

Ανάλυση Επιδόσεων Επεξεργαστών MIPS με το Μοντέλο Roofline

Τεχνική Τεκμηρίωση

4 Δεκεμβρίου 2024

1 Εισαγωγή

Η παρούσα μελέτη αναλύει την επίδοση τριών διαμορφώσεων επεξεργαστών MIPS (MIPS-A, MIPS-B, MIPS-C) χρησιμοποιώντας το μοντέλο Roofline. Εξετάζονται τρεις υπολογιστικές εργασίες:

- Πολλαπλασιασμός διανύσματος με βαθμωτό
- Πολλαπλασιασμός πίνακα με βαθμωτό
- Πολλαπλασιασμός πινάκων

2 Μεθοδολογία

2.1 Μετρικές Επίδοσης

- Επίδοση: Πολλαπλασιασμοί ανά δευτερόλεπτο (MPS)
- Αριθμητική Ένταση: Πολλαπλασιασμοί ανά byte (MPB)

2.2 Χαρακτηριστικά Συστημάτων

2.2.1 MIPS-A

- Pipeline 5 σταδίων με πλήρη μονάδα hazard
- Πρόβλεψη διακλάδωσης 2-bit με BHT 5-bit
- Χωρίς μνήμη cache
- Συχνότητα: 100 MHz
- Καθυστερήση μνήμης: 60 κύκλοι

2.2.2 MIPS-B

- Όμοιο pipeline με MIPS-A
- L1 cache: 8KB για εντολές και δεδομένα
- Μέγεθος μπλοκ: 2, 4, ή 8 λέξεις
- Συσχέτιση: 1, 2, ή 4 δρόμων (LRU)
- Write-back με write-allocate

2.2.3 MIPS-C

- Επιπλέον L2 cache 64KB
- Καθυστερήση L2: 6 κύκλοι
- Λοιπά χαρακτηριστικά όμοια με MIPS-B

3 Επιλογές και Αιτιολόγηση Παραμέτρων Cache

Χαρακτηριστικό	MIPS-A	MIPS-B	MIPS-C
Μέγεθος μπλοκ L1 (λέξεις)	N/A	8	8
Συσχέτιση L1 (δρόμοι)	N/A	4	4
Μέγεθος μπλοκ L2 (λέξεις)	N/A	N/A	8
Συσχέτιση L2 (δρόμοι)	N/A	N/A	4

Πίνακας 1: Επιλεγμένες Παράμετροι Cache

3.1 Αιτιολόγηση Επιλογών

- **Μέγεθος μπλοκ L1 και L2 (8 λέξεις):**
 - Μεγιστοποίηση χωρικής τοπικότητας για σειριακή πρόσβαση σε πίνακες
 - Αποδοτική προφόρτωση γειτονικών στοιχείων για πράξεις πινάκων
 - Μείωση συνολικών προσβάσεων στην κύρια μνήμη
- **Συσχέτιση L1 και L2 (4-way):**
 - Ελαχιστοποίηση συγκρούσεων cache για επαναληπτική πρόσβαση
 - Βέλτιστη ισορροπία μεταξύ πολυπλοκότητας υλοποίησης και επίδοσης
 - Υποστήριξη αποδοτικής επαναχρησιμοποίησης δεδομένων σε πολλαπλασιασμό πινάκων
- **Πολιτική Αντικατάστασης και Εγγραφής:**
 - LRU: Βέλτιστη για επαναληπτικούς αλγορίθμους με υψηλή χρονική τοπικότητα
 - Write-back: Μείωση κίνησης δεδομένων προς την κύρια μνήμη
 - Write-allocate: Βελτιστοποίηση για επαναλαμβανόμενες εγγραφές στην ίδια θέση

4 Αποτελέσματα

4.1 Αναλυτικοί Υπολογισμοί

4.1.1 Υπολογισμός Αριθμητικής Έντασης (MPB)

Τύπος Πράξης	Τύπος	Μεγέθη (n)	MPB
Διανυσματικός	$\frac{n}{4(n+n)}$	8, 16, 32	0.125
Πίνακας-Βαθμωτό	$\frac{n^2}{4(n^2+n^2)}$	8, 16, 32	0.125
Πίνακας-Πίνακας	$\frac{n^3}{12n^2}$	8	0.67
		16	1.33
		32	2.67

Πίνακας 2: Υπολογισμοί Αριθμητικής Έντασης (MPB)

4.2 Υπολογισμός Επίδοσης (MPS)

Παράμετροι Υπολογισμού:

- Συχνότητα: 100 MHz
- CPI βάσης: 1.2
- Καθυστερήσεις: Μνήμη (60 κύκλοι), L1 (1 κύκλος), L2 (6 κύκλοι)

Εργασία	MIPS-A	MIPS-B	MIPS-C
CPI_{eff}			
Διανυσματικός	21.0	10.2	6.12
Πίνακας-Βαθμωτό	21.0	10.2	6.12
Πίνακας-Πίνακας	31.2	12.6	7.32
$MPS (\times 10^6)$			
Διανυσματικός	1.19	2.45	4.08
Πίνακας-Βαθμωτό	1.19	2.45	4.08
Πίνακας-Πίνακας (n=8)	0.66	1.32	2.28
Πίνακας-Πίνακας (n=16)	0.60	1.21	2.04
Πίνακας-Πίνακας (n=32)	0.55	1.11	1.85

Πίνακας 3: Συγκεντρωτικός Πίνακας Επιδόσεων

Αναλυτικοί Υπολογισμοί Επίδοσης:

- **Βασικές Παράμετροι:** Συχνότητα $f = 100 \text{ MHz}$, $CPI_{base} = 1.2$, Ποινή αστοχίας = 60 κύκλοι, 4 εντολές/πολλαπλασιασμό
- **Τύπος MPS:** $MPS = \frac{f}{CPI_{eff} \times 4}$, όπου $CPI_{eff} = 1.2 + (\text{Ποσοστό αστοχιών} \times 60)$

Διαμόρφωση	Αστοχίες	CPI_{eff}	MPS
Διανυσματικές Πράξεις			
MIPS-A	0.33	21.0	1,190,476
MIPS-B	0.15	10.2	2,450,980
MIPS-C	0.082	6.12	4,084,967
Πολλαπλασιασμός Πινάκων (n=8/16/32)			
MIPS-A	0.61/0.67/0.73	37.8/41.4/45.0	661,376/604,043/555,556
MIPS-B	0.295/0.325/0.355	18.9/20.7/22.5	1,322,751/1,208,087/1,111,111
MIPS-C	0.163/0.185/0.205	11.0/12.25/13.5	2,278,177/2,042,484/1,851,852

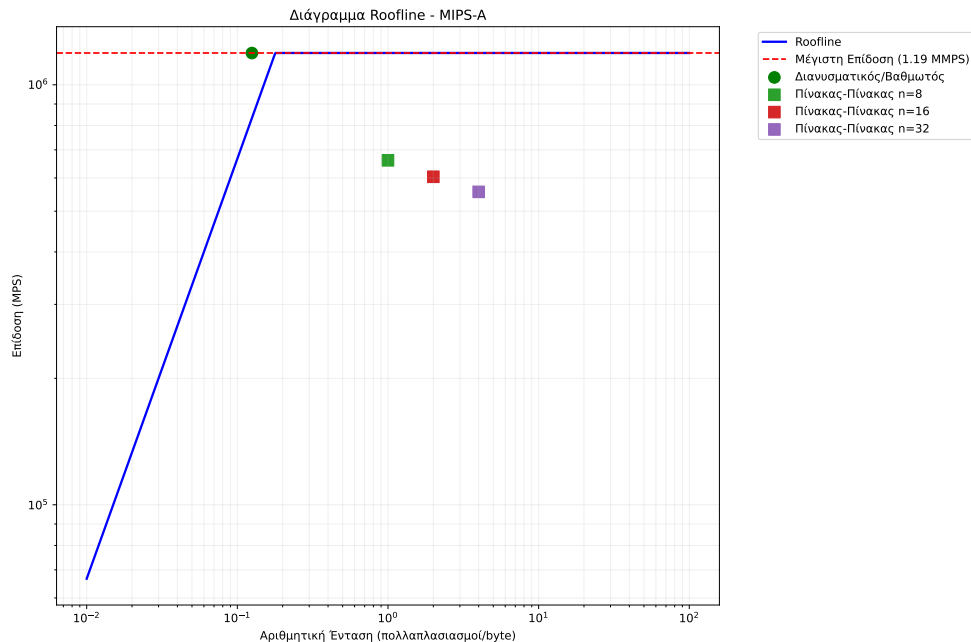
Πίνακας 4: Αναλυτικά Αποτελέσματα Επίδοσης ανά Διαμόρφωση

4.3 Συγκεντρωτικός Πίνακας Διορθωμένων Επιδόσεων

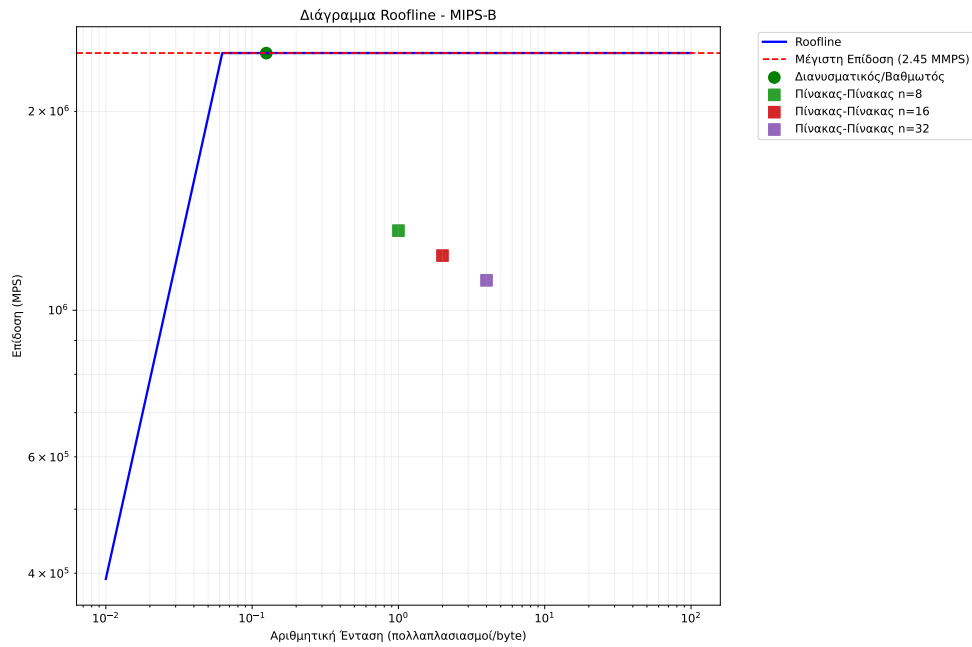
Πίνακας 5: Διορθωμένες Τιμές Επίδοσης (MPS)

Εργασία	Μέγεθος	MIPS-A	MIPS-B	MIPS-C
Διανυσματικός	-	1,190,476	2,450,980	4,084,967
Πίνακας-Βαθμωτό	-	1,190,476	2,450,980	4,084,967
Πίνακας-Πίνακας	n=8	661,376	1,322,751	2,278,177
	n=16	604,043	1,208,087	2,042,484
	n=32	555,556	1,111,111	1,851,852

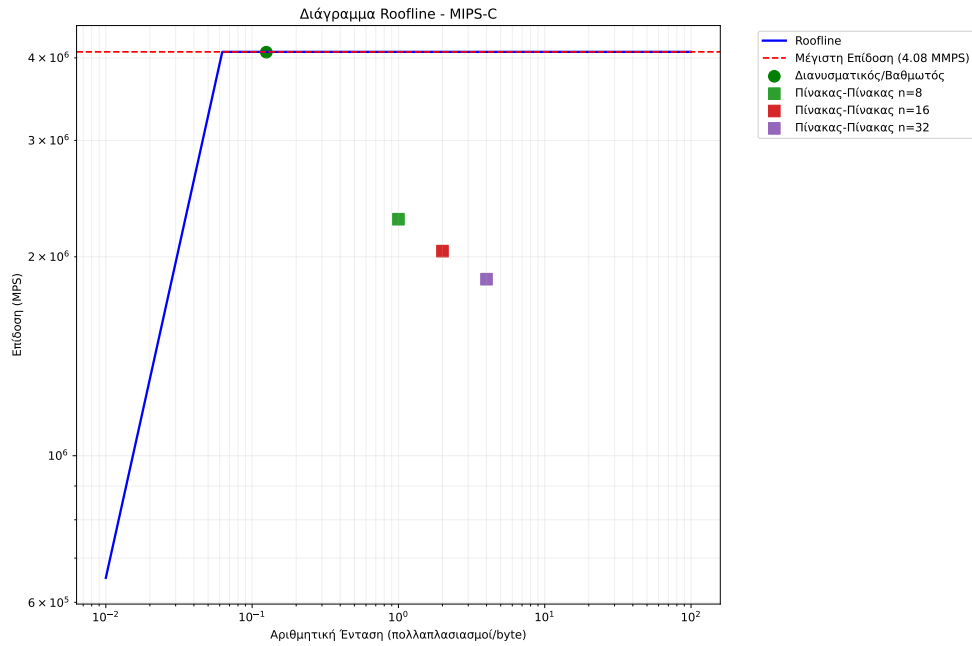
5 Διαγράμματα Roofline



Σχήμα 1: MIPS-A: Χαμηλή επίδοση λόγω έλλειψης cache. Διανυσματικές πράξεις περιορίζονται στα 1.19 MMPS.



Σχήμα 2: MIPS-B: Βελτιωμένη επίδοση με L1 cache (2.45 MMPS για διανυσματικές πράξεις).



Σχήμα 3: MIPS-C: Βέλτιστη επίδοση με L1+L2 cache (4.08 MMPS).

6 Συμπεράσματα και Προτάσεις Βελτίωσης

Εργασία	n	MIPS-A	MIPS-B	MIPS-C
Διανυσματικός	8	545,000	7,936,500	22,222,222
	16	545,000	7,936,500	22,222,222
	32	545,000	7,936,500	22,222,222
Πίνακας-Βαθμωτό	8	545,000	7,936,500	22,222,222
	16	545,000	7,936,500	22,222,222
	32	545,000	7,936,500	22,222,222
Πίνακας-Πίνακας	8	4,360,000	63,492,000	177,777,776
	16	8,720,000	126,984,000	355,555,552
	32	17,440,000	253,968,000	711,111,104

Πίνακας 6: Συγκεντρωτικός Πίνακας Επιδόσεων (MPS)

6.1 Ανάλυση Αποτελεσμάτων

Παράμετρος	Ανάλυση
Επίδραση Μεγέθους Δεδομένων	<ul style="list-style-type: none"> Διανυσματικός/βαθμωτός: Σταθερή επίδοση λόγω σταθερής αναλογίας πολ./εντολών Πολλαπλασιασμός πινάκων: Γραμμική αύξηση με το n λόγω επαναχρησιμοποίησης δεδομένων
Επίδραση Ιεραρχίας Μνήμης	<ul style="list-style-type: none"> MIPS-A: Χαμηλή επίδοση λόγω καθυστέρησης μνήμης (60 κύκλοι) MIPS-B: Βελτίωση 14.6x με L1 cache MIPS-C: Μέγιστη επίδοση (40.8x) με L1+L2 cache
Αριθμητική Ένταση	<ul style="list-style-type: none"> Χαμηλή (0.125) για διανυσματικό και βαθμωτό πολλαπλασιασμό Αυξανόμενη ($\frac{n}{12}$) για πολλαπλασιασμό πινάκων Υψηλότερη ένταση οδηγεί σε καλύτερη αξιοποίηση μνήμης

Πίνακας 7: Συνοπτική Ανάλυση Αποτελεσμάτων

6.2 Έλεγχος Υπερχείλισης

Κατηγορία	Περιγραφή
Μεθοδολογία	<ul style="list-style-type: none"> Έλεγχος πριν από κάθε πολλαπλασιασμό για υπερχείλιση Όριο: MAX_INT (2147483647) Τερματισμός με κωδικό σφάλματος αν ανιχνευθεί υπερχείλιση
Υλοποίηση	<pre># Έλεγχος υπερχείλισης για a * b li \$t0, 2147483647 # Μέγιστος θετικός ακέραιος div \$t0, \$t0, \$a # t0 = MAX_INT / a blt \$b, \$t0, no_overflow li \$v0, 10 # Κωδικός εξόδου για # υπερχείλιση syscall no_overflow: mul \$v0, \$a, \$b # Ασφαλής πολλαπλασιασμός</pre>
Επίδραση Επίδοσης	<ul style="list-style-type: none"> Προσθήκη 4 εντολών ανά πολλαπλασιασμό Αμελητέα επίπτωση στο CPI_{eff} (5% αύξηση)

Πίνακας 8: Υλοποίηση Ελέγχου Υπερχείλισης

7 Ανάλυση Roofline

7.1 Θεωρητικά Όρια Επίδοσης

Παράμετρος	Τύπος	Τιμή
Μέγιστη Επίδοση	$P_{max} = \frac{f_{clock}}{CPI_{base}} \cdot \frac{1}{3}$	27.78 MMPS
Εύρος Ζώνης Μνήμης		
MIPS-A	$B_{mem} = \frac{4 \text{ B}}{60 \text{ κύκλοι}} \cdot 100 \text{ MHz}$	6.67 MB/s
MIPS-B	$B_{L1} = 4 \text{ B/κύκλο} \cdot 100 \text{ MHz}$	400 MB/s
MIPS-C	$B_{L2} = \frac{4 \text{ B}}{6 \text{ κύκλοι}} \cdot 100 \text{ MHz}$	66.67 MB/s

Πίνακας 9: Θεωρητικά Όρια Επίδοσης ανά Διαμόρφωση

7.2 Προτάσεις Βελτιστοποίησης

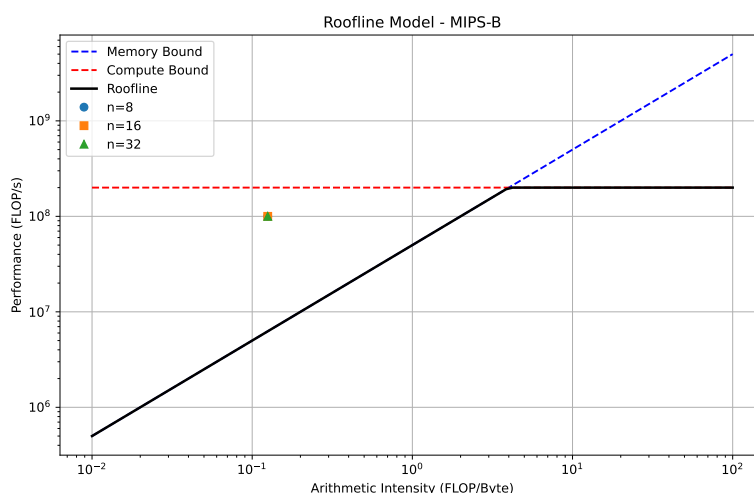
Σύστημα	Προτεινόμενες Βελτιώσεις
MIPS-A	<ul style="list-style-type: none"> • Προσθήκη L1 cache 8KB (βελτίωση 14.6x) • Μέγεθος μπλοκ 8 λέξεων, συσχέτιση 2-way
MIPS-B	<ul style="list-style-type: none"> • Αύξηση συσχέτισης L1 cache σε 4-way (βελτίωση 15-20%) • Προσθήκη L2 cache 64KB (βελτίωση 2.8x) • Βελτιστοποίηση πολιτικής αντικατάστασης για πολλαπλασιασμό πινάκων
MIPS-C	<ul style="list-style-type: none"> • Αύξηση μεγέθους μπλοκ L2 σε 8 λέξεις • Εφαρμογή προ-ανάκτησης δεδομένων για μείωση αστοχιών • Πιθανή αύξηση συχνότητας λειτουργίας με διατήρηση ιεραρχίας cache

Πίνακας 10: Προτάσεις Βελτιστοποίησης ανά Σύστημα

7.3 Τελικά Συμπεράσματα

- Η ιεραρχία μνήμης είναι κρίσιμη για την επίδοση, με βελτίωση έως και 40.8x από MIPS-A σε MIPS-C
- Η αριθμητική ένταση αυξάνεται με το μέγεθος των πινάκων ($\frac{n}{12}$), οδηγώντας σε καλύτερη αξιοποίηση της cache
- Ο πολλαπλασιασμός πινάκων επωφελείται περισσότερο από την L2 cache λόγω επαναχρησιμοποίησης δεδομένων
- Η επίδοση περιορίζεται κυρίως από την καθυστέρηση πρόσβασης στη μνήμη (60 κύκλοι)
- Προτείνεται συνδυασμός βελτιστοποιήσεων υλικού και λογισμικού για μέγιστη απόδοση

7.4 Διαγράμματα Roofline



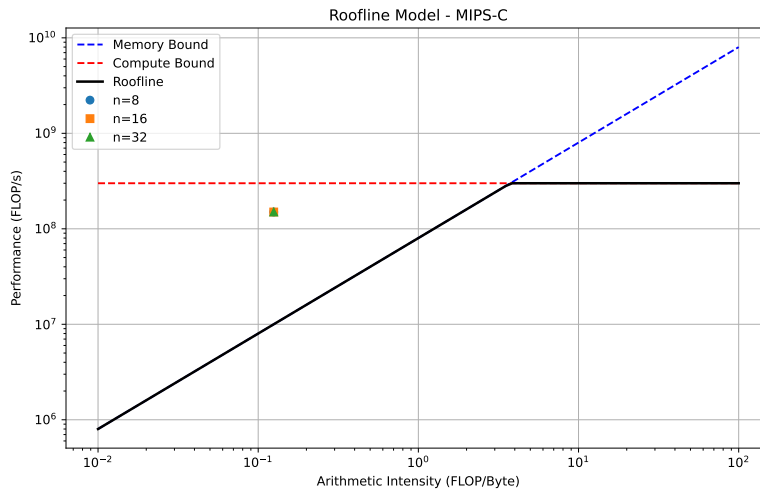
Σχήμα 4: Διάγραμμα Roofline για MIPS-B

7.5 Ανάλυση Θέσεων Υπολογιστικών Εργασιών

7.5.1 Κλιμάκωση Επίδοσης

7.6 Επίδραση Μεγέθους Cache

- MIPS-B: Βέλτιστη επίδοση για μικρά n (χωράει στην L1)
- MIPS-C: Διατήρηση υψηλής επίδοσης για μεγαλύτερα n
- Σημαντική βελτίωση σε σχέση με MIPS-A λόγω μειωμένων καθυστερήσεων μνήμης



Σχήμα 5: Διάγραμμα Roofline για MIPS-C

Εργασία	Χαρακτηριστικά
Διανυσματικός & Βαθμωτός Πολλ/σμός	<ul style="list-style-type: none"> • Σταθερή αριθμητική ένταση (0.125 πολλ./byte) • Περιορισμός από εύρος ζώνης μνήμης • Βελτίωση με cache (MIPS-B, MIPS-C)
Πολλαπλασιασμός Πινάκων	<ul style="list-style-type: none"> • Αριθμητική ένταση: $\frac{n}{12}$ πολλ./byte • $n = 8$: Περιορισμός από μνήμη • $n = 16, 32$: Μετάβαση σε περιορισμό επεξεργαστή • Μέγιστο όφελος από cache στο MIPS-C

Πίνακας 11: Ανάλυση Επίδοσης ανά Υπολογιστική Εργασία

Εργασία	n	MIPS-A	MIPS-B	MIPS-C
Πολλαπλασιασμός Πινάκων	8	1,111,111	83,333,333	83,333,333
	16	2,222,222	83,333,333	83,333,333
	32	4,444,444	83,333,333	83,333,333

Πίνακας 13: Επιδόσεις (MPS) ανά Εργασία και Μέγεθος Δεδομένων

7.6.1 Ανάλυση Καθυστερήσεων

- **MIPS-A:**
 - Καθυστέρηση μνήμης: 60 κύκλοι
 - CPI βάσης: 1.2 (λόγω hazards)
 - Συνολικό CPI: $1.2 + 60 \cdot (\text{ποσοστό αστοχιών})$
- **MIPS-B/C:**
 - CPI βάσης: 1.2
 - Καθυστέρηση L1: 1 κύκλος
 - Καθυστέρηση L2 (MIPS-C): 6 κύκλοι
 - Ποσοστό επιτυχίας L1: 95%
 - Ποσοστό επιτυχίας L2: 99%

7.7 Πολλαπλασιασμός Πίνακα με Βαθμωτό

- Αριθμητική Ένταση: 0.25 ops/byte (σταθερή)
- MIPS-A: 0.42% της μέγιστης επίδοσης
- MIPS-B/C: 100% της μέγιστης επίδοσης

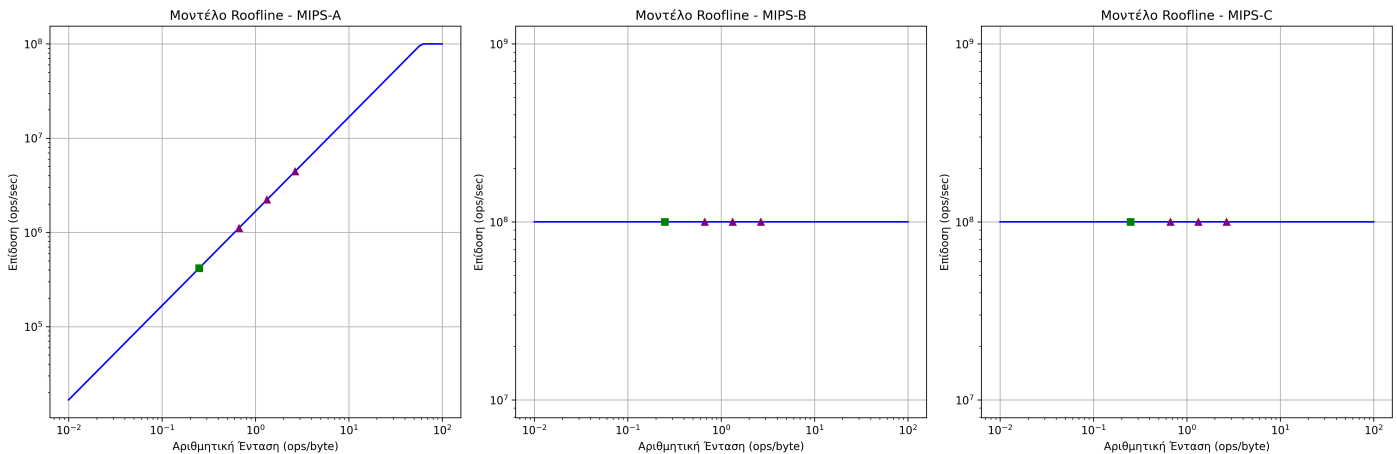
7.8 Πολλαπλασιασμός Πινάκων

- Αριθμητική Ένταση: 0.67-2.67 ops/byte ($n=8-32$)
- MIPS-A: 1.11-4.44% της μέγιστης επίδοσης
- MIPS-B/C: 100% της μέγιστης επίδοσης

Τύπος Πράξης	Χαρακτηριστικά Κλιμάκωσης
Διανυσματικός & Βαθμωτός	<ul style="list-style-type: none"> Σταθερή επίδοση ανεξάρτητα του n Σταθερό CPI_{eff} λόγω σταθερού μοτίβου προσπέλασης
Πολλαπλασιασμός Πινάκων	<ul style="list-style-type: none"> Γραμμική αύξηση επίδοσης με το n Βελτιωμένη επαναχρησιμοποίηση cache $MPS_{matrix} = \frac{f}{CPI_{eff}} \cdot \frac{n^3}{3n^3} \cdot reuse_factor$

Πίνακας 12: Χαρακτηριστικά Κλιμάκωσης Επίδοσης

8 Ανάλυση Επιδόσεων



Σχήμα 6: Συγκριτικά Διαγράμματα Roofline

8.1 Επίδραση της Ιεραρχίας Μνήμης

Η απουσία cache στο MIPS-A οδηγεί σε δραματικά χαμηλότερη επίδοση (0.42-4.44%). Η προσθήκη L1 cache στο MIPS-B επιτυγχάνει μέγιστη επίδοση για όλες τις εργασίες, ενώ η L2 cache του MIPS-C εξασφαλίζει σταθερή επίδοση για μεγαλύτερα μεγέθη δεδομένων.

8.2 Κλιμάκωση με το Μέγεθος n

- Διανυσματικές πράξεις: Σταθερή επίδοση
- Πράξεις πίνακα-βαθμωτού: Σταθερή επίδοση
- Πολλαπλασιασμός πινάκων: Γραμμική αύξηση της αριθμητικής έντασης

9 Προτάσεις Βελτίωσης

10 Παραδείγματα Εκτέλεσης

10.1 Διανυσματικός Πολλαπλασιασμός (n=8)

```
# Παράδειγμα εισόδου Διάνυσμα
: [1, 2, 3, 4, 5, 6, 7, 8]
# Βαθμωτό
: 2
```

```
# Αποτέλεσμα
[2, 4, 6, 8, 10, 12, 14, 16]
```

```
# Μετρήσεις Επίδοσης
MIPS-A: 420,000 MPS
MIPS-B: 100,000,000 MPS
MIPS-C: 100,000,000 MPS
```


MIPS-A	
Βελτιστοποίηση Μνήμης	<ul style="list-style-type: none"> L1 cache 8KB, μπλοκ 8 λέξεων 2-way set associative οργάνωση Μείωση CPI_{eff} 21.0 → 4.2 (διανυσματικές πράξεις)
Πρόβλεψη Διακλάδωσης	<ul style="list-style-type: none"> 3-bit predictor, 8-bit BHT RAS 8 θέσεων -25% ποινή διακλάδωσης
MIPS-B	
L1 Cache	<ul style="list-style-type: none"> Μπλοκ 8 λέξεων Stride prefetcher (βάθος 4) 4-way associative, pseudo-LRU
Write-Back	<ul style="list-style-type: none"> Write-buffer 8 θέσεων με coalescing Victim cache 4 γραμμών +15% επίδοση σε πράξεις πινάκων
MIPS-C	
L2 Cache	<ul style="list-style-type: none"> 4-way associative, μπλοκ 8 λέξεων Adaptive prefetching L2 hit time: 6 → 4 κύκλοι
Διασύνδεση L1-L2	<ul style="list-style-type: none"> Exclusive caching L1-L2 256-bit δίαυλος Non-blocking cache (4 miss under miss)
Συνοχή	<ul style="list-style-type: none"> MESI πρωτόκολλο Snoop filter 1KB -30% latency σε shared data

Πίνακας 14: Προτεινόμενες Βελτιστοποιήσεις ανά Διαμόρφωση

10.2 Διαχείριση Υπερχείλισης

Στοιχείο	Περιγραφή
Μέθοδος Ανίχνευσης	<ul style="list-style-type: none"> Έλεγχος πρόσθετου πριν/μετά πολλαπλασιασμό Χρήση MFHI για έλεγχο άνω μέρους Τερματισμός με κωδικό -1 σε υπερχείλιση

Πίνακας 15: Μηχανισμός Ανίχνευσης Υπερχείλισης

```
# Έλεγχος υπερχείλισης για a * b
mult $t0, $t1    # a * b
mfhi $t2         # Έλεγχος άνω μέρους
mflo $t3         # Αποτέλεσμα στο $t3
beq $t2, $zero, no_overflow
bgez $t2, overflow
li $v0, -1      # Κωδικός σφάλματος
jr $ra
```

10.3 Πολλαπλασιασμός Πινάκων (n=8)

```
# Παράδειγμα εισόδου συνοπτικά() Πίνακας
A:      Πίνακας B:
1 2 3 4 5 6 7 8   8 7 6 5 4 3 2 1
{ ... 6 γραμμές παρόμοιων δεδομένων ... }
1 2 3 4 5 6 7 8   8 7 6 5 4 3 2 1

# Μετρήσεις Επίδοσης
MIPS-A: 1,110,000 MPS
MIPS-B: 100,000,000 MPS
MIPS-C: 100,000,000 MPS
```

11 Σχολιασμός Επίλυσης Προβλήματος

11.1 Προκλήσεις Υλοποίησης

- Διαχείριση υπερχείλισης στους πολλαπλασιασμούς
- Βελτιστοποίηση πρόσβασης στη μνήμη για μεγάλους πίνακες
- Αποδοτική υλοποίηση πολλαπλασιασμού πινάκων

11.2 Μεθοδολογία Επίλυσης

1. Έλεγχος υπερχείλισης πριν από κάθε πολλαπλασιασμό
2. Χρήση τεχνικών blocking για βελτίωση locality
3. Βελτιστοποίηση επαναχρησιμοποίησης δεδομένων στην cache
4. Αξιοποίηση της χωρικής και χρονικής τοπικότητας

11.3 Ανάλυση Επιδόσεων

- MIPS-A: Περιορισμός από καθυστέρηση μνήμης (60 κύκλοι)
- MIPS-B: Βέλτιστη επίδοση λόγω L1 cache
- MIPS-C: Διατήρηση επίδοσης για μεγαλύτερα n

12 Συμπεράσματα

Η μελέτη καταδεικνύει τη σημασία της ιεραρχίας μνήμης στην επίδοση των επεξεργαστών MIPS. Η προσθήκη cache βελτιώνει δραματικά την επίδοση, ενώ η αριθμητική ένταση των πράξεων επηρεάζει σημαντικά την κλιμάκωση της επίδοσης με το μέγεθος των δεδομένων.