



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ
www.cslab.ece.ntua.gr

ΣΥΣΤΗΜΑΤΑ ΠΑΡΑΛΛΗΛΗΣ ΕΠΕΞΕΡΓΑΣΙΑΣ

<http://www.cslab.ece.ntua.gr/courses/pps/>

8-1-2004

Ακαδημαϊκό Έτος 2003-2004, 9ο Εξάμηνο ΗΜ&ΜΥ

Προτεινόμενα Θέματα για παρουσιάσεις

A) Cache Coherent Non-Uniform Memory Access Architectures

1) Flat, Memory-Based Directory:

a) Stanford FLASH (πρώην DASH)

papers: [The Stanford DASH Multiprocessor](#), [The DASH prototype: Implementation and performance](#),
ή: [The Stanford FLASH Multiprocessor](#)

Η αρχιτεκτονική FLASH ήταν η πρώτη που χρησιμοποίησε το 1992 (DASH) το σχήμα της συνάφειας με κατάλογο (directory based coherence) για κατανεμημένη μνήμη, και μάλιστα στην κύρια μνήμη (flat memory-based coherence protocol). Οι επεξεργαστές διασυνδέονται σε τοπολογία πλέγματος (mesh). Το πρωτόκολλο συνάφειας προέρχεται από την μηχανή DASH που ήταν πρόγονος του FLASH.

b) MIT Alewife

paper: [The MIT Alewife Machine: Architecture and Performance](#)

Η αρχιτεκτονική Alewife υποστηρίζει μέχρι 512 επεξεργαστές διασυνδεδεμένους σε πλέγμα (mesh connected) Full map directory (up to 5 readers per cache line)

c) SGI Origin 2000

paper: [The SGI Origin: A ccNUMA Highly Scalable Server](#)

Η αρχιτεκτονική Origin αποτελείται από 2-512 κόμβους με 1-2 επεξεργαστές R10000 για τον κάθε κόμβο. Το ιδιαίτερο χαρακτηριστικό της είναι ότι σε κάθε κόμβο οι 2 επεξεργαστές δε λειτουργούν με πρωτόκολλο snoopy αλλά ανεξάρτητα. Για τη διασύνδεση μεταξύ διαφορετικών κόμβων χρησιμοποιείται ένα ειδικό κύκλωμα δρομολόγησης που κατασκευάστηκε από την SGI και λέγεται SPIDER. Πολλά τέτοια κυκλώματα δρομολόγησης διασυνδέουν κόμβους σχηματίζοντας τοπολογίες υπερκύβων. Κάθε δύο κόμβοι διασυνδέονται σε ένα κύκλωμα δρομολόγησης ενώ τα κυκλώματα δρομολόγησης αποτελούν τους κόμβους ενός υπερκύβου. Έτσι, για 32 επεξεργαστές χρησιμοποιείται υπερκύβος 3^{ης} τάξεως, άρα $2^3=8$ δρομολογητές, άρα 8×2 κόμβοι/δρομολ. \times επεξεργ/κόμβο = 32 επεξεργαστές. (memory-based directory coherence protocol). Το πρωτόκολλο συνάφειας γρήγορης μνήμης βασίζεται σε αυτό της μηχανής Stanford DASH.

2) Flat, Cache-Based Directory:

a) IBM-Sequent NUMA-Q (πρώην StING)

papers: [Sequent's NUMA-Q SMP Architecture](#),

ή: [StING: A ccNUMA Coputer System for the Commercial Marketplace](#)

Η αρχιτεκτονική NUMA-Q αποτελείται από τετράδες συμμετρικής πολυεπεξεργασίας διασυνδεδεμένες σε απλό δακτύλιο με πρωτόκολλο SCI. Οι επεξεργαστές κάθε τετράδας διασυνδέονται μέσω διαδρόμου, ενώ υπάρχει και ένας ελεγκτής (controller) που βασίζεται στο πρωτόκολλο SCI για την συνάφεια γρήγορης μνήμης μεταξύ μνημών που ανήκουν σε διαφορετικές τετράδες. (cache-based directory coherence protocol)

b) The NUMachine Multiprocessor

papers: [Design and Implementation of the NUMachine Multiprocessor](#), [The NUMachine Multiprocessor](#)
Η αρχιτεκτονική NUMachine προέρχεται από το πανεπιστήμιο του Τορόντο. Είναι μια αρχιτεκτονική συνάφειας γρήγορης μνήμης βασισμένη σε κατάλογο (directory based cache coherency) και αποτελείται από κόμβους που ανά τέσσερις συνδέονται σε δακτύλιο (ring), ενώ κάθε τέσσερις δακτύλιοι συνδέονται επίσης σε ανώτερης τάξης δακτύλιο (hierarchical). Ο κάθε κόμβος αποτελείται από τέσσερις επεξεργαστές σε διάδρομο (4 procs-bus connected-SMP node) και έχει ένα ελεγκτή μνήμης για διασύνδεση με το δακτύλιο (hierarchical directory controllers). Κάθε δακτύλιος έχει $4 \times 4 = 16$ επεξεργαστές και ο δακτύλιος ανώτερης τάξης διασυνδέει τέσσερις δακτυλίους φτάνοντας στους $4 \times 16 = 64$ επεξεργαστές.

c) Data General-Avion 25000 ccNUMA architecture (NUMALiNE)

link: [Data General Avion NUMALine](#)

Η αρχιτεκτονική Avion 25000 ccNUMA αποτελείται από προκατασκευασμένες τετράδες (commodity quads) επεξεργαστών Pentium Xeon της Intel. Οι τετράδες αυτές ονομάζονται Standard High Volume (SHV) server nodes. Κάθε τετράδα συνδέεται μέσω ενός συνδέσμου SCI σε διπλό δακτύλιο (SCI-based dual rings). (cache-based flat directory-based coherence protocol)

3) [A Survey of Parallel Computer Architectures](#)

B) Πρωτόκολλα Συνάφειας Μνήμης

1) σε Διάδρομο (Bus Snooping Cache Coherency Protocols)

- a) Πρωτόκολλα MESI (Modified-Exclusive-Shared-Invalid) και MSI (Modified-Shared-Valid) (υλικό από βιβλίο Culler 5.3)
- b) [Cache Coherence Protocols: Evaluation using a Multiprocessor Simulation Model](#)

2) με Κατάλογο (Directory Based- Cache Coherency Protocols)

- a) (Flat-limited) memory-based directory coherence (υλικό από βιβλίο Culler 8.5)
paper: [Directory-Based Cache Coherence in Large-Scale Multiprocessors](#)
- b) (linked list-chained)-cache-based directory coherence (SCI)
paper: [Scalable Shared-Memory Multiprocessor Architectures](#)
- c) [Evaluation of Directory Schemes for Cache Coherence](#)

3) Γενικά

- a) [A Survey of Cache Coherence schemes for multiprocessors](#)
- b) [Managing Cache Coherence in Multiprocessor Computer Systems](#)
- c) [Compiler-Directed Cache Management in Multiprocessors](#)

C) Cluster Computing – Modern Cluster Interconnects

- a) High Performance web servers using clusters
papers: [Dynamic Load Balancing on Web-Server Systems](#)
[The Anatomy of a Large-Scale Hypertextual Web Search Engine \(google\)](#)
[High Performance Web Site Design Techniques](#)
[Linux Virtual Server for Scalable Network Services](#)
[On Balancing the Load in a Clustered Web Farm](#)
[Petaops and Exaops: Supercomputing on the Web](#)
[The State of the Art in Locally Distributed Web-Server Systems](#)
- b) Scalable Coherent Interface as System Area Network (SAN) for Clusters
paper: [The Scalable Coherent Interface and Related Standards Projects](#)
- c) [Utilizing Network Cache on an SCI-Based PC Cluster](#)
- d) User Level Networking-Virtual Interface Architecture (VIA)
paper: [An SCI-based Software VIA System for PC Clustering](#)
- e) User-level Network Interface Protocols
papers: [Design Issues for User-Level Network Interface Protocols on Myrinet](#), [User-level Network Interface Protocols](#)

- f) [Myrinet - A Gigabit-per-Second Local Area Network](#)
- g) [OPIOM: Off-processor IO with Myrinet - READ: Put Disks at Network Level](#)
- h) [Boosting the Performance of Myrinet Networks](#)
- i) [NIC-Based Atomic Memory Operations in Myrinet/GM - NIC-Based Reduction in Myrinet Clusters: Is It Beneficial](#)
- j) [An Evaluation of some Beowulf clusters](#)
- k) [Overview of recent Supercomputers](#)