובות הבאות:

11110000001111111000 10110100 00

11110000001111111000 10110101 00

11110000001111111000 10110110 00

11110000001111111000 10110111 00

00001111000000000111 10110100 00

00001111000000000111 10110101 00

00001111000000000111 10110110 00

00001111000000000111 10110111 00

00111001100011111111 10110100 00

00111001100011111111 10110101 00

00111001100011111111 10110110 00

00111001110011111111 10110111 00

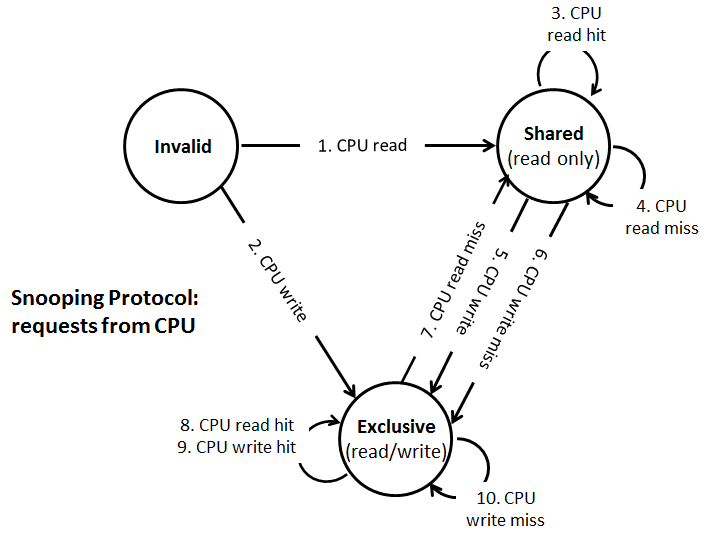
מה יהיה מספר ה misses ?

1. (10 נק') הוצע להפוך את ה cache ל 2-way set associative וזאת מבלי לשנות את גדלו. האם עבור הלולאה שבסעיף הקודם הדבר יועיל? ואם כן, מה יהיה מספר ה ?misses וזאת בהנחה שבתחילת הלולאה אף אחת מהכתובות הנ"ל לא הופיע ב cache.
2. (5 נק') כעת הוצע להפוך את ה cache ל 4-way set associative וזאת מבלי לשנות את גדלו. האם עבור הלולאה שבסעיף הקודם הדבר יועיל? ואם כן, מה יהיה מספר ה ?misses וזאת בהנחה שבתחילת הלולאה אף אחת מהכתובות הנ"ל לא הופיע ב cache.

**שאלה ב' (55 נקודות)**

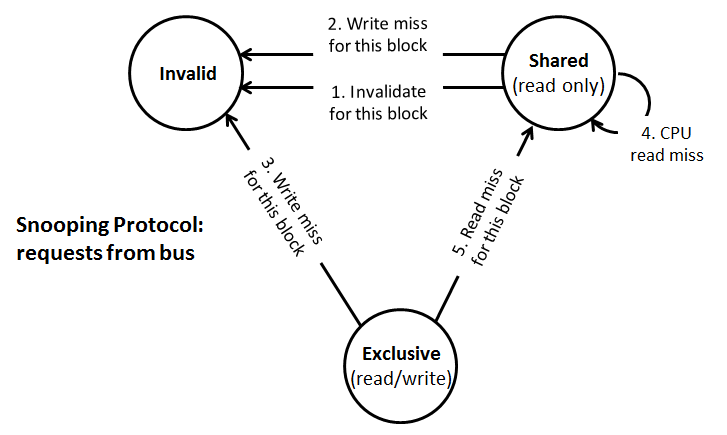
נתון מעבד בעל ארבע ליבות בארכיטקטורת 32 סיביות ו shared memory. לכל ליבה זכרון מטמון משלה עם bus משותף. פרוטוקול הקוהרנטיות של הזכרונות הינו snooping protocol,write invalidate. וזכרונות המטמון פועלים בשיטת writeback.

1. (15 נקודות) מכונת המצבים הנענית לבקשות של המעבד מתוארת להלן. לכל אחד מהמעברים רשום בטבלה בכן/לא האם נדרשת פעולה על ה bus. במדה וכן פרט במקום המתאים בטבלה מהי בדיוק הפעולה הנדרשת. במדה ולא נמק במקום המתאים בטבלה מדוע.



|  |  |  |
| --- | --- | --- |
| Detailed operation | Bus required? |  |
| 1. במדה וה block נמצא, מאחר והמצב הנוכחי invalid, קריאת נתונים איננה יכולה להיענות מהcache המקומי. 2. יתכן גם והבלוק איננו נמצא.  יש להודיע read miss על ה bus. זה גורר פעולתcache רגילה של החלפת הבלוק. 1. הבאתו מה cache שנמצא במצב exclusive, או 2. הבאתו מהזיכרון הראשי. בכל מקרה מצבו של ה block משתנה ל shared. |  | 1 |
| 1. במדה וה block נמצא, מאחר והמצב הנוכחי invalid, כתיבת נתונים איננה יכולה להיעשות ב cache המקומי. 2. יתכן גם והבלוק איננו נמצא.  יש להודיע write miss על ה bus. זה גורר פעולתcache רגילה של החלפת הבלוק. 1. הבאתו מה cache שנמצא במצב exclusive, או 2. הבאתו מהזיכרון הראשי. בכל מקרה מצבו של ה block משתנה ל exclusive. | כן | 2 |
| קרא מהזיכרון המקומי. מצב הזכרון נשאר shared. | לא | 3 |
| מאחר ומצב הblock הינו shared, פירוש הדבר שזהו miss רגיל כמו ב uniprocessor. צריך להודיע read miss על ה bus, להביא את הבלוק מהזיכרון הראשי, ולבצע replacement רגיל. מצב ה block עובר ל shared. | כן | 4 |
| מאחר והייתה כתיבה ל block (עקב hit), מצבו הופך ל exclusive, ויש להודיע על ה bus לכל הזכרונות האחרים המחזיקים באותו block (במדה וישנם כאלה) שעלים להפוך ל invalid. | כן | 5 |
| מאחר וה block במצב shared, פירוש הדבר שזהו miss רגיל. יש להודיע miss write על ה bus ולהביא את ה block המבוקש מהזיכרון. מאחר ומדובר בכתיבה, מצבו נקבע ל exclusive. | כן | 6 |
| מאחר ומצב הblock הינו exclusive, פירוש הדבר שזהו miss רגיל כמו ב uniprocessor. צריך להודיע write miss על ה bus, לבצע writeback לזיכרון הראשי (במדה וה block מוחלף) ולהביא את הבלוק מהזיכרון הראשי, ולבצע replacement רגיל. מצב ה block עובר ל shared. | כן | 7 |
| זהו read hit רגיל. הבקשה נענית מה block המקומי משום שהוא במצב exclusive. מצבו נשאר exclusive. | לא | 8 |
| זהו write hit רגיל. הכתיבה נעשית ל block המקומי משום שהוא במצב exclusive. מצבו נשאר exclusive. | לא | 9 |
| מאחר וה block במצב exclusive והתרחש write miss, יש להביא את הblock מהזיכרון. יש כמובן לבצע לפניכן writeback לזיכרון הראשי (במדה וה block מוחלף). מצב ה block נקבע exclusive. | כן | 10 |

1. (10 נקודות) נתונה מכונת מצבים הנענית לבקשות של הbus, בנתונים של סעיף 1. רשום בטבלה שלהלן לכל אחד מהמעברים בכן/לא האם נדרשות פעולות הקשורות לזכרון הן **למעבד הנוכחי** והן **לזכרון הראשי**. במדה וכן, פרט במקום המתאים בטבלה מהי בדיוק הפעולה הנדרשת. במדה ולא, נמק במקום המתאים בטבלה מדוע.



בטבלה השתמשנו במושג cache ו block במשמעות דומה.

|  |  |  |
| --- | --- | --- |
| Detailed operation | Mem operation required? |  |
| זו פעולת coherence. ה block היה במצב shared. עקב write hit ב block אחר, מצב הבלוק הנוכחי הופך ל invalid. | לא | 1 |
| זו פעולת coherence. לblock הנוכחי נכתב ב cache אחר עובדה ששודרה על ה bus בהודעת ה write miss. מצבו על כן הופך ל invalid. | לא | 2 |
| זו פעולת coherence. על הbus מופיע write miss שקרה ב cache אחר עבור ה block הנוכחי (במקום האחר ה block היה invalid). ה cache הנוכחי מספק את ה miss ע"י writeback והפניה לזיכרון הראשי שהופיעה על ה bus מבוטלת. מאחר והכתיבה נעשתה בcache אחר, ה cache הנוכחי הופך ל invalid. | כן | 3 |
| על ה bus מופיע read miss שקרה ב cache אחר ל block שאיננו הנוכחי. ה miss טופל במקום האחר וזה לא נוגע ל block הנוכחי, כך שלא נדרשת שום פעולת זכרון וגם לא שינוי מצב. | לא | 4 |
| זו פעולת coherence. על הbus מופיע read miss שקרה ב cache אחר עבור ה block הנוכחי (במקום האחר ה block היה invalid). ה cache הנוכחי מספק את ה miss ע"י writeback (כל ה caches האחרים מסונכרנים) והפניה לזיכרון הראשי שהופיע על ה bus מבוטלת. מצב כל ה caches הופך ל shared. | כן | 5 |

1. (5 נקודות) נתון מעבד בעל ארבע ליבות, ולכל ליבה זכרון מטמון משלה עם bus משותף. זכרון המטמון הינו 4KByte מסוג direct map, גודל בלוק 8 בתים. המערכת שומרת על קוהרנטיות הזכרון ע"י שימוש בפרוטוקול שבסעיפים א' וב' .בזמן מסויים מעבדים P1 ו P2 מכילים את הבלוק שכתובתו0xAB7121C0 ואילו המעבדים P3 ו P4 מכילים את הבלוק שכתובתו 0xAC5121C0. מצב מכונות המצבים ככתוב בטבלה שלהלן כאשר לכל הכניסות אינדקס משותף.

* השלימו את ערכי ה tag בטבלה **והסבירו את תשובתכם**.
* כמה סיביות בסה"כ נדרשות למימוש מכונות המצבים הנ"ל?

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P4 cache | | P3 cache | | P2 cache | | P1 cache | |
| tag | state | tag | state | tag | state | tag | state |
| 0xAC512 | shared | 0xAC512 | shared | 0xAB712 | shared | 0xAB712 | shared |
| **הסבר:** 8 בתים בבלוק דורשים 3 סיביות כתובת. לזיכרון 4KByte וזה גורר 512 בלוקים, כלומר 9 סיביות לאינדקס. זה מותיר 2032-(9+3)= סיביות ל tag. | | | | | | | |

לפניכם סדרה של פניות לזיכרון. יש לפרט את המתרחש במערכת. בכל אחד מהסעיפים נדרש:

* למלא את הטבלה המתארת את מצב הזכרונות בסיום כל פעולות הזכרון הנדרשות בהתאם ולנמק בשורות שמתחת לטבלה את התשובה.
* לציין במפורש האם נדרש write-back ואם כן, על ידי איזה מעבד.

1. (5 נקודות) בהמשך פנה מעבד P3 וביקש לכתוב את המילה בכתובת0xAB7121C4.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P4 cache | | P3 cache | | P2 cache | | P1 cache | |
| tag | state | tag | state | tag | state | tag | state |
| 0xAC512 | shared | 0xAB712 | exclusive | 0xAB712 | invalid | 0xAB712 | invalid |
| **הסבר:** הבלוק הנדרש כתוצאה מהמלה בכתובת0xAB7121C4 איננו ב P3, אבל פנייה על ה bus מוצאת אותה ב P1 ו P2. מאחר והבלוק שהיה ב P3 איננו "מלוכלך", לא נדרש write-back. הבלוק שהיה ב P3 מועף ואת מקומו תופס הבלוק שכתובתו 0xAB7121C0 וה tag מעודכן בהתאם. מאחר ו P3 ליכלך אותו, מצבו עובר ל exclusive, ואילו העותקים ב P1 ו P2 הופכים ל invalid. | | | | | | | |

1. (5 נקודות) בהמשך פנה מעבד P2 וביקש לכתוב את המילה ב כתובת 0xAB7121C8.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P4 cache | | P3 cache | | P2 cache | | P1 cache | |
| tag | state | tag | state | tag | state | tag | state |
| 0xAC512 | shared | 0xAB712 | invalid | 0xAB712 | exclusive | 0xAB712 | invalid |
| **הסבר:** כתובת 0xAB7121C8 אמנם נמצאת ב P2 אבל הבלוק invalid. הבלוק המעודכן נמצא ב P3, שחייב כעת לבצע write-back ל bus בלבד (שימו לב שאין צורך לעדכן את הזכרון הראשי, מדוע?) ועובר למצב invalid. P2 מקבל את הבלוק המעודכן על ה bus, כותב לתוכו והופך לבעל הבית, ולכן מצבו עובר ל exclusive. | | | | | | | |

1. (5 נקודות) בהמשך פנה מעבד P1 וביקש לקרא את המילה ב כתובת 0xAB7121F0 .

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P4 cache | | P3 cache | | P2 cache | | P1 cache | |
| tag | state | tag | state | tag | state | tag | state |
| 0xAC512 | shared | 0xAB712 | invalid | 0xAB712 | exclusive | 0xAB712 | invalid |
| **הסבר:** מאומה לא יקרה! מדובר בכלל באינדקס אחר. | | | | | | | |

1. (5 נקודות) בהמשך פנה מעבד P2 וביקש לכתוב את המילה ב כתובת 0xCC7121C8.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P4 cache | | P3 cache | | P2 cache | | P1 cache | |
| tag | state | tag | state | tag | state | tag | state |
| 0xAC512 | shared | 0xAB712 | invalid | 0xCC712 | exclusive | 0xAB712 | invalid |
| **הסבר**: יש מצב של write miss ונדרש לפנות מקום בcache למידע החדש, ולכן מתבצע back של המידע לזיכרון הראשי כיוון שזה המידע המעודכן שלא נמצא במקום אחר. נשים לב שהמצב של הזיכרונות נשאר זהה כיוון שמידע החדש מופיע עדיין רק ב P2. | | | | | | | |

1. (5 נקודות) בהמשך פנה מעבד P4 וביקש לקרוא את המילה ב כתובת 0xCC7121C8 .

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P4 cache | | P3 cache | | P2 cache | | P1 cache | |
| tag | state | tag | state | tag | state | tag | state |
| 0xCC712 | shared | 0xAB712 | invalid | 0xCC712 | shared | 0xAB712 | invalid |
| **הסבר**: עכשיו המידע של הכתובת 0xCC7121C8 מופיע בשני מקומות ולכן לא ניתן לעשות בו שינוי ללא הודעה לbus על כך ולכן מעבירים גם את P2 למצב של קריאה בלבד .shared | | | | | | | |