

ΗΥ-335 : Δίκτυα Υπολογιστών

Επίπεδο Δικτύου

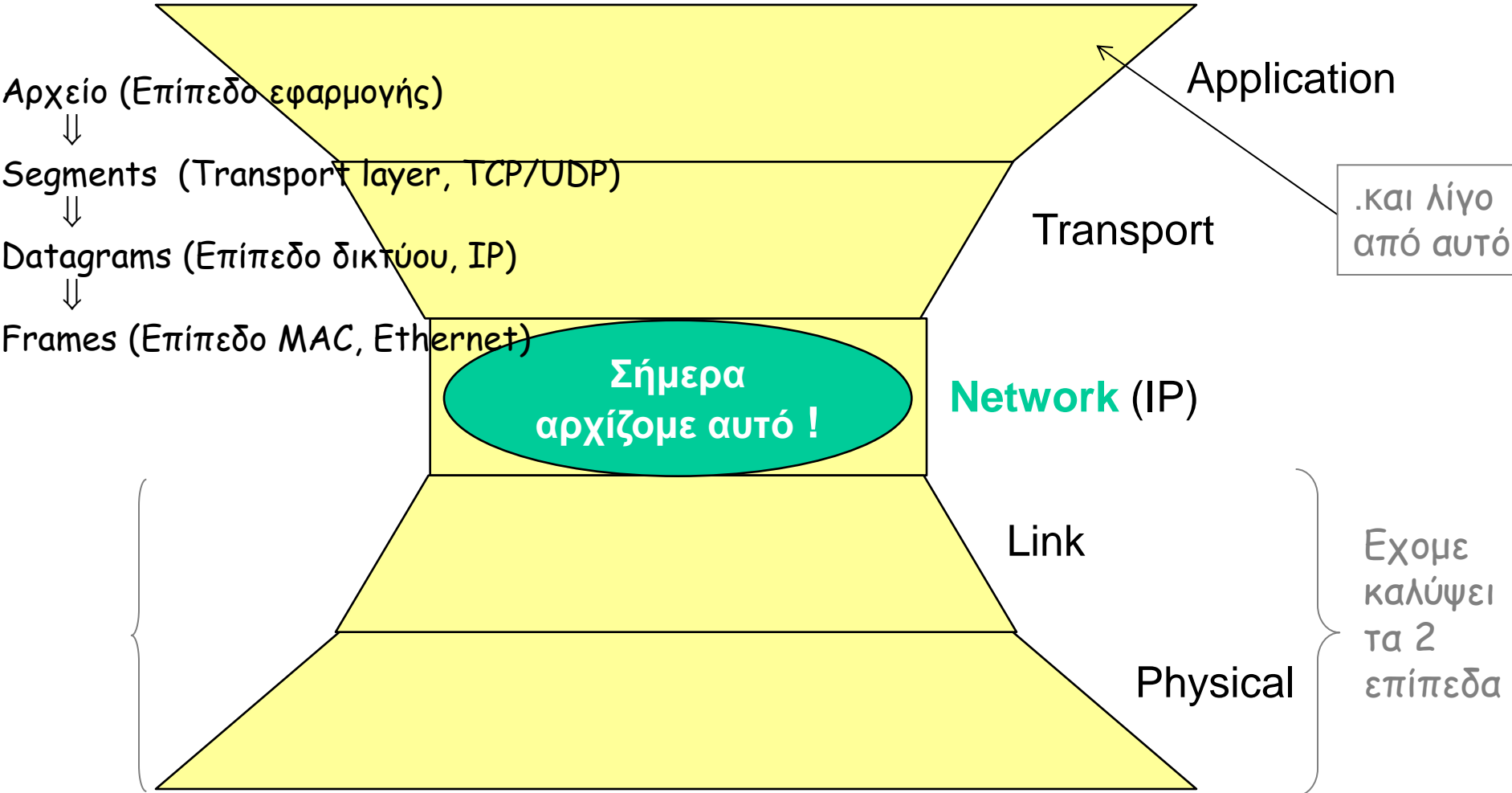
Δρομολογητές, αρχιτεκτονική, διευθυνσιοδότηση

Μαρία Παπαδοπούλη

Τμήμα Επιστήμης Υπολογιστών

Πανεπιστήμιο Κρήτης

Χειμερινό εξάμηνο 2014-2015



Επίπεδο Δικτύου (βασικά θέματα)

Στόχοι:

Κατανόηση των βασικών αρχών σχεδιασμού του :

- network layer service μοντέλο
- Δρομολόγηση (routing) & προώθηση (forwarding)
- Πως δουλεύει ένας δρομολογητής (router)
- Δρομολόγηση (επιλογή μονοπατιού)
- Επίδραση των μεγάλων μεγεθών στην απόδοση των αλγορίθμων δρομολόγησης

Γενικοί Μηχανισμοί Επιπέδου Δικτύου

- **Δρομολόγηση (routing)**

Καθορισμός του μονοπατιού για την αποστολή (δρομολόγηση) ενός πακέτου, από την πηγή/αποστολέα end-host στον τελικό παραλήπτη/προορισμό end-host

Οι αλγόριθμοι λέγονται αλγόριθμοι δρομολόγησης (routing algorithms)

- **Πρωώθηση (forwarding)**

Όταν ένα πακέτο φτάνει στην διεπαφή "εισόδου" (input) του δρομολογητή κι αυτός πρέπει να το προωθήσει στη σωστή διεπαφή εξόδου (output link)

- **Call setup**

Χρειάζονται οι δρομολογητές σε ένα επιλεγμένο μονοπάτι από την πηγή-αποστολέα στον παραλήπτη/τελικό προορισμό να συμμετέχουν σε ένα «handshake» πριν την αποστολή των πακέτων της εφαρμογής

VC setup, Data transfer & VC teardown

Προσοχή: το Call setup ΔΕΝ υπάρχει στο Διαδίκτυο (Internet)!

Επισκόπηση της αρχιτεκτονικής του δρομολογητή

Δύο σημαντικές λειτουργίες του δρομολογητή:

- Τρέχει αλγορίθμους/πρωτόκολλα δρομολόγησης (πχ RIP, OSPF, BGP)
- Προωθεί τα datagrams από την εισερχόμενη στην εξερχόμενη σύνδεση

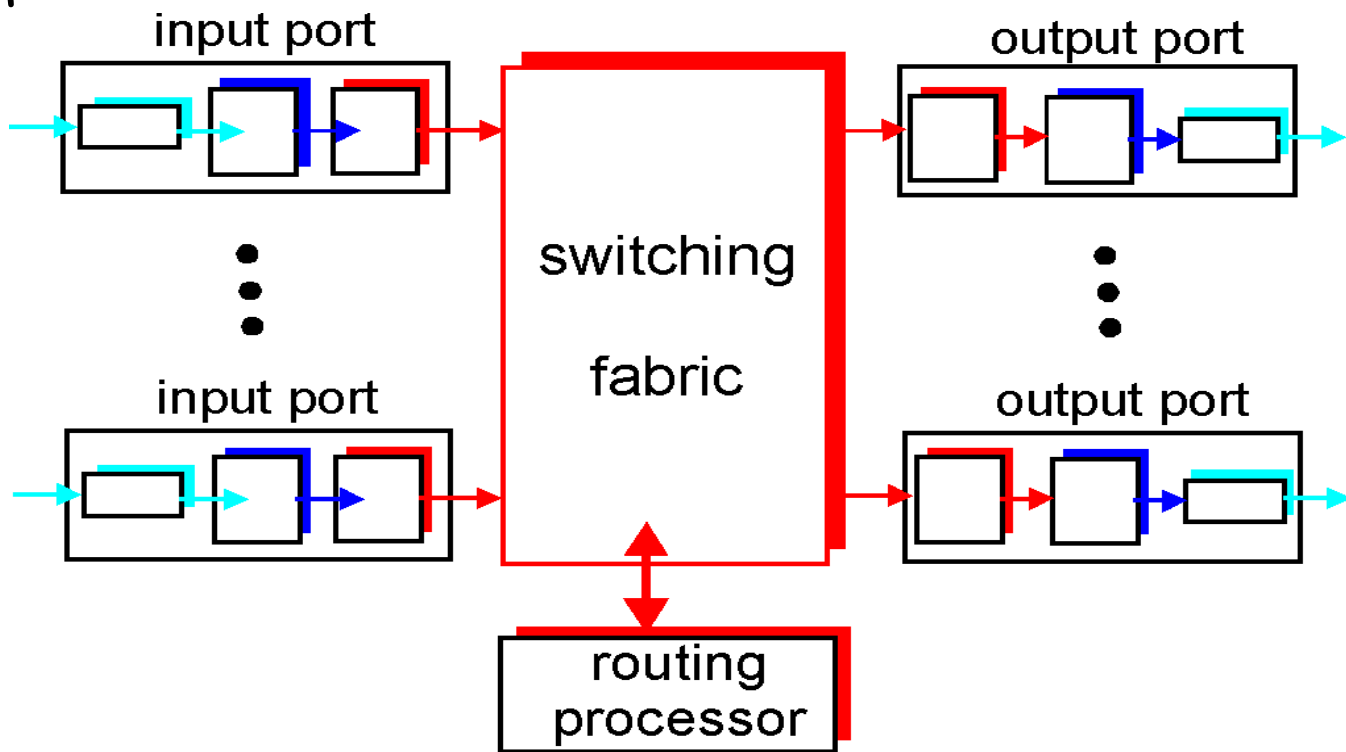
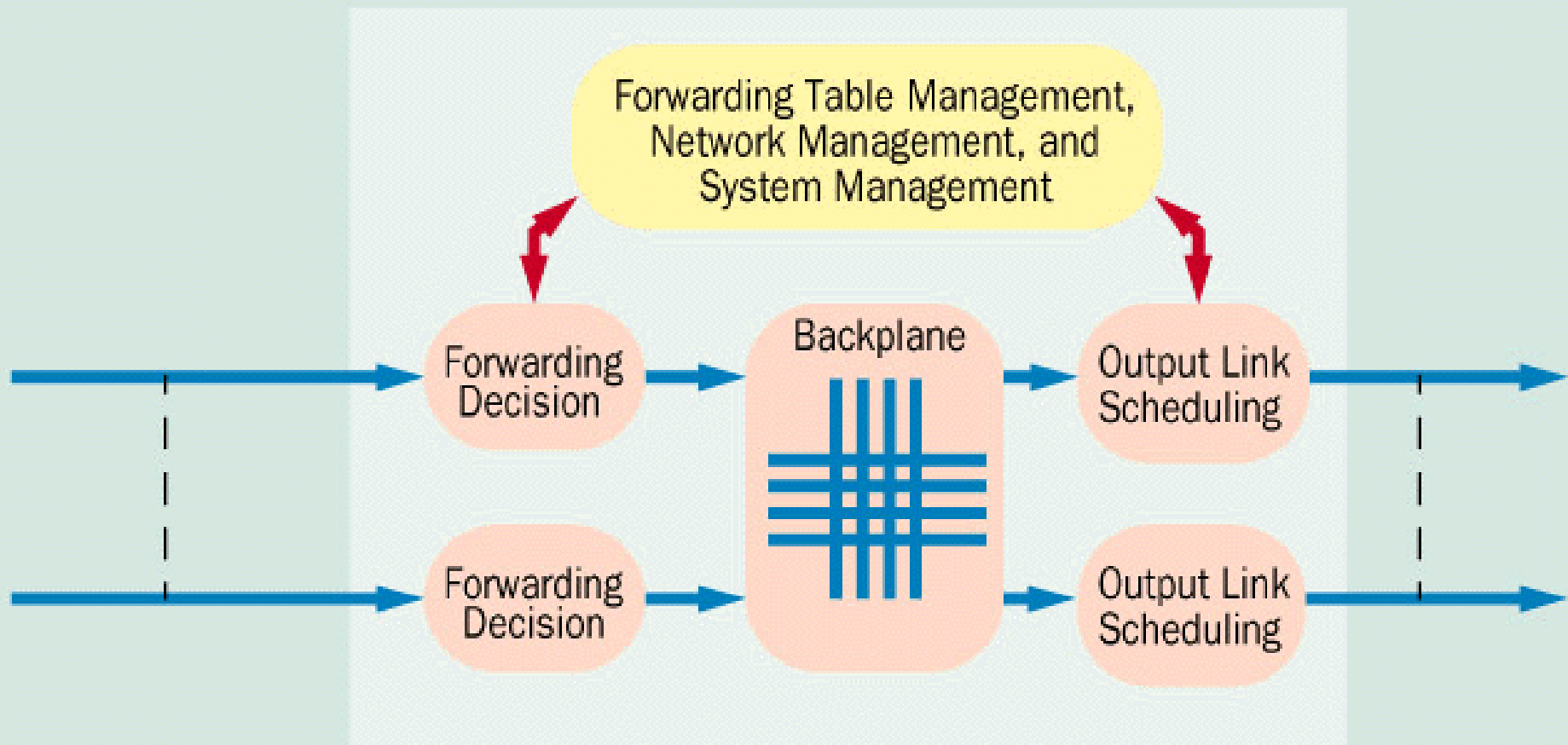


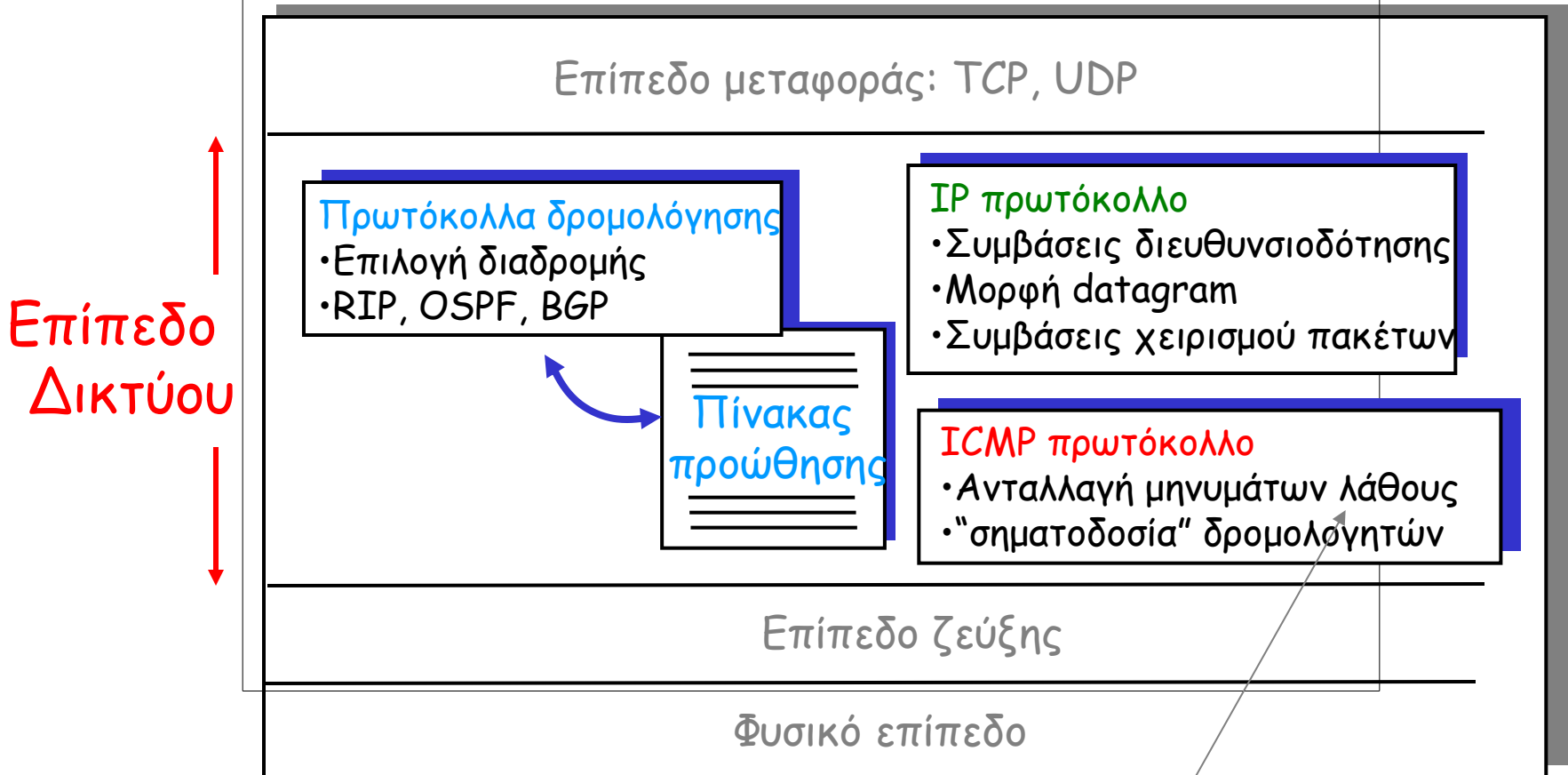
FIGURE 1 The Key Common Architectural Components of a Router



In high-performance systems, the forwarding decision, backplane and output link scheduling must be performed in hardware, while the less timely management and maintenance functions are performed in software.

Παράδειγμα: το επίπεδο δικτύου στο Διαδίκτυο

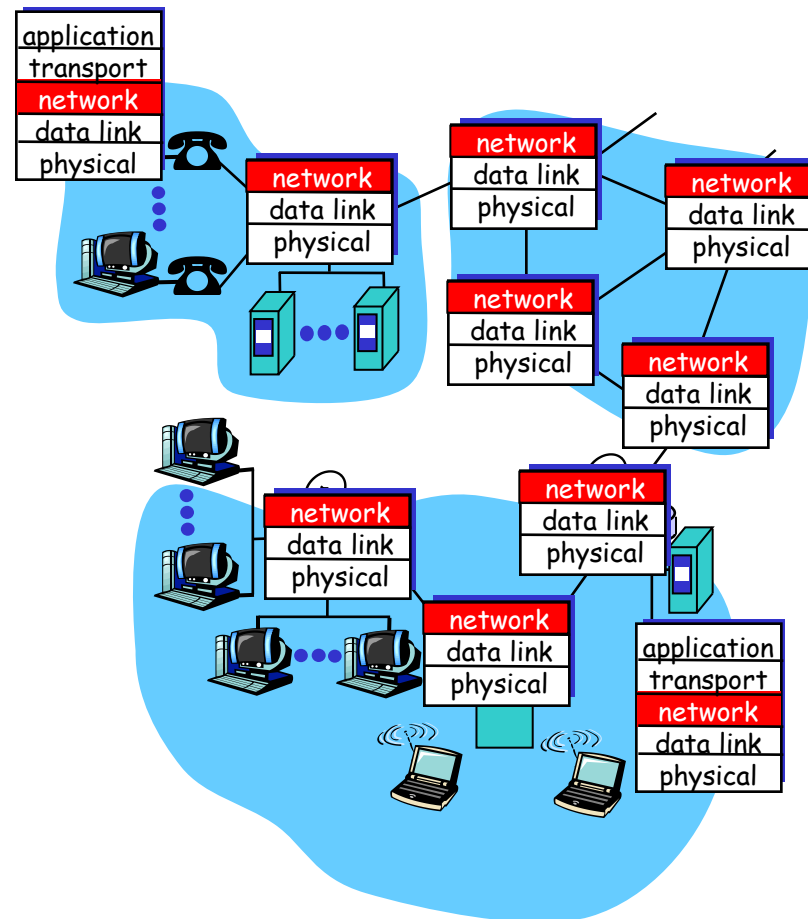
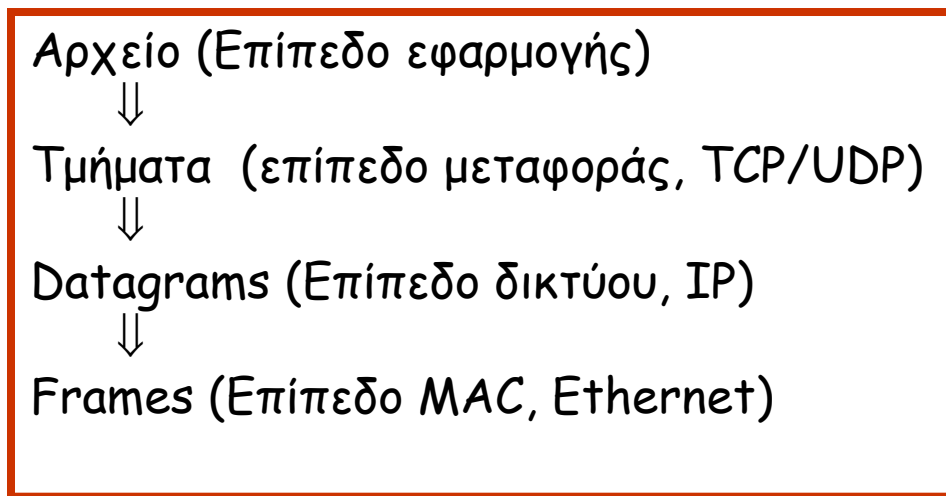
Λειτουργίες του επιπέδου δικτύου host και δρομολογητές:



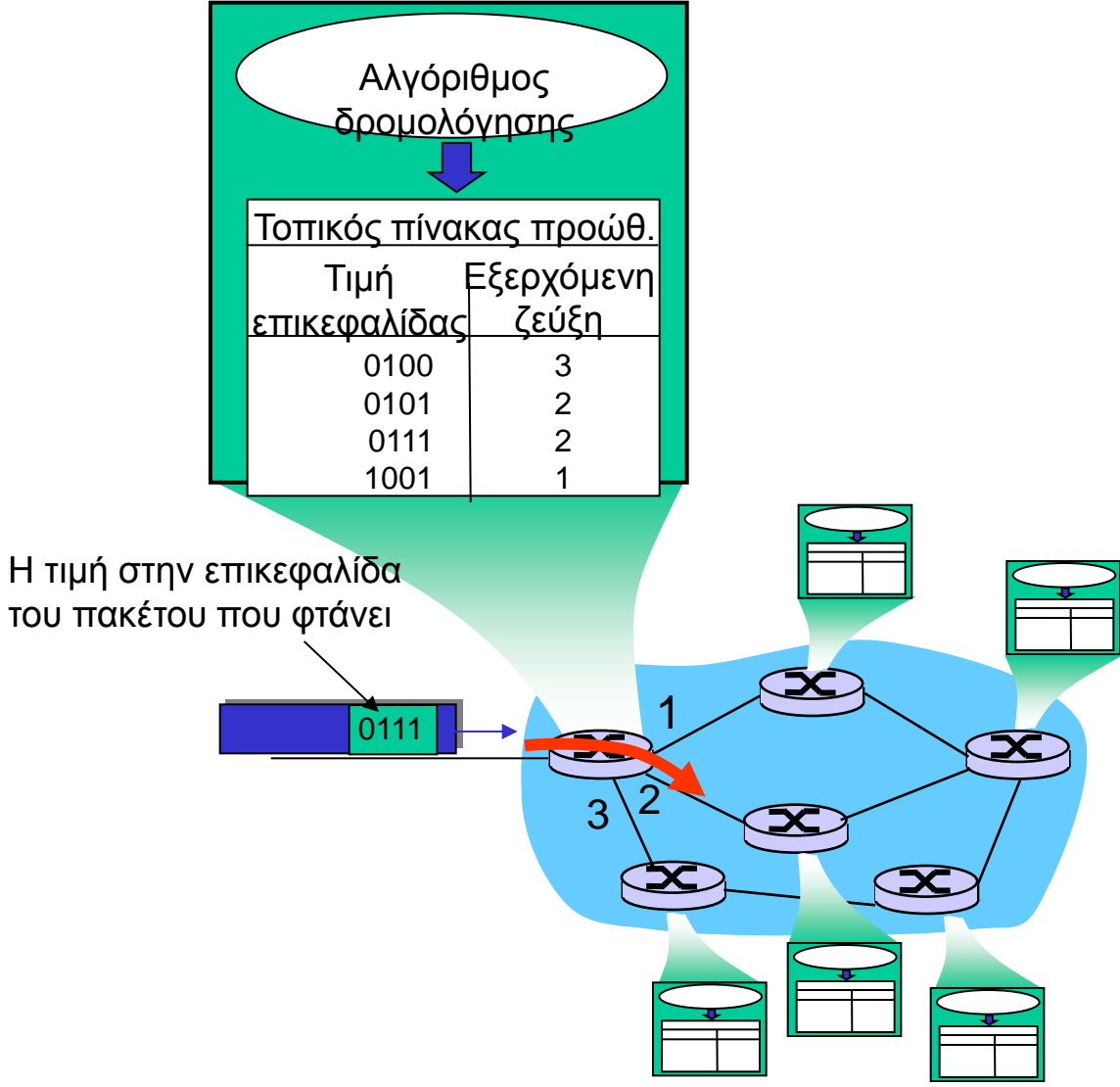
πχ έλλειψη κάποιας υπηρεσίας από έναν server,
απουσία ενός υπολογιστή από το δίκτυο

Επίπεδο δικτύου

- Πρωτόκολλα επιπέδου δικτύου σε **κάθε host**, και **δρομολογητή**
- Ο δρομολογητής εξετάζει τα **πεδία επικεφαλίδας** σε όλα τα **IP datagrams** που περνάνε από αυτόν

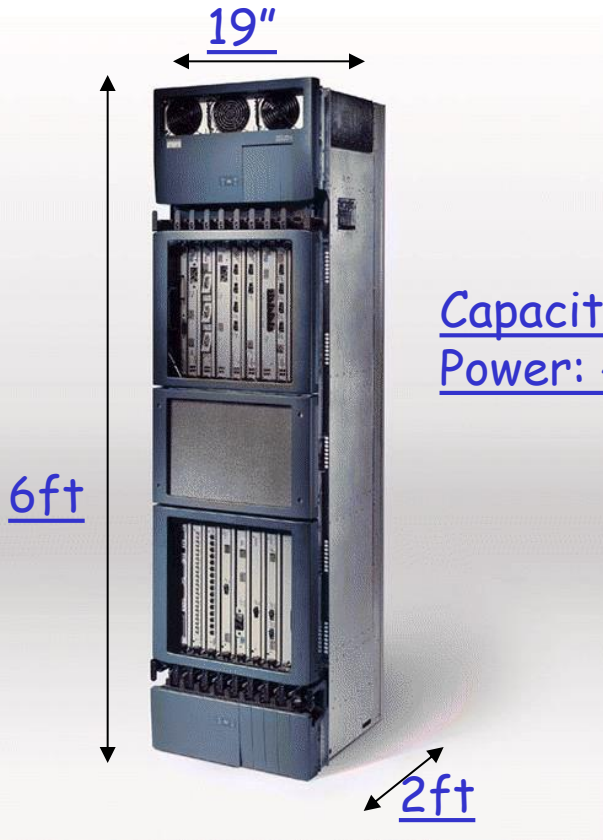


Αλληλεπίδραση μεταξύ δρομολόγησης & προώθησης



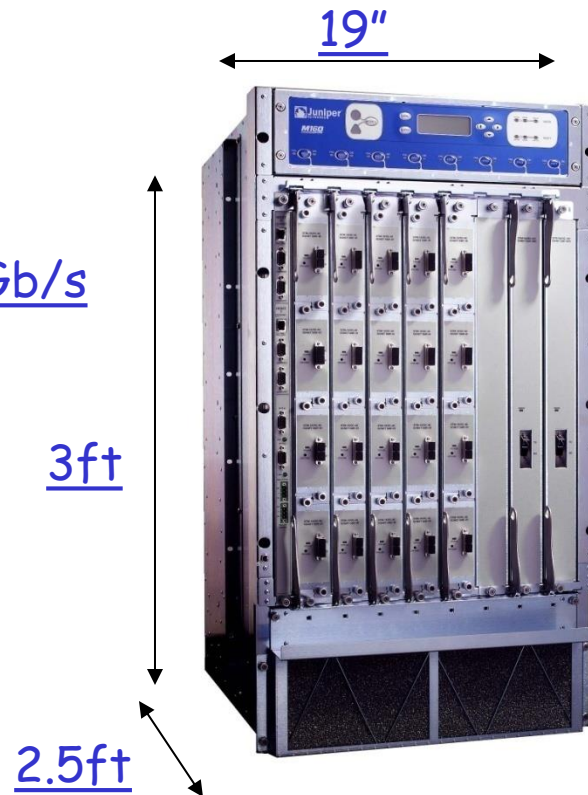
Πως μοιάζει ένας δρομολογητής

Cisco GSR 12416



Capacity: 160Gb/s
Power: 4.2kW

Juniper M160



Capacity: 80Gb/s
Power: 2.6kW

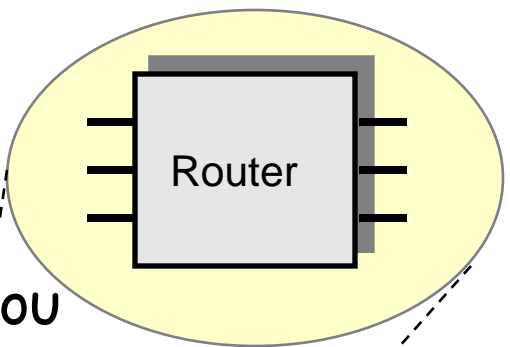
Big, Fast Routers: Why Bother?

- r Faster link bandwidths
- r Increasing demands
- r Larger network size (hosts, routers, users)
- r More cost effective

IP Routers

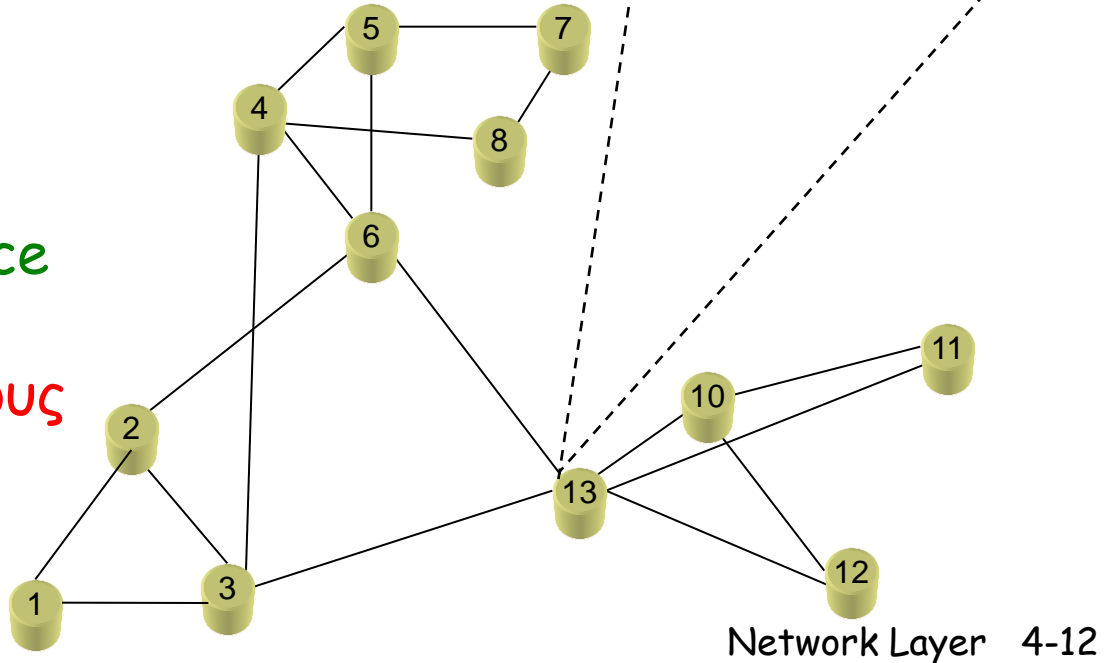
Αποτελούνται από

- Διεπαφές εισόδου interfaces όπου τα πακέτα φτάνουν
- Διεπαφές εξόδου από όπου τα πακέτα προωθούνται στο δίκτυο
- τρόπος σύνδεσης των διεπαφών εισόδου και εξόδου



Υλοποιεί:

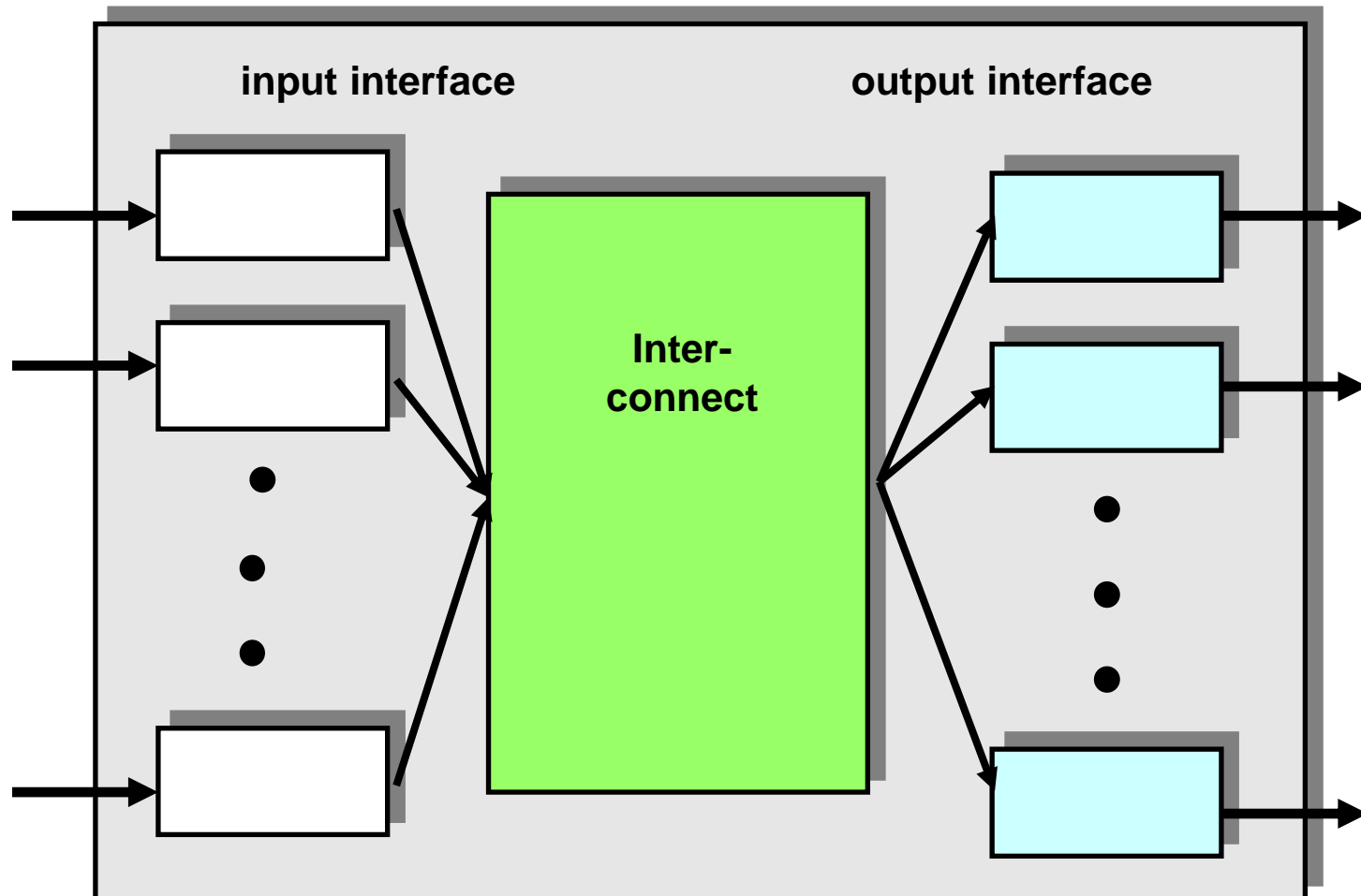
- *Πρώθηση* πακέτων στο κατάλληλο output interface
- Χειρισμό των πόρων **εύρους ζώνης & πινάκων** χώρου



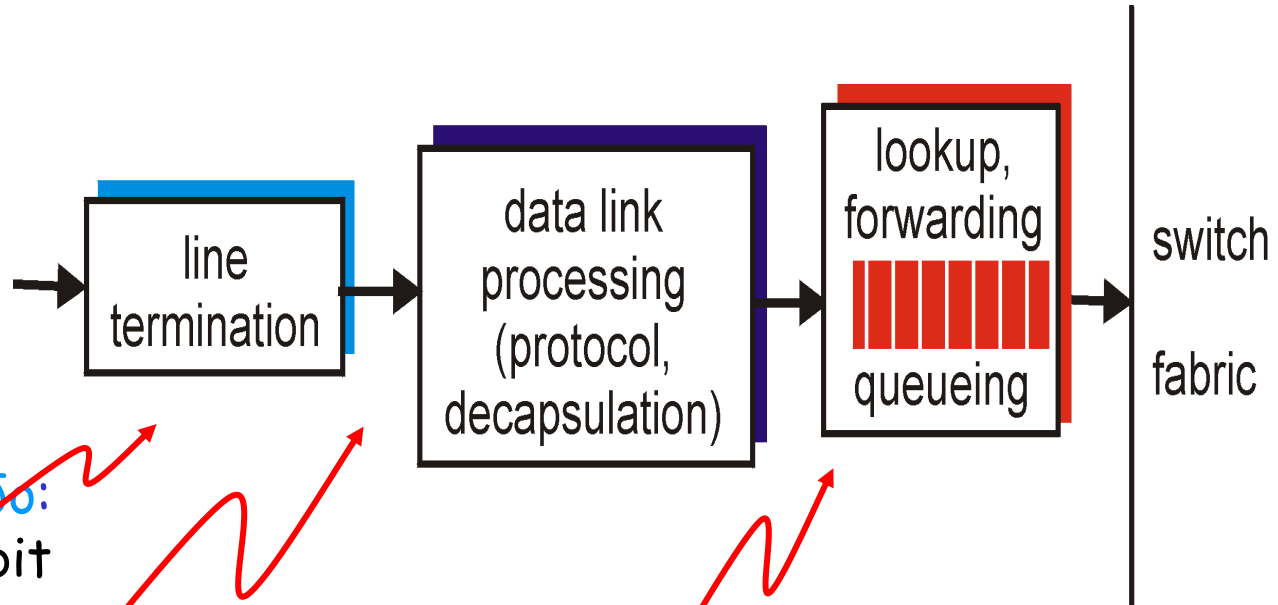
Θα το ακούσετε κι ως level-3 switch

Γενική αρχιτεκτονική ενός δρομολογητή

Οι διεπαφές εισόδου & εξόδου συνδέονται μέσω μιας διασύνδεσης



Λειτουργίες θύρας εισόδου



φυσικό επίπεδο:

Λήψη σε επίπεδο bit

επίπεδο ζεύξης
δεδομένων:

π.χ., Ethernet
(βλ. Κεφ. 5)

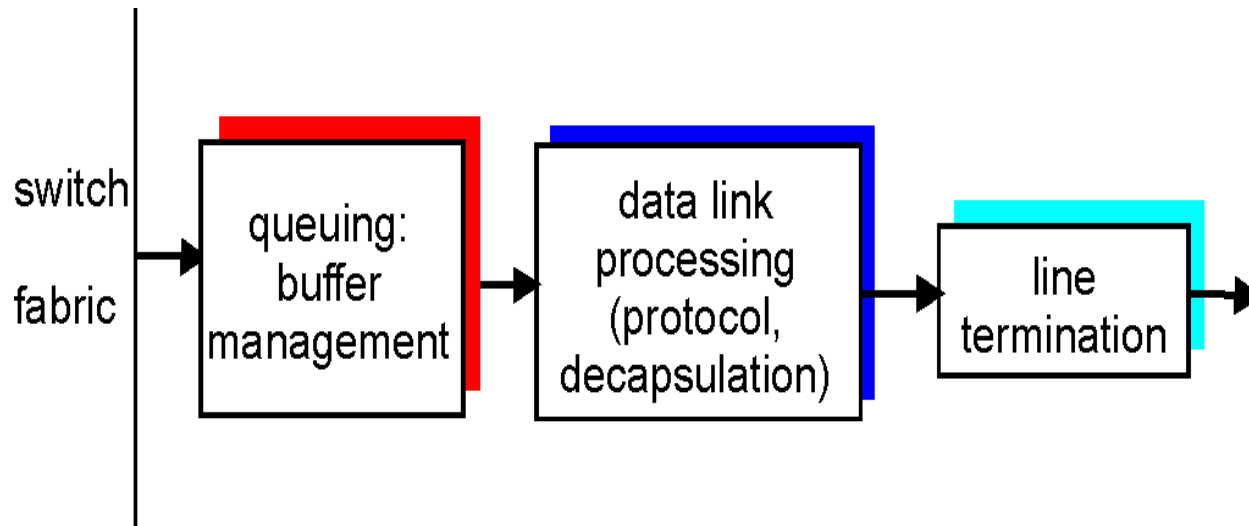
Αποκεντριοποιημένο switching:

Με βάση τον προορισμό του datagram, ψάχνει για την θύρα εξόδου χρησιμοποιώντας τον πίνακα προώθησης στη μνήμη της θύρας εισόδου

Στόχος: να ολοκληρωθεί η επεξεργασία της θύρας εισόδου στην ταχύτητα της γραμμής

☞ **queueing**: όταν τα **datagrams φτάνουν γρηγορότερα από το ρυθμό προώθησης** στο switch

Λειτουργίες Θύρας Εξόδου



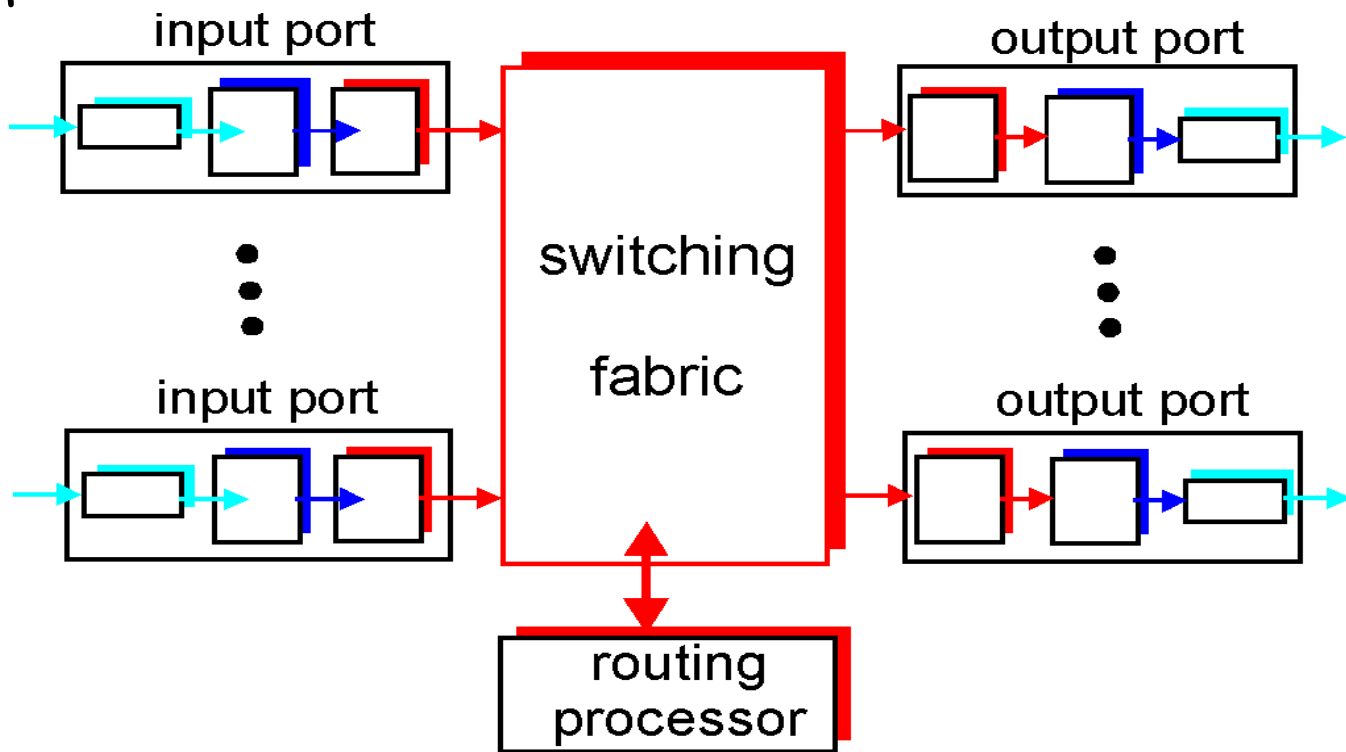
Αποθήκευση (buffering) χρειάζεται όταν τα datagrams φτάνουν από το fabric γρηγορότερα από τον ρυθμό μετάδοσης

Η προγραμματιστική μέθοδος επιλέγει ανάμεσα στα datagrams της ουράς για μετάδοση

Επισκόπηση της αρχιτεκτονικής του δρομολογητή

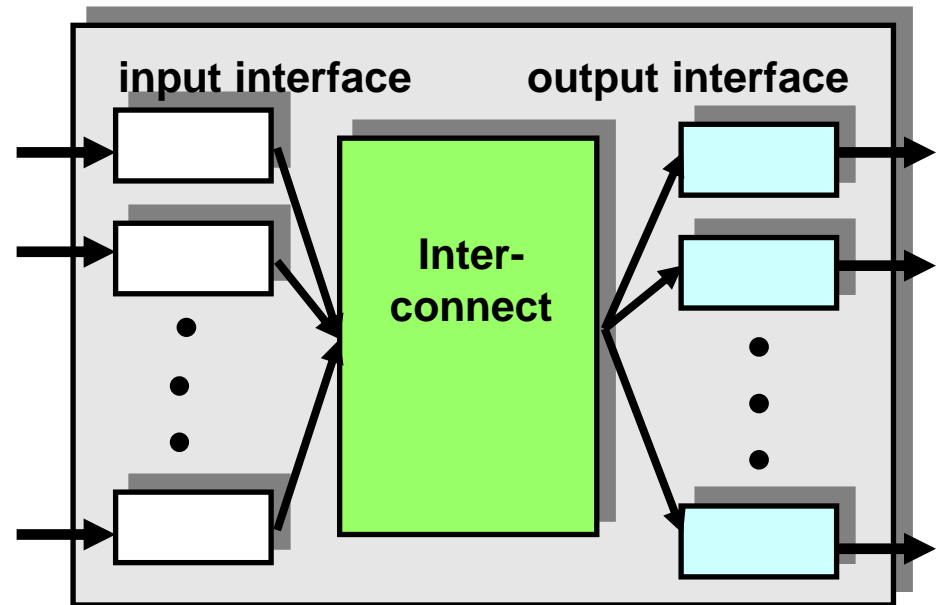
Δύο σημαντικές λειτουργίες του δρομολογητή:

- Τρέχει αλγορίθμους/πρωτόκολλα δρομολόγησης (πχ RIP, OSPF, BGP)
- Προωθεί τα datagrams από την εισερχόμενη στην εξερχόμενη σύνδεση

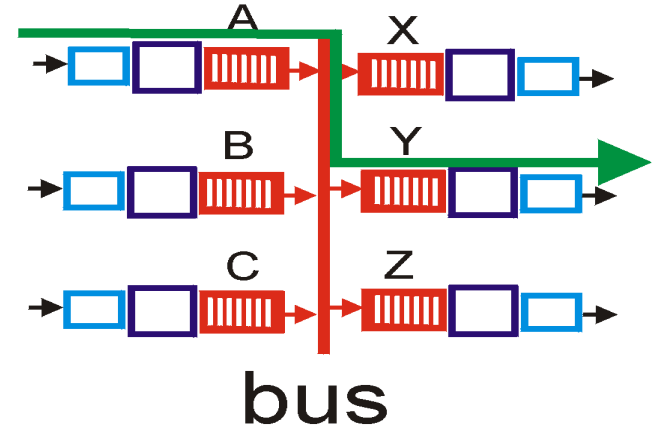
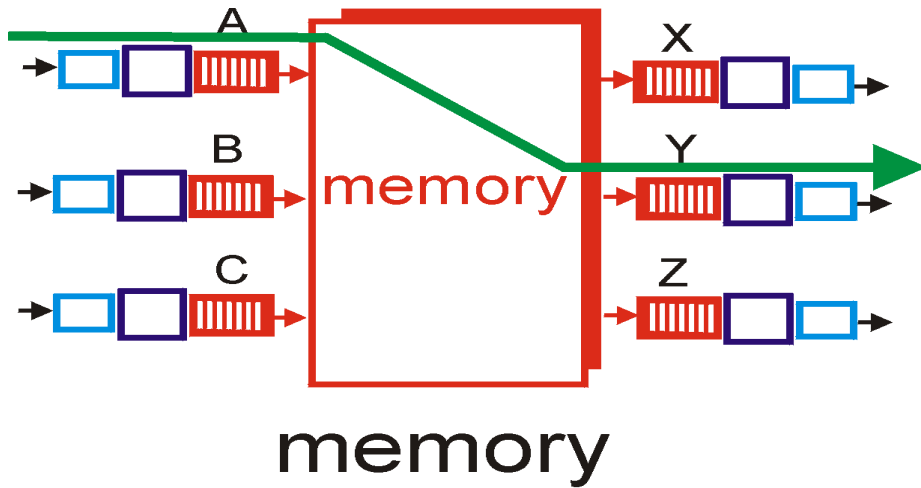


Γενική αρχιτεκτονική ενός δρομολογητή

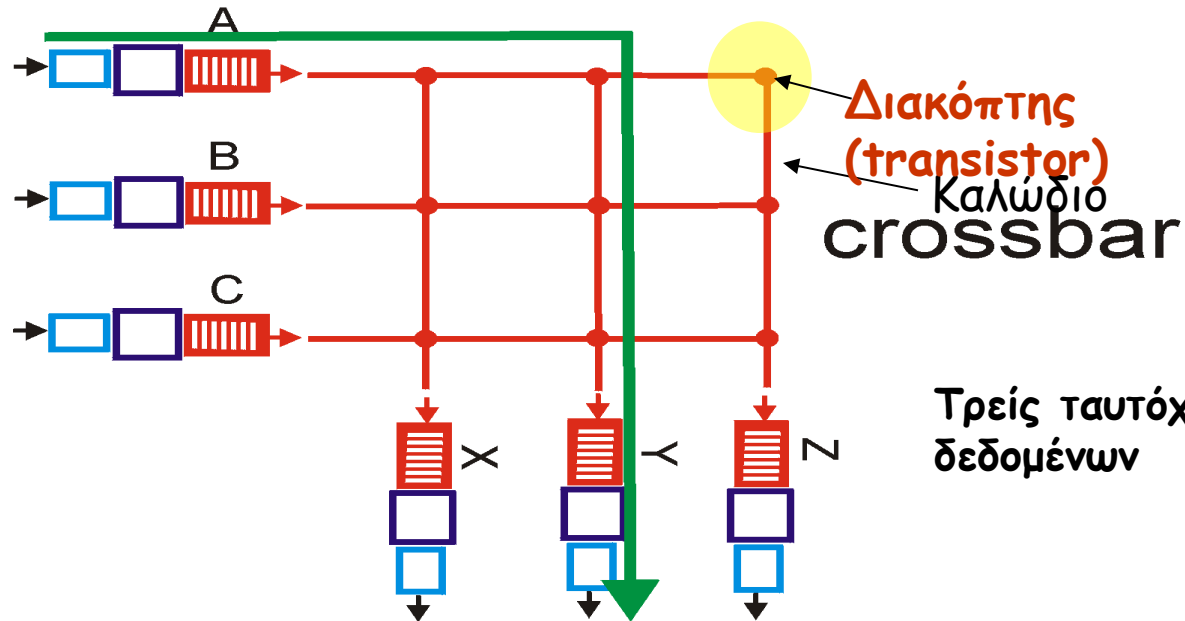
- Οι διεπαφές εισόδου & εξόδου συνδέονται μέσω μιας διασύνδεσης
- Η διασύνδεση μπορεί να υλοποιηθεί μέσω
 1. **Κοινής μνήμης**
δρομολογητές χαμηλής χωρητικότητας, π.χ., PC-based δρομολογητές
 2. **Διαμοιραζόμενος δίαυλος (shared bus)**
Δρομολογητές μεσαίας χωρητικότητας
 3. **Σημείο-προς-σημείο (switched) δίαυλος**
Δρομολογητές υψηλής χωρητικότητας



Βασικοί τύποι τεχνολογίας δρομολογητών

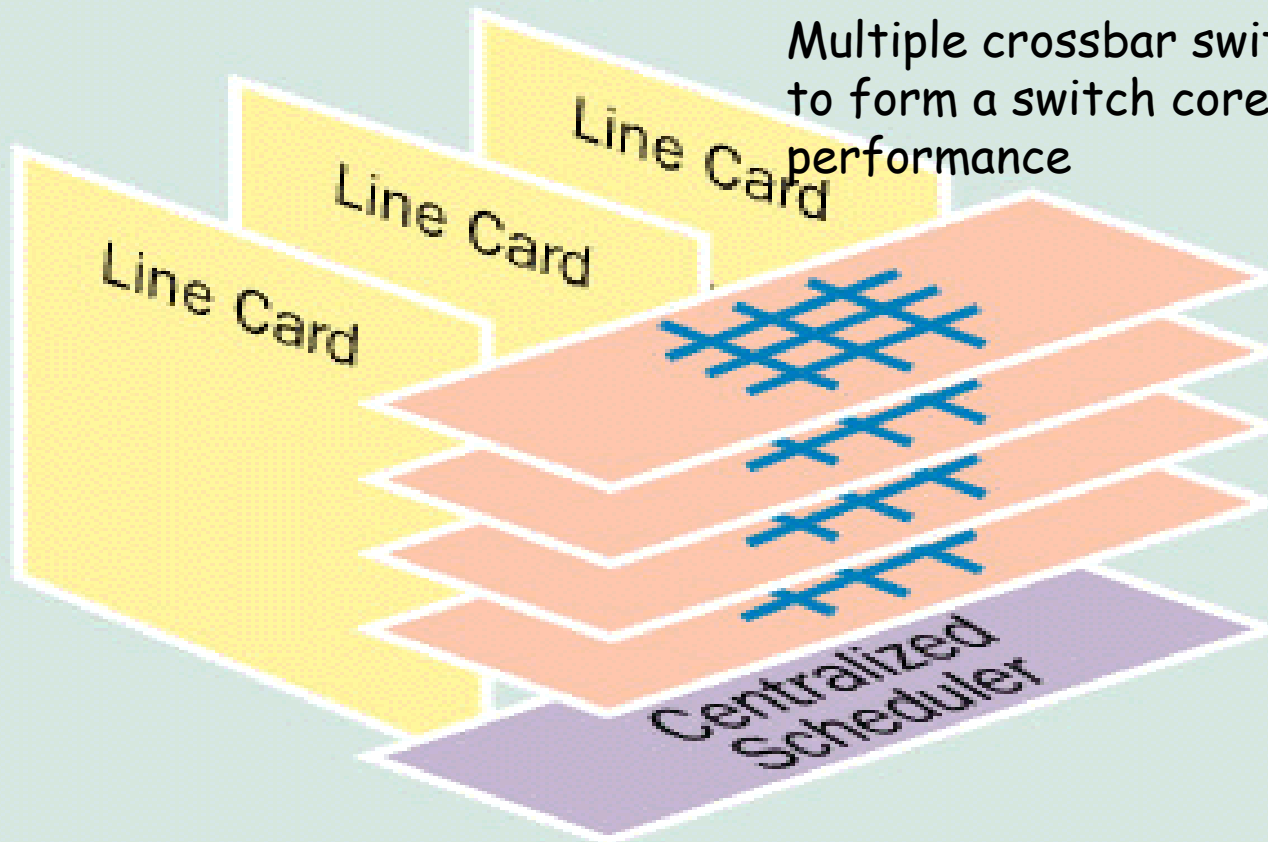


Μεγαλύτερος
παραλληλισμός



Τρεις ταυτόχρονες ροές
δεδομένων

FIGURE 5 A Four-Way Parallel Crossbar Switch, Interconnecting Three Line Cards



Multiple crossbar switches in parallel to form a switch core for boosting performance

Note: A centralized scheduler connects to each line card, and determines the configuration of the crossbar switch for each time slot.

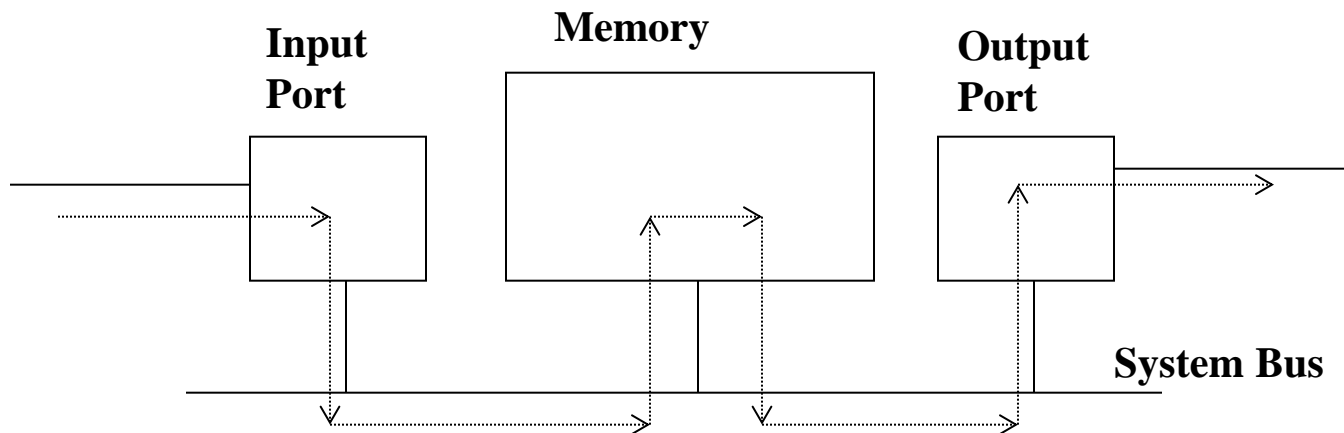
Switching Memory

Δρομολογητές πρώτης γενιάς

Παραδοσιακοί υπολογιστές με το **switching** κάτω από τον άμεσο έλεγχο της CPU

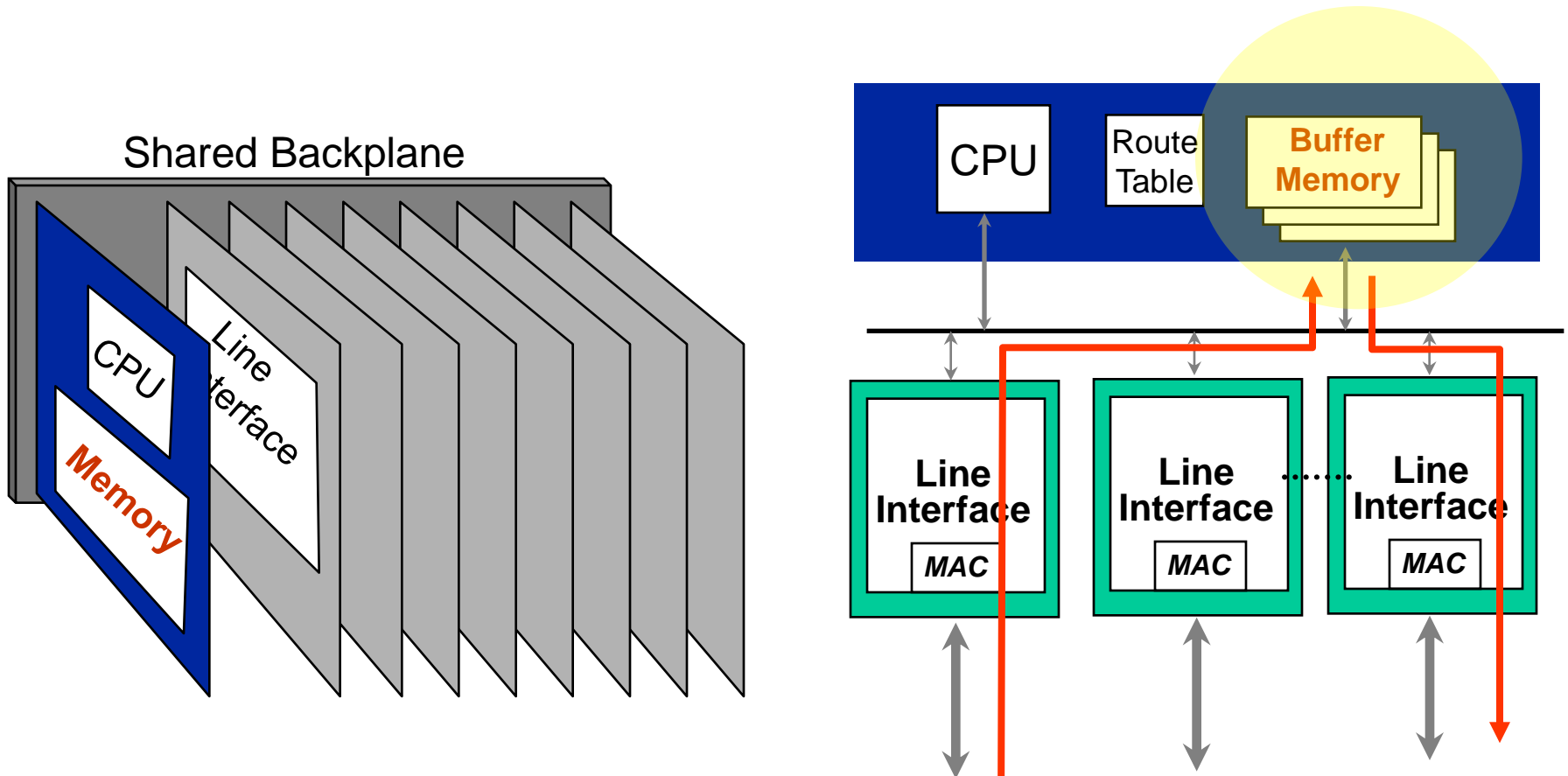
☞ Το πακέτο αντιγράφεται στη μνήμη του συστήματος

- η ταχύτητα μειώνεται από το bandwidth της μνήμης: 2 περάσματα από το δίαυλο ανά datagram



Κάθε χρονική στιγμή **μία** input port μπορεί να γράψει στη μνήμη & την επόμενη στιγμή **μία άλλη input port** μπορεί να γράψει **ή** μια output port να διαβάσει

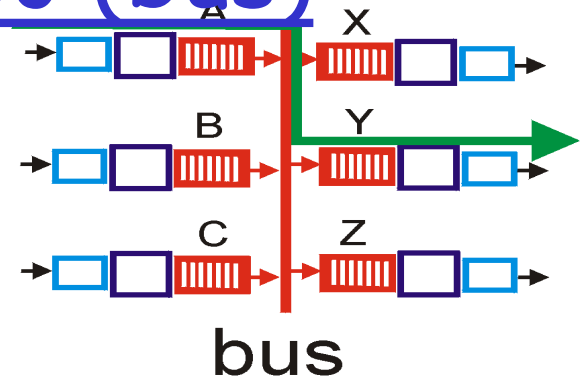
Διαμοιραζόμενη μνήμη (1η γενιά)



Τυπικά < 0.5Gbps συνολική χωρητικότητα

☞ Περιορίζεται από τον ρυθμό της διαμοιραζόμενης μνήμης

Switching μέσω ενός διαύλου (bus)



Το datagram από την θύρα εισόδου της μνήμης \Rightarrow θύρα εξόδου της μνήμης μέσω ενός διαμοιραζόμενου διαύλου

bus contention: Η ταχύτητα του switching *περιορίζεται* από το bandwidth του διαύλου

1 Gbps bus, Cisco 1900: *επαρκής ταχύτητα για access & enterprise routers* (όχι τοπικούς ή backbone)

Κάθε χρονική στιγμή *μία θύρα εισόδου* μπορεί να γράψει στο bus & την *επόμενη στιγμή μία άλλη θύρα εισόδου* μπορεί να γράψει ή *μία θύρα εξόδου να διαβάσει*

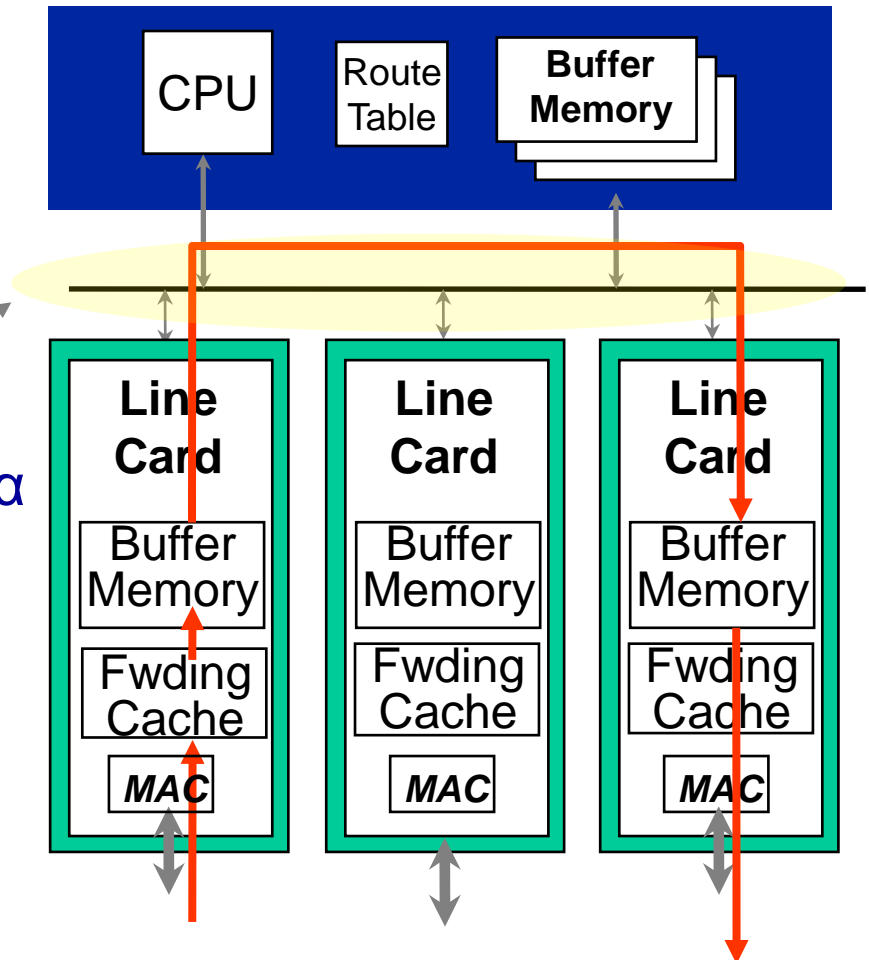
Διαμοιραζόμενος δίαυλος (shared bus) (2η γενιά)

Bypasses memory bus with direct transfer over bus between line cards

Moves forwarding decisions local to card to reduce CPU pain

Τυπικά < 5Gb/s συνολική χωρητικότητα

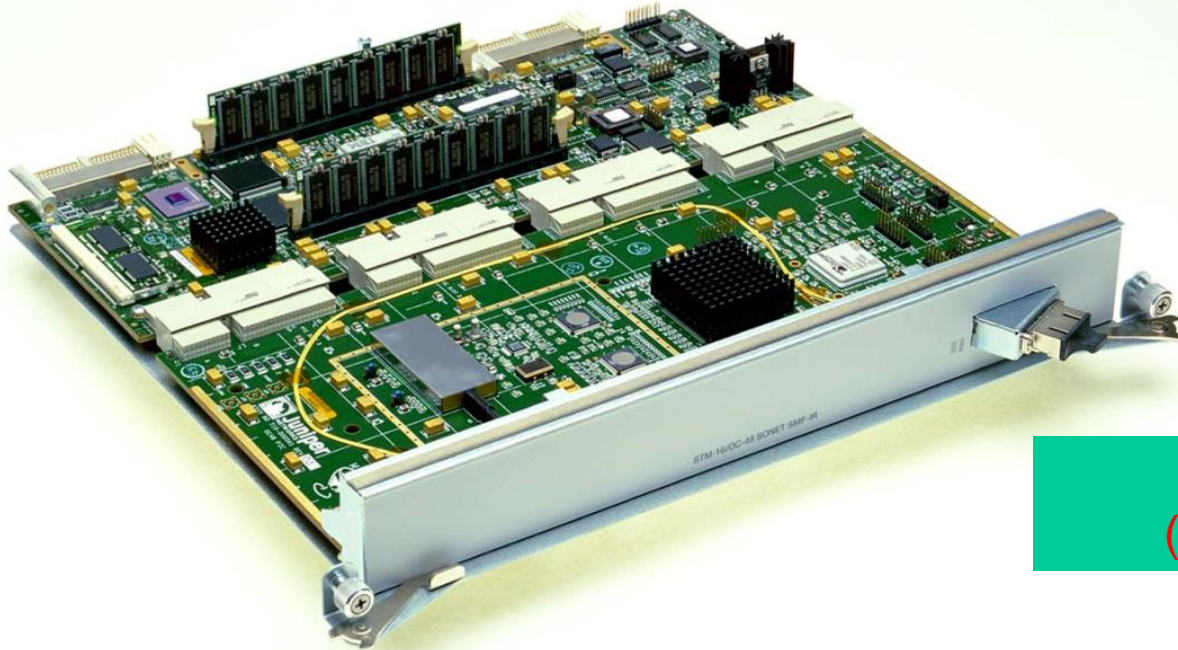
👉 Περιορίζεται από τον
διαμοιραζόμενο δίαυλο



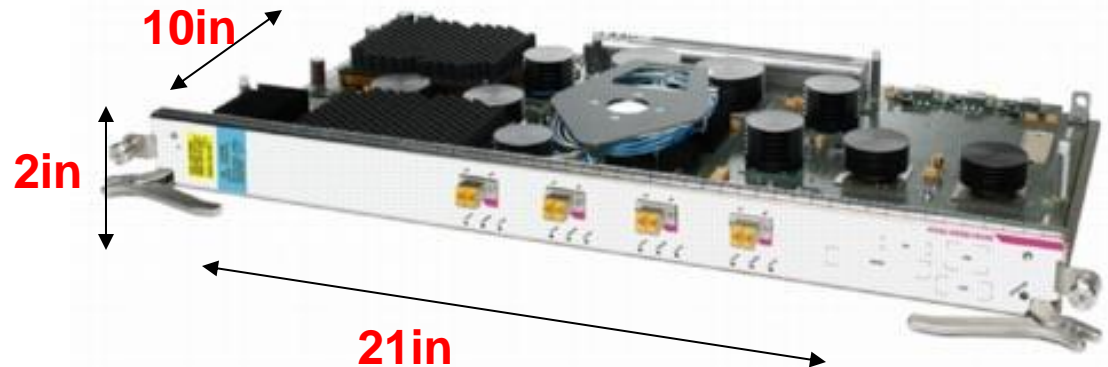
(* Slide by Nick McKeown)

What a Router Line Card Looks Like

1-Port OC48 (2.5 Gb/s)
(for Juniper M40)



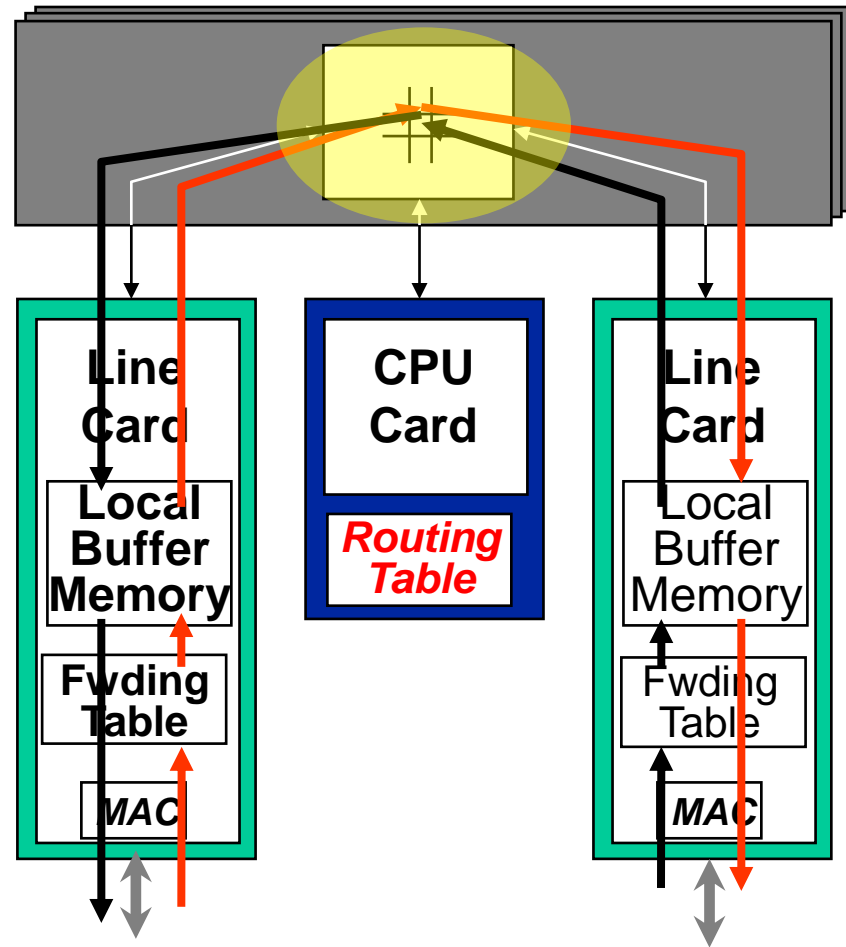
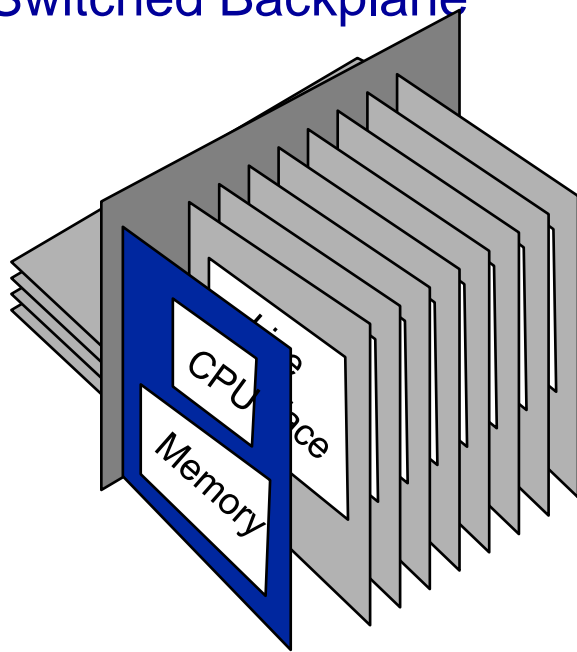
4-Port 10 GigE
(for Cisco CRS-1)



Power: about 150 Watts

Σημείο-προς-σημείο Switch (3η γενιά)

Switched Backplane



Τυπικά < 50Gbps συνολική χωρητικότητα

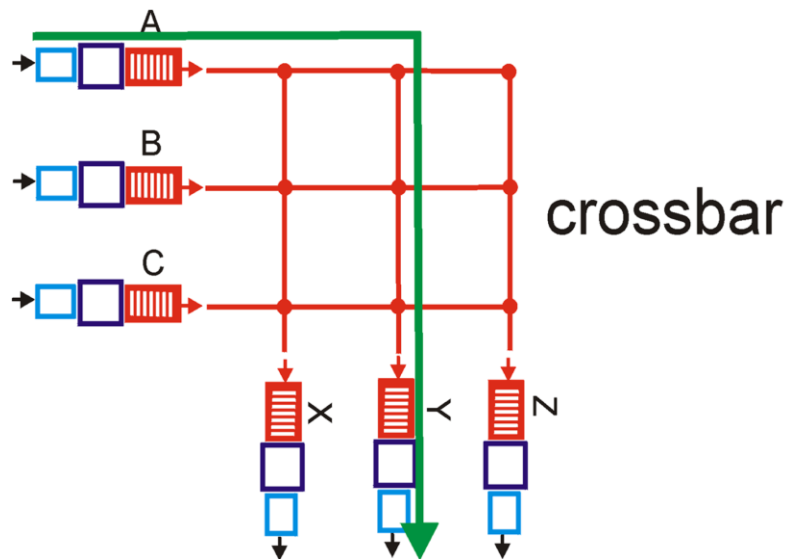
(*Slide by Nick McKeown)

Switched Backplane

- Every input port has a connection to every output port
- During each timeslot, each input connected to zero or one outputs

Advantage: Exploits parallelism

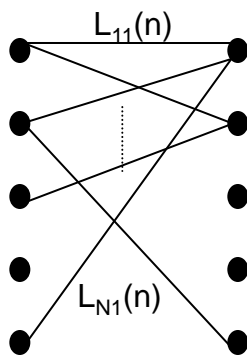
Disadvantage: Need scheduling algorithm



Crossbar Switching

- In each timeslot, one-to-one mapping between inputs and outputs
- Crossbar constraint: If input I is connected to output j , no other input connected to j , no other output connected to input I
- **Goal:** Maximal matching

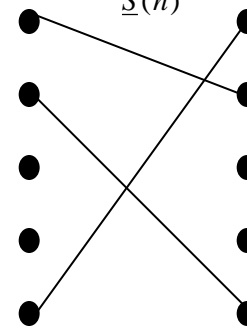
Traffic Demands



Maximum
Weight Match \rightarrow

Bipartite Match

$$S^*(n) = \arg \max_{\underline{S}(n)} (\underline{L}^T(n) \cdot \underline{S}(n))$$



Point-to-Point Switch - Διασύνδεση (Interconnect)

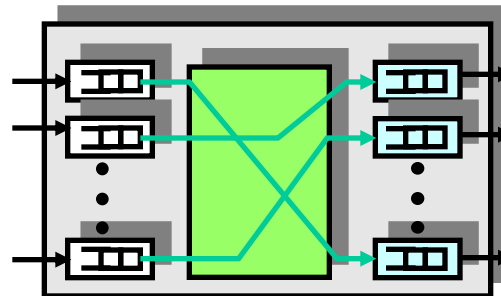
Point-to-point switch επιτρέπει παράλληλες μεταφορές πακέτων μεταξύ οποιοδήποτε δύο ζευγών input-output interfaces αρκεί να *μην* περιλαμβάνουν *κοινό interface* τα δύο αυτά ζεύγη

- ☞ Στόχος: σχεδιασμός ενός προγράμματος μετάδοσης των πακέτων που να
- Παρέχει ποιότητα υπηρεσιών (*quality of service*)
 - Βελτιστοποιεί το throughput του δρομολογητή

Προκλήσεις:

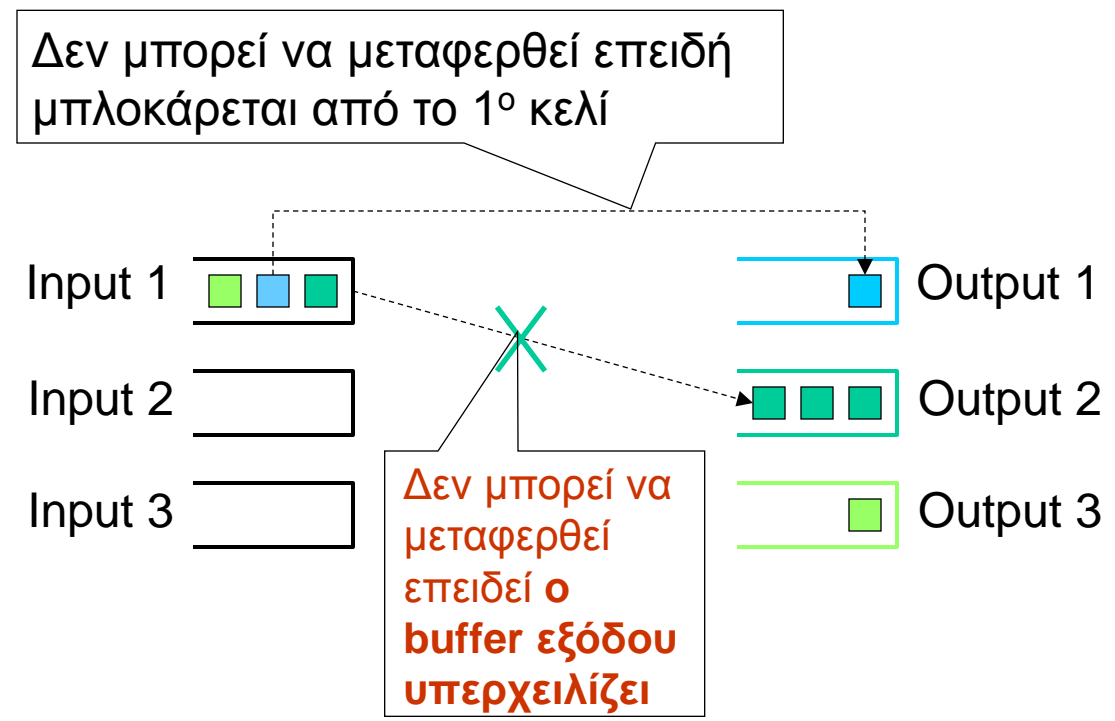
- Αντιμετώπιση του προβλήματος μπλοκαρίσματος στην αρχή της ουράς (*head-of-line blocking*) στις εισόδους
- Επίλυση input/output speedups contention
- Αποφυγή του "πετάματος" (drop) των πακέτων στην έξοδο λόγω υπερχείλισης (όταν είναι δυνατό)

✍ Σημείωση: τα πακέτα είναι κατακερματισμένα σε κελιά (cells) σταθερού μεγέθους στις εισόδους & επανασυναρμολογούνται στις εξόδους

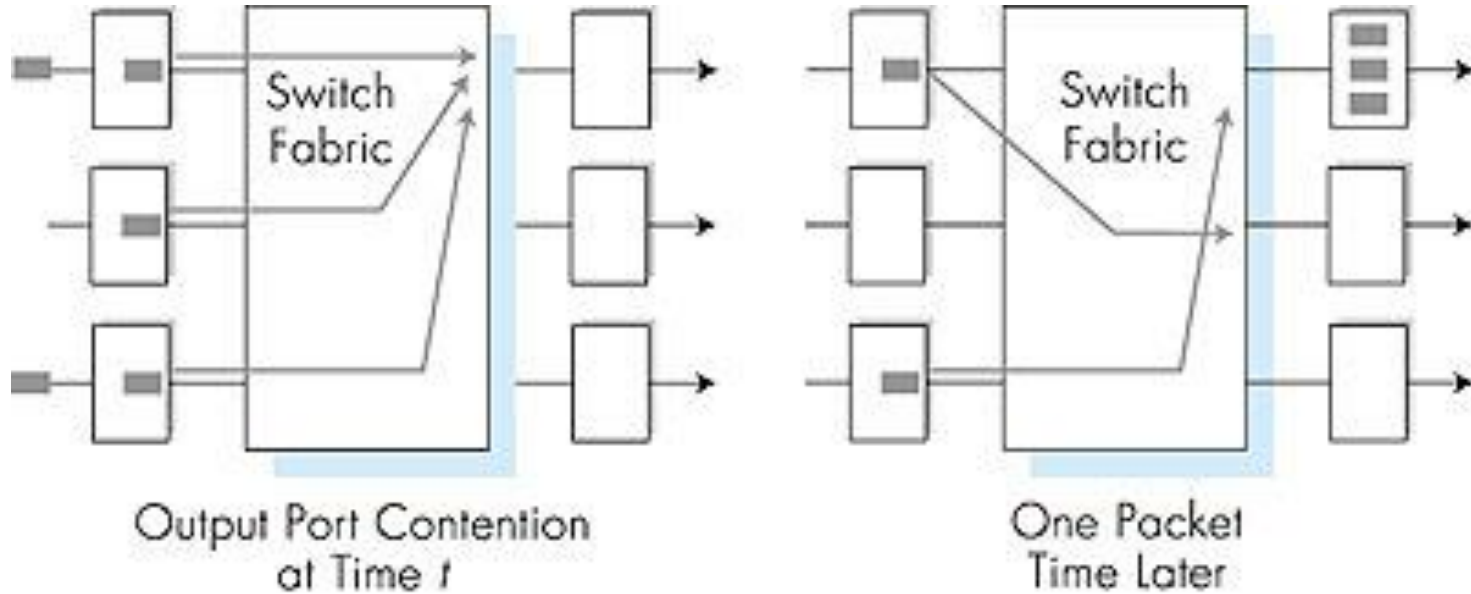


Μπλοκάρισμα της αρχής της ουράς (Head-of-line Blocking)

Το κελί στην **κεφαλή μίας ουράς εισόδου** δεν μπορεί να μεταφερθεί, και έτσι μπλοκάρει τα επόμενα κελιά



Αναμονή στην ουρά της θύρας εξόδου



Αποθήκευση όταν ο ρυθμός άφιξης μέσω του switch ξεπερνάει την ταχύτητα της γραμμής εξόδου

Αναμονή στην ουρά (καθυστέρηση) και απώλειες λόγω υπερχείλησης του buffer εξόδου!

Router buffer

The performance of a router is affected by the size of router buffer and the memory bandwidth

Switching μέσω ενός δικτύου διασύνδεσης

- Ξεπερνάει τους περιορισμούς στο bandwidth του διαύλου
- Άλλα δίκτυα διασύνδεσης σχεδιάστηκαν αρχικά για να συνδέουν επεξεργαστές μέσα σε πολυεπεξεργαστές
- Ανεπτυγμένος σχεδιασμός: με τη διάσπαση του datagram σε κελιά σταθερού μεγέθους, τα κελιά γίνονται switch μέσω του fabric
πχ Cisco 12000: switches Gbps μέσω του δικτύου διασύνδεσης

Βασικές Λειτουργίες του Επιπέδου Δικτύου

Πρώθηση (forwarding): προωθεί τα πακέτα από το input του δρομολογητή στο κατάλληλο output

Δρομολόγηση (routing): καθορίζει το μονοπάτι που θα πάρουν τα πακέτα από τον αποστολέα στον παραλήπτη

Αλγόριθμοι δρομολόγησης

↳ **δρομολόγηση:** διαδικασία σχεδιασμού/οργάνωσης του ταξιδιού από τον αποστολέα στον παραλήπτη

⇒ **Πρώθηση:** διαδικασία μίας μετάβασης από τον ένα σταθμό στον άλλο

Εγκαθίδρυση σύνδεσης

- Η 3^η πιο σημαντική λειτουργία σε κάποιες αρχιτεκτονικές δικτύου:
ATM, frame relay, X.25
- Πριν σταλούν τα datagrams , οι δύο hosts & οι παρεμβαλλόμενοι δρομολογητές εγκαθιδρύουν μία εικονική σύνδεση
Συμμετοχή των δρομολογητών
- Υπηρεσία σύνδεσης σε επίπεδο δικτύου και μεταφοράς:
Δίκτυο: μεταξύ των δύο hosts
Μεταφορά: μεταξύ των δύο διεργασιών

Μοντέλο υπηρεσίας δικτύου

Q: Τι *μοντέλο υπηρεσίας* για το "κανάλι" που μεταφέρει datagrams από τον αποστολέα στον παραλήπτη?

Παράδειγμα υπηρεσιών για κάθε datagram:

- Εγγυημένη παράδοση
- Εγγυημένη παράδοση με καθυστέρηση <40 msec

Παράδειγμα υπηρεσιών για μία ροή από datagrams:

- Πράδοση των datagrams σε σειρά
- Εγγυημένο ελάχιστο bandwidth για τη ροή
- Περιορισμοί σε αλλαγές και κενά μεταξύ των πακέτων

Μοντέλα υπηρεσιών επιπέδου δικτύου:

Αρχιτεκτονική δικτύου	Μοντέλο υπηρεσίας	Εγγυάται ?			Ένδειξη συμφόρησης	
		Bandwidth	Απώλεια	Σειρά Χρονισμός		
Διαδίκτυο	Βέλτιστης προσπάθειας	κανένα	όχι	όχι	όχι	Όχι (υπονοείται από απώλεια)
ATM	CBR	Σταθερός ρυθμός	ναι	ναι	ναι	καμία συμφόρηση
ATM	VBR	Εγγυημένος ρυθμός	ναι	ναι	ναι	καμία συμφόρηση
ATM	ABR	Εγγυημένο ελάχιστο	όχι	ναι	όχι	ναι
ATM	UBR	κανένα	όχι	ναι	όχι	όχι

Υπηρεσία σύνδεσης και ασυνδειστρεφής υπηρεσία επιπέδου δικτύου

Δίκτυο από αυτοδύναμα πακέτα (datagram) παρέχει ασυνδειστρεφή υπηρεσία επιπέδου δικτύου

Δίκτυο ιδεατών κυκλωμάτων (VC) παρέχει υπηρεσία σύνδεσης επιπέδου δικτύου

Ανάλογο με την υπηρεσία επιπέδου μεταφοράς, αλλά:

- **Υπηρεσία:** από host-σε-host
- **Καμία επιλογή:** το δίκτυο παρέχει ή το ένα ή το άλλο
- **Υλοποίηση:** στον πυρήνα

Ιδεατά κυκλώματα (virtual circuits)

“source-to-dest μονοπάτι συμπεριφέρεται όπως μια τηλεφωνική σύνδεση (circuit)”

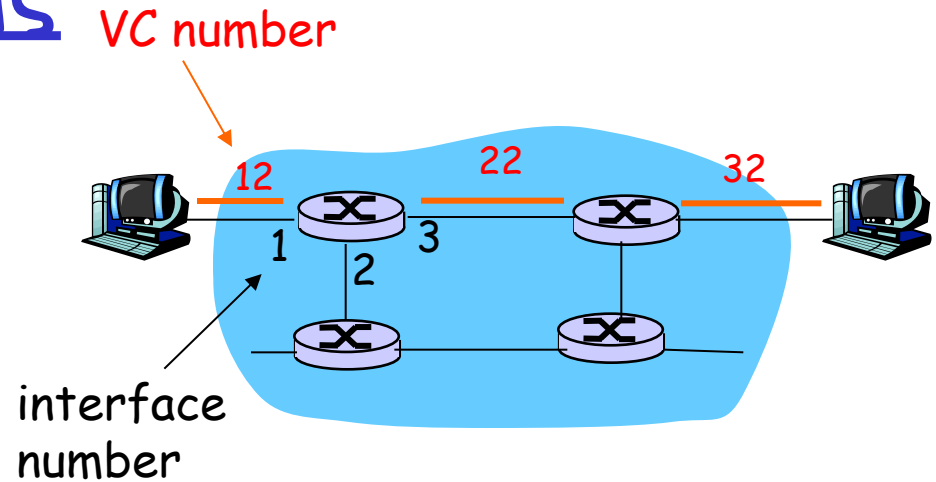
- Εγκαθίδρυση κλήσης πριν τη μεταφορά δεδομένων και τερματισμός της μετά τη μεταφορά
- Κάθε πακέτο έχει έναν VC identifier (όχι διεύθυνση host προορισμού)
- *Κάθε δρομολογητής στο μονοπάτι πηγής-προορισμού διατηρεί κατάσταση για κάθε σύνδεση που περνάει από αυτόν*
- Ζεύξεις, πόροι δρομολογητών (bandwidth, buffers) μπορούν να απονεμηθούν στο VC

VC υλοποίηση

Ένα VC αποτελείται από:

1. Μονοπάτι από την πηγή στον προορισμό
 2. VC αριθμούς, ένας αριθμός για κάθε ζεύξη κατά μήκος του μονοπατιού
 3. Εγγραφές στους πίνακες προώθησης των δρομολογητών κατά μήκος του μονοπατιού
- Το πακέτο που ανήκει στο VC μεταφέρει έναν VC αριθμό
 - Ο VC αριθμός πρέπει να αλλάζει σε κάθε ζεύξη
Ο νέος VC αριθμός προέρχεται από τον πίνακα προώθησης

Πίνακας προώθησης



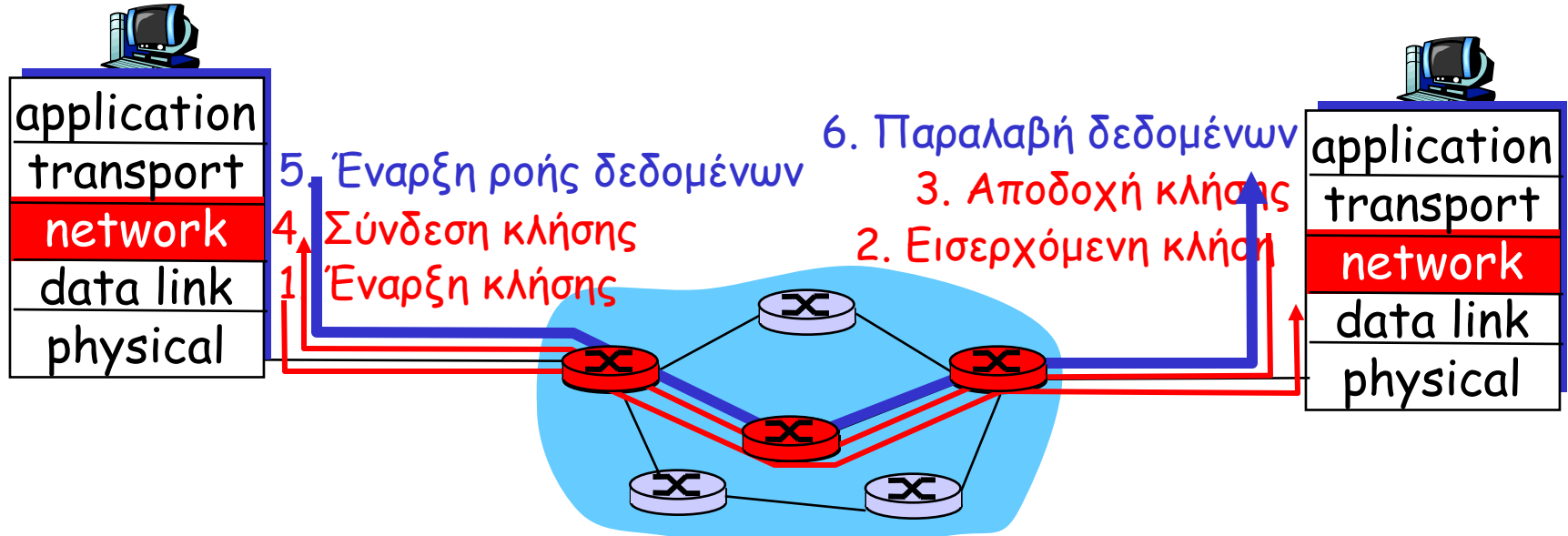
Πίνακας προώθησης για τον δρομολογητή:

εισερχόμενο interface	εισερχόμενος VC #	εξερχόμενο interface	εξερχόμενος VC #
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
...

Οι δρομολογητές διατηρούν πληροφορία για την κατάσταση της σύνδεσης!

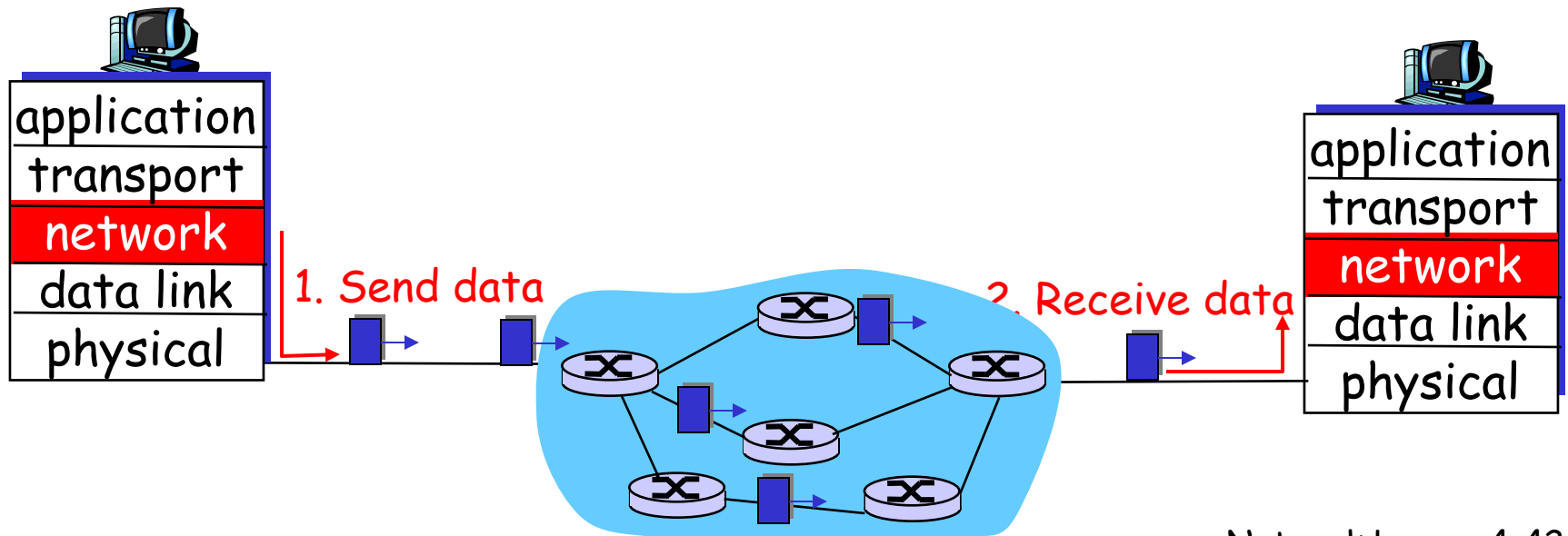
Ιδεατά κυκλώματα: πρωτόκολλα σηματοδosis

- χρησιμοποιούνται για εγκαθίδρυση, διατήρηση και τερματισμό των VC
 - Χρησιμοποιούνται στο ATM, frame-relay, X.25
- 📢 Δεν χρησιμοποιούνται στο σημερινό διαδίκτυο (Internet)



Δίκτυα αυτοδύναμων πακέτων (datagram networks)

- Καμία εγκαθίδρυση κλήσης στο επίπεδο δικτύου
- δρομολογητές: καμία κατάσταση για τις συνδέσεις από άκρη σε άκρη
 - Δεν υπάρχει η έννοια της σύνδεσης στο επίπεδο δικτύου
- Τα πακέτα προωθούνται με βάση τη διεύθυνση του host προορισμού
- ☞ Τα πακέτα μεταξύ του ίδιου ζευγαριού πηγής-προορισμού μπορεί να πάρουν διαφορετικά μονοπάτια



Δίκτυο ιδεατών κυκλωμάτων ή αυτοδύναμων πακέτων

Διαδίκτυο

- Ανταλλάγη δεδομένων μεταξύ υπολογιστών
 - «ελαστική» υπηρεσία, χωρίς αυστηρές απαιτήσεις χρονισμού
- “ευφυή” τερματικά συστήματα (computers)
 - Μπορούν να προσαρμόζονται, να ελεγχουν, να επανέρχονται από λαθη
 - Απλότητα στο εσωτερικό του δικτύου
 - πολυπλοκότητα “άκρα”
- Πολλοί τύποι ζεύξεων
 - Διαφορετικά χαρακτηριστικά
 - Δύσκολα έχουμε ομοιόμορφη υπηρεσία

ΑΤΜ

- Εξελίχθηκε από το τηλεφωνικό δίκτυο
- Ανθρώπινη συνομιλία:
 - Αυστηρές απαιτήσεις χρονισμού και αξιοπιστίας
- Ανάγκη για εγγυημένη υπηρεσία
- “χαζά” τερματικά συστήματα
 - Τηλεφωνα
 - Πολυπλοκότητα μέσα στο δίκτυο

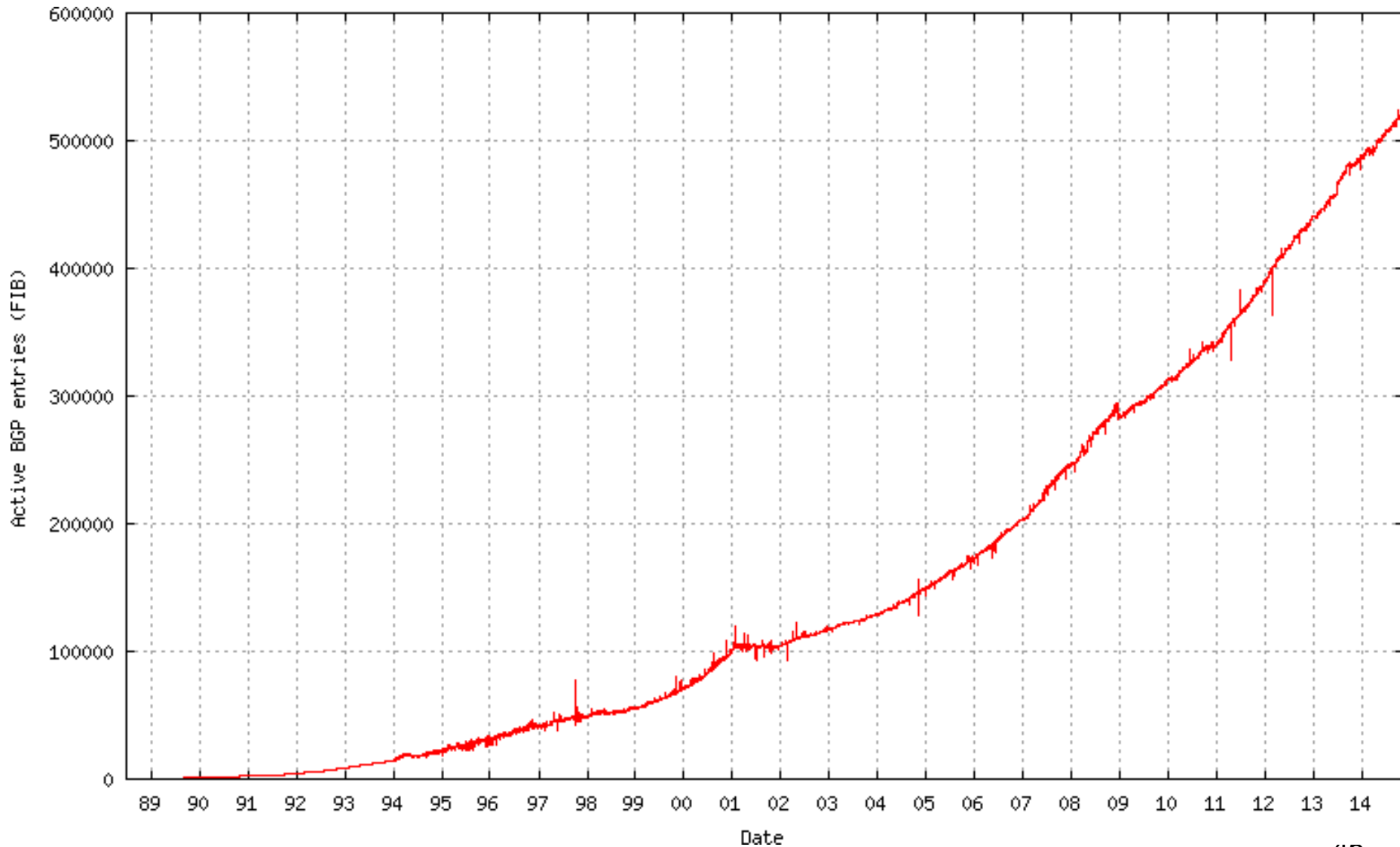
Προσέξτε τους όρους: “τερματικά συστήματα”, “άκρα”, ελαστική υπηρεσία

IP Address Lookup

Challenges:

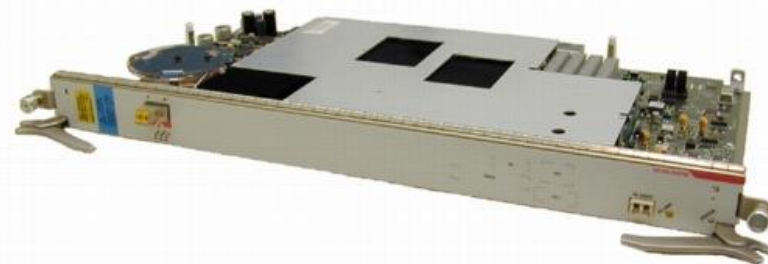
1. Longest-prefix match (not exact).
2. Tables are large and growing.
3. Lookups must be fast.

Address Tables are Large



Lookups Must be Fast

Year	Line	40B packets (Mpkt/s)
1997	622Mb/s	1.94
1999	2.5Gb/s	7.81
2001	10Gb/s	31.25
2003	40Gb/s	125



**Cisco CRS-1 1-Port OC-768C
(Line rate: 42.1 Gb/s)**

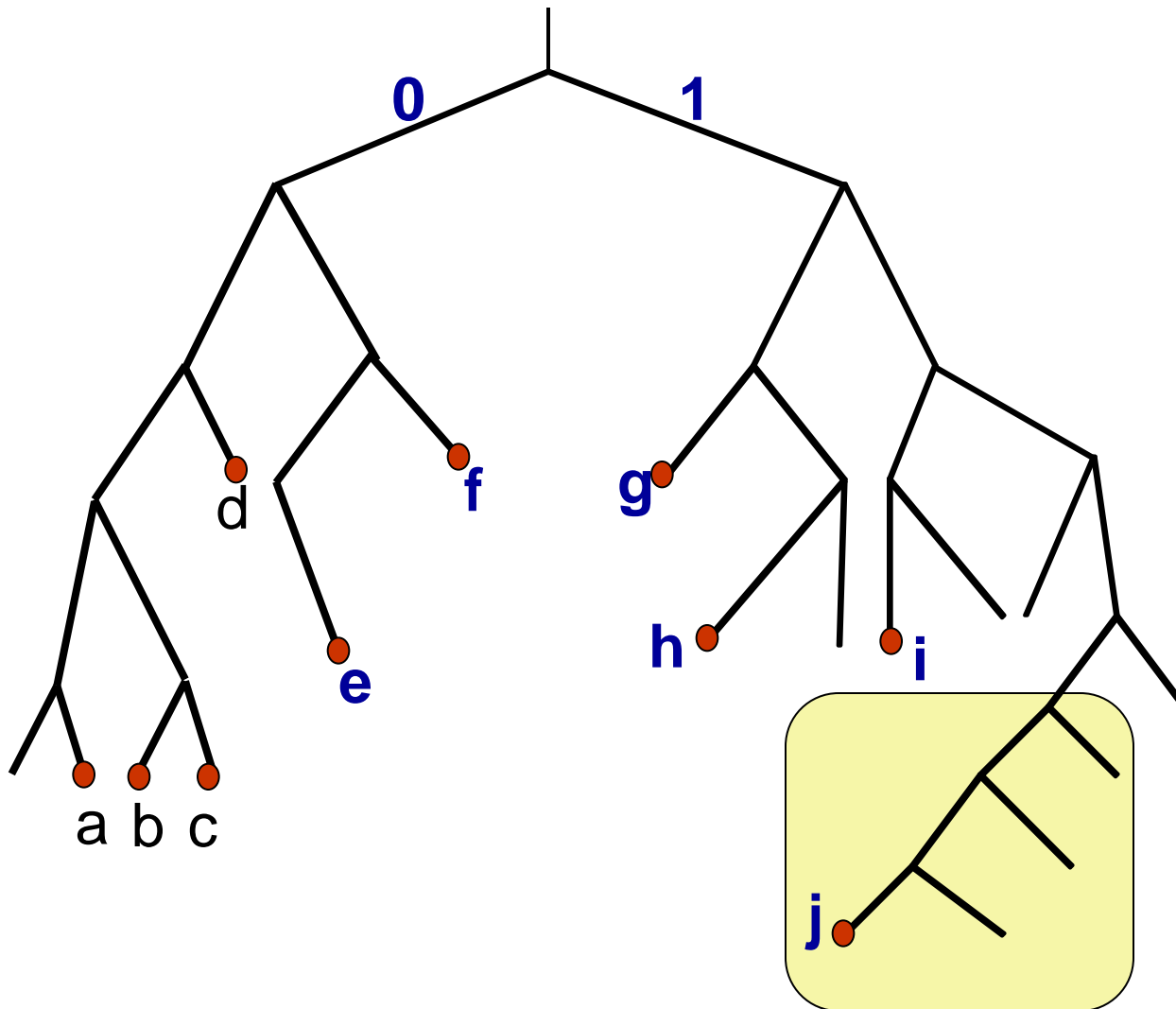
OC-12

OC-48

OC-192

OC-768

IP Address Lookup: Binary Tries

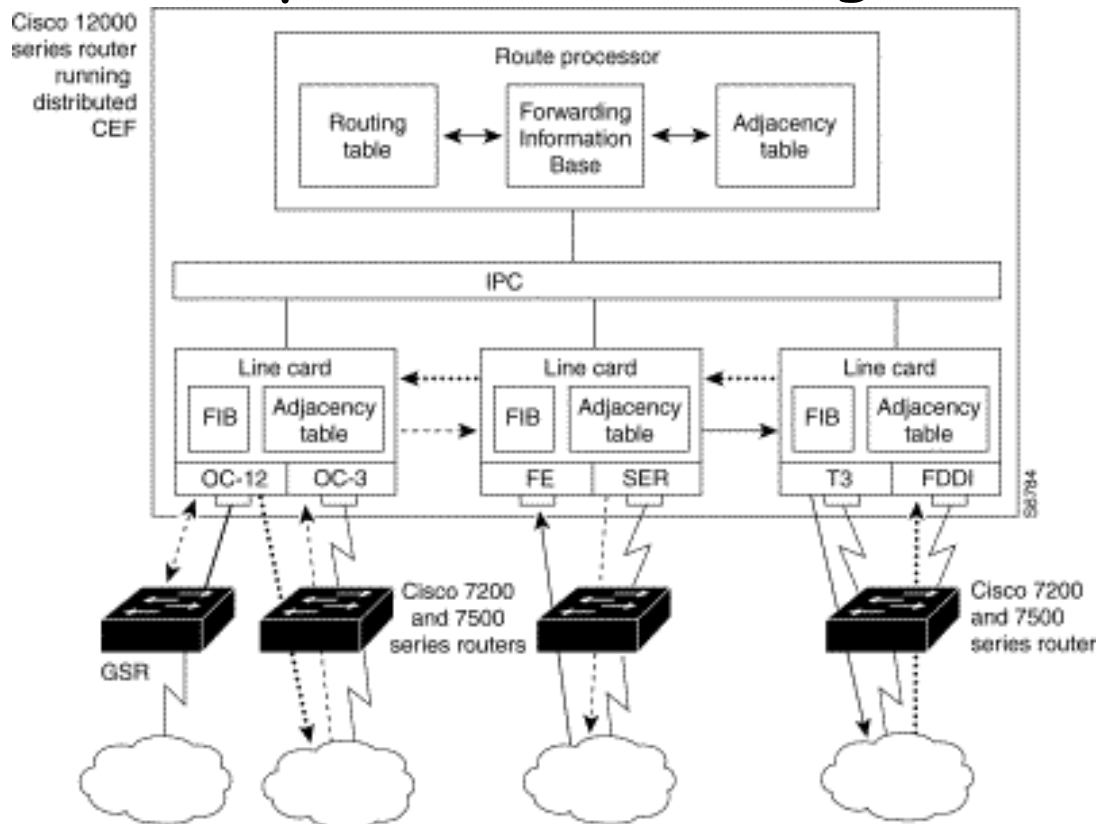


Example Prefixes:

- a) 00001
- b) 00010
- c) 00011
- d) 001
- e) 0101
- f) 011
- g) 100
- h) 1010
- i) 1100
- j) 11110000

Faster Lookup: Alternatives

- Caching
 - Packet trains exhibit temporal locality
 - Many packets to same destination
- Cisco Express Forwarding



Πίνακας προώθησης

Διευθύνσεις προορισμού

Διεπαφή ζεύξης

11001000 00010111 00010000 00000000

ως

0

11001000 00010111 00010111 11111111

11001000 00010111 00011000 00000000

ως

1

11001000 00010111 00011000 11111111

11001000 00010111 00011001 00000000

ως

2

11001000 00010111 00011111 11111111

αλλιώς

Longest prefix matching

<u>Prefix Match</u>	<u>Διεπαφή ζεύξης</u>
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
otherwise	3

Παραδείγματα

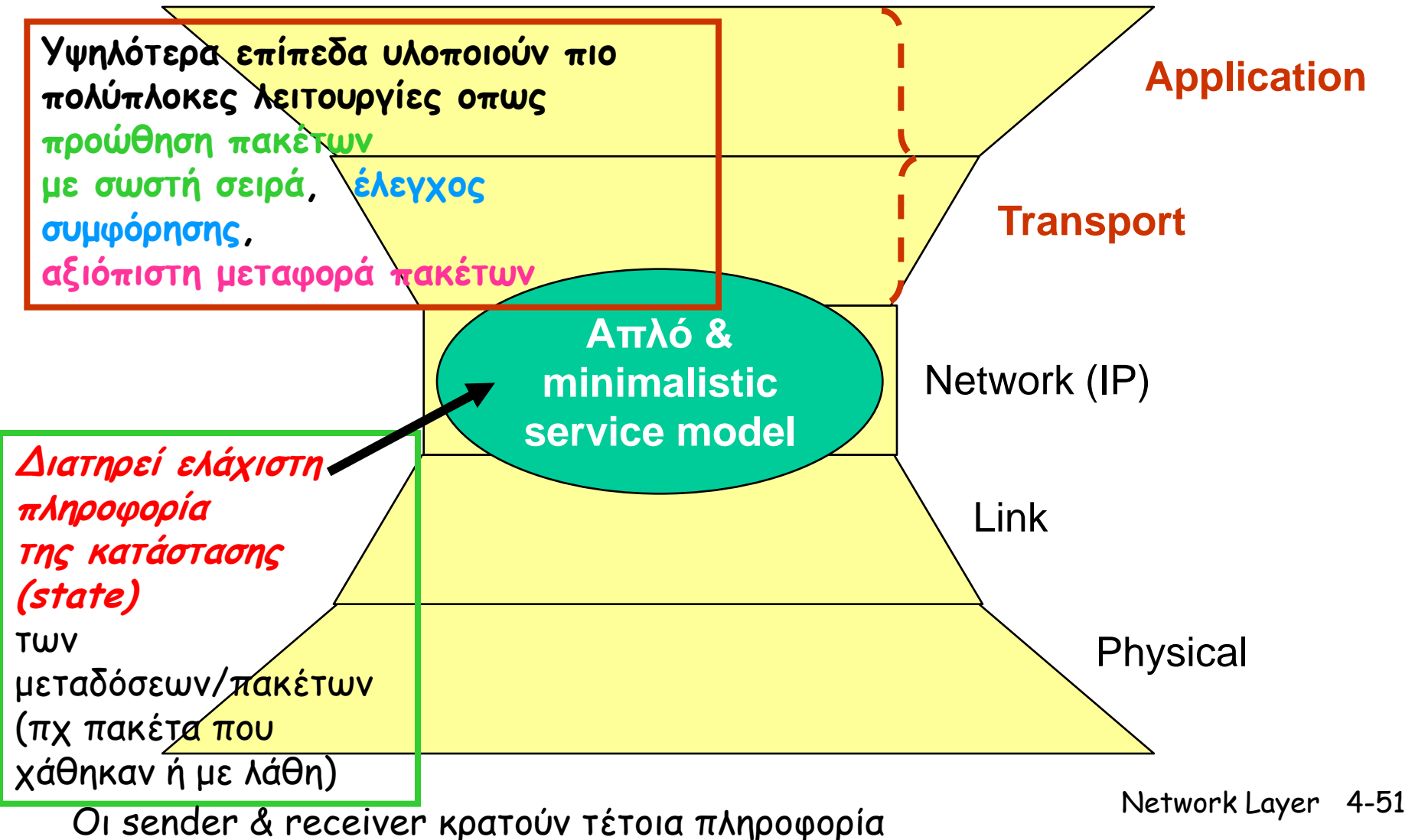
DA: 11001000 00010111 00010110 10100001

Ποιά διεπαφή?

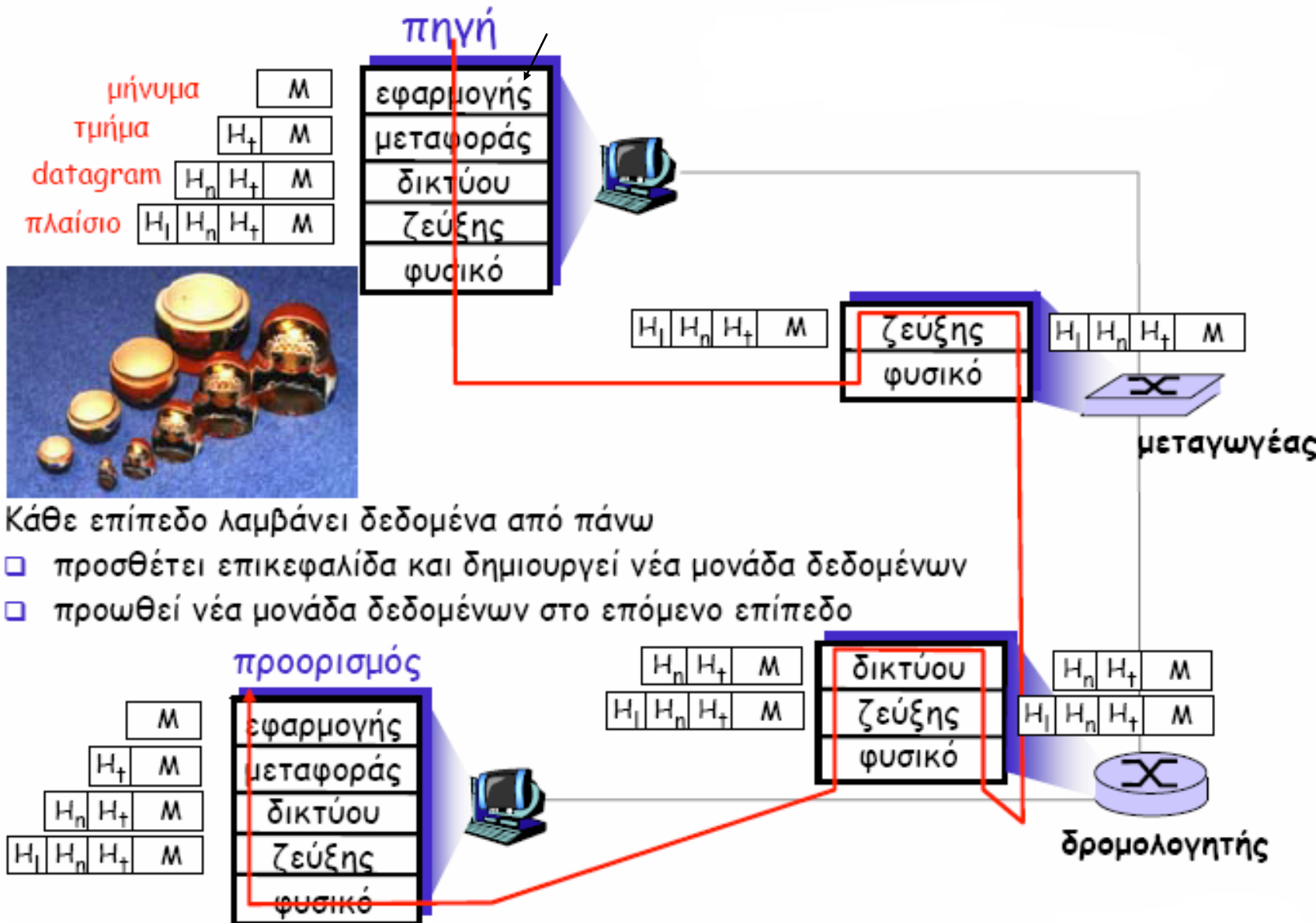
DA: 11001000 00010111 00011000 10101010

Ποια διεπαφή?

Complexity at the end-hosts

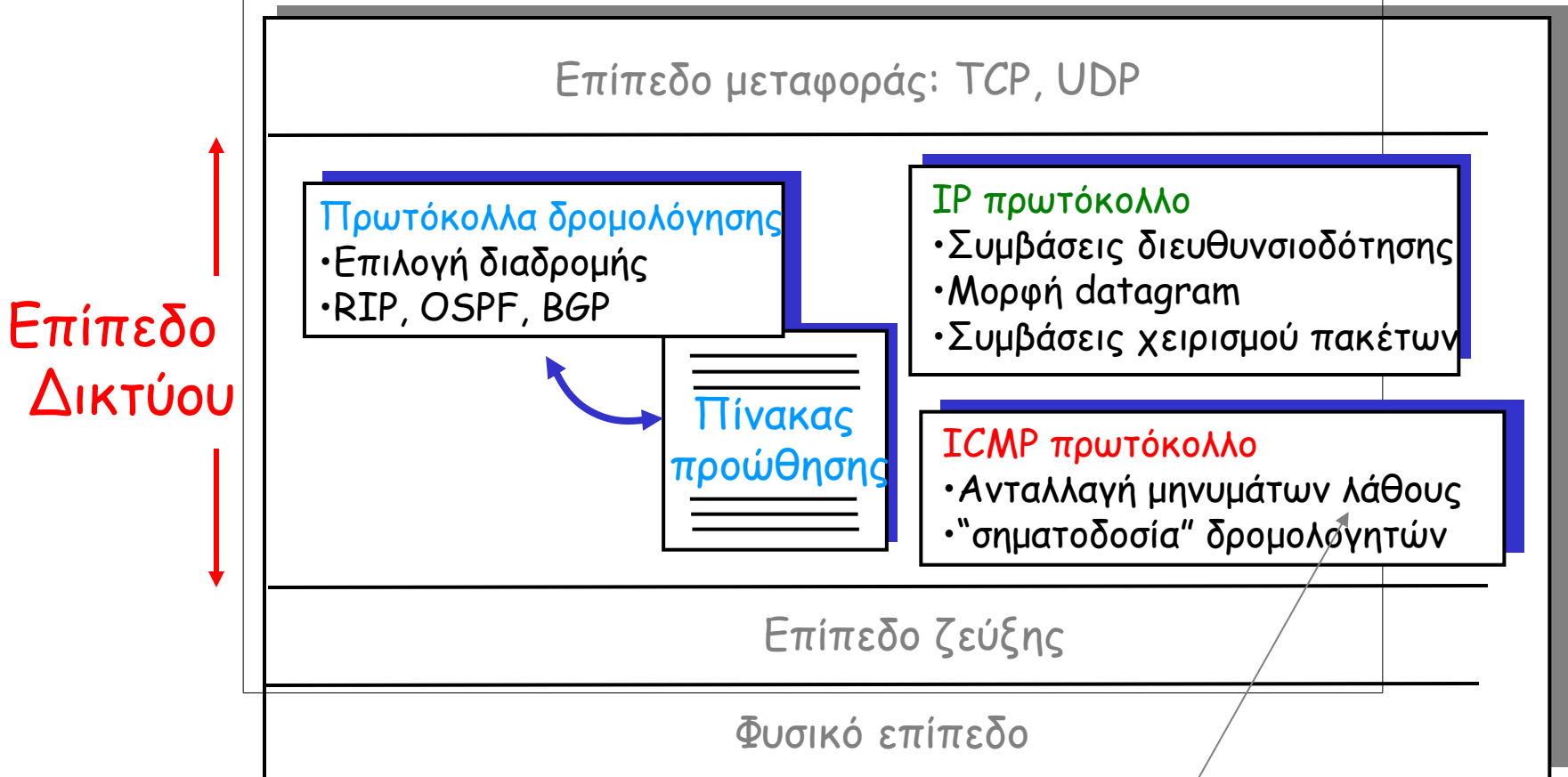


Ενθυλάκωση



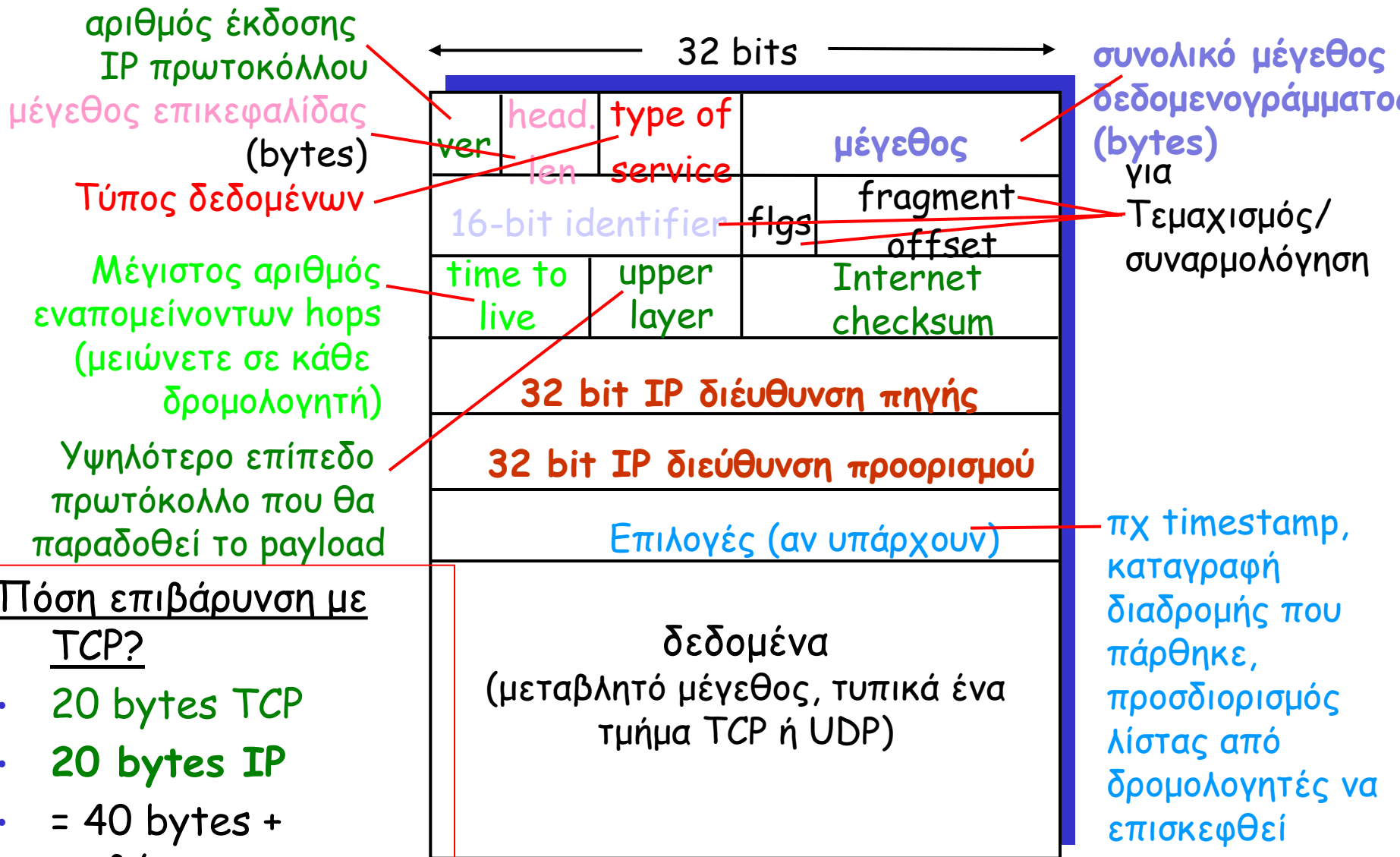
Το επίπεδο δικτύου στο Διαδίκτυο

Λειτουργίες του επιπέδου δικτύου host και δρομολογητές:



πχ έλλειψη κάποιας υπηρεσίας από έναν server,
απουσία ενός υπολογιστή από το δίκτυο

Διάταξη IP δεδομενογράμματος



Πόση επιβάρυνση με TCP?

- 20 bytes TCP
- 20 bytes IP
- = 40 bytes + επιβάρυνση

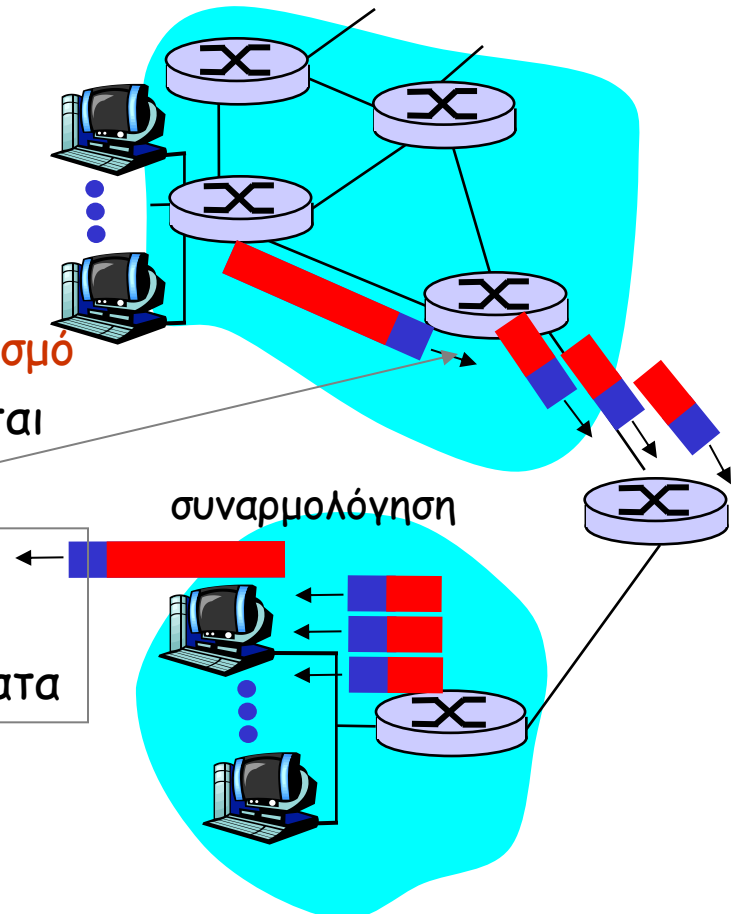
ανώτερου επιπέδου

IP Τεμαχισμός & Συναρμολόγηση

- Οι ζεύξεις δικτύου έχουν MTU (μέγιστο μέγεθος μεταφοράς): **μεγαλύτερο πιθανό πλαίσιο επιπέδου ζεύξης**
διαφορετικοί τύποι ζεύξης, διαφορετικά MTUs

- μεγάλα IP δεδομογράμματα χωρίζονται ("τεμαχίζονται") μέσα στο δίκτυο
 - Ένα δεδομένογραμμα σπάει σε μικρότερα
 - "συναρμολογούνται" **μόνο** στον τελικό προορισμό
 - Τα bits της IP επικεφαλίδας χρησιμοποιούνται για να αναγνωρίσουν, κομμάτια σε σειρά

τεμαχισμός:
in: ένα μεγάλο δεδομένογραμμα
out: 3 μικρότερα δεδομογράμματα



IP Τεμαχισμός & Συναρμολόγηση

	length	ID	fragflag	offset	
	=4000	=x	=0	=0	

Παράδειγμα

- 4000 byte δεδομένογραμμα
- MTU = 1500 bytes

Ένα μεγάλο δεδομένογραμμα γίνεται μερικά μικρότερα δεδομένογραμματα

1480 bytes στο πεδίο δεδομένων

offset =
1480/8

	length	ID	fragflag	offset	
	=1500	=x	=1	=0	

	length	ID	fragflag	offset	
	=1500	=x	=1	=185	

	length	ID	fragflag	offset	
	=1040	=x	=0	=370	

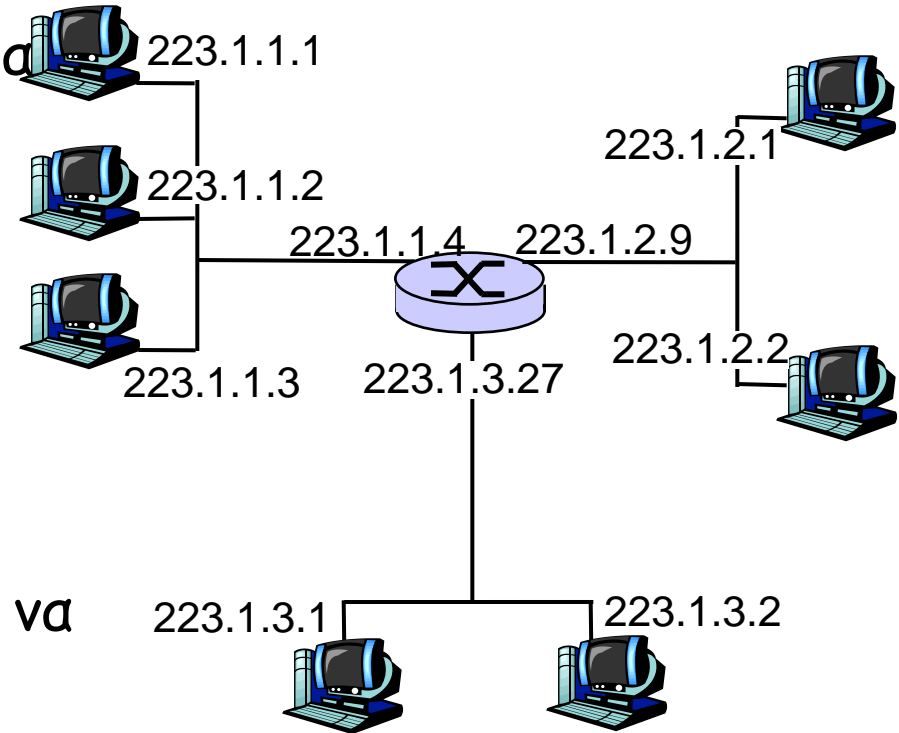
IP διευθυνσιοδότηση: εισαγωγή

IP διεύθυνση: **32-bit ταυτότητα** για τη διεπαφή host/δρομολογητή

Διεπαφή: σύνδεση μεταξύ host/δρομολογητή & φυσικής ζεύξης

Οι δρομολογητές γενικά έχουν **πολλές διεπαφές**

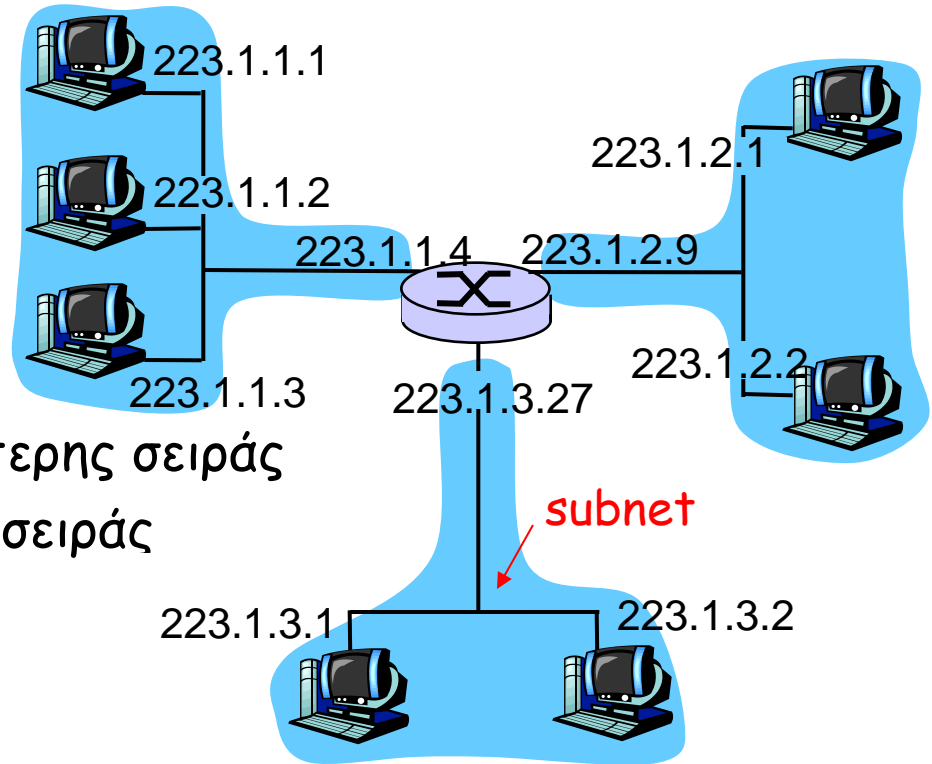
☞ **Αλλά και ένα host μπορεί να έχει πολλές διεπαφές** (πχ IEEE802.11, Bluetooth, Ethernet)



223.1.1.1 = $\underbrace{11011111}_{223} \underbrace{00000001}_1 \underbrace{00000001}_1 \underbrace{00000001}_1$

IP διευθύνσεις σχετίζονται με κάθε διεπαφή

Υποδίκτυα (subnets)



IP διεύθυνση:

- **τμήμα υποδικτύου:** bits υψηλότερης σειράς
- **τμήμα host:** bits χαμηλότερης σειράς

❖ Τι είναι υποδίκτυο ?

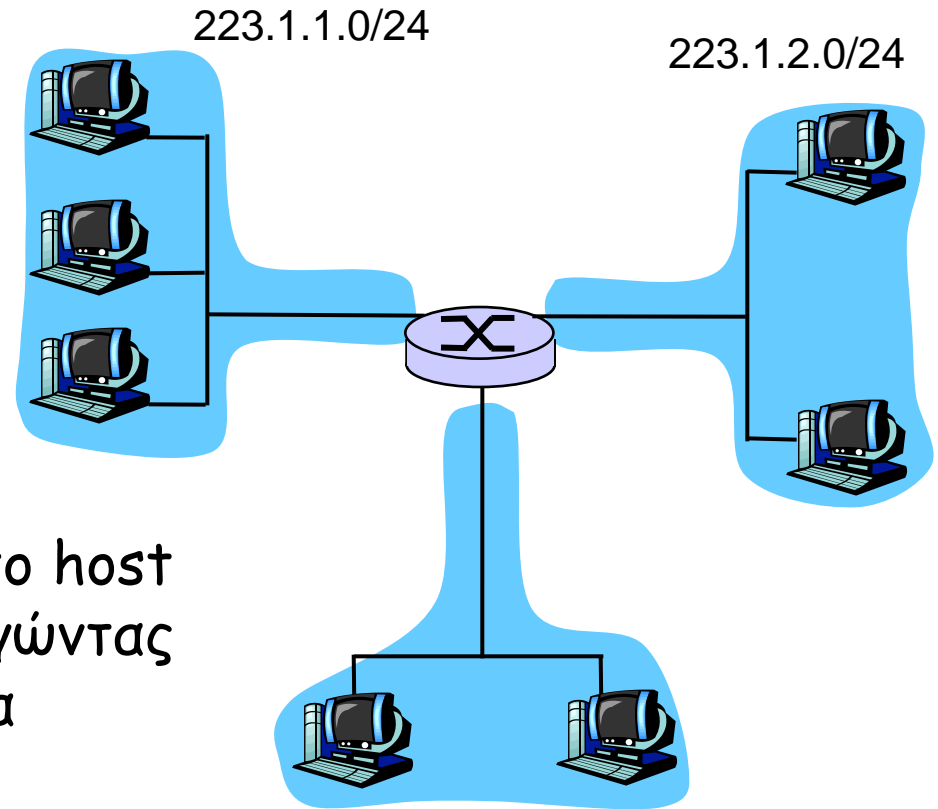
δίκτυο που αποτελείται από 3 υποδίκτυα

- διεπαφές συσκευών με ίδιο τμήμα υποδικτύου της IP διεύθυνσης
- **μπορούν να φτάσουν φυσικά η μία την άλλη χωρίς να παρεμβάλλεται κάποιος δρομολογητής**

Υποδίκτυα

Συνταγή

☞ Για να καθορίσεις τα υποδίκτυα, αποσύνδεσε κάθε διεπαφή από το host του ή τον δρομολογητή, δημιουργώντας "νησιά" από απομονωμένα δίκτυα



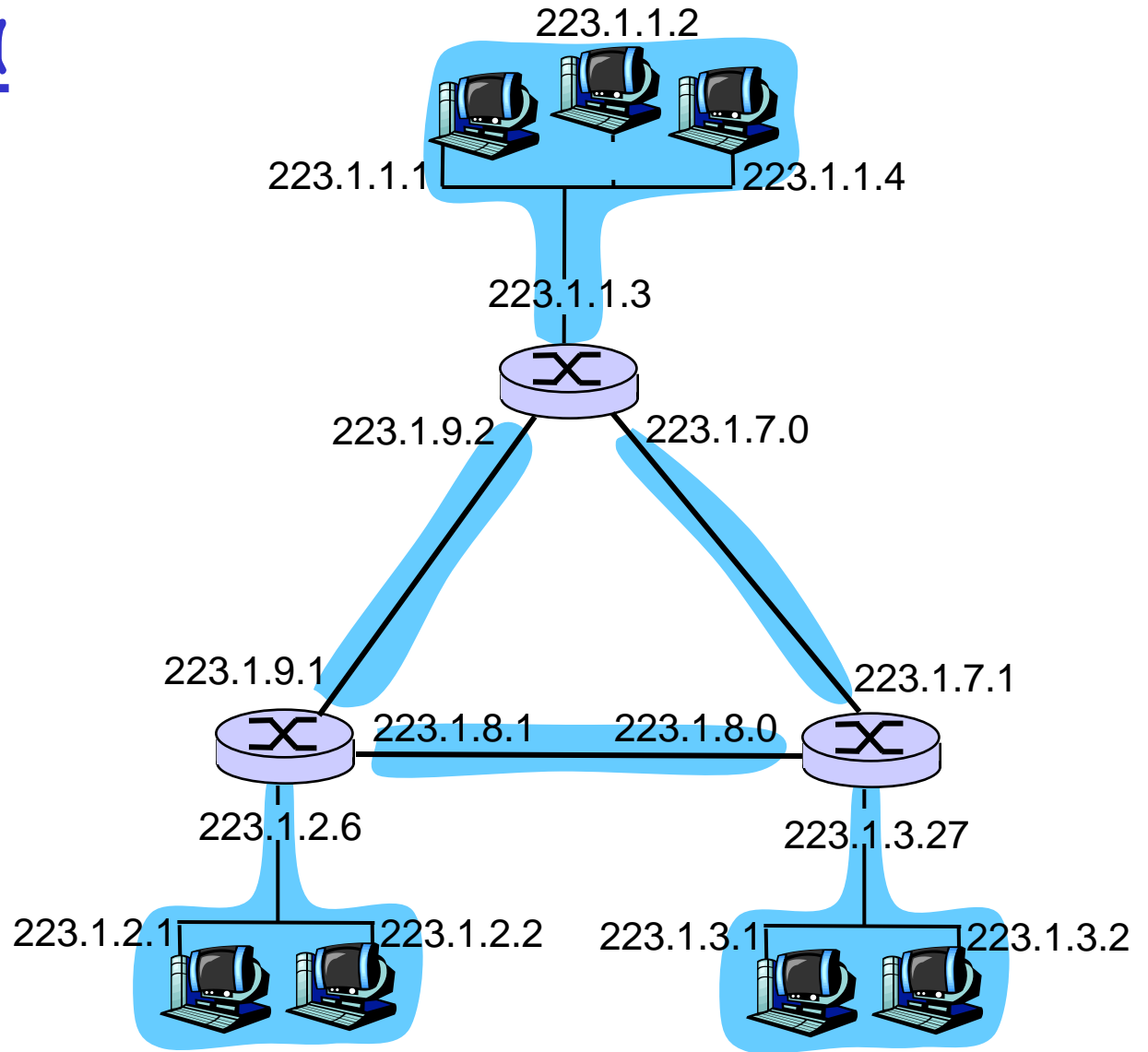
223.1.3.0/24

Subnet mask: /24

Κάθε απομονωμένο δίκτυο ονομάζεται
υποδίκτυο

Υποδίκτυα

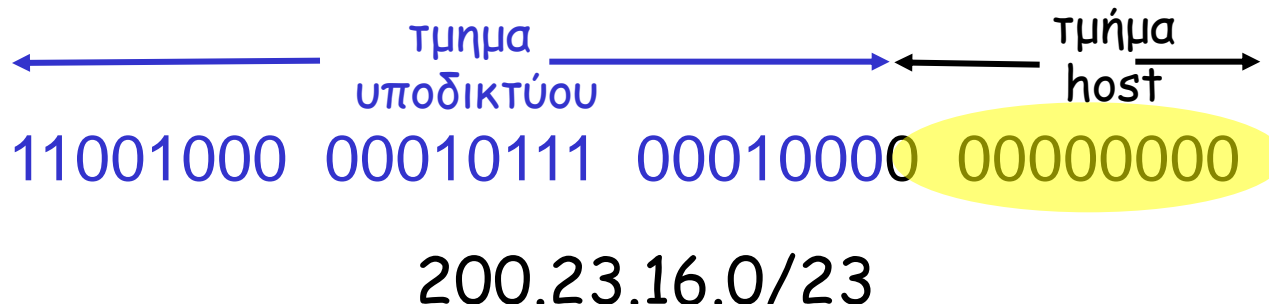
Πόσα είναι?



IP διευθύνσεις: CIDR

CIDR: Classless InterDomain Routing

- Τμήμα υποδικτύου διεύθυνσης αυθαίρετου μεγέθους
- διάταξη διεύθυνσης: $a.b.c.d/x$, όπου x είναι ο αριθμός των bits στο **τμήμα υποδικτύου της διεύθυνσης**



IP διευθύνσεις: πώς να πάρετε μία?

❖ Πως το *host* παίρνει IP διεύθυνση?

- hard-coded (δεν αλλάζει) από τον admin συστήματος σε αρχείο
 - Wintel: control-panel->network->configuration->tcp/ip->properties
 - UNIX: /etc/rc.config
- **D**ynamic **H**ost **C**onfiguration **P**rotocol (**DHCP**): δυναμικά παίρνει διεύθυνση από τον εξυπηρετητή
 - 👍 "plug-and-play" (το βάζεις και παίζει)

IP διευθύνσεις: πώς να πάρετε μία?

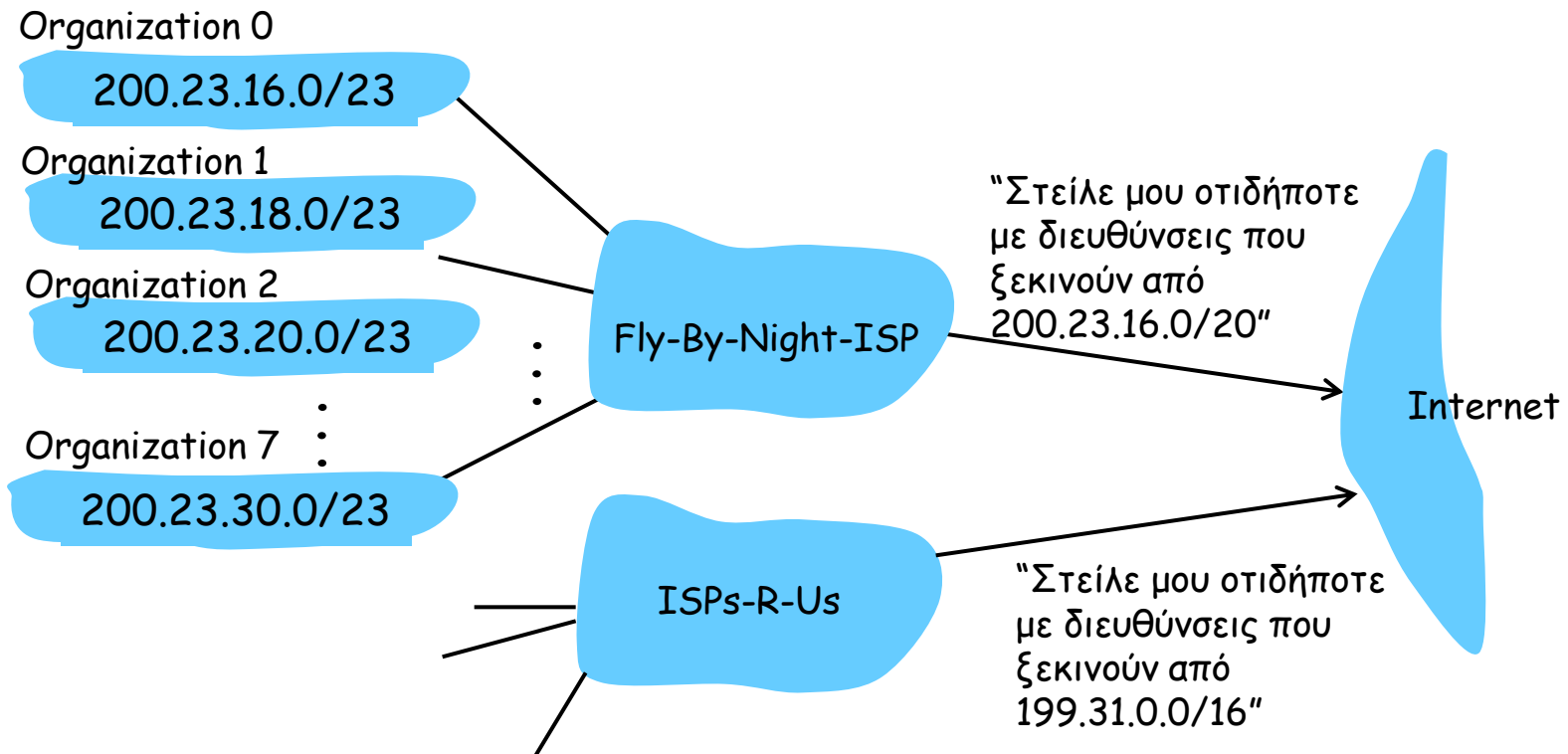
Q: πώς το δίκτυο παίρνει το τμήμα υποδικτύου μιας IP διεύθυνσης?

A: παίρνει το δεσμευμένο τμήμα από τον χώρο διευθύνσεων του **παροχέα του (ISP)**

ISP's block	<u>11001000</u>	<u>00010111</u>	<u>00010000</u>	00000000	200.23.16.0/20
Organization 0	<u>11001000</u>	<u>00010111</u>	<u>00010000</u>	00000000	200.23.16.0/23
Organization 1	<u>11001000</u>	<u>00010111</u>	<u>00010010</u>	00000000	200.23.18.0/23
Organization 2	<u>11001000</u>	<u>00010111</u>	<u>00010100</u>	00000000	200.23.20.0/23
...
Organization 7	<u>11001000</u>	<u>00010111</u>	<u>00011110</u>	00000000	200.23.30.0/23

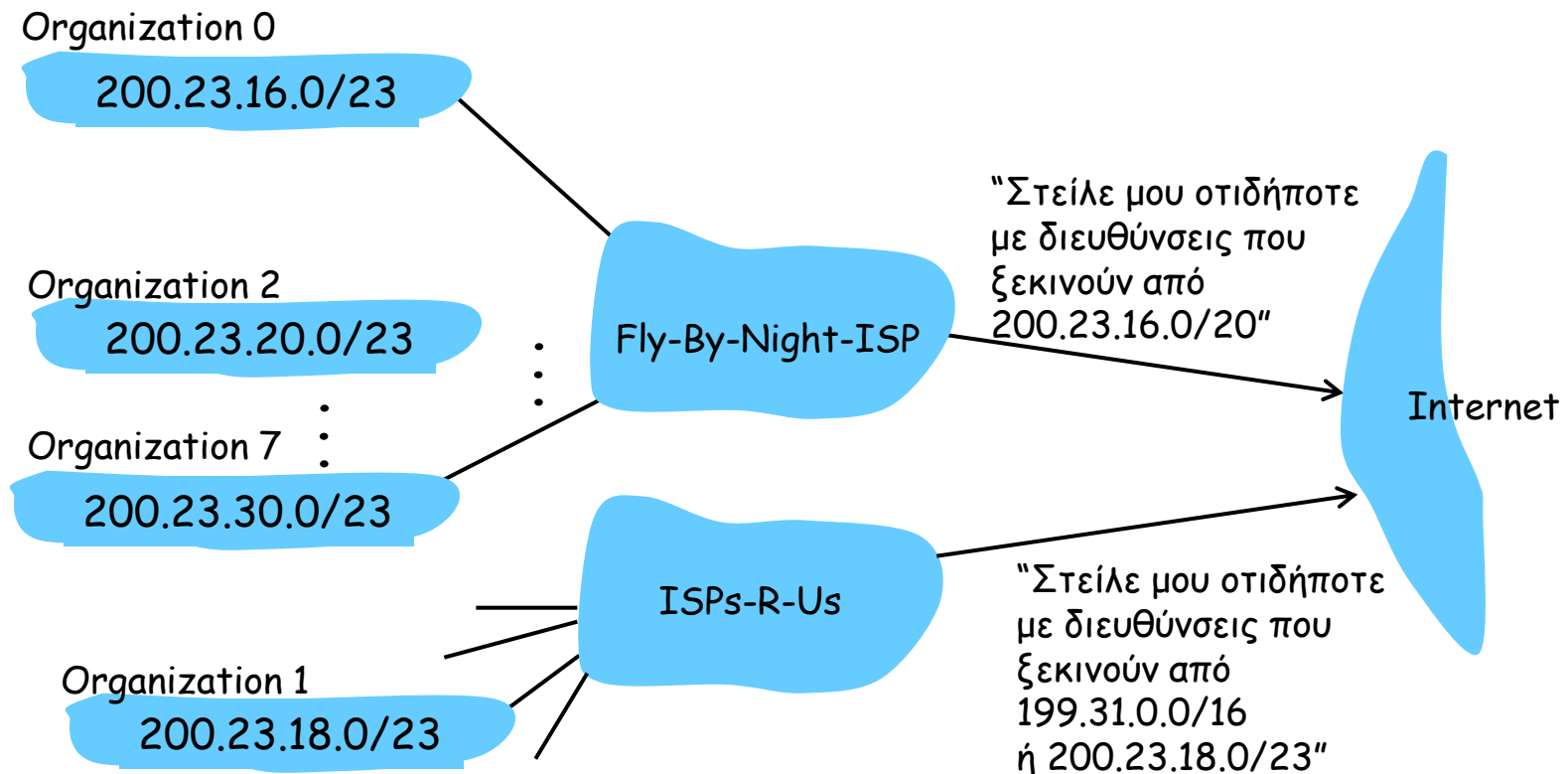
Ιεραρχική διευθυνσιοδότηση: συνάθροιση διαδρομής (route aggregation)

Η ιεραρχική διευθυνσιοδότηση επιτρέπει την αποτελεσματική διάδοση πληροφοριών δρομολόγησης:



Ιεραρχική διευθυνσιοδότηση: περισσότερες συγκεκριμένες διαδρομές

ISPs-R-Us έχει μία πιο συγκεκριμένη διαδρομή προς Organization 1



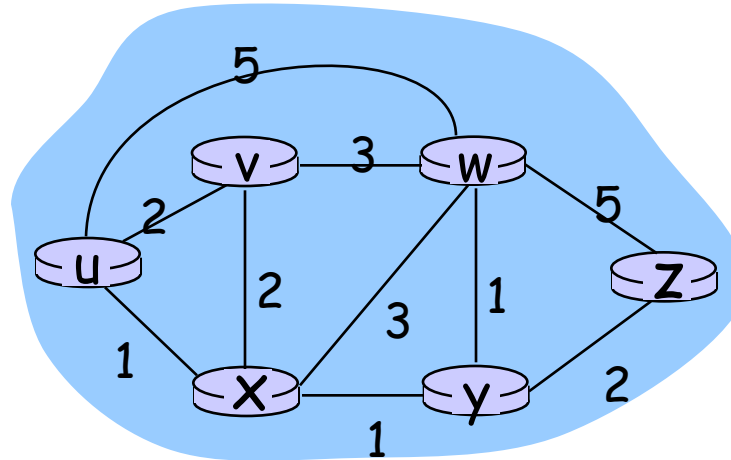
IP διευθυνσιοδότηση: τελευταία λόγια...

Q: Πώς ένας παροχέας παίρνει ένα block από διευθύνσεις?

A: **ICANN**: Internet Corporation for Assigned Names & Numbers

- δεσμεύει διευθύνσεις
- ελέγχει το DNS
- αναθέτει ονόματα σε περιοχές (domains), επιλύει συγκρούσεις

Αφηρημένη έννοια του γράφου



Graph: $G = (N, E)$

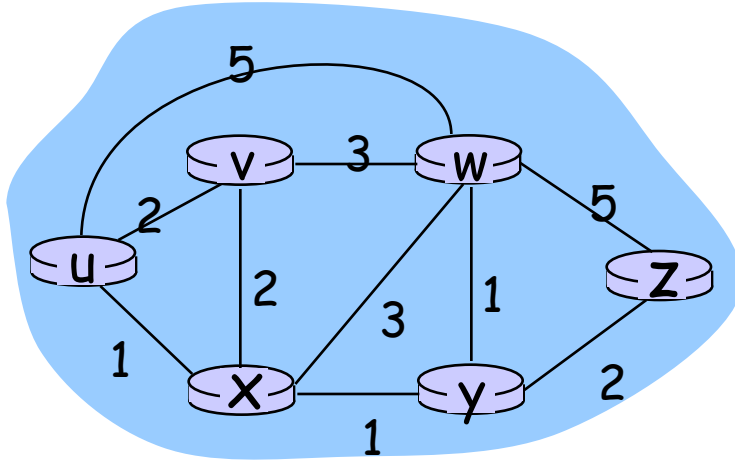
$N =$ σύνολο δρομολογητών = $\{u, v, w, x, y, z\}$

$E =$ σύνολο ζεύξεων = $\{(u, v), (u, x), (v, x), (v, w), (x, w), (x, y), (w, y), (w, z), (y, z)\}$

Παρατήρηση: Η αφηρημένη έννοια του γράφου είναι χρήσιμη σε άλλα ευρύτερα πλαίσια δικτύων

Παράδειγμα: P2P, όπου N είναι το σύνολο των peers και E σύνολο των TCP συνδέσεων

Αφηρημένη έννοια του γράφου: κόστος



- $c(x,x')$ = κόστος ζεύξης (x,x')
- πχ, $c(w,z) = 5$
- το κόστος θα μπορούσε να είναι πάντα 1, η αντιστρόφως ανάλογο του bandwidth, ή αντιστρόφως ανάλογο της συμφόρησης

Κόστος μονοπατιού $(x_1, x_2, x_3, \dots, x_p) = c(x_1, x_2) + c(x_2, x_3) + \dots + c(x_{p-1}, x_p)$

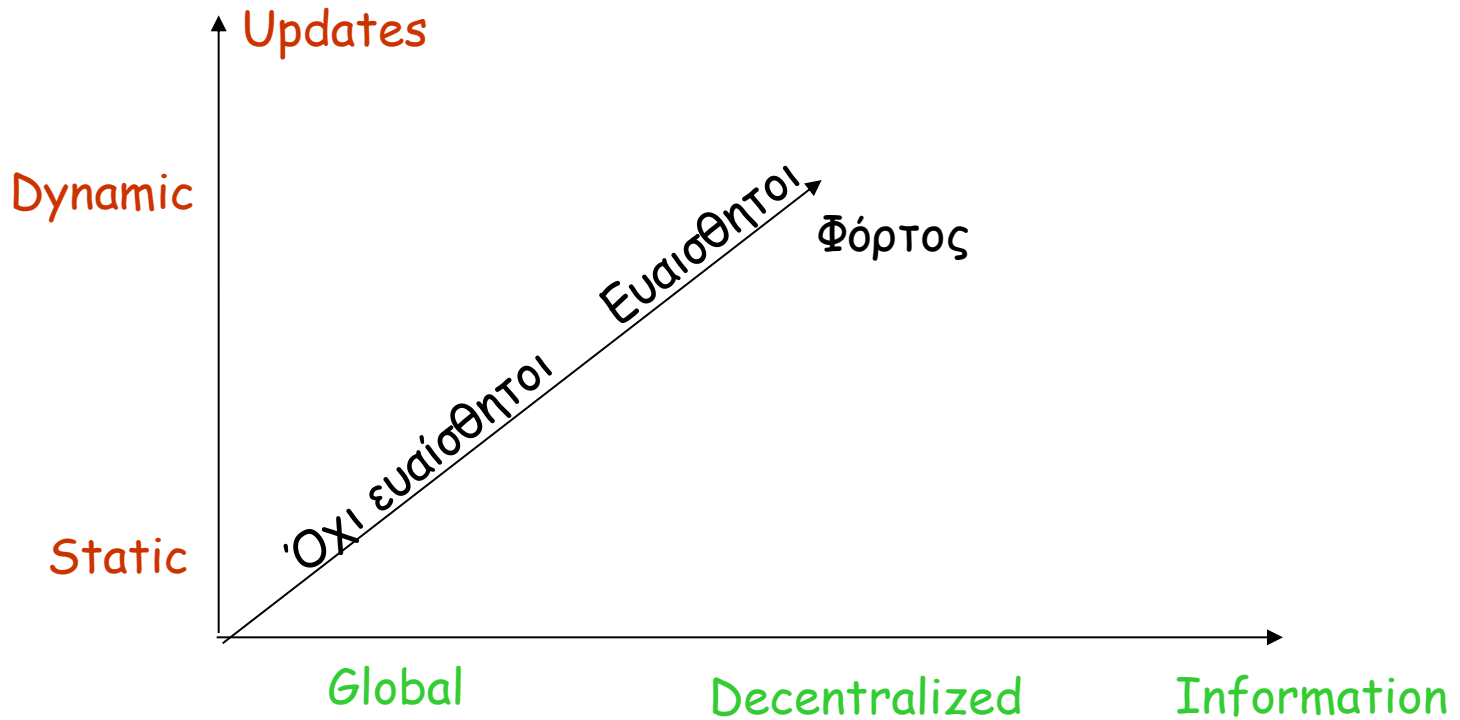
Ερώτηση: Ποιο είναι το μονοπάτι ελάχιστου κόστους μεταξύ u και z ?

Αλγόριθμος δρομολόγησης: αλγόριθμος που βρίσκει το μονοπάτι ελάχιστου κόστους

Αλγόριθμοι δρομολόγησης

- **Dynamic αλγόριθμοι**: άλλαξε τα μονοπάτια δρομολόγησης καθώς το φορτίο κίνησης του δικτύου ή η τοπολογία αλλάζει
- Μπορεί να τρέξει είτε
 - ✓ **περιοδικά** είτε
 - ✓ σε **άμεση απάντηση** στις αλλαγές τοπολογίας & κόστος ζεύξης
- ☹ Ευάλωτοι σε προβλήματα όπως βρόχοι δρομολόγησης & μεταπτώσεις στις διαδρομές
- **Ευαίσθητοι στον φόρτο** : το κόστος ζεύξης ποικίλει δυναμικά για να αντικατοπτρίσει το τρέχον επίπεδο συμφόρησης στην υποκείμενη ζεύξη
- ☞ Οι σημερινοί αλγόριθμοι δρομολόγησης του Διαδικτύου (πχ, RIP, OSPF, BGP) δεν είναι ευαίσθητοι στο φόρτο κίνησης

Ταξινόμηση Αλγορίθμων Δρομολόγησης



Ταξινόμηση Αλγορίθμων Δρομολόγησης

Αντιπαράθεση: Δρομολόγηση με Καθολική vs. Τοπική πληροφορίας

Πλήρη (καθολική) (Global) εικόνα του δικτύου:

- Όλοι οι δρομολογητές γνωρίζουν **πλήρως την τοπολογία** και πληροφορίες για το κόστος των ζεύξεων
- “κατάστασης ζεύξης” (link state) αλγόριθμοι

Τοπική εικόνα του δικτύου (Decentralized):

- Ο δρομολογητής ξέρει τους **φυσικά συνδεδεμένους γείτονες**, και το κόστος ζεύξης προς τους γείτονες
- Επαναληπτική διεργασία υπολογισμών, ανταλλαγή πληροφοριών με τους γείτονες
- Κεντριοποιημένος
- “Διανύσματος απόστασης(distance vector)” αλγόριθμοι

Ταξινόμηση Αλγορίθμων Δρομολόγησης

Αντιπαράθεση: Static vs. dynamic?

Στατικοί (Static): οι διαδρομές αλλάζουν αργά με τον χρόνο

Δυναμικοί (Dynamic): οι διαδρομές αλλάζουν πιο γρήγορα

- Περιοδική ενημέρωση
- σε απάντηση στις αλλαγές κόστους ζεύξης

Router Design

- Many trade-offs: power, \$\$\$, throughput, reliability, flexibility
- Move towards distributed architectures
 - Line-cards have forwarding tables
 - Switched fabric between cards
 - Separate Network processor for "slow path" & control
- Important bottlenecks on fast path
 - Longest prefix match
 - Cross-bar scheduling
- Beware: lots of feature creep

Στοχασμός της ημέρας

Αντιγράφων από τη συνέντευξη του κ. Στέφανου Τραχανά, Διευθυντή των Πανεπιστημιακών Εκδόσεων Κρήτης (ΠΕΚ) στο Βήμα Science:

"... Κι από αυτή την πλευρά η επιμονή στην ποιότητα είναι από μόνη της ένα θετικό στοίχημα για το μέλλον. Δείχνει έμπρακτα ότι πιστεύεις στην καλύτερη πλευρά της χώρας σου και των ανθρώπων της."