

Απόδοση KME

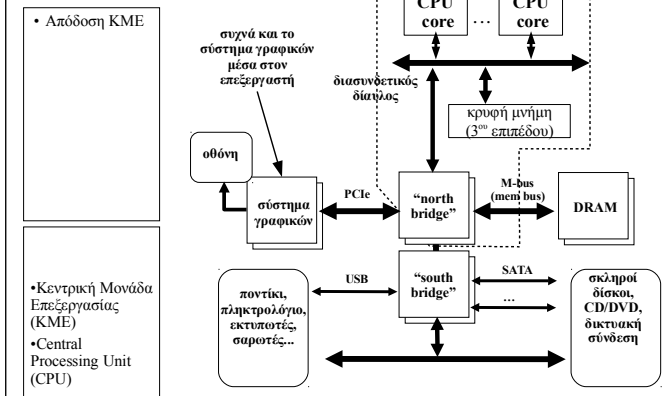
(Μέτρηση και τεχνικές βελτίωσης απόδοσης)

<http://mixstef.github.io/courses/comparch/>



Μ.Στεφανιδάκης

Κεντρική Μονάδα Επεξεργασίας



Αρχιτεκτονική Υπολογιστών – “Απόδοση KME”

2

Απόδοση KME

• Απόδοση KME

• Υπολογιστικό σύστημα

- Η απόδοση εξαρτάται από όλα τα επιμέρους τμήματά του
 - Υλικό και λογισμικό

• Κεντρική Μονάδα Επεξεργασίας

- Πόσο γρήγορα εκτελείται ένα πρόγραμμα;
 - Σε αρχιτεκτονικές ειδικού σκοπού είναι επιπλέον επιθυμητά: αξιοπιστία, κατανάλωση ενέργειας
- Πώς επηρεάζει η αρχιτεκτονική την απόδοση;
 - Πόσο γρηγορότερα εκτελείται ένα πρόγραμμα μετά από μια αρχιτεκτονική αλλαγή;

Αρχιτεκτονική Υπολογιστών – “Απόδοση KME”

3

Εκτέλεση προγραμμάτων

• Απόδοση KME

i
Response time:
Ο συνολικός χρόνος εκτέλεσης ενός προγράμματος
Μαζί με επιβάρυνση λειτουργικού συστήματος
Throughput:
Ο ρυθμός ολοκλήρωσης έργου σε συγκεκριμένο χρόνο

• Χρόνος εκτέλεσης (execution time)

- Αύξηση απόδοσης \Leftrightarrow Μείωση χρόνου εκτέλεσης

• Για υπολογιστή X:

$$\text{Απόδοση}(X) = \frac{1}{\text{Χρόνος Εκτέλεσης}(X)}$$

Αρχιτεκτονική Υπολογιστών – “Απόδοση KME”

4

Συγκρίνοντας δύο υπολογιστές

• Απόδοση ΚΜΕ

• Συγκρίνοντας αποδόσεις

- Έστω υπολογιστές X και Y
- Εάν:

$$\text{Απόδοση}(X) > \text{Απόδοση}(Y)$$

- Τότε (και αντίστροφα):

$$\text{Χρόνος Εκτέλεσης}(X) < \text{Χρόνος Εκτέλεσης}(Y)$$

;

Ο X εκτελεί ένα πρόγραμμα σε 10 sec και ο Y σε 15 sec. Πόσο πιο γρήγορος είναι ο X;

$$\frac{\text{Απόδοση}(X)}{\text{Απόδοση}(Y)} = \frac{\text{Χρόνος Εκτέλεσης}(Y)}{\text{Χρόνος Εκτέλεσης}(X)} = n$$

Ο X είναι n φορές γρηγορότερος από τον Y

Χρόνος Εκτέλεσης (Execution Time)

• Απόδοση ΚΜΕ

• Χρόνος εκτέλεσης στην ΚΜΕ

- Ο χρόνος για τον οποίο η ΚΜΕ εκτελεί εντολές του προγράμματος
 - Όχι χρόνος για αναμονή E/E ή για άλλες διεργασίες

• Συνιστώσες

- Χρόνος προγράμματος χρήστη
 - Για το πρόγραμμα καθεαυτό
- Χρόνος συστήματος
 - Λειτουργίες ΛΣ για την εξυπηρέτηση του προγράμματος

Βασικά μεγέθη μέτρησης χρόνου εκτέλεσης

• Απόδοση ΚΜΕ

• Κύκλος ρολογιού (περίοδος)

- Clock Cycle (CC)
 - Η διάρκεια ενός κύκλου ρολογιού (περίοδος) κατά τον οποίο η ΚΜΕ εκτελεί τις μικρότερες βασικές λειτουργίες
 - Σταθερό μέγεθος

• Κύκλοι ρολογιού ανά εντολή

- Clocks Per Instruction (CPI)
 - Οι απαιτούμενοι κύκλοι ρολογιού για την ολοκλήρωση μιας εντολής
 - Ενδεχομένως διαφορετικό μέγεθος ανά τύπο εντολής

;

Πώς υπολογίζεται το CPI όταν υπάρχουν διαφορετικοί τύποι εντολών;

• Αριθμός εντολών

- Instruction Count (IC)
 - Ο αριθμός των εντολών ενός προγράμματος

Χρόνος εκτέλεσης στην ΚΜΕ

• Απόδοση ΚΜΕ

• Χρόνος Εκτέλεσης για ένα πρόγραμμα

$$\text{ExecTime} = \text{IC} \times \text{CPI} \times \text{CC}$$

- Τι μπορεί να κάνει ο σχεδιαστής ΚΜΕ για να βελτιώσει την απόδοση;
 - Να μειώσει τον κύκλο ρολογιού (CC)
 - Να μειώσει τον αριθμό κύκλων ανά εντολή (CPI)
 - Ο αριθμός εντολών δεν αλλάζει

Παράδειγμα

• Απόδοση KME

Τύπος εντολής	A	B	C
CPI	1	2	3

Ακολουθία κώδικα	A	B	C
1	2	1	2
2	4	1	1

- Επιλογή μεταξύ 2 ακολουθιών εντολών
 - Ποια ακολουθία εκτελεί τις περισσότερες εντολές;
 - Ποια είναι ταχύτερη;
 - Ποιο το μέσο CPI σε κάθε περίπτωση;

[Patterson-Hennessy "Computer Organization and Design", 3rd ed]

Συσχέτιση με λογισμικό

• Απόδοση KME

- Αλγόριθμος
 - Καθορίζει το IC
 - Ενδεχομένως καθορίζει το CPI, ευνοώντας ορισμένους τύπους εντολών (π.χ. κινητής υποδιαστολής)
- Γλώσσα προγραμματισμού - Μεταγλωττιστής
 - Καθορίζει το IC (μετάφραση εντολών υψηλού επιπέδου)
 - Καθορίζει το CPI απαιτώντας/χρησιμοποιώντας συγκεκριμένους τύπους εντολών



Το υλικό (ISA) καθορίζει και τα τρία μεγέθη (IC, CPI και CC)

Μετροπρογράμματα

• Απόδοση KME

- Benchmarks**
 - Για τη μέτρηση της απόδοσης
 - Και τη σύγκριση μεταξύ υπολογιστών
 - Θα πρέπει να αντιπροσωπεύουν τις πραγματικές εφαρμογές
 - Υπό ρεαλιστικές συνθήκες εκτέλεσης και δεδομένα εισόδου
 - Χωρίς "εσωτερικές" ειδικές βελτιστοποιήσεις
 - Δυνατότητα επανάληψης μέτρησης
 - Διαφορετικά για ανόμοιες κλάσεις υπολογιστών
 - PCs, servers, embedded systems...

Ο "νόμος" του Amdahl

• Απόδοση KME

- "Η βελτίωση της συνολικής απόδοσης ενός συστήματος μέσω της εισαγωγής ενός νέου χαρακτηριστικού, περιορίζεται από το βαθμό χρήσης αυτού του νέου χαρακτηριστικού"
- Ερμηνεία – συνέπειες
 - Οι περισσότεροι χρησιμοποιούμενες περιπτώσεις πρέπει να είναι γρήγορες
 - Δεν ωφελεί η βελτιστοποίηση των σπάνιων περιπτώσεων
 - Η μη χρήση του νέου χαρακτηριστικού εμποδίζει να επιτύχουμε την «τέλεια» απόδοση

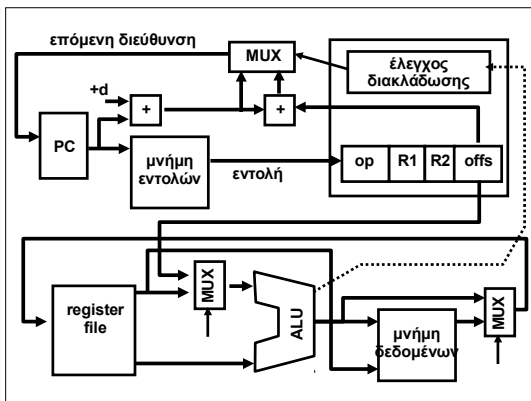


Ένα πρόγραμμα τρέχει για 100 sec σε έναν υπολογιστή και εκτελεί πολλαπλασιασμούς για 80 sec. Πόσο πρέπει να βελτιώσω τη ταχύτητα του πολλαπλασιασμού για να πενταπλασιάσω τη συνολική απόδοση;

KME ενός κύκλου (single-cycle)

- Απόδοση KME
- KME ενός κύκλου

Κάθε εντολή ολοκληρώνεται σε έναν κύκλο ρολογιού



Αρχιτεκτονική Υπολογιστών – “Απόδοση KME”

13

KME ενός κύκλου (single-cycle)

- Απόδοση KME
- KME ενός κύκλου

• CPI = 1

- Σε κάθε έναν κύκλο ρολογιού ολοκληρώνεται μια εντολή ή
- κάθε εντολή απαιτεί έναν κύκλο ρολογιού
- Πόσο πρέπει να είναι το CC;
 - Όσο με τη διάρκεια της μεγαλύτερης λειτουργίας
 - Μη αποδοτικό σχήμα
 - Όλες οι εντολές δεν απαιτούν τον ίδιο χρόνο

Αρχιτεκτονική Υπολογιστών – “Απόδοση KME”

14

Παράδειγμα

- Απόδοση KME
- KME ενός κύκλου

Εντολή	IF	ID	EX	DM	WB	Σύνολο
Αριθμητική	200	50	100	0	50	400 ps
Διακλάδωση	200	50	100	0	0	350 ps
Ανάγνωση μνήμης	200	50	100	200	50	600 ps
Εγγραφή μνήμης	200	50	100	200	0	550 ps

- CC πρέπει να είναι 600 ps (single cycle CPU)
 - Αν ήταν δυνατή η χρήση με μεταβλητό CC (προσοχή: ρακτικά αδύνατο!)
 - Ποια η βελτίωση της απόδοσης;
 - 25% ανάγνωση, 10% εγγραφή, 45% αριθμητικές, 20% διακλάδωση

[Patterson-Hennessy “Computer Organization and Design”, 3rd ed]

Αρχιτεκτονική Υπολογιστών – “Απόδοση KME”

15

KME πολλαπλών κύκλων (multi-cycle)

- Απόδοση KME
- KME ενός κύκλου
- KME πολλαπλών κύκλων

• CPI > 1

- Κάθε εντολή χωρίζεται σε έναν μεταβλητό αριθμό βημάτων
- Κάθε βήμα απαιτεί έναν κύκλο ρολογιού
- Πόσο πρέπει να είναι το CC;
 - Όσο με τη διάρκεια ολοκλήρωσης του μεγαλύτερου βήματος
- Καταχωρητές για τη συγκράτηση αποτελεσμάτων μεταξύ βημάτων
- Μέρη της KME μπορούν να χρησιμοποιηθούν για περισσότερες από μία φορές κατά την εκτέλεση μιας εντολής

Αρχιτεκτονική Υπολογιστών – “Απόδοση KME”

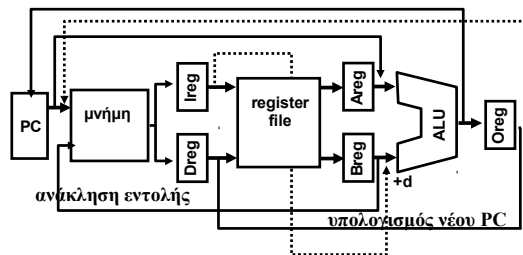
16

Παράδειγμα: Εντολή load (βήμα IF)

- Απόδοση KME
- KME ενός κύκλου
- KME πολλαπλών κύκλων



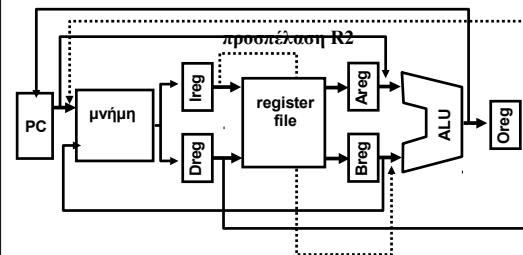
Η εντολή load απαιτεί τα περισσότερα βήματα (κύκλους ρολογιού) για να εκτελεστεί



$$R1 \leftarrow \text{mem}[R2 + \text{offset}]$$

Παράδειγμα: Εντολή load (βήμα ID)

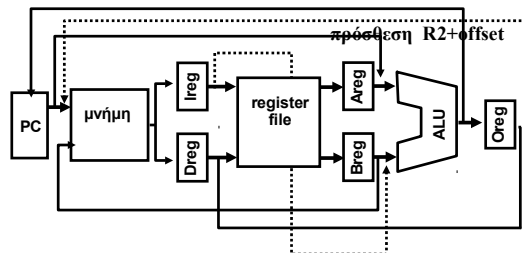
- Απόδοση KME
- KME ενός κύκλου
- KME πολλαπλών κύκλων



$$R1 \leftarrow \text{mem}[R2 + \text{offset}]$$

Παράδειγμα: Εντολή load (βήμα EX)

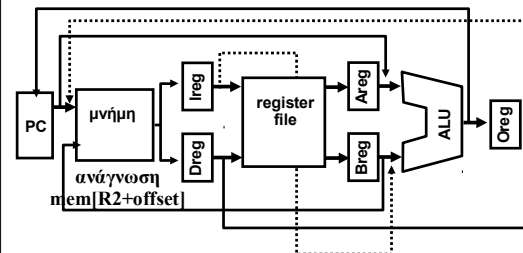
- Απόδοση KME
- KME ενός κύκλου
- KME πολλαπλών κύκλων



$$R1 \leftarrow \text{mem}[R2 + \text{offset}]$$

Παράδειγμα: Εντολή load (βήμα DM)

- Απόδοση KME
- KME ενός κύκλου
- KME πολλαπλών κύκλων



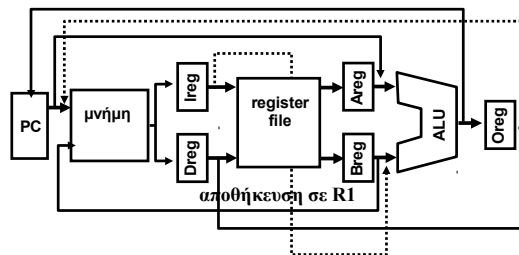
$$R1 \leftarrow \text{mem}[R2 + \text{offset}]$$

Παράδειγμα: Εντολή load (βήμα WB)

- Απόδοση KME
- KME ενός κύκλου
- KME πολλαπλών κύκλων

;

Σε κάθε βήμα, μερικά τμήματα μένουν ανενεργά. Πώς θα μπορούσαμε να τα εκμεταλλευτούμε;



$R1 \leftarrow \text{mem}[R2 + \text{offset}]$

Μονάδα Ελέγχου KME πολλαπλών κύκλων

- Απόδοση KME
- KME ενός κύκλου
- KME πολλαπλών κύκλων

- Πολυπλοκότητα σημάτων ελέγχου
 - Δημιουργία σημάτων σε κάθε βήμα εκτέλεσης
 - Ανάλογα με το είδος της εντολής
 - Διατήρηση προηγούμενης κατάστασης
- Μέθοδοι υλοποίησης
 - Αυτόματα πεπερασμένων καταστάσεων
 - Ακολουθιακά λογικά κυκλώματα
 - Παραγωγή σημάτων ελέγχου ανάλογα με εισόδους και τρέχουσα κατάσταση
 - Μικροπρόγραμμα
 - Καθορισμός σημάτων μέσω μικροεντολών
 - Εσωτερικά στην KME
 - Για υλοποίηση σύνθετων εντολών με πολλά βήματα και πολλαπλά περάσματα από το datapath
 - Μερικές φορές είναι εγγράψιμο (updates, patches..)

Απόδοση KME πολλαπλών κύκλων

- Απόδοση KME
- KME ενός κύκλου
- KME πολλαπλών κύκλων

- Πλεονεκτήματα
 - Δεν απαιτείται ο μέγιστος χρόνος για το CC
 - Μέρη της KME μπορούν να χρησιμοποιηθούν με πολλαπλό τρόπο κατά την εκτέλεση μιας εντολής
- Μειονεκτήματα
 - Η μονάδα ελέγχου γίνεται πολυπλοκότερη
 - Η πολυπλοκότητα πιθανόν να ακυρώνει τα πλεονεκτήματα
- Σήμερα
 - Οι KME υψηλής απόδοσης χρησιμοποιούν πρόσθετες τεχνικές παραλληλισμού σε επίπεδο εντολών (instruction level parallelism – ILP)
 - (στο επόμενο μάθημα...)