

Αρχιτεκτονική Υπολογιστών



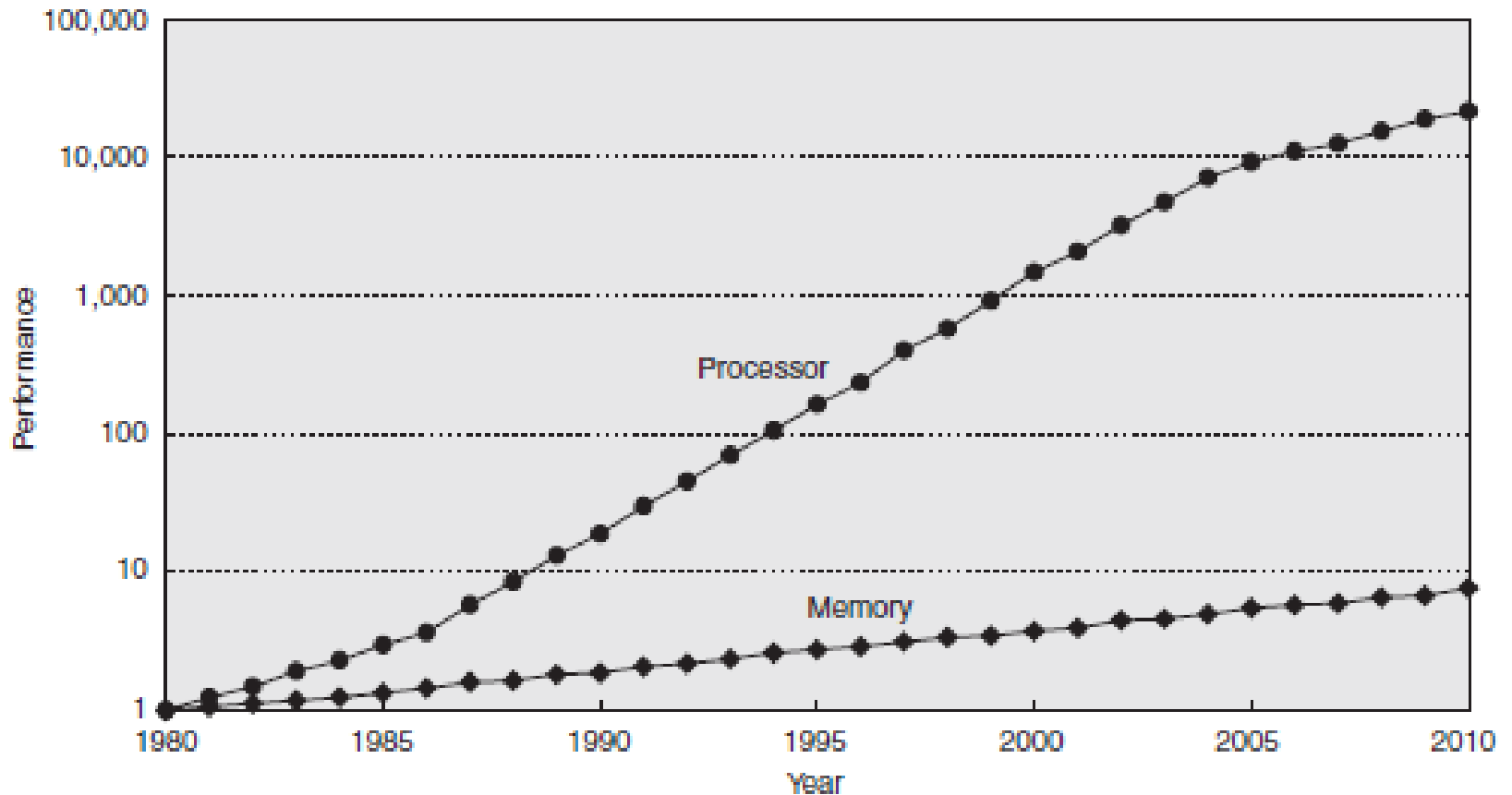
Γιώργος Δημητρίου

Ενότητα 10^η:
Ιεραρχία Μνήμης, Αποθήκευση
και Αρτηρίες

Θέματα Ιεραρχίας Μνήμης & Αρτηριών

- Κρυφή μνήμη ↔ Κύρια μνήμη
 - Βελτιστοποιήσεις κρυφής μνήμης
- Κύρια μνήμη ↔ Ιδεατή μνήμη
 - Βελτιστοποιήσεις κύριας και ιδεατής μνήμης
- Θέματα δευτερεύουσας μνήμης
 - Συστήματα δίσκων
- Αρτηρίες διασύνδεσης
 - Αρτηρίες διαχωρισμένης προσπέλασης

Απόδοση CPU έναντι DRAM



© Hennessy-Patterson

Κρυφή Μνήμη

- Μέσος χρόνος προσπέλασης =
Χρόνος επιτυχούς προσπέλασης +
Λόγος αστοχίας ×
Χρόνος προσπέλασης επόμενου επιπέδου
- Η σχεδίαση της κρυφής μνήμης γίνεται με σκοπό την ελαχιστοποίηση του μέσου χρόνου προσπέλασης
- Αυξημένη πολυπλοκότητα όταν υποστηρίζεται εκτέλεση εκτός σειράς

Μείωση Κόστους Αστοχίας

- Πολλαπλά επίπεδα κρυφής μνήμης
 - L2 και L3 επίπεδα
- Επιλογή δεδομένου από το μπλοκ
 - Προηγείται το δεδομένο που ζητείται
- Προτεραιότητα στις εντολές φόρτωσης
 - Έλεγχος για επικάλυψη διευθύνσεων
- Συγχώνευση στις αποθηκεύσεις
- Κρυφή μνήμη θυμάτων (victim cache)

Μείωση Λόγου Αστοχίας

- Αύξηση μεγέθους μπλοκ
- Αύξηση μεγέθους μνήμης
- Αύξηση βαθμού συσχέτισης
- Ψευδοσυσχετιστική κρυφή μνήμη (pseudo-associative)
- Βελτιστοποιήσεις μεταγλωττιστή

Παραλληλισμός στην Κρυφή Μνήμη

- Nonblocking caches
 - Οι προσπελάσεις συνεχίζονται σε μια αστοχία μέχρι ενός αριθμού αστοχιών
- Εκ των προτέρων προσκόμιση από το υλικό (hardware prefetching)
 - Συνήθως σε εξωτερικό χώρο (stream buffer)
- Εκ των προτέρων προσκόμιση από το λογισμικό (software prefetching)

Βελτίωση Χρόνου Προσπέλασης σε Ευστοχία

- Μικρές και απλές κρυφές μνήμες
- Λογικές διευθύνσεις αντί για φυσικές
 - Μεταγωγές περιβάλλοντος (context switches):
 - Ακύρωση
 - Διατήρηση κατάλληλης ετικέτας
- Πρόβλεψη δρόμου (way prediction)
- Μερική επικάλυψη προσπελάσεων
- Αποθήκευση χρήσιμων εντολών: κρυφές μνήμες ίχνους (trace caches)

Σύνοψη Βελτιστοποιήσεων

Technique	Hit time	Bandwidth	Miss penalty	Miss rate	Hardware cost/complexity	Comment
Small and simple caches	+			-	0	Trivial; widely used
Way-predicting caches	+				1	Used in Pentium 4
Trace caches	+				3	Used in Pentium 4
Pipelined cache access	-	+			1	Widely used
Nonblocking caches		+	+		3	Widely used
Banked caches		+			1	Used in L2 of Opteron and Niagara
Critical word first and early restart			+		2	Widely used
Merging write buffer			+		1	Widely used with write through
Compiler techniques to reduce cache misses				+	0	Software is a challenge; some computers have compiler option
Hardware prefetching of instructions and data			+	+	2 instr., 3 data	Many prefetch instructions; Opteron and Pentium 4 prefetch data
Compiler-controlled prefetching			+	+	3	Needs nonblocking cache; possible instruction overhead; in many CPUs

Κύρια Μνήμη

- Οργάνωση χώρου διευθύνσεων
 - Διαφύλλωση χαμηλής τάξης
- Τεχνολογία κατασκευής και διασύνδεσης με ΚΜΕ
 - DIMM / SIMM
 - SDRAM / DRAM
 - DDR SDRAM
 - RAMBUS

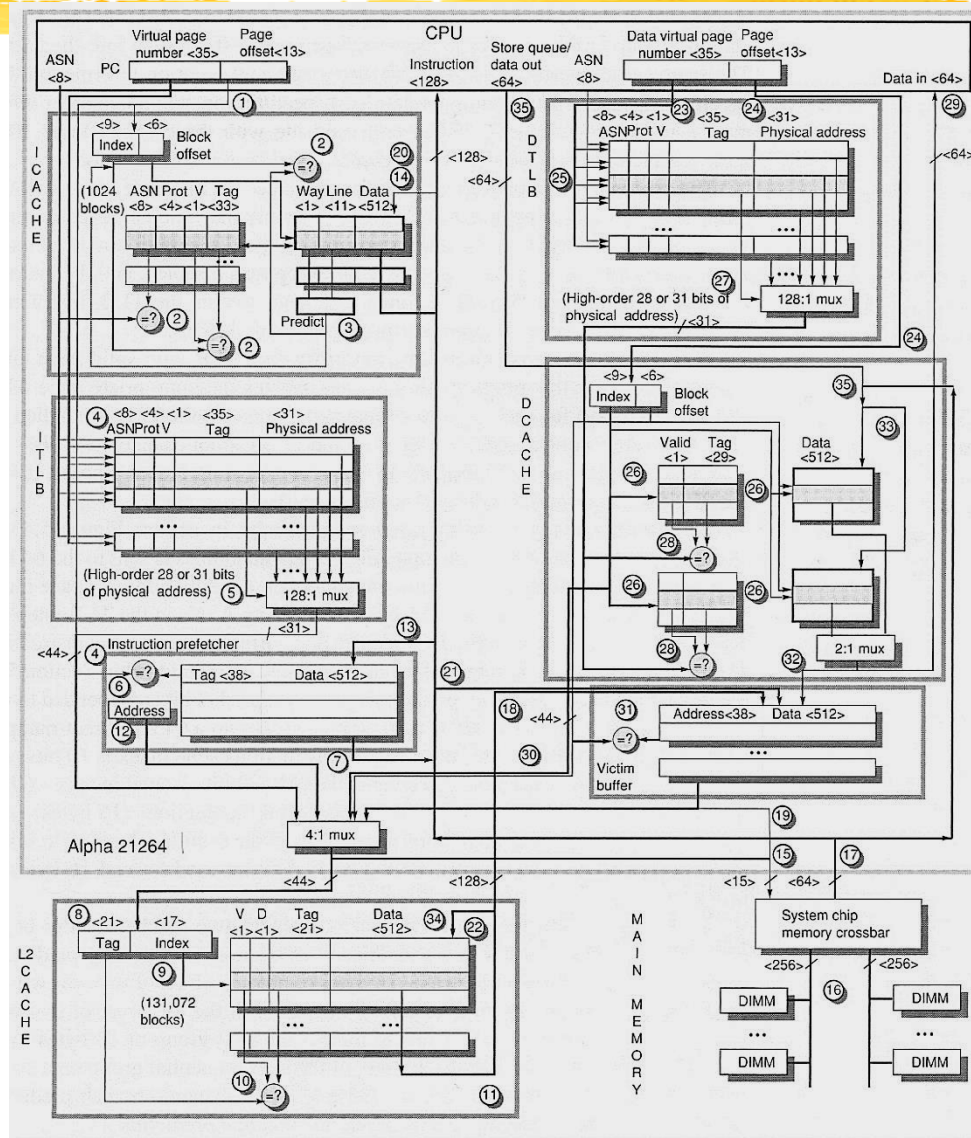
Εξέλιξη στην Κύρια Μνήμη

Production year	Chip size	DRAM type	Row access strobe (RAS)		Column access strobe (CAS)/ data transfer time (ns)	Cycle time (ns)
			Slowest DRAM (ns)	Fastest DRAM (ns)		
1980	64K bit	DRAM	180	150	75	250
1983	256K bit	DRAM	150	120	50	220
1986	1M bit	DRAM	120	100	25	190
1989	4M bit	DRAM	100	80	20	165
1992	16M bit	DRAM	80	60	15	120
1996	64M bit	SDRAM	70	50	12	110
1998	128M bit	SDRAM	70	50	10	100
2000	256M bit	DDR1	65	45	7	90
2002	512M bit	DDR1	60	40	5	80
2004	1G bit	DDR2	55	35	5	70
2006	2G bit	DDR2	50	30	2.5	60
2010	4G bit	DDR3	36	28	1	37
2012	8G bit	DDR3	30	24	0.5	31

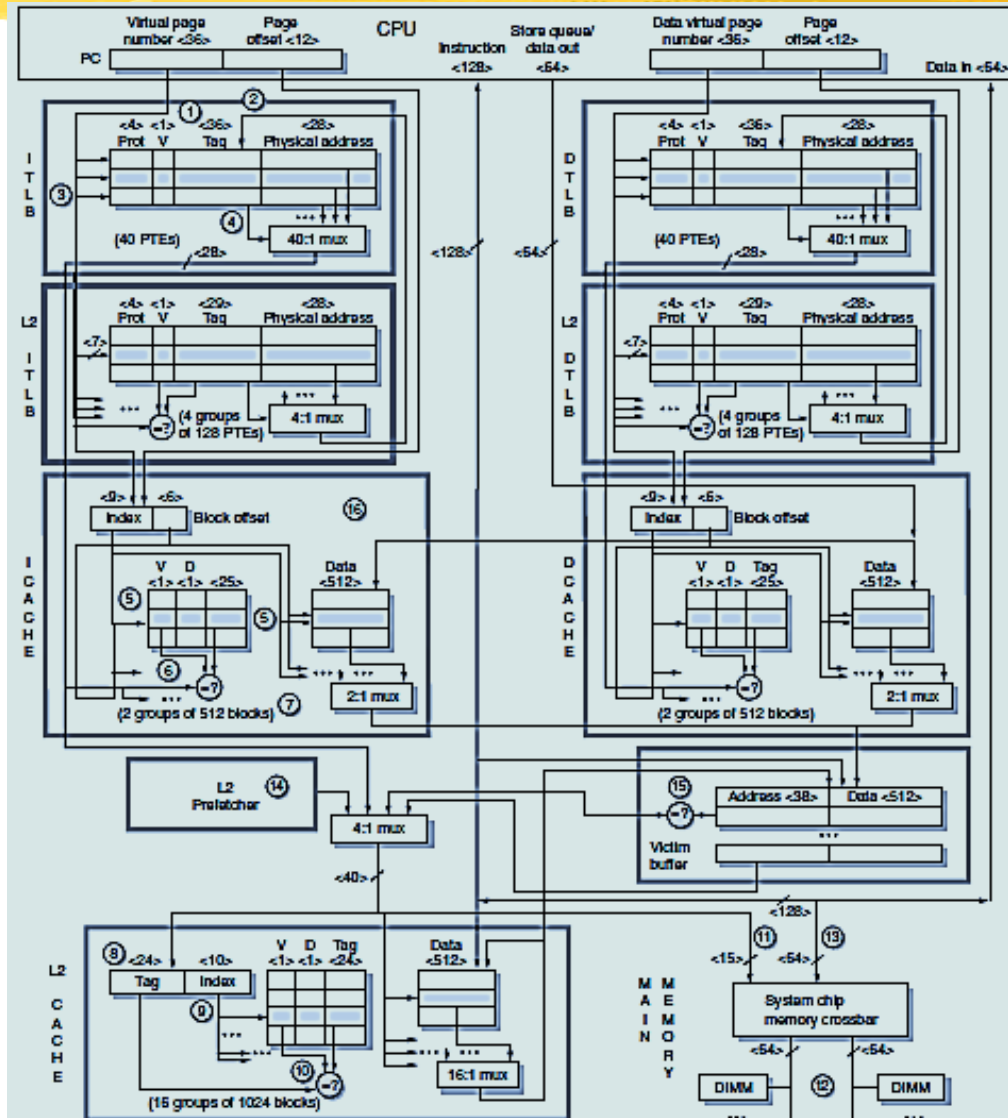
Ιδεατή Μνήμη & Προστασία

- Μέγεθος σελίδας
 - Πρόβλημα κατακερματισμού (fragmentation)
- Κρυφή μνήμη TLB
 - Οργάνωση πλήρους συσχέτισης
- Ψηφία προστασίας τόσο στην TLB όσο και στους πίνακες σελίδων
 - $(\text{Base} + \text{Address}) \leq \text{Bound}$

Παράδειγμα: Alpha 21264



Παράδειγμα: AMD Opteron



Συστήματα Δευτερεύουσας Αποθήκευσης

- Εξέλιξη τεχνολογίας δίσκων
- Συστήματα RAID: Redundant Arrays of Inexpensive (Independent) Disks

Θέματα Αρτηριών



- Αρτηρία διαχωρισμένης προσπέλασης (split transaction bus)