

CS-534: Packet Switch Architecture

Εισαγωγικό Σημείωμα (Introductory Note):

Το μάθημα αυτό βρίσκεται ανάμεσα στο "hardware" και στα "δίκτυα". Ασχολείται με την αρχιτεκτονική των μεταγωγέων (switches) και δρομολογητών (routers) πακέτων, των δομικών λίθων δηλαδή από τους οποίους κατασκευάζονται τα σημερινά δίκτυα υψηλής ταχύτητας, είτε δίκτυα επικοινωνίας είτε δίκτυα διασύνδεσης συστοιχίων υπολογιστών (clusters) και πολυεπεξεργαστών. Ενώ παλαιότερα οι δομικοί αυτοί λίθοι ήταν συνήθως υπολογιστές γενικού σκοπού που επετύγχαναν τη δρομολόγηση μέσω λογισμικού, σήμερα --στις ψηλές ταχύτητες-- η υποστήριξη από ειδικευμένο υλικό (hardware) έχει καταστεί αναγκαία. Ένα σημαντικό μέρος του μαθήματος έχει μιάν αναλογία προς την "αρχιτεκτονική υπολογιστών" (HY-425): εξετάζει την οργάνωση των ψηφιακών κυκλωμάτων που υλοποιούν τους μεταγωγείς, την αλληλεπίδραση των κυκλωμάτων αυτών με το λογισμικό και με τα πακέτα που δρομολογούν, τις εναλλακτικές αρχιτεκτονικές και το κόστος και την επίδοση καθεμιάς. Ένα άλλο μέρος του μαθήματος εξετάζει θέματα τοπολογίας του δικτύου διασύνδεσης. Το μάθημα μελετά με όσο το δυνατόν πιό ενιαίο τρόπο τη δομή των ενεργών στοιχείων των δικτύων SAN, LAN, MAN, WAN, του Internet, των δικτύων ψηφιακής τηλεφωνίας, και των δικτύων διασύνδεσης πολυεπεξεργαστών, αναδεικνύοντας την κοινή βάση πίσω απ' όλες αυτές τις φαινομενικά ανόμοιες --αλλά συγκλίνουσες-- τεχνολογίες επικοινωνίας. Γίνονται επίσης αναφορές στην πρόσφατη μετεξέλιξη των συνδέσεων μέσα στα chips σε Network-on-Chip (NoC). Μελετώνται πάντως μόνον τα δίκτυα που αποτελούνται από συνδέσμους σημείου-προς-σημείο, μίας κατεύθυνσης (unidirectional, point-to-point links), και όχι δίκτυα κοινόχρηστου μέσου (shared medium) όπως οι ασύρματες επικοινωνίες. Επίσης, η έμφαση είναι κυρίως στη μεταγωγή πακέτων, και λιγότερο στην (σχετικά απλούστερη) μεταγωγή κυκλωμάτων.

Course Content -- Περιεχόμενο του Μαθήματος:

[(a) in Greek:] Η αρχιτεκτονική του υλικού (hardware) των μεταγωγέων (switches) και δρομολογητών (routers) των δικτύων πακέτων υψηλής ταχύτητας.

- Σύνδεσμοι και η παροχή τους: σημείο-προς-σημείο αντί κοινόχρηστης αρτηρίας, σειριακοί, παράλληλοι.
- Μεταγωγή κυκλωμάτων: πολύπλεξη διαίρεσης χρόνου, μεταγωγή χρόνου/χώρου, πολυπλέκτες προσθήκης-αφαίρεσης, αντίστροφη πολύπλεξη, πολυεπίπεδα δίκτυα, μη-μπλοκάρισμα καθαρά ή με αναδιάταξη.
- Έννοιες μεταγωγής πακέτων: κοινή αντί διαμοιρασμένης παροχή, στατιστική πολύπλεξη, ανταγωνισμός εξόδου, ενταμίευση, εσωτερικό μπλοκάρισμα, ουρές, μπλοκάρισμα σε ουρά (HOL), χρονοδρομολόγηση, έλεγχος ροής, κόψιμο δρόμου αντί αποθήκευσης-και-προώθησης.
- Γενεές μεταγωγέων.
- Αρχιτεκτονικές ενταμιευτών: φαρδιές μνήμες, διαφυλλωμένες μνήμες, μνήμες με ομοχειρία, πολλαπλές ουρές σε κοινό ενταμιευτή, ουρές για multicast. Τεμαχισμός και ανασυγκόλληση: παροχή κυττάρων, ουρές, απόρριψη πακέτων.
- Αρχιτεκτονικές ουρών: ουρές εξόδων ή σημείων διασταύρωσης, κοινόχρηστος ενταμιευτής, ουρές εισόδων, ουρές εικονικά εξόδων, εσωτερική επιτάχυνση (CIOQ), επιδόσεις, χρονοδρομολόγηση crossbar με ουρές εικονικά εξόδων.
- Αρχιτεκτονικές δικτύων μεταγωγής: σταυραγωγός (crossbar), μεταγωγέας παράλληλων επιπέδων, πολυβάθμια δίκτυα, υπερκύβος, banyan, Benes, Clos, παχέα δένδρα, προσαρμοστική δρομολόγηση, αναδιάταξη πακέτων, δίκτυα με ή χωρίς ενταμιευτές.
- Έλεγχος ροής: στατικός/δυναμικός, με/άνευ απωλειών, ρητός/υπονοούμενος, άκρη-με-άκρη/κομάτι-κομάτι, ρυθμός/πιστώσεις, αδιάκριτος/ανά-ροή, δρομολόγηση wormhole, QFC, δίκτυα μεταγωγής με εσωτερική οπισθοπίεση.
- Χρονοδρομολόγηση για ποιότητα υπηρεσιών: ταχεία υλοποίηση προτεραιοτήτων, κυκλικής εξυπηρέτησης, και κυκλικής με συντελεστές βάρους.
- Βιβλιογραφική μελέτη υλικού αναζήτησης σε πίνακες δρομολόγησης και κατηγοριοποίησης ροών.

[(b) in English:] The architecture of the hardware of switches and routers for high speed packet networks.

- Links and their throughput: point-to-point versus shared medium, serial, parallel.
- Circuit switching: time-division multiplexing, time/space switching, add-drop multiplexors, inverse

multiplexing, multi-stage fabrics, strictly or rearrangeably non-blocking.

- Packet switching concepts: shared versus partitioned throughput, statistical multiplexing, output contention, buffering, internal blocking, queueing, head-of-line blocking, scheduling, flow control, cut-through versus store-and-forward.
- Switch generations.
- Buffer architectures: wide/interleaved/pipelined memory, multiple queues in shared buffer, multicast queues. Segmentation and reassembly: cell throughput, queueing, packet dropping.
- Queueing architectures: crosspoint/output queueing, shared buffering; input queueing, virtual output queueing, internal speed-up (CIOQ); performance, crossbar scheduling under virtual output queueing.
- Switching fabric architectures: crossbars, parallel-plane, multi-stage networks, hypercube, banyan, Benes, Clos, fat trees, adaptive routing, packet resequencing, buffered versus bufferless fabrics.
- Flow control: static/dynamic, lossy/lossless, implicit/explicit, end-to-end/hop-by-hop, rate/credit, indiscriminate/per-flow; wormhole routing, QFC, switching fabrics with internal backpressure.
- Scheduling for QoS: fast implementations of priority, round-robin, and weighted round-robin.
- Bibliographic readings on routing table lookup and flow classification hardware.

Position in the Curriculum:

- Computer Science Graduate Course.
- **Thematic Areas:**
 - A: Microelectronic Systems Architecture
 - B: Computer Networks and Digital Communications
- **Prerequisites:**
 - CS-225 "Computer Organization", and CS-335 "Computer Networks"
- Other useful background:
 - CS-425 "Computer Architecture"
- **Four (4) Credit Units.**

Instructor, TA, Schedule:

- Instructor: Manolis G.H. Katevenis, Professor. Ηλ-Ταχ: kateveni
- Teaching Assistant: Georgios Passas, Graduate Student. Ηλ-Ταχ: passas
- Lectures: Mondays and Wednesdays, 5-7 (17:15 - 19:00), in L.K. B-211
- Assistive Sessions (Φροντιστήριο): Fridays, 3-5, in L.K. B-211 (will convene by invitation).

Electronic Communication:

- Primary Web Page: <http://archvlsi.ics.forth.gr/~kateveni/534/>
- ... to be mirrored, sometime, to: <http://www.csd.uoc.gr/~hy534/>
- Κατάλογος Ηλ-Ταχ: **hy534-list** (subscribe by e-mailing to "Majordomo@csd.uoc.gr" placing the line "subscribe hy534-list" in the *body* (not header) of the message).

Exercises, Bibliography

- Weekly Exercise Sets.
- Paper Presentations: study a couple of individually assigned recent research papers during the semester, and present their basic ideas orally to the class.
- Transparencies by the instructor, available on the web.
- Some reading material from the books: **W. Dally, B. Towles:** "Principles and Practices of Interconnection Networks", *Elsevier - Morgan Kaufmann*, 2004, ISBN: 978-0-12-200751-4; **J. Duato, S. Yalamanchili, L. Ni:** "Interconnection Networks: an Engineering Approach" *Morgan Kaufmann - Elsevier*, second printing, 2003, ISBN 1-55860-852-4; **S. Keshav:** "An Engineering Approach to Computer Networking", *Addison Wesley*, 1997, ISBN 0-201-63442-2.
- Other bibliography: <http://archvlsi.ics.forth.gr/~kateveni/534/09a/bibliography.html>

Grading:

- **25 %** from the **exercises** (mostly) and from the paper presentations and the general class participation (to a lesser degree);
- **25 %** from the **midterm exam**; and
- **50 %** from the **final exam**.