



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Δίκτυα Επικοινωνιών Ι

Ενότητα 4: Επίπεδο Δικτύου

Διδάσκοντες:
Άννα Τζανακάκη

Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών

Δίκτυα Επικοινωνιών Ι

Τμήμα Πληροφορικής και Τηλεπικοινωνιών



Εθνικό & Καποδιστριακό
Πανεπιστήμιο Αθηνών

Θεματικές Ενότητες (ΘΕ) μαθήματος:

ΘΕ1: Εισαγωγή

(Κεφ. 1 του βιβλίου)

ΘΕ2: Επίπεδο Εφαρμογής

(Κεφ. 2 του βιβλίου)

ΘΕ3: Επίπεδο Μεταφοράς

(Κεφ. 3 του βιβλίου)

ΘΕ4: Επίπεδο Δικτύου

(Κεφ. 4 του βιβλίου)

**ΘΕ5: Επίπεδο Ζεύξης: Ζεύξεις, Δίκτυα
Πρόσβασης, Δίκτυα Τοπικής Περιοχής**

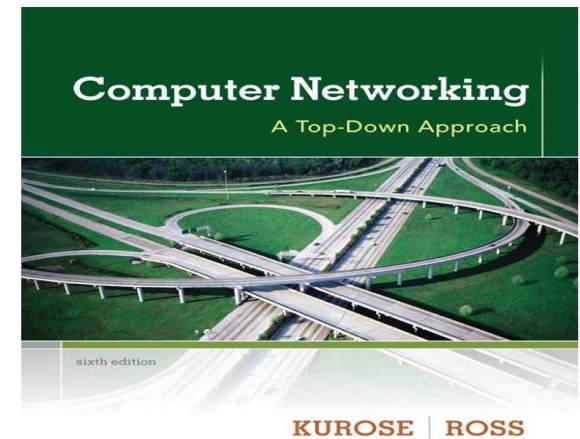
(Κεφ. 5 του βιβλίου)

Συνιστώμενο Βιβλίο:

**Computer Networking: A Top-Down
Approach, by Kurose & Ross,
Addison-Wesley**

Ελληνική Μετάφραση:

Εκδόσεις : Μ. Γκιούρδας



Οι περισσότερες από τις διαφάνειες αυτής της ενότητας αποτελούν προσαρμογή και απόδοση στα ελληνικά των διαφανειών που συνοδεύουν το βιβλίο *Computer Networking: A Top-Down Approach*, J.F. Kurose and K.W. Ross, 6/E, Addison-Wesley.

All material copyright 1996-2012
J.F. Kurose and K.W. Ross, All Rights Reserved

Προσαρμογή και επιμέλεια της απόδοσης των πρωτότυπων διαφανειών στα ελληνικά :
Λάζαρος Μεράκος

Κεφάλαιο 4: Επίπεδο Δικτύου

Στόχοι κεφαλαίου:

- Κατανόηση των βασικών αρχών πίσω από τις υπηρεσίες του επιπέδου δικτύου:
 - Μοντέλα υπηρεσιών του επιπέδου δικτύου
 - Προώθηση vs δρομολόγηση (forwarding vs routing)
 - Πως δουλεύει ένας δρομολογητής (router)
 - Δρομολόγηση (επιλογή διαδρομής)
 - (Ευρυ)εκπομπή, πολυεκπομπή (Broadcast, multicast)
- πραγμάτωση, υλοποίηση στο Διαδίκτυο

Κεφάλαιο 4: Επίπεδο Δικτύου

4.1 Εισαγωγή

4.2 Δίκτυα εικονικού κυκλώματος και δεδομενογράμματος

4.3 Τι βρίσκεται μέσα σ' ένα δρομολογητή

4.4 IP: Πρωτόκολλο Διαδικτύου (Internet Protocol)

- Μορφή δεδομενογράμματος
- Διευθυνσιοδότηση IPv4
- ICMP
- IPv6

4.5 Αλγόριθμοι δρομολόγησης

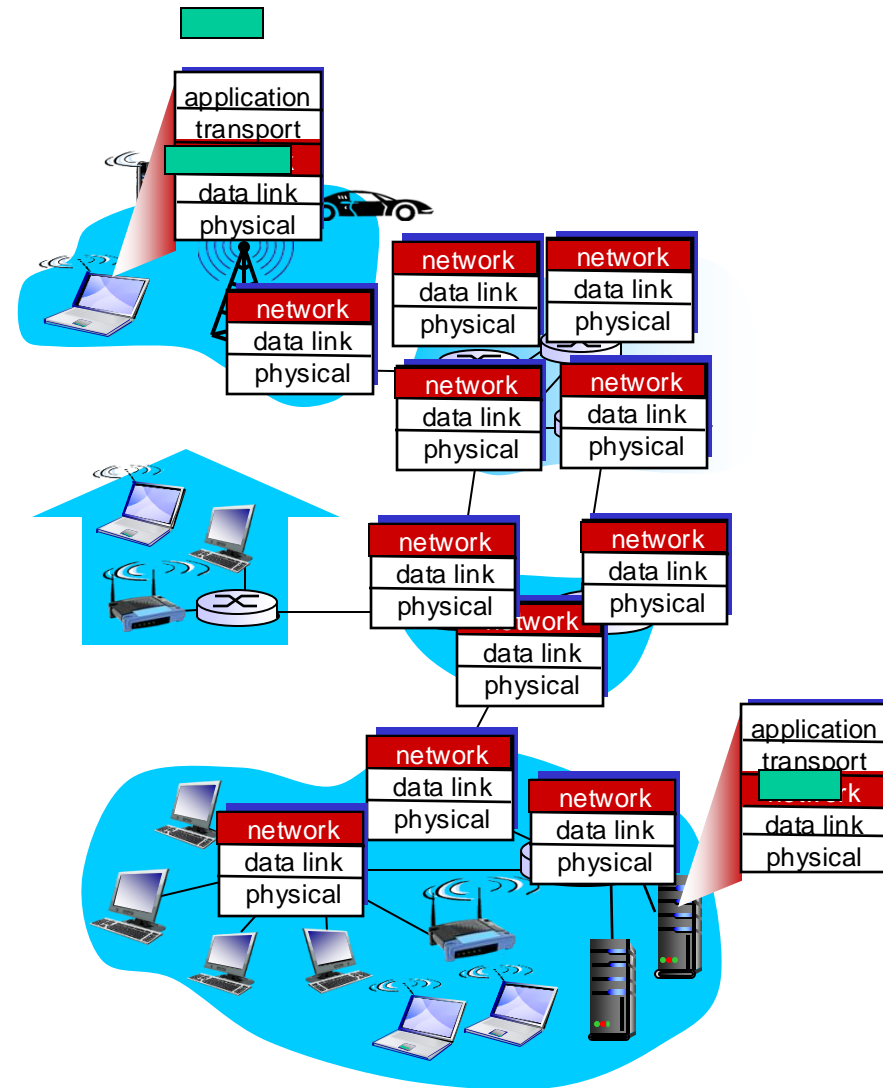
- Κατάστασης ζεύξης (Link State)
- Διανύσματος απόστασης (Distance Vector)
- Ιεραρχική δρομολόγηση

4.6 Δρομολόγηση στο Διαδίκτυο

- RIP
- OSPF
- BGP

Επίπεδο Δικτύου

- ❑ Μεταφορά τμήματος από τον υπολογιστή αποστολέα στον υπολογιστή δέκτη
- ❑ Στην πλευρά αποστολής ενθυλακώνει τα τμήματα σε datagrams (δεδομενογράμματα)
- ❑ Στην πλευρά του δέκτη, παραδίδει τα τμήματα στο επίπεδο μεταφοράς
- ❑ Πρωτόκολλα επιπέδου δικτύου σε **κάθε** υπολογιστή, δρομολογητή
- ❑ Ο δρομολογητής εξετάζει πεδία της κεφαλίδας όλων των IP datagrams που περνούν από αυτόν



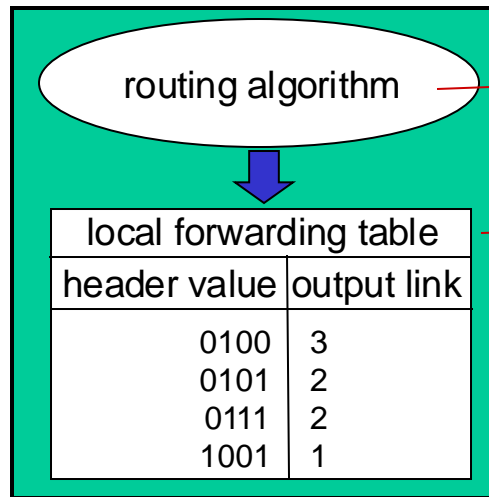
Δύο λειτουργίες κλειδιά του επιπέδου δικτύου

- **Πρώθηση (forwarding):** μετακίνηση πακέτων από την είσοδο του δρομολογητή στην κατάλληλη έξοδο του δρομολογητή
- **Δρομολόγηση (routing):** καθορισμός διαδρομής που ακολουθούν τα πακέτα από την προέλευση στον προορισμό
 - Αλγόριθμοι δρομολόγησης

αναλογία:

- **δρομολόγηση:** διαδικασία σχεδιασμού ταξιδιού από την προέλευση στον προορισμό
- **πρώθηση:** διαδικασία περάσματος από μία διασταύρωση

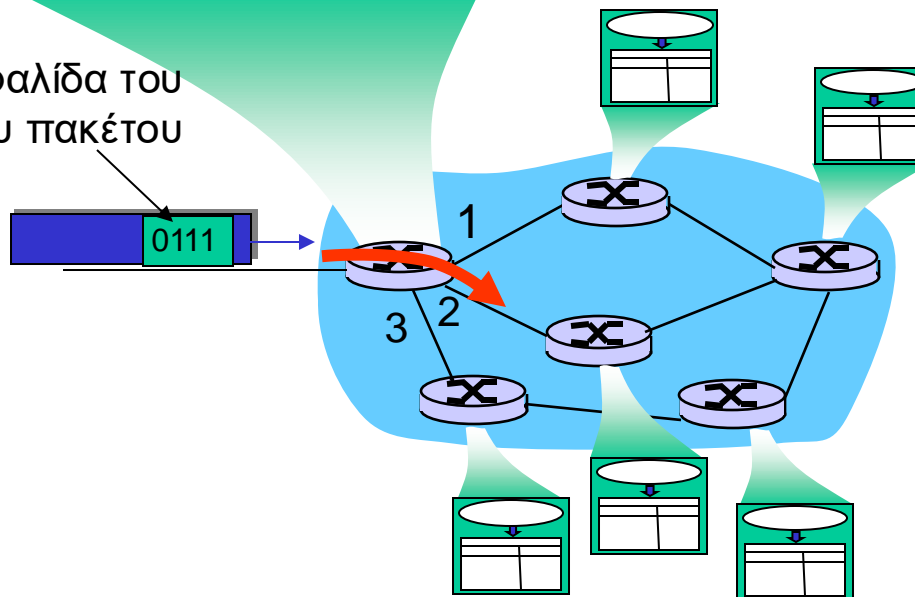
Αλληλεπίδραση μεταξύ δρομολόγησης και προώθησης



Ο αλγόριθμος δρομολόγησης καθορίζει το μονοπάτι από άκρο σε άκρο διαμέσου του δικτύου

Ο πίνακας προώθησης καθορίζει την προώθηση τοπικά σε αυτόν τον δρομολογητή

τιμή στην κεφαλίδα του αφικνούμενου πακέτου

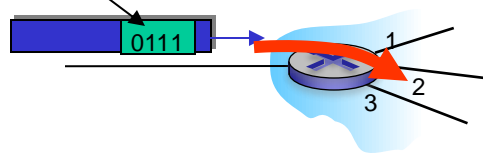


Network layer: data plane, control plane

Data plane (Επίπεδο Δεδομένων)

- *Local*, per-router function (τοπική λειτουργία ανά δρομολογητή)
- Καθορίζει πως ένα datagram που φτάνει στο input port του δρομολογητή προωθείτε στο output port του δρομολογητή

values in arriving packet header



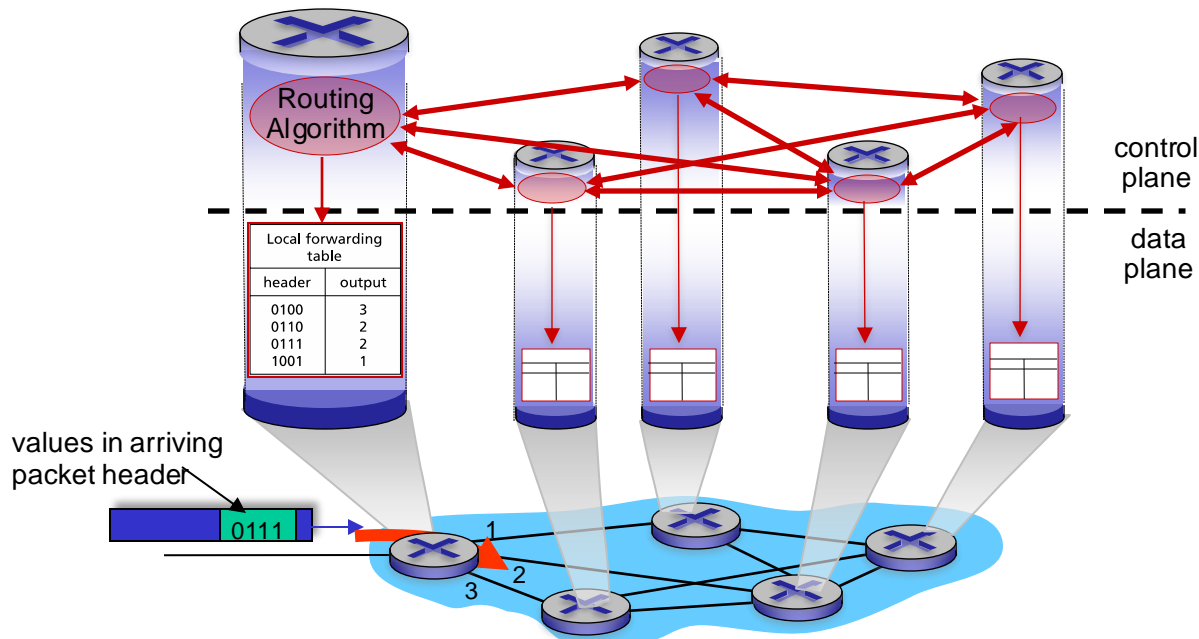
Control plane (Επίπεδο Ελέγχου)

- *network-wide* logic (λογική σε όλη την έκταση του δικτύου)
 - Καθορίζει πως ένα datagram δρομολογείται ανάμεσα σε δρομολογητές κατά μήκος ενός μονοπατιού από άκρη-σε-άκρη (end-end path) από την πηγή προς τον προορισμό
-
- Δύο προσεγγίσεις επιπέδου ελέγχου :
 - *Παραδοσιακοί αλγόριθμοι δρομολόγησης (traditional routing algorithms)*: υλοποιημένοι σε δρομολογητές (implemented in routers)
 - *software-defined networking (SDN)*: υλοποιημένοι σε servers που βρίσκονται σε απόσταση (remote servers)

Υλοποίηση: Per-router control plane

Μεμονωμένα στοιχεία αλγορίθμου δρομολόγησης σε *κάθε δρομολογητή* αλληλεπιδρούν στο επίπεδο ελέγχου

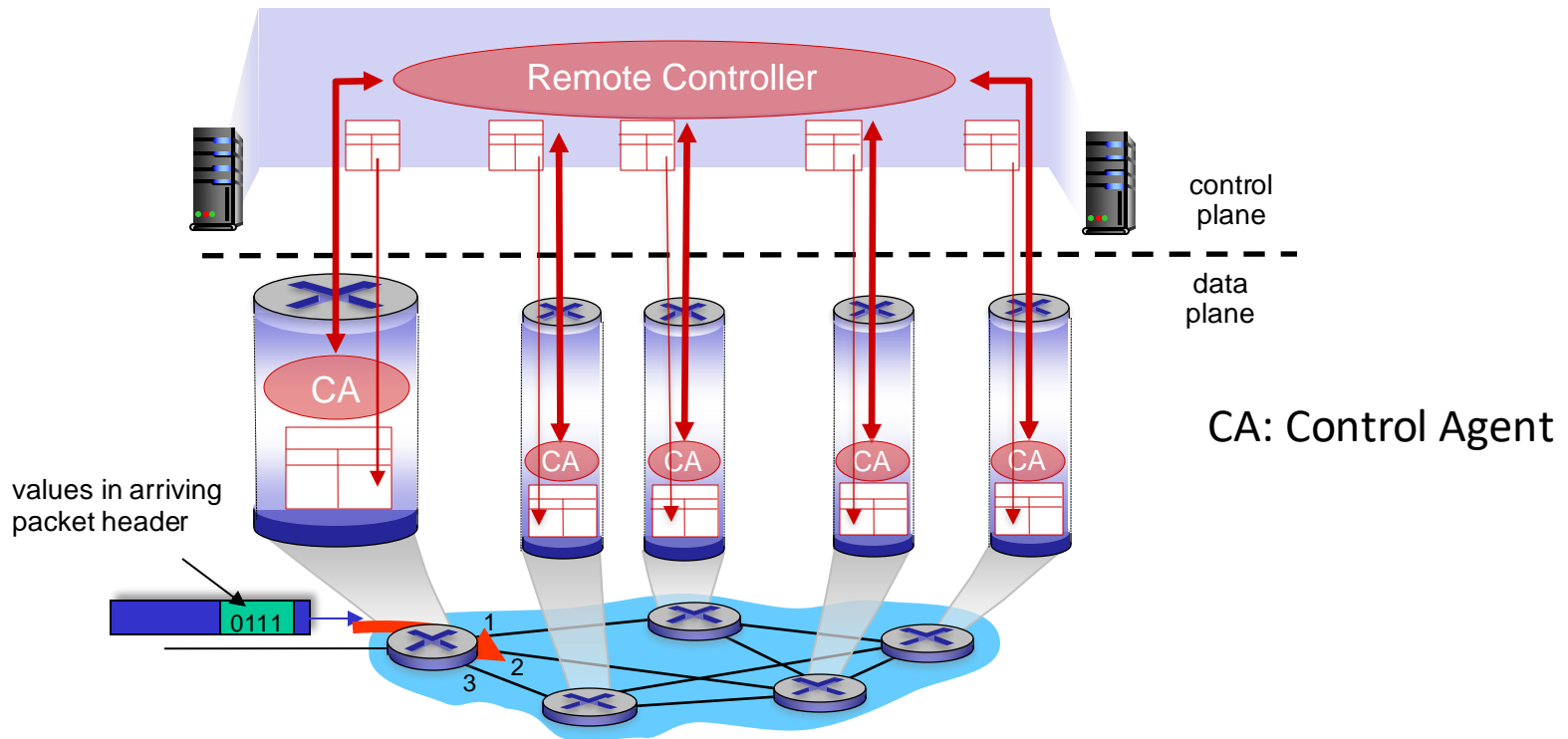
(Individual routing algorithm components *in each and every router* interact in the control plane)



Υλοποίηση: Software-Defined Networking (SDN) control plane

The remote controller computes, installs forwarding tables in routers

A distinct (typically remote) controller interacts with local control agents (CAs)



Software-Defined Networking: Οριζόμενη μέσω λογισμικού δικτύωση
Συχνά open source υλοποιήσεις

Δημιουργία σύνδεσης

- ❑ Η 3^η σημαντική λειτουργία σε ορισμένες δικτυακές αρχιτεκτονικές:
 - ❑ ATM, frame relay, X.25, MPLS
- ❑ Πριν τη ροή των datagrams, οι δύο τερματικοί υπολογιστές και οι δρομολογητές που μεσολαβούν εγκαθιστούν εικονική σύνδεση
 - ❑ συμμετέχουν οι δρομολογητές
- ❑ Υπηρεσία σύνδεσης επιπέδου δικτύου έναντι μεταφοράς:
 - ❑ **δικτύου**: μεταξύ δύο υπολογιστών (μπορεί επίσης να συμμετέχουν δρομολογητές που μεσολαβούν στην περίπτωση των εικονικών κυκλωμάτων (VCs))
 - ❑ **μεταφοράς**: μεταξύ δύο διεργασιών

Μοντέλο υπηρεσιών δικτύου

Ε: Ποιό *μοντέλο υπηρεσίας* για το «κανάλι» που μεταφέρει datagrams από τον αποστολέα στο δέκτη?

Παράδειγμα υπηρεσιών για ξεχωριστά datagrams:

- Εγγυημένη παράδοση
- Εγγυημένη παράδοση με καθυστέρηση μικρότερη από 40 msec

Παράδειγμα υπηρεσιών για ροή datagrams:

- Σε σειρά παράδοση των datagrams
- Εγγυημένο ελάχιστο εύρος ζώνης στη ροή
- Περιορισμοί στις αλλαγές των χρονικών αποστάσεων των πακέτων

Μοντέλα υπηρεσιών επιπέδου δικτύου

Αρχιτεκτονική δικτύου	Μοντέλο υπηρεσίας	Εγγυήσεις ;				Ανάδραση συμφόρησης
		Εύρος ζώνης	Απώλειες	Σειρά	Χρονισμός	
Διαδίκτυο	Βέλτιστης προσπάθειας	καμία	όχι	όχι	όχι	όχι (συνάγεται από απώλειες)
ATM	CBR	σταθερός ρυθμός	ναι	ναι	ναι	χωρίς συμφόρηση
ATM	VBR	εγγυημένος ρυθμός	ναι	ναι	ναι	χωρίς συμφόρηση
ATM	ABR	εγγυημένος ελάχιστος	όχι	ναι	όχι	ναι
ATM	UBR	καμία	όχι	ναι	όχι	όχι

C: constant, V: variable, A: available, U: unspecified

Κεφάλαιο 4: Επίπεδο Δικτύου

4.1 Εισαγωγή

4.2 Δίκτυα εικονικού κυκλώματος και δεδομενογράμματος

4.3 Τι βρίσκεται μέσα σ' ένα δρομολογητή

4.4 IP: Πρωτόκολλο Διαδικτύου (Internet Protocol)

- ❑ Μορφή δεδομενογράμματος
- ❑ Διευθυνσιοδότηση IPv4
- ❑ ICMP
- ❑ IPv6

4.5 Αλγόριθμοι δρομολόγησης

- ❑ Κατάστασης ζεύξης (Link State)
- ❑ Διανύσματος απόστασης (Distance Vector)
- ❑ Ιεραρχική δρομολόγηση

4.6 Δρομολόγηση στο Διαδίκτυο

- ❑ RIP
- ❑ OSPF
- ❑ BGP

4.7 Δρομολόγηση (ευρυ)εκπομπής και πολυεκπομπής

Υπηρεσία επιπέδου δικτύου με & χωρίς σύνδεση

- ❑ Το δίκτυο datagram παρέχει υπηρεσία επιπέδου δικτύου χωρίς σύνδεση
- ❑ Το δίκτυο εικονικού κυκλώματος (Virtual Circuit -VC) παρέχει υπηρεσία επιπέδου δικτύου με σύνδεση
- ❑ Ανάλογο με τις υπηρεσίες επιπέδου μεταφοράς, αλλά
 - ❑ **υπηρεσία:** υπολογιστής-προς-υπολογιστή (host-to-host)
 - ❑ **χωρίς επιλογή:** το δίκτυο παρέχει τη μία ή την άλλη
 - ❑ **υλοποίηση:** στον πυρήνα του δικτύου

Εικονικά κυκλώματα (Virtual circuits)

“Η διαδρομή από την προέλευση στον προορισμό συμπεριφέρεται αρκετά σαν τηλεφωνικό κύκλωμα”

- ως προς την απόδοση
- ενέργειες του δικτύου κατά μήκος της διαδρομής από την προέλευση στον προορισμό

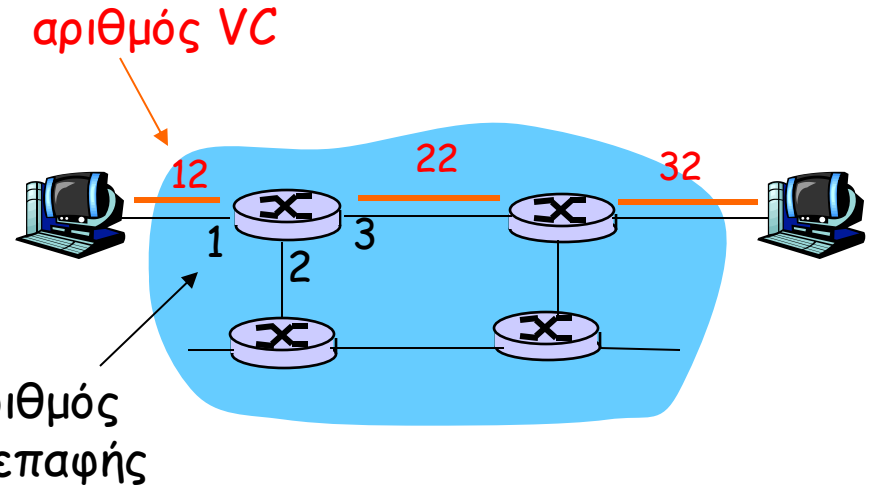
- Εγκαθίδρυση και τερματισμός για κάθε κλήση πριν τα δεδομένα αρχίσουν να ρέουν
- Κάθε πακέτο φέρει αναγνωριστικό του VC (όχι τη διεύθυνση του υπολογιστή προορισμού)
- Κάθε δρομολογητής στη διαδρομή προέλευσης-προορισμού διατηρεί «κατάσταση» για κάθε σύνδεση που περνά
- Οι πόροι ζεύξης, δρομολογητή (εύρος ζώνης, ενταμιευτές) μπορούν να αποδοθούν σε VC (αποκλειστικοί πόροι = προβλέψιμη υπηρεσία)

Υλοποίηση VC

Ένα VC αποτελείται από

1. **διαδρομή** από την προέλευση στον προορισμό
 2. **αριθμούς VC**, ένας αριθμός για κάθε ζεύξη κατά μήκος της διαδρομής
 3. **καταχωρίσεις σε πίνακες προώθησης** στους δρομολογητές κατά μήκος της διαδρομής
- ❑ Το πακέτο που ανήκει στο VC φέρει τον αριθμό του VC (αντί για διεύθυνση προορισμού)
 - ❑ Ο αριθμός VC ενδέχεται να αλλάζει σε κάθε ζεύξη.
 - ❑ Ο νέος αριθμός VC προέρχεται από τον πίνακα προώθησης

Πίνακας προώθησης



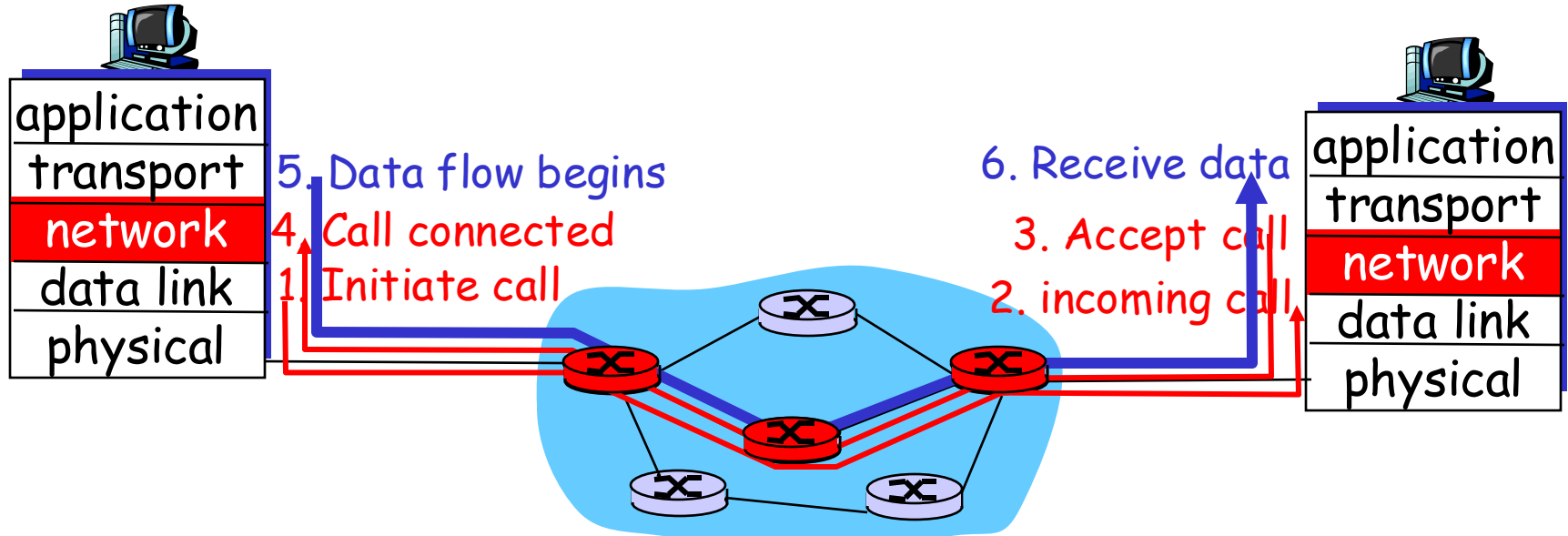
Πίνακας προώθησης στον
πάνω αριστερά δρομολογητή:

Incoming interface	Incoming VC #	Outgoing interface	Outgoing VC #
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
...

Οι δρομολογητές διατηρούν πληροφορίες κατάστασης σύνδεσης!

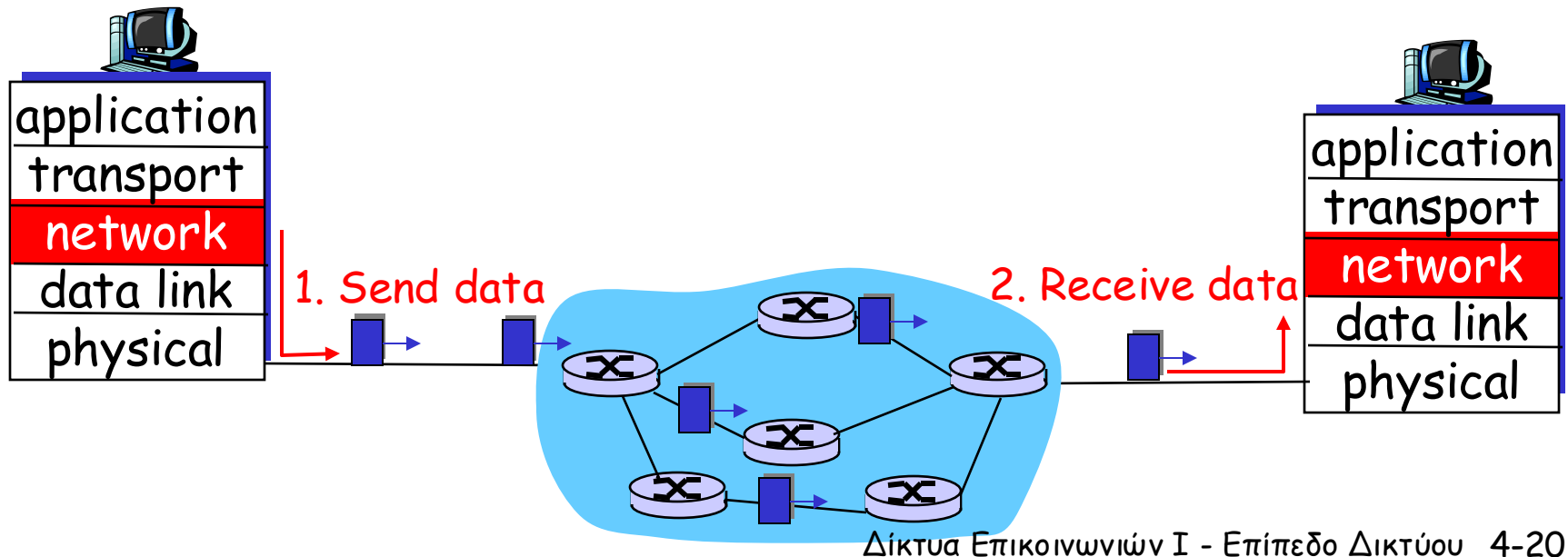
Εικονικά κυκλώματα: πρωτόκολλα σηματοδότησης

- ❑ Χρησιμοποιούνται για την εγκαθίδρυση, συντήρηση και τερματισμό του VC
- ❑ Χρησιμοποιούνται στα ATM, frame-relay, X.25
- ❑ Δε χρησιμοποιούνται στο σημερινό Διαδίκτυο

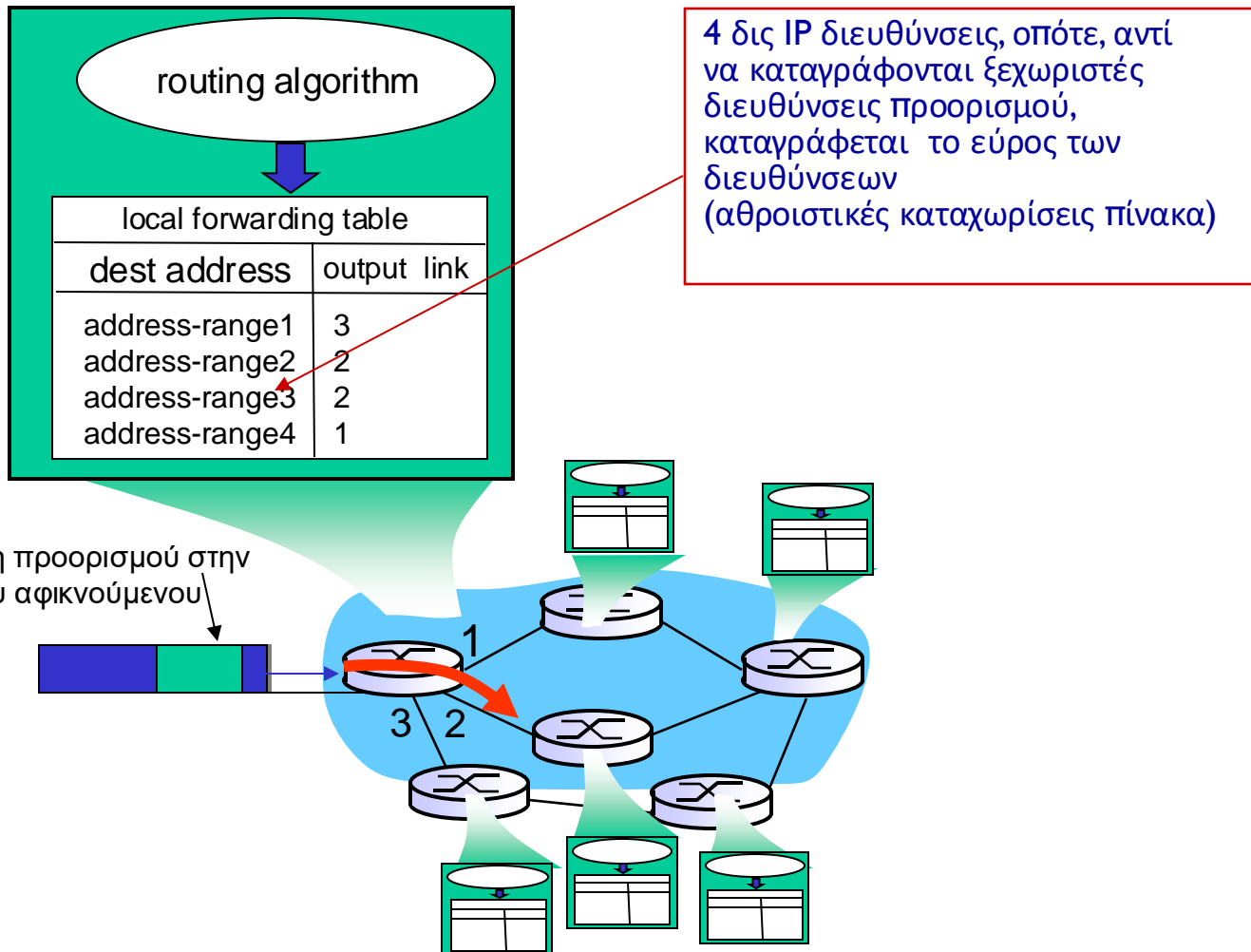


Δίκτυα Δεδομενογράμματος (Datagram)

- ❑ Χωρίς εγκαθίδρυση κλήσης στο επίπεδο δικτύου
- ❑ δρομολογητές: χωρίς κατάσταση για τις από-άκρο-σε-άκρο συνδέσεις
 - ❑ Χωρίς την έννοια της «σύνδεσης» στο επίπεδο δικτύου
- ❑ Τα πακέτα προωθούνται χρησιμοποιώντας τη διεύθυνση του υπολογιστή προορισμού
 - Τα πακέτα μεταξύ του ίδιου ζεύγους προέλευσης προορισμού ενδέχεται να ακολουθήσουν διαφορετικές διαδρομές



Πίνακας Προώθησης Δεδομενογραμμάτων



Πίνακας προώθησης

4 δισεκατομμύρια
πιθανές καταχωρίσεις

Εύρος διευθύνσεων προορισμού

Διεπαφή Ζεύξης

11001000 00010111 00010000 00000000

έως

0

11001000 00010111 00010111 11111111

11001000 00010111 00011000 00000000

έως

1

11001000 00010111 00011000 11111111

11001000 00010111 00011001 00000000

έως

2

11001000 00010111 00011111 11111111

διαφορετικά

3

E: Τι συμβαίνει αν το εύρος δεν κατανέμεται τόσο βολικά;

Ταίριασμα μεγαλύτερου προθέματος (Longest prefix matching)

ταίριασμα μεγαλύτερου προθέματος

κατά την αναζήτηση καταχώρισης στον πίνακα προώθησης για δοσμένη διεύθυνση προορισμού, χρησιμοποιείται το **μεγαλύτερο** πρόθεμα διεύθυνσης που ταιριάζει στη διεύθυνση προορισμού

Εύρος Διεύθυνσης Προορισμού	Διεπαφή Ζεύξης
11001000 00010111 00010*** *****	0
11001000 00010111 00011000 *****	1
11001000 00010111 00011*** *****	2
διαφορετικά	3

Παραδείγματα

DA: 11001000 00010111 00010110 10100001

Ποιά διεπαφή;

DA: 11001000 00010111 00011000 10101010

Ποιά διεπαφή;

Δίκτυο datagram ή VC: Γιατί;

Διαδίκτυο (datagram)

- Ανταλλαγή δεδομένων μεταξύ υπολογιστών
 - "ελαστική" υπηρεσία, χωρίς αυστηρές χρονικές απαιτήσεις
- "Έξυπνα" τερματικά συστήματα (υπολογιστές)
 - Μπορούν να προσαρμόζονται, να επιτελούν έλεγχο, διόρθωση σφαλμάτων
 - **Απλό εντός του δικτύου, πολυπλοκότητα στα "άκρα"**
- Πολλά είδη ζεύξεων
 - Διαφορετικά χαρακτηριστικά
 - Δύσκολα ομοιόμορφη υπηρεσία

ATM (VC)

- Εξελίχθηκε από την τηλεφωνία
- Ανθρώπινες συζητήσεις:
 - Αυστηρές χρονικές απαιτήσεις, απαιτήσεις αξιοπιστίας
 - Ανάγκη για εγγυημένη υπηρεσία
- "Χαζά" τερματικά συστήματα
 - Τηλέφωνα
 - **Πολυπλοκότητα εντός του δικτύου**

Κεφάλαιο 4: Επίπεδο Δικτύου

4.1 Εισαγωγή

4.2 Δίκτυα εικονικού κυκλώματος και δεδομενογράμματος

4.3 Τί βρίσκεται μέσα σ' ένα δρομολογητή

4.4 IP: Πρωτόκολλο Διαδικτύου (Internet Protocol)

- ❑ Μορφή δεδομενογράμματος
- ❑ Διευθυνσιοδότηση IPv4
- ❑ ICMP
- ❑ IPv6

4.5 Αλγόριθμοι δρομολόγησης

- ❑ Κατάστασης ζεύξης (Link State)
- ❑ Διανύσματος απόστασης (Distance Vector)
- ❑ Ιεραρχική δρομολόγηση

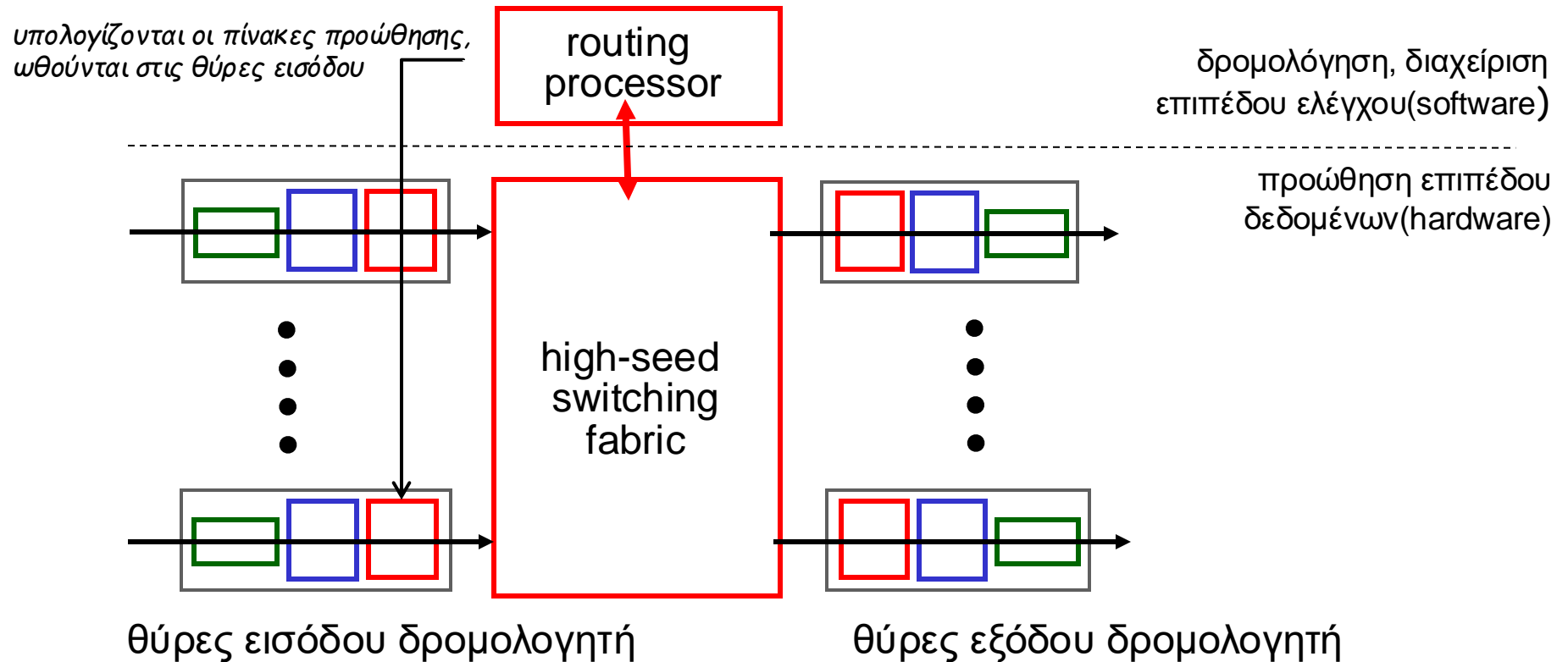
4.6 Δρομολόγηση στο Διαδίκτυο

- ❑ RIP
- ❑ OSPF
- ❑ BGP

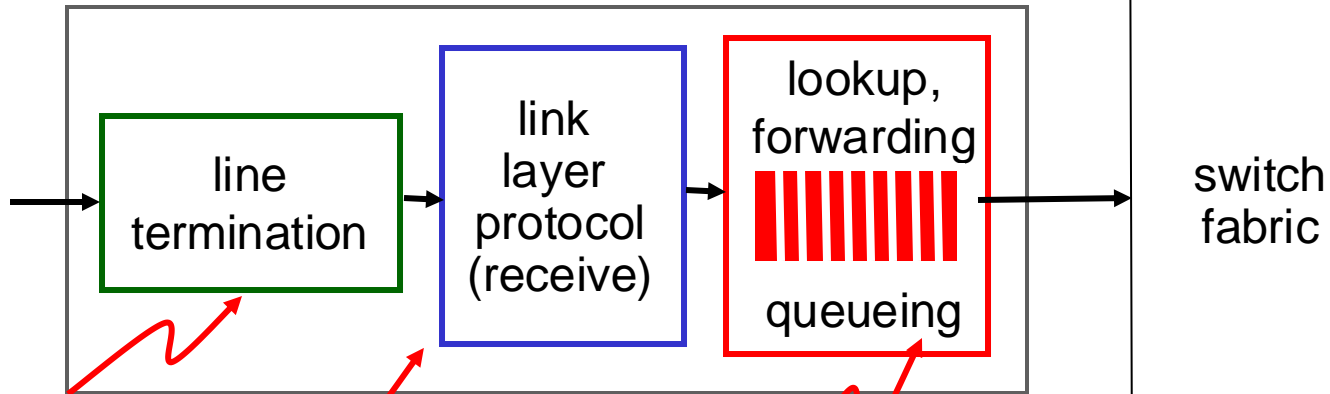
Επισκόπηση αρχιτεκτονικής δρομολογητή

Δύο βασικές λειτουργίες του δρομολογητή:

- Εκτέλεση αλγορίθμων/πρωτοκόλλων δρομολόγησης (RIP, OSPF, BGP)
- Προώθηση datagrams από εισερχόμενη σε εξερχόμενη ζεύξη



Λειτουργίες θύρας εισόδου



Φυσικό επίπεδο:
Λήψη σε επίπεδο bit

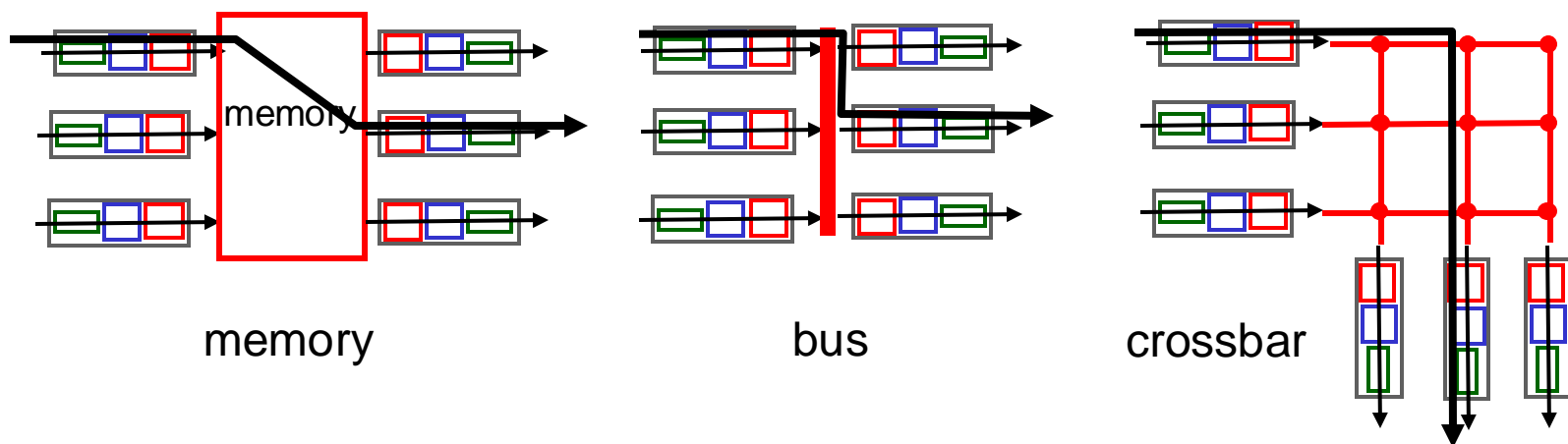
Επίπεδο ζεύξης δεδομένων:
π.χ., Ethernet
βλ. Κεφάλαιο 5

Αποκεντρωμένη μεταγωγή:

- Δεδομένου του προορισμού του datagram, αναζήτηση της θύρας εξόδου με χρήση του πίνακα προώθησης στη μνήμη της θύρας εισόδου ("match plus action")
- Σκοπός: ολοκλήρωση της επεξεργασίας της θύρας εισόδου με «ταχύτητα γραμμής» ('line speed')
- Αναμονή: αν τα datagrams φτάνουν ταχύτερα από το ρυθμό προώθησης στο δόμημα μεταγωγής

Τύποι δομήματος μεταγωγής

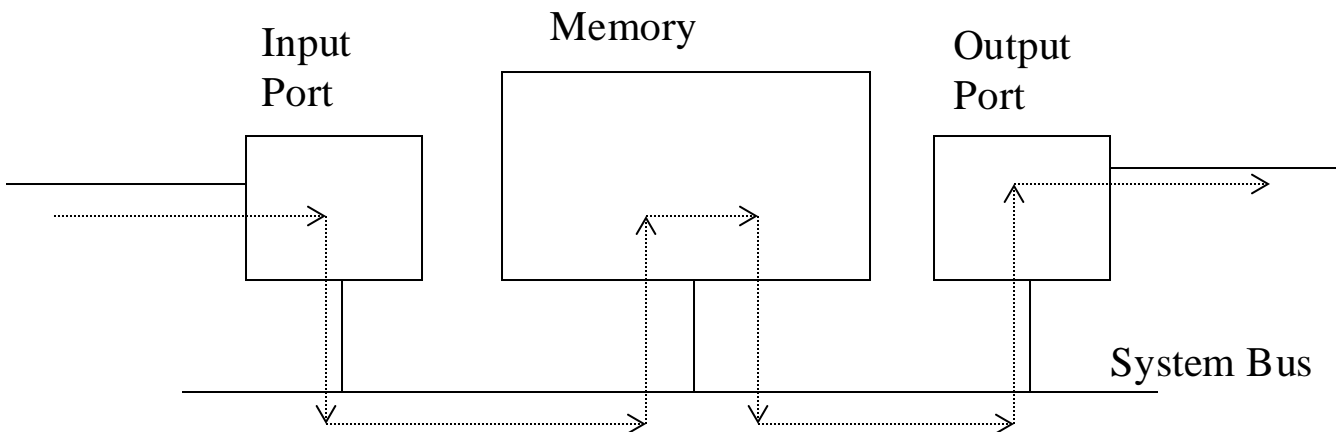
- ❑ Μεταφορά πακέτου από τον ενταμιευτή (buffer) εισόδου στον κατάλληλο ενταμιευτή εξόδου
- ❑ Ρυθμός μεταγωγής: ρυθμός στον οποίο τα πακέτα μπορούν να μεταφερθούν από τις εισόδους στις εξόδους
 - συχνά μετριέται ως πολλαπλάσιο του ρυθμού εισόδου/εξόδου της γραμμής
 - N εισοδοί: ρυθμός μεταγωγής N φορές ο επιθυμητός ρυθμός της γραμμής
- ❑ 3 τύποι δομημάτων μεταγωγής



Μεταγωγή μέσω μνήμης

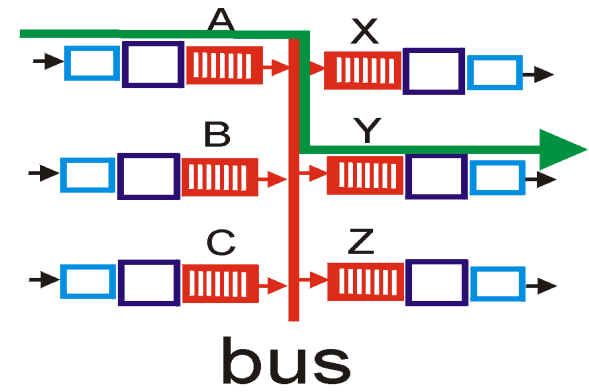
Δρομολογητές πρώτης γενιάς:

- ❑ παραδοσιακοί υπολογιστές με μεταγωγή υπό τον άμεσο έλεγχο της CPU
- ❑ το πακέτο αντιγράφεται στη μνήμη του συστήματος
- ❑ η ταχύτητα περιορίζεται από το εύρος ζώνης της μνήμης (2 διασχίσεις του διαύλου ανά datagram)



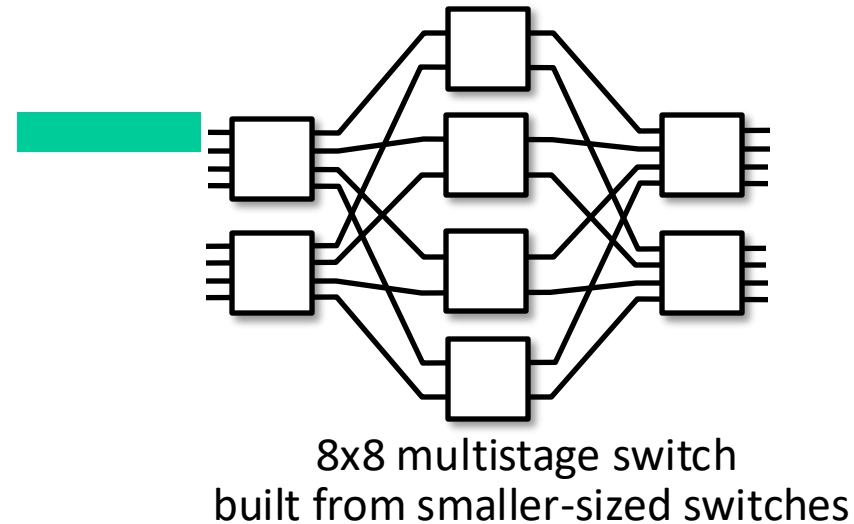
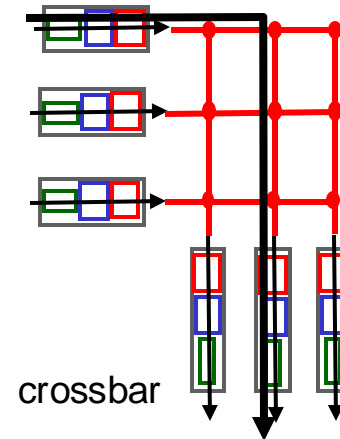
Μεταγωγή μέσω διαύλου

- Το datagram από τη θύρα εισόδου της μνήμης στη θύρα εξόδου της μνήμης μέσω διαμοιραζόμενου διαύλου (bus)
- **Ανταγωνισμός διαύλου:** η ταχύτητα μεταγωγής περιορίζεται από το εύρος ζώνης του διαύλου
- Διαύλος 32 Gbps, Cisco 5600: επαρκής ταχύτητα για δρομολογητές πρόσβασης και εταιρικούς δρομολογητές



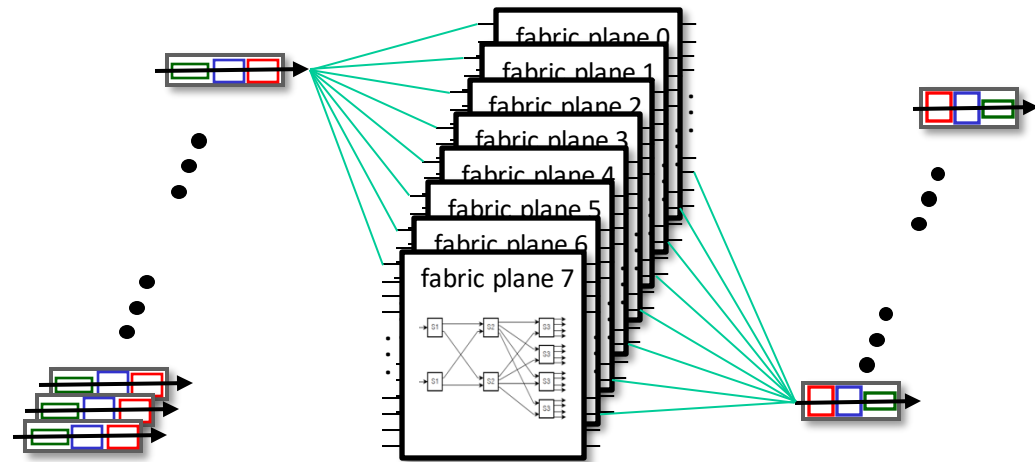
Μεταγωγή μέσω δικτύου διασύνδεσης

- ❑ Ξεπερνά τους περιορισμούς εύρους ζώνης του διαύλου
- ❑ Δίκτυα Banyan, crossbar δίκτυα, άλλα δίκτυα διασύνδεσης που αρχικά αναπτύχθηκαν για τη διασύνδεση επεξεργαστών σε πολυεπεξεργαστικά συστήματα
- ❑ Προηγμένη σχεδίαση: κατάτμηση του datagram σε σταθερού μήκους κελιά (cells), μεταγωγή των κελιών διαμέσω του δομήματος
- ❑ Cisco 12000: μετάγει 60 Gbps μέσω του δικτύου διασύνδεσης

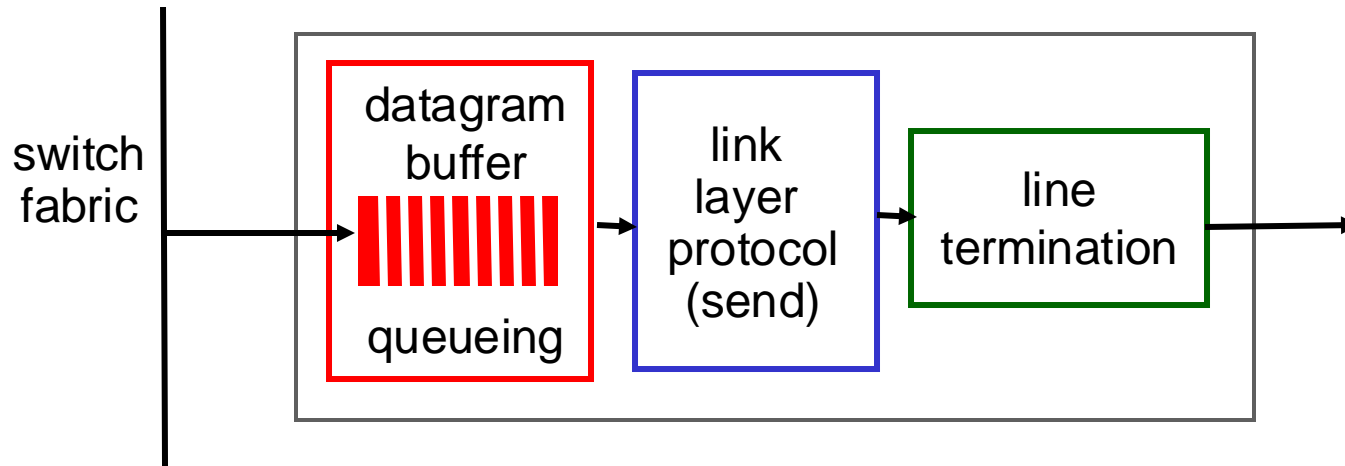


Μεταγωγή μέσω δικτύου διασύνδεσης

- Κλιμάκωση μέσω πολλαπλών παράλληλων:
 - speedup, scaleup via parallelism
- Cisco CRS router:
 - basic unit: 8 switching planes
 - each plane: 3-stage interconnection network
 - up to 100's Tbps switching capacity

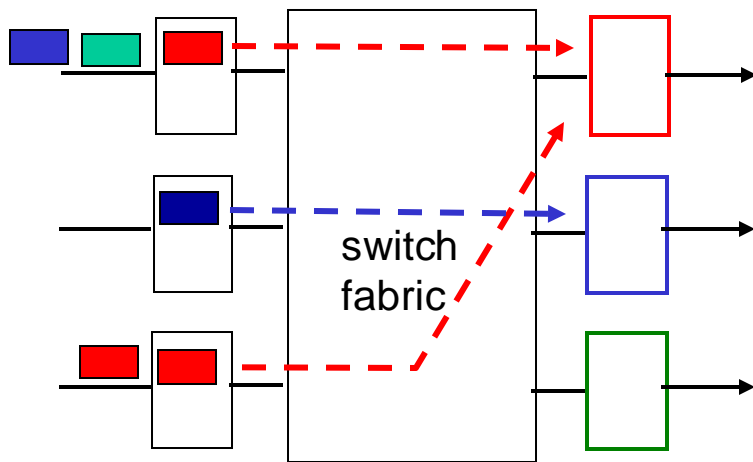


Θύρες εξόδου

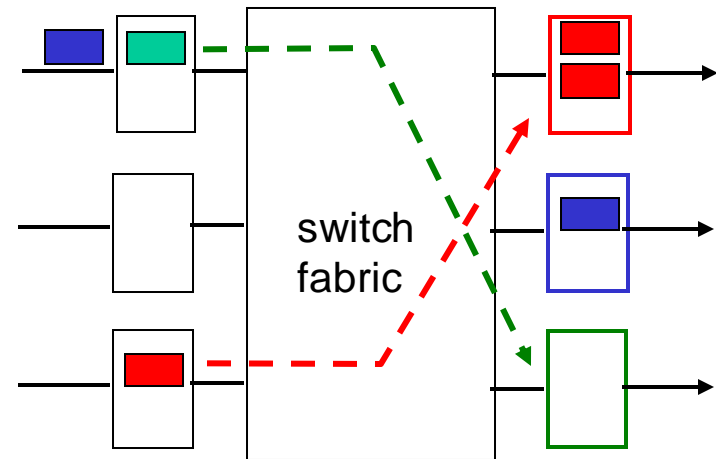


- ❑ Απαιτείται **ενταμίευση (buffering)** όταν τα datagrams φτάνουν από το δόμημα μεταγωγής ταχύτερα από το ρυθμό μετάδοσης
- ❑ Η **πολιτική χρονοπρογραμματισμού (scheduling discipline)** επιλέγει κάποιο από τα ενταμιευμένα datagrams για μετάδοση

Ενταμίευση στη θύρα εξόδου



τη στιγμή t , περισσότερα
πακέτα από την είσοδο στην
έξοδο



ένα πακέτο αργότερα

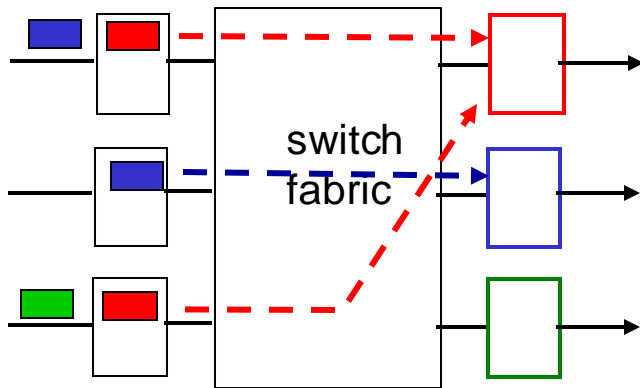
- ❑ Ενταμίευση απαιτείται όταν ο ρυθμός άφιξης μέσω του μεταγωγού υπερβαίνει την ταχύτητα της γραμμής εξόδου
- ❑ *Καθυστέρηση αναμονής και απώλειες λόγω υπερχειλίσις του ενταμιευτή της θύρας εξόδου!*

Πόση ενταμίευση;

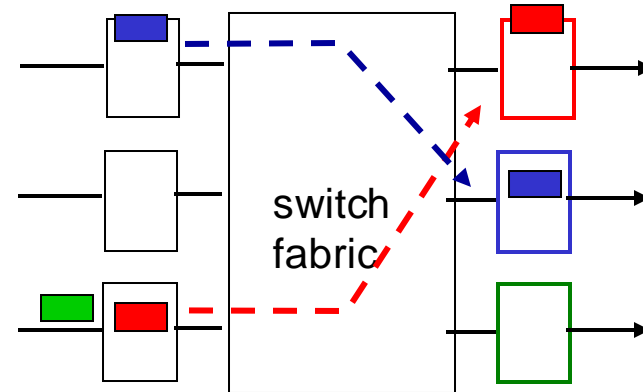
- Πρακτικός κανόνας του RFC 3439: μέση ενταμίευση ίση με το «τυπικό» RTT (π.χ. 250 msec) επί τη χωρητικότητα της ζεύξης C
 - π.χ., $C = 10$ Gps ζεύξη : ενταμιευτής 2.5 Gbit
- Πρόσφατη σύσταση: με N ροές, ενταμίευση ίση με $\frac{RTT \cdot C}{\sqrt{N}}$

Ενταμίευση στη θύρα εισόδου

- Το δόμημα μεταγωγής πιο αργό από το συνδυασμό των θυρών εισόδου -> ενδέχεται να εμφανιστεί αναμονή στις ουρές εισόδου
 - *Καθυστέρηση αναμονής και απώλειες λόγω υπερχειλίσσης του ενταμιευτή της θύρας εισόδου!*
- *Μπλοκάρισμα κεφαλής γραμμής (Head-of-the-Line (HOL) blocking):* ενταμιευμένο datagram στην κορυφή της ουράς εμποδίζει άλλα datagrams από το να προωθηθούν



συναγωνισμός στη θύρα εξόδου:
μόνο ένα κόκκινο datagram μπορεί να μεταφερθεί
το χαμηλότερο κόκκινο πακέτο μπλοκάρεται



ένα πακέτο αργότερα:
το πράσινο πακέτο αντιμετωπίζει
μπλοκάρισμα κεφαλής γραμμής

Κεφάλαιο 4: Επίπεδο Δικτύου

4.1 Εισαγωγή

4.2 Δίκτυα εικονικού κυκλώματος και δεδομενογράμματος

4.3 Τι βρίσκεται μέσα σ' ένα δρομολογητή

4.4 IP: Πρωτόκολλο Διαδικτύου (Internet Protocol)

- ❑ Μορφή δεδομενογράμματος
- ❑ Διευθυνσιοδότηση IPv4
- ❑ ICMP
- ❑ IPv6

4.5 Αλγόριθμοι δρομολόγησης

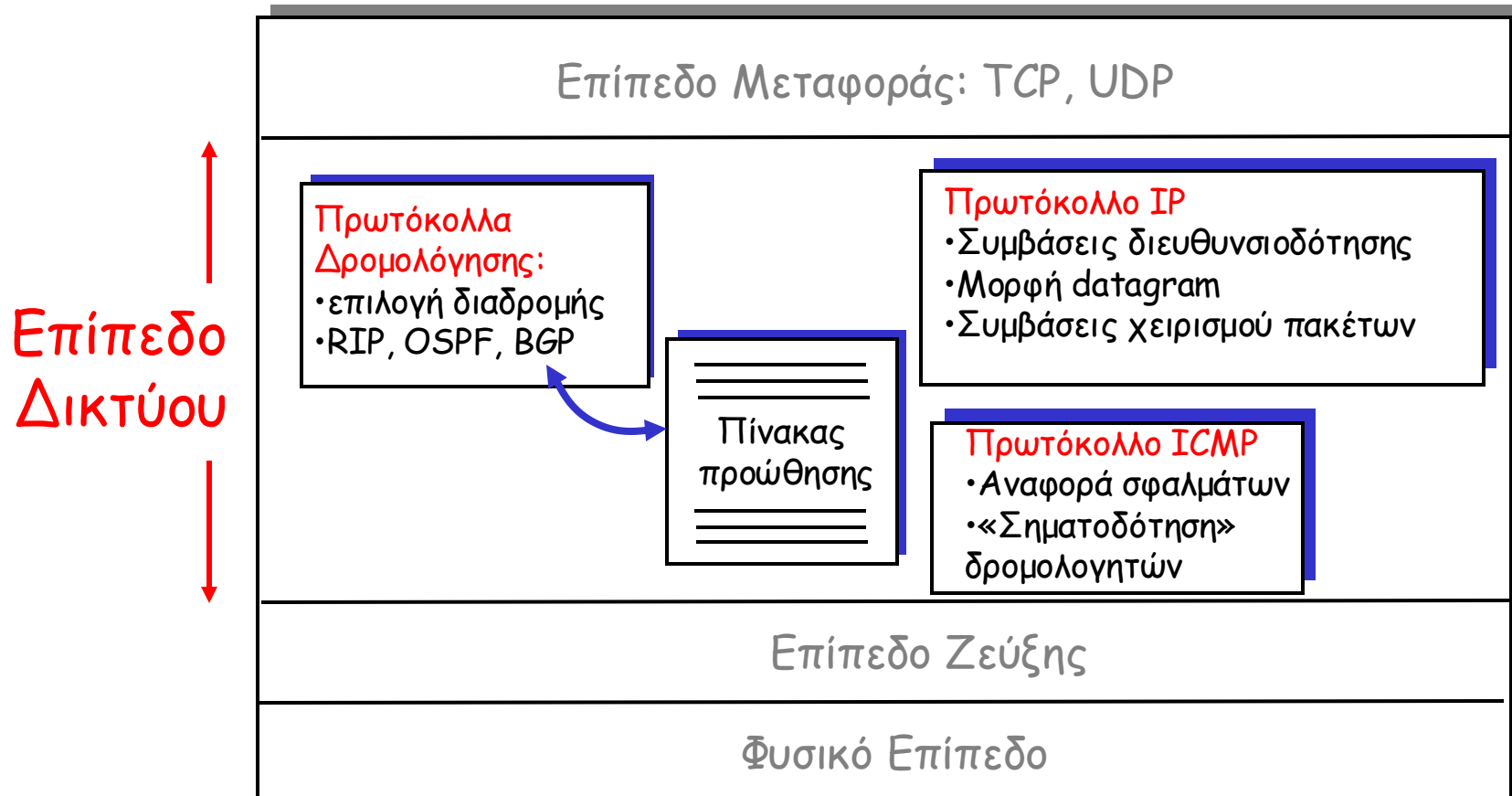
- ❑ Κατάστασης ζεύξης (Link State)
- ❑ Διανύσματος απόστασης (Distance Vector)
- ❑ Ιεραρχική δρομολόγηση

4.6 Δρομολόγηση στο Διαδίκτυο

- ❑ RIP
- ❑ OSPF
- ❑ BGP

Το επίπεδο δικτύου του Διαδικτύου

Λειτουργίες επιπέδου δικτύου υπολογιστή, δρομολογητή:



Δομή IP datagram

αριθμός έκδοσης
IP πρωτοκόλλου
μήκος κεφαλίδας
(bytes)

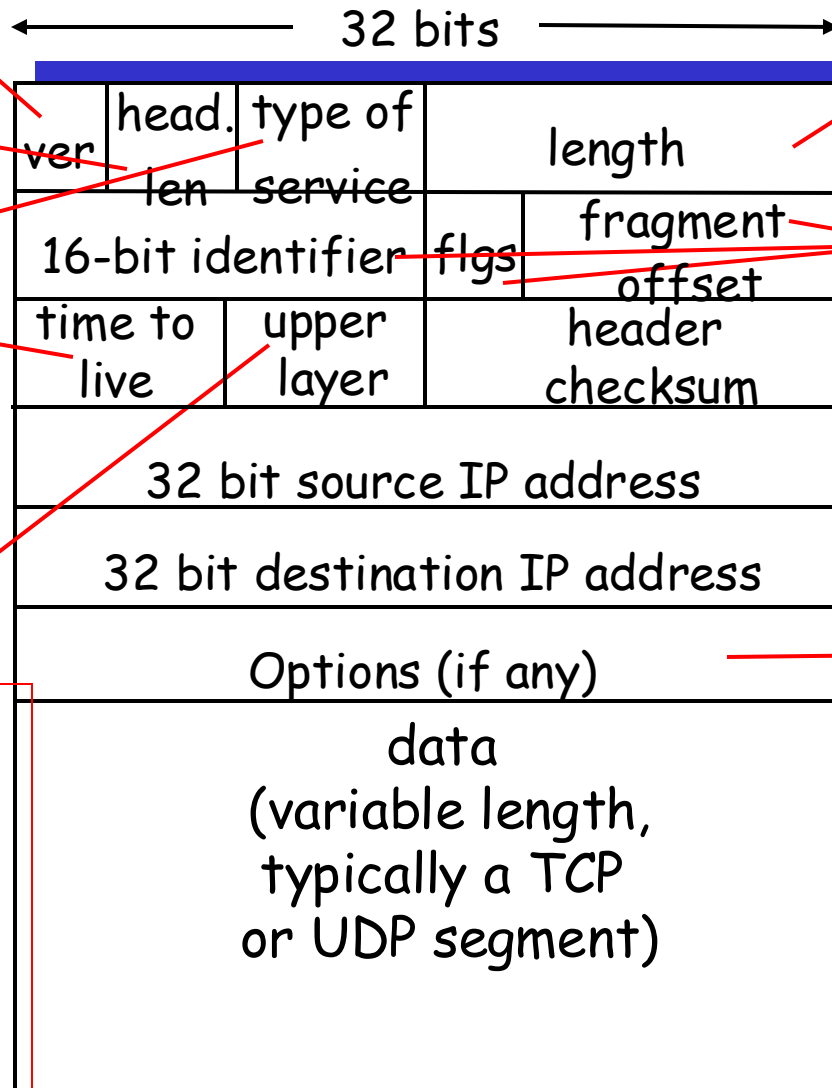
"τύπος" δεδομένων

μέγιστος αριθμός
αλμάτων (hops) που απομένουν
(μειώνεται κατά 1 σε
κάθε δρομολογητή)

Πρωτόκολλο ανώτερου επιπέδου
που θα παραδοθεί το φορτίο

TCPoverhead

- 20 bytes του TCP
- 20 bytes του IP
- = 40 bytes +
overhead επιπέδου
εφαρμογής

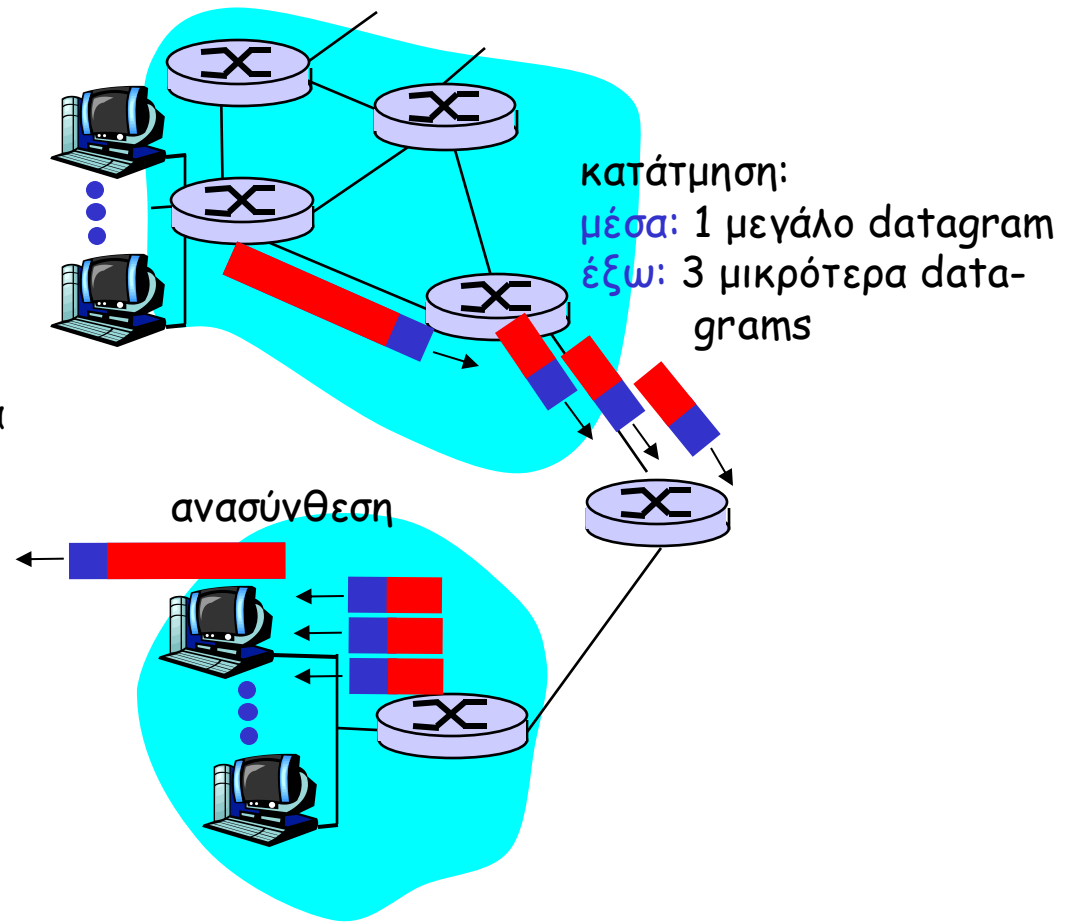


Συνολικό μήκος
datagram (bytes)
(max 65535 byte)
για
κατάτμηση/
ανασύνθεση

Π.χ. χρονοσφραγίδα,
καταγραφή διαδρομής
που ακολουθείται,
καθορισμός λίστας
δρομολογητών
που θα επισκεφτεί

Κατάτμηση και Ανασύνθεση του IP (Fragmentation & Reassembly)

- Οι ζεύξεις του δικτύου έχουν μέγιστη μονάδα μεταφοράς (MTU (max.transfer size)) - μέγιστο δυνατό πλαίσιο επιπέδου ζεύξης
 - Διαφορετικοί τύποι ζεύξης, διαφορετικά MTUs
- μεγάλο IP datagram τεμαχίζεται εντός του δικτύου
 - ένα datagram γίνεται πολλαπλά datagrams
 - "ανασυντίθενται" μόνο στον τελικό προορισμό
 - Τα bits της κεφαλίδας IP χρησιμοποιούνται για την ταυτοποίηση, διάταξη των σχετικών τεμαχίων (fragments)



Κατάτμηση και Ανασύνθεση του IP

Παράδειγμα

- 4000 byte datagram = 3980 bytes ωφέλιμο φορτίο
- MTU = 1500 bytes

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

Ένα μεγάλο datagram γίνεται πολλά μικρότερα datagrams

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

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

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

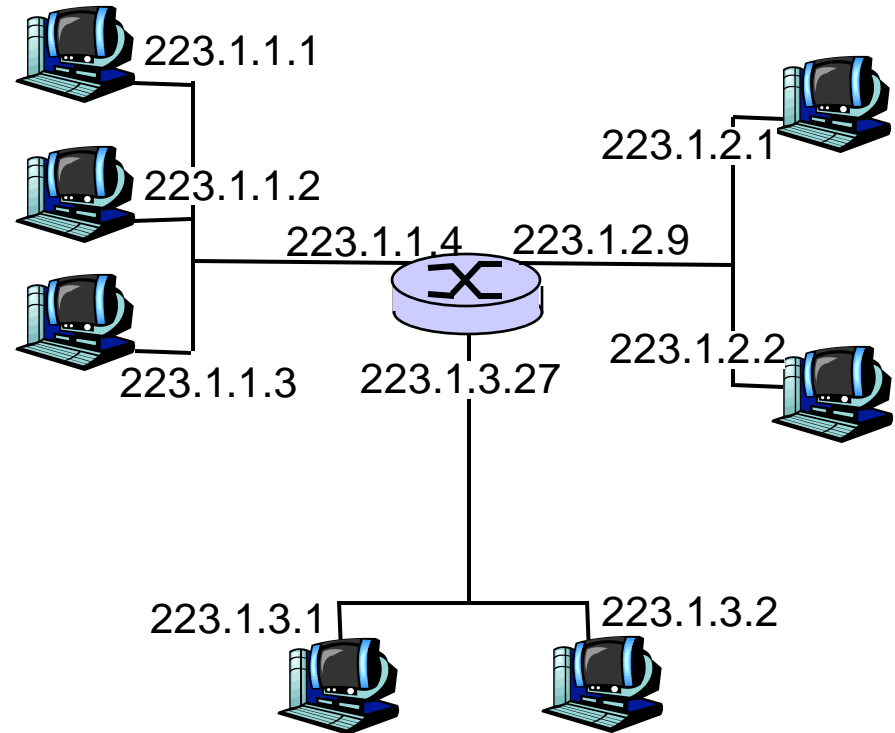
1480 bytes
στο πεδίο δεδομένων
(data field)

Μετατόπιση (offset) =
1480/8

Μετατόπιση (offset) =
2960/8

Διευθυνσιοδότηση IP

- Διεύθυνση IP (IP address):
32-bit αναγνωριστικό της διασύνδεσης του υπολογιστή, δρομολογητή
- Διασύνδεση (interface):
σύνδεση μεταξύ υπολογιστή/δρομολογητή και φυσικής ζεύξης
 - Οι δρομολογητές τυπικά έχουν πολλές διασυνδέσεις
 - Ένας υπολογιστής τυπικά έχει μία ή δύο διασυνδέσεις (π.χ. ενσύρματο Ethernet, ασύρματο 802.11)
- Διευθύνσεις IP σχετίζονται με κάθε διεπαφή



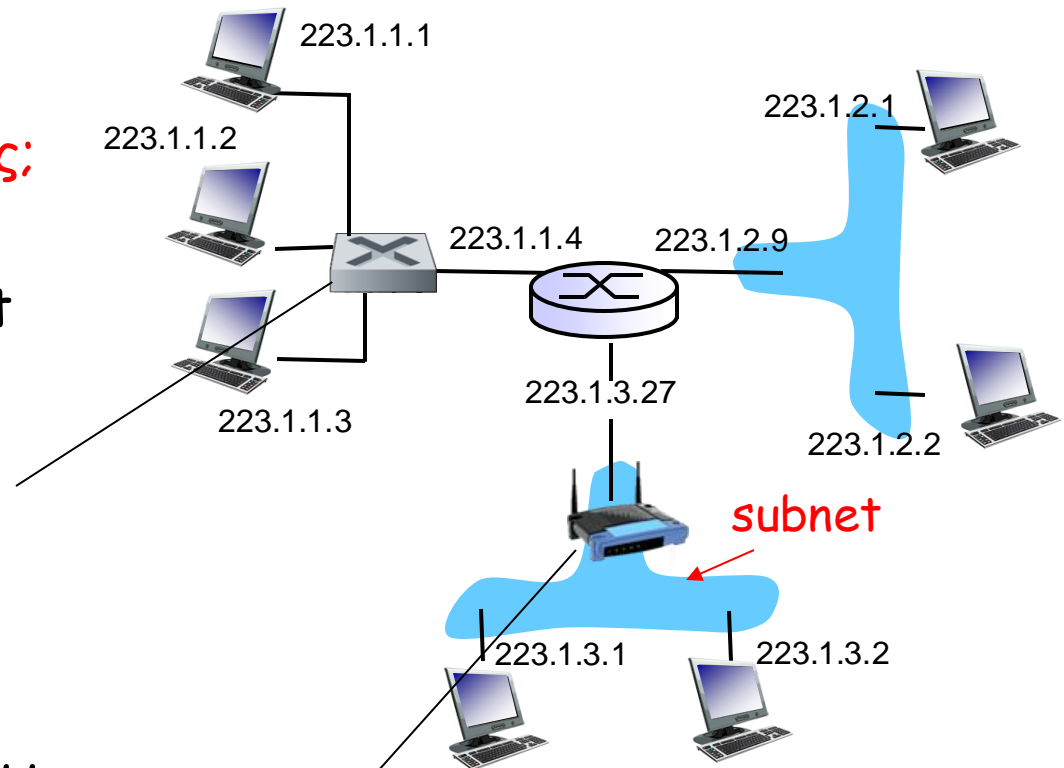
$$223.1.1.1 = \underbrace{11011111}_{223} \underbrace{00000001}_1 \underbrace{00000001}_1 \underbrace{00000001}_1$$

Διευθυνσιοδότηση IP: εισαγωγή

Ε: πώς συνδέονται στην πραγματικότητα οι διεπαφές;

Α: οι ενσύρματες Ethernet διεπαφές συνδέονται μέσω Ethernet μεταγωγών (switches)

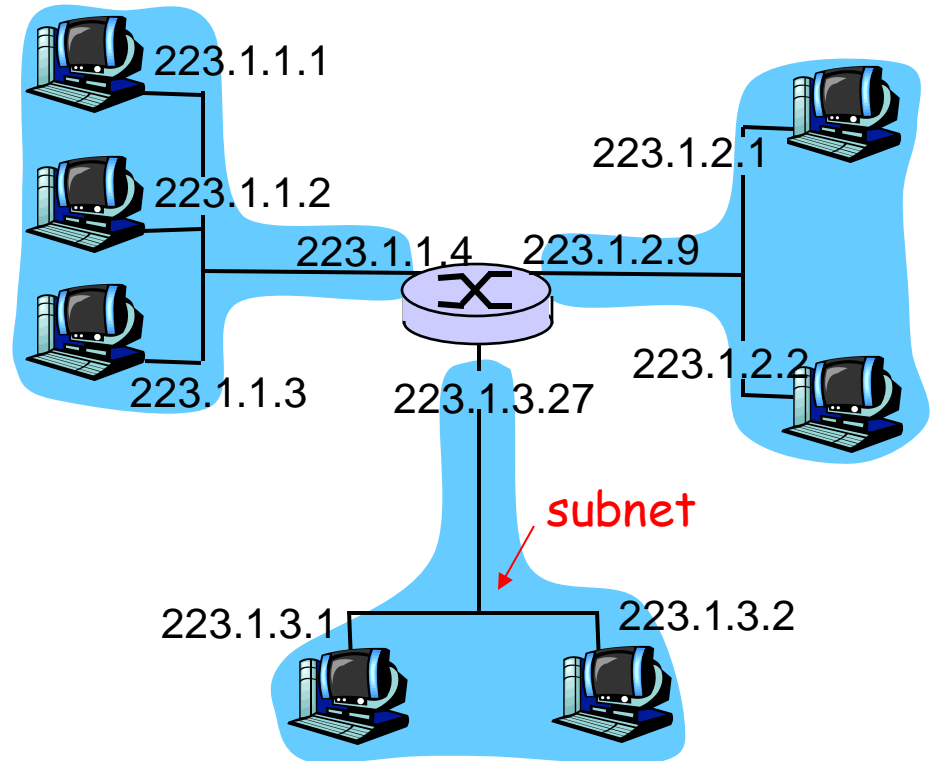
Για τώρα: δε χρειάζεται να ανησυχείτε για το πώς μία διεπαφή συνδέεται με μία άλλη (χωρίς δρομολογητή να παρεμβάλλεται)



Α: οι ασύρματες WiFi διεπαφές συνδέονται μέσω WiFi σταθμών βάσης (base station)

Υποδίκτυα (Subnets)

- Διεύθυνση IP:
 - Τμήμα υποδικτύου (subnet part) - bits υψηλής τάξης
 - Τμήμα υπολογιστή (host part) - bits χαμηλής τάξης
- Τι είναι ένα υποδίκτυο;
 - Διασυνδέσεις συσκευών με ίδιο τμήμα υποδικτύου στην IP διεύθυνση
 - Έχουν φυσική πρόσβαση το ένα στο άλλο **χωρίς μεσολάβηση δρομολογητή**

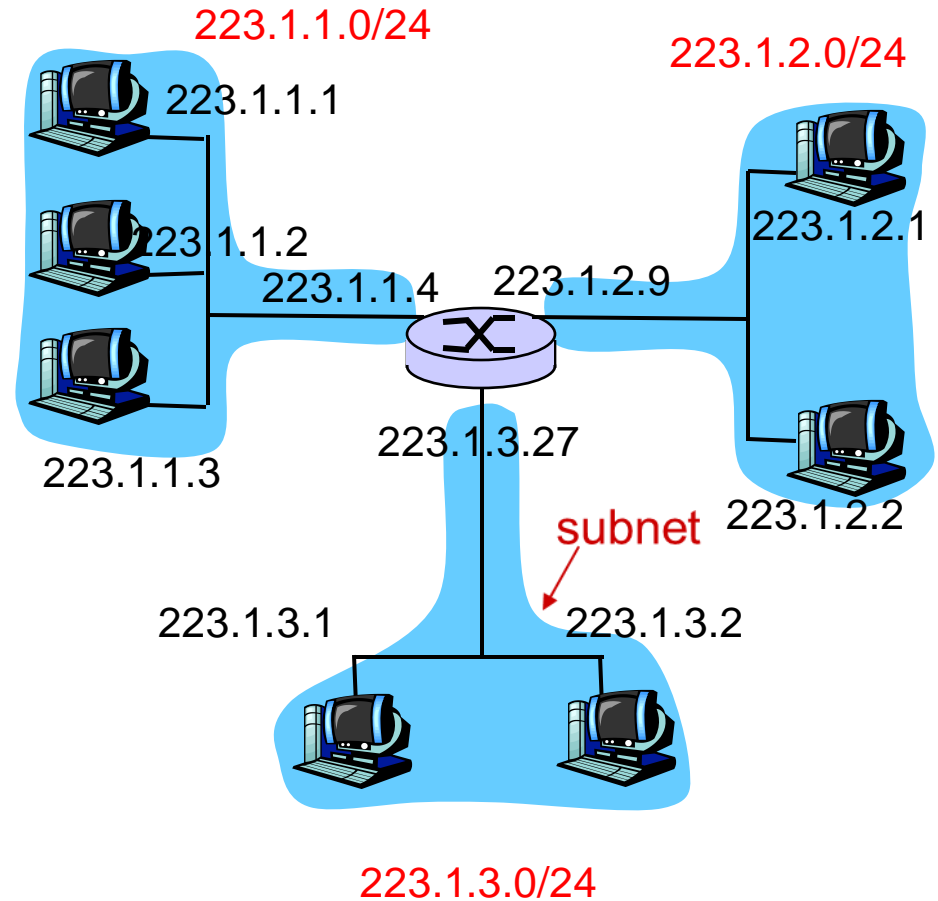


Δίκτυο που αποτελείται από 3 υποδίκτυα

Υποδίκτυα

Συνταγή

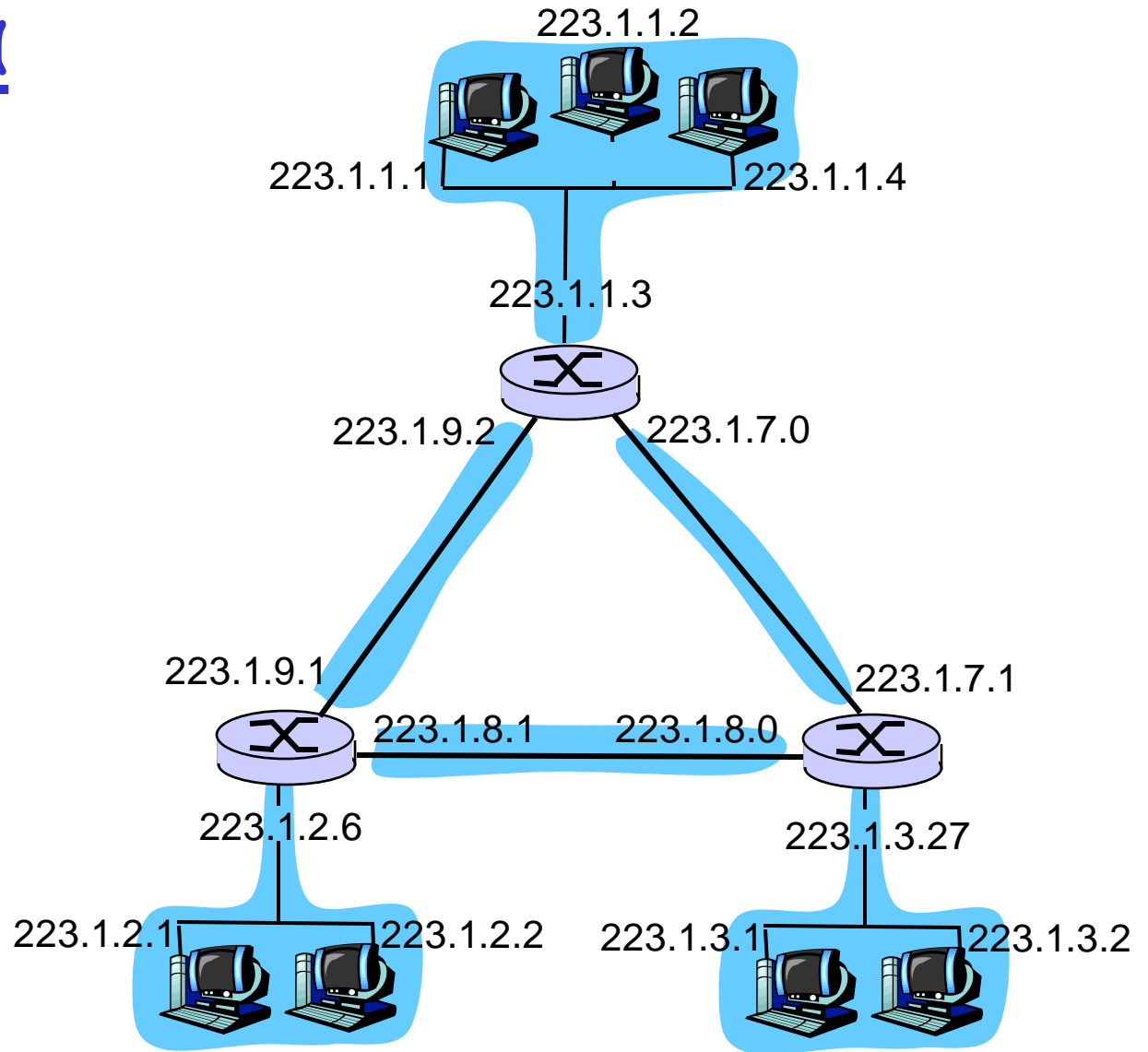
- Για τον καθορισμό των υποδικτύων, αποσυνδέστε κάθε διεπαφή από τον υπολογιστή ή το δρομολογητή, δημιουργώντας νησίδες απομονωμένων δικτύων
- Κάθε απομονωμένο δίκτυο καλείται **υποδίκτυο**



Μάσκα υποδικτύου
(Subnet mask): /24

Υποδίκτυα

Πόσα;



Διευθυνσιοδότηση IP: CIDR

CIDR: Classless InterDomain Routing

(Αταξική Διατομεακή Δρομολόγηση)

- Το τμήμα υποδικτύου (subnet part) της διεύθυνσης έχει αυθαίρετο μήκος
- Δομή διεύθυνσης: $a.b.c.d/x$, όπου x είναι ο # bits στο τμήμα υποδικτύου της διεύθυνσης



200.23.16.0/23

Διευθύνσεις IP: πώς αποδίδονται;

Ε: Πώς παίρνει ένας υπολογιστής διεύθυνση IP?

- Προσδιορισμένο σε ένα αρχείο από το διαχειριστή του συστήματος
 - Windows: control-panel->network->configuration->tcp/ip->properties
 - UNIX: /etc/rc.config
- **DHCP: Dynamic Host Configuration Protocol:**
δυναμική απόδοση διεύθυνσης από έναν εξυπηρέτη
 - "plug-and-play"

DHCP: Dynamic Host Configuration Protocol

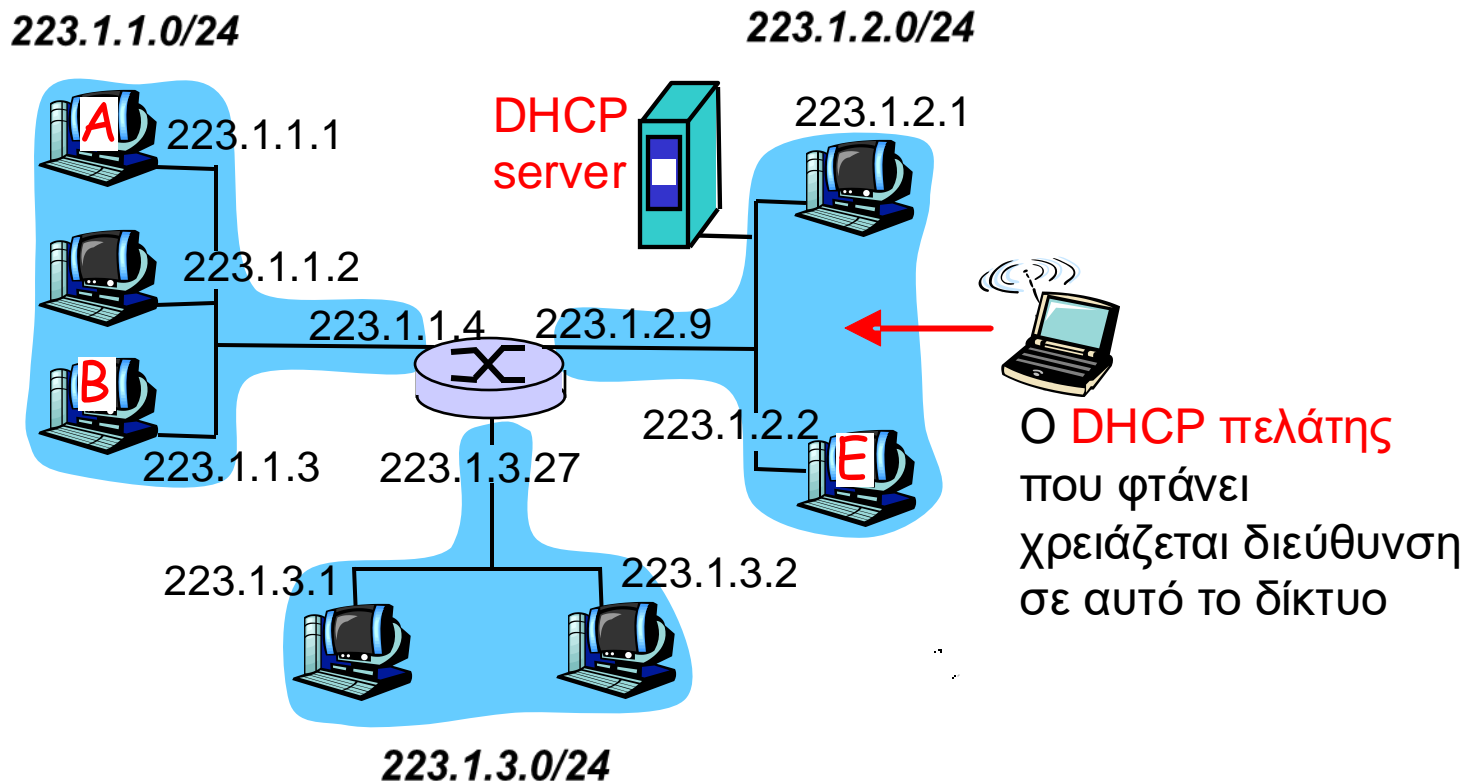
Σκοπός: να επιτρέπει στον υπολογιστή να αποκτήσει δυναμικά διεύθυνση IP από τον εξυπηρέτη του δικτύου όταν συνδέεται στο δίκτυο

- Μπορεί να ανανεώσει τη μίσθωση της διεύθυνσης που χρησιμοποιείται
- Επιτρέπει την επαναχρησιμοποίηση των διευθύνσεων (κρατά τη διεύθυνση μόνο όσο είναι συνδεδεμένος και «ενεργός»)
- Υποστήριξη για κινητούς χρήστες που θέλουν να συνδεθούν στο δίκτυο

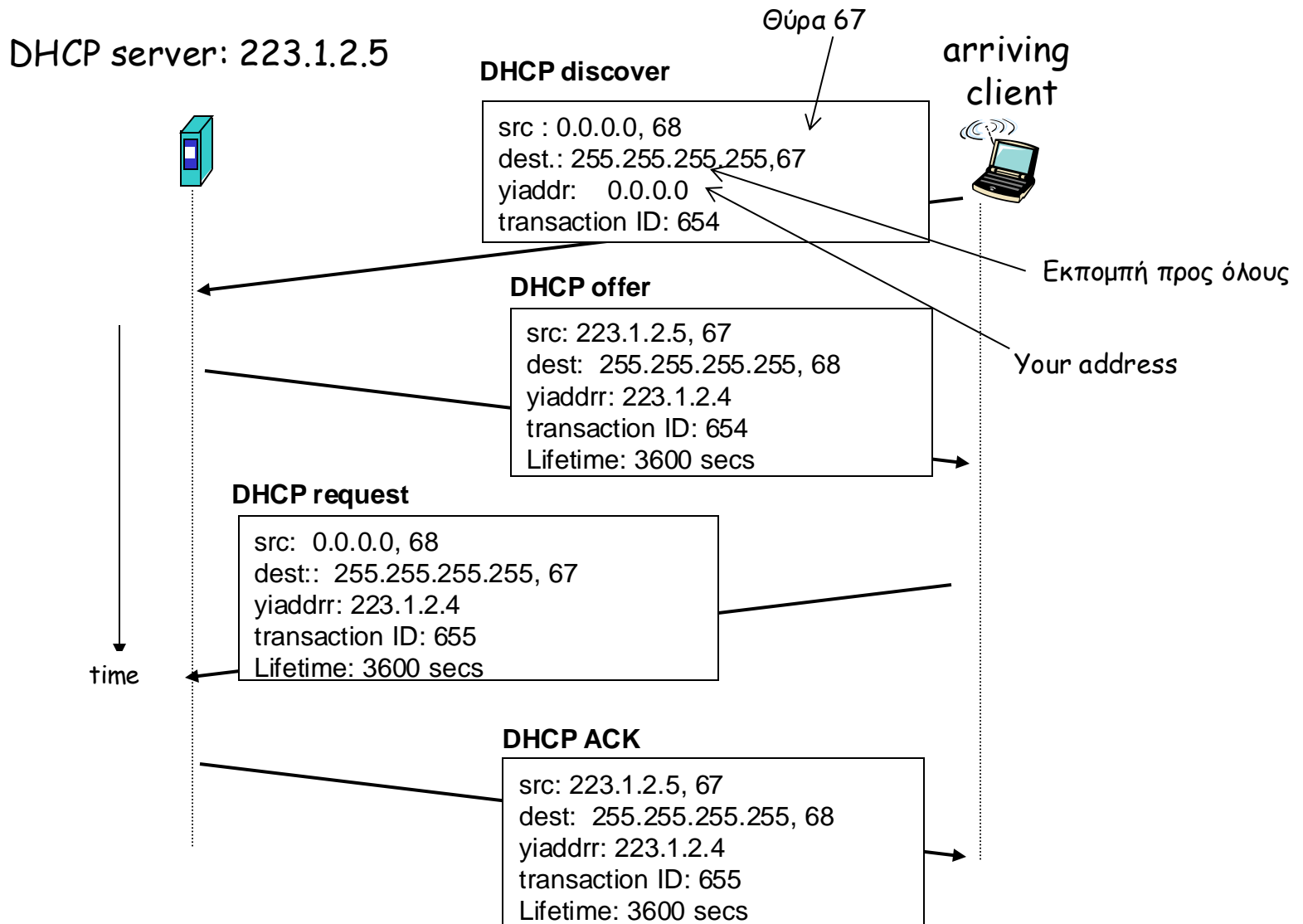
Επισκόπηση DHCP:

- Ο υπολογιστής εκπέμπει (broadcasts) "DHCP discover" msg
- Ο εξυπηρέτης DHCP αποκρίνεται με ένα μήνυμα "DHCP offer" msg
- Ο υπολογιστής ζητά διεύθυνση IP: "DHCP request" msg
- Ο εξυπηρέτης DHCP στέλνει τη διεύθυνση: "DHCP ack" msg

Σενάριο πελάτη-εξυπηρετή DHCP



Σενάριο πελάτη-εξυπηρετή DHCP

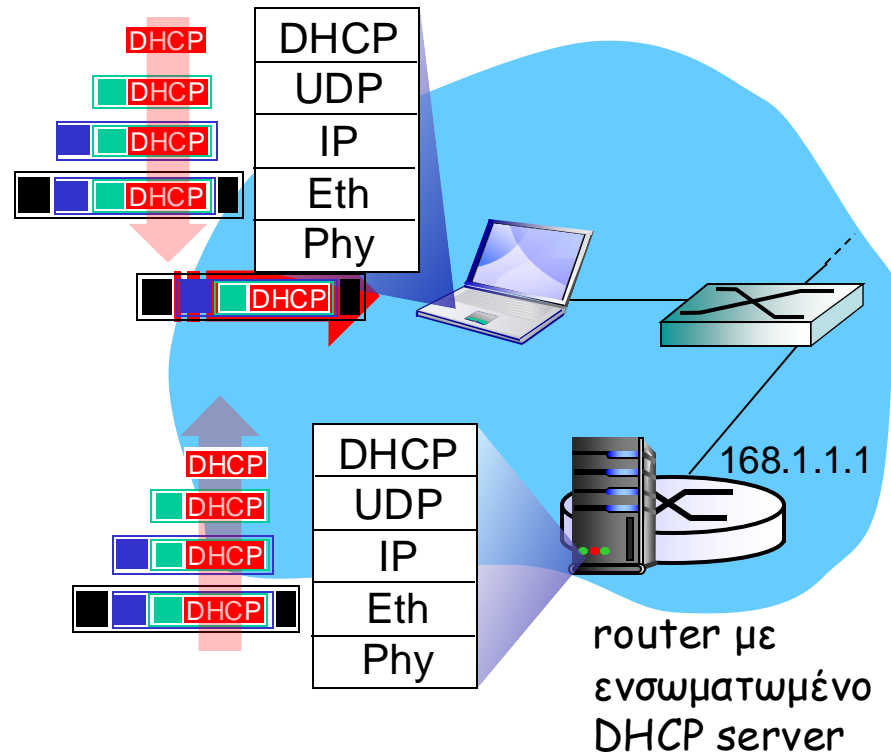


DHCP: περισσότερα

Το **DHCP** μπορεί να κάνει περισσότερα από απλά να κατανέμει μια IP διεύθυνση στο υποδίκτυο:

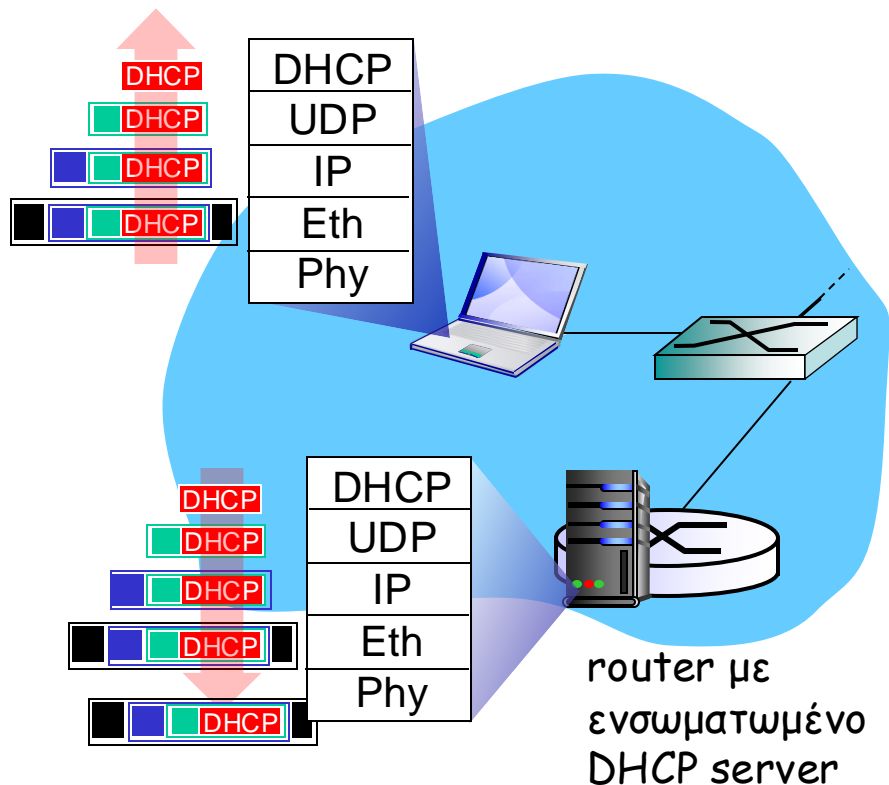
- Διεύθυνση του δρομολογητή πρώτου άλματος (first-hop router) για τον πελάτη
- Όνομα και IP διεύθυνση του DNS server
- Μάσκα δικτύου (υποδεικνύοντας το τμήμα δικτύου σε σχέση με το τμήμα υπολογιστή της διεύθυνσης)

DHCP: παράδειγμα



- ❖ Για τη σύνδεση του laptop απαιτείται η IP διεύθυνσή του, η διεύθυνση του δρομολογητή πρώτου άλματος, η διεύθυνση του DNS server: χρησιμοποίησε DHCP
- ❖ Το μήνυμα DHCP request ενθυλακώνεται στο UDP, που ενθυλακώνεται σε IP, που ενθυλακώνεται στο Ethernet 802.1
- ❖ Εκπέμπεται ένα broadcast Ethernet πλαίσιο στο LAN (προορισμός: FF-FF-FF-FF-FF-FF), που λαμβάνεται στο δρομολογητή ο οποίος τρέχει τον DHCP server
- ❖ Το Ethernet frame αποθυλακώνεται σε IP, το IP αποθυλακώνεται σε UDP και το UDP σε DHCP.

DHCP: παράδειγμα



- Ο DHCP server σχηματίζει το DHCP ACK μήνυμα εμπεριέχοντας την IP διεύθυνση του πελάτη, την IP διεύθυνση του δρομολογητή πρώτου άλματος για τον πελάτη, καθώς και το όνομα και IP διεύθυνση του DNS server.
- Ενθυλακώνεται το πακέτο του DHCP server, το πλαίσιο προωθείται στον πελάτη, όπου και αποθυλακώνεται ως το επίπεδο του DHCP πακέτου.
- Ο πελάτης πλέον γνωρίζει την IP διεύθυνσή του, το όνομα και IP διεύθυνση του DNS server και την IP διεύθυνση του δρομολογητή πρώτου άλματος.

Διευθύνσεις IP: πώς αποδίδονται;

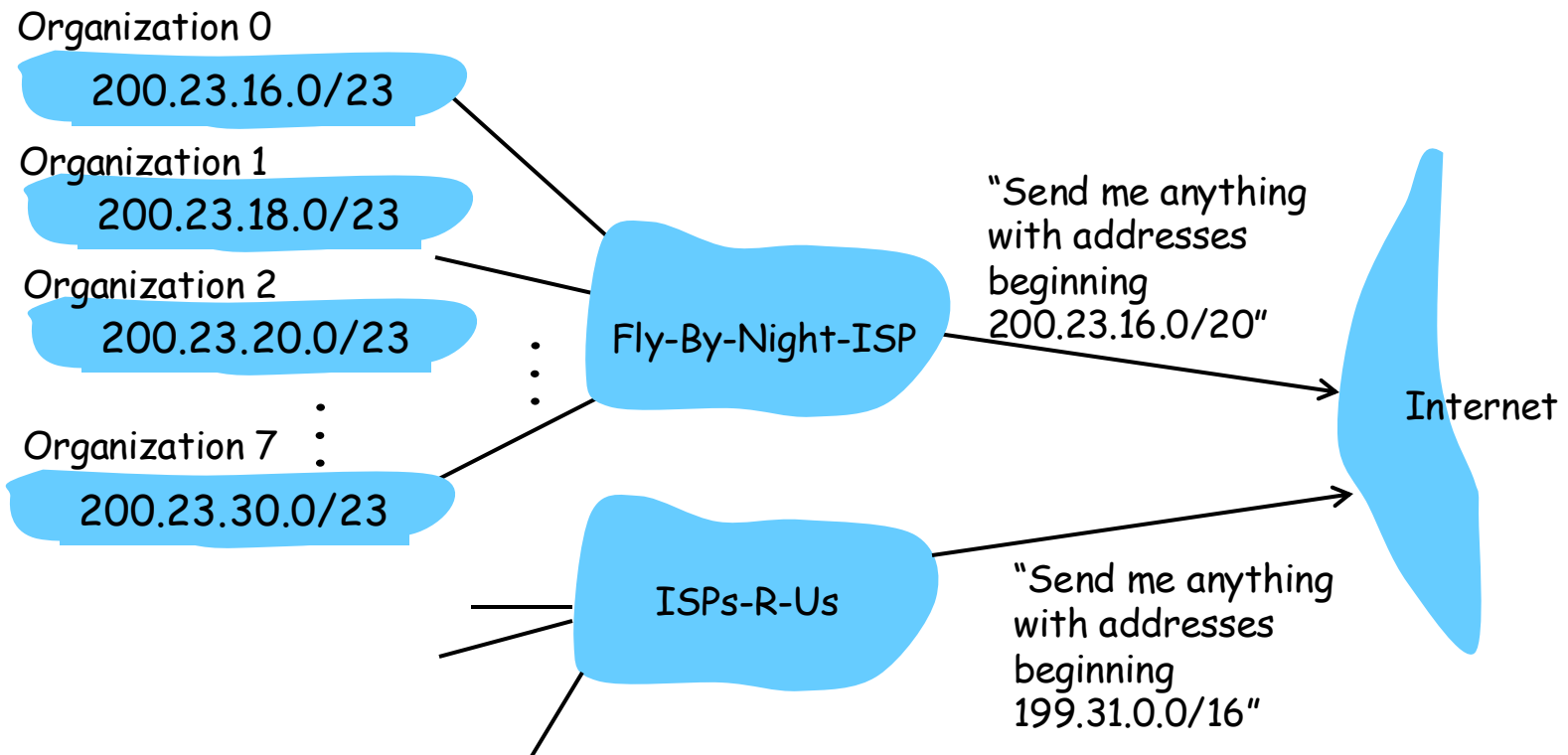
Ε: Πώς εκχωρείται στο δίκτυο το τμήμα υποδικτύου της διεύθυνσης IP;

Α: Λαμβάνει μερίδιο από τον χώρο διευθύνσεων που έχει εκχωρηθεί στον ISP του

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

Ιεραρχική διευθυνσιοδότηση (hierarchical addressing): συνάθροιση διαδρομών

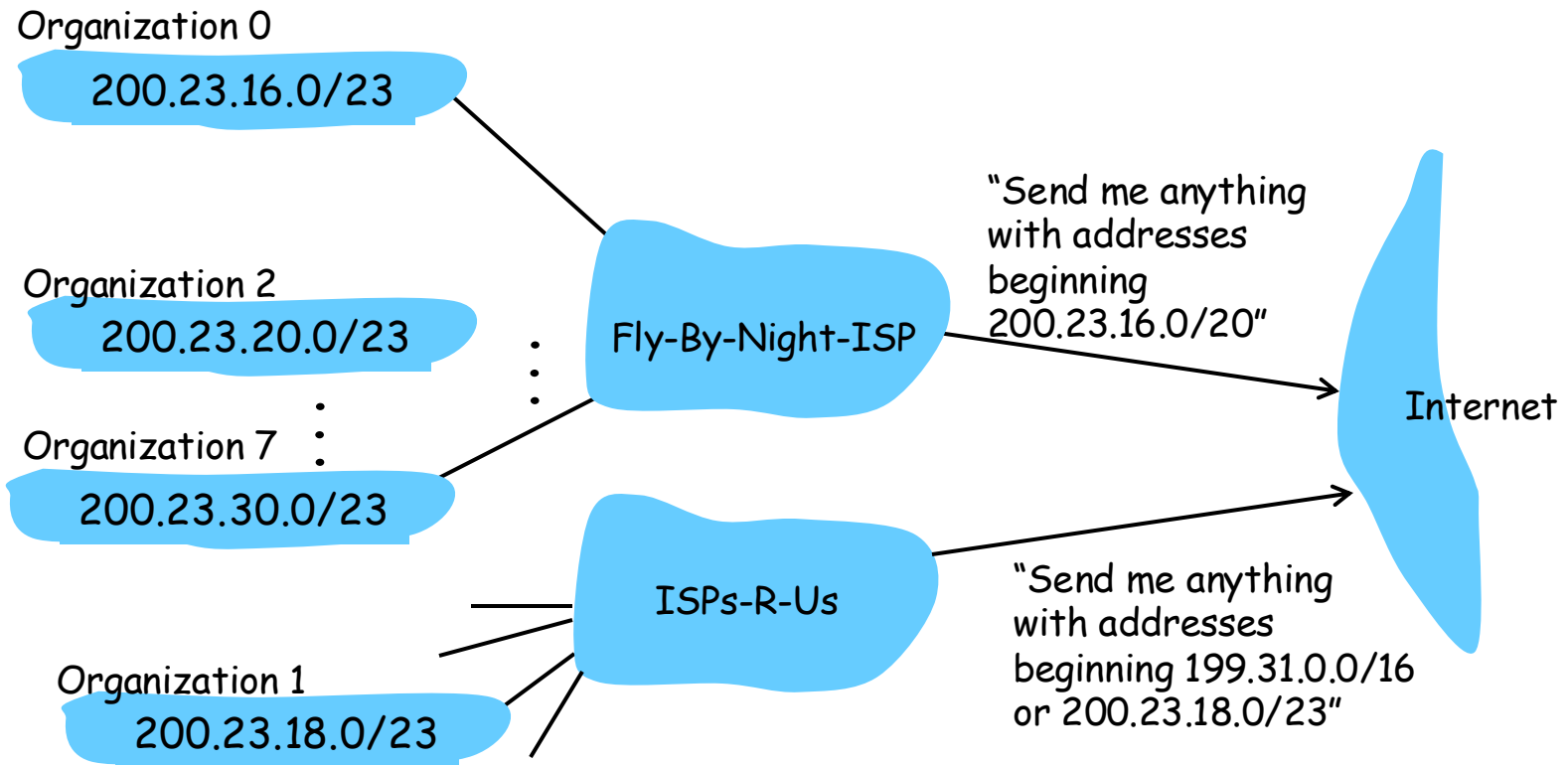
Η ιεραρχική διευθυνσιοδότηση επιτρέπει την αποδοτική διαφήμιση (advertisement) των πληροφοριών δρομολόγησης



Ιεραρχική διευθυνσιοδότηση : επιλέγονται οι πιο συγκεκριμένες διαδρομές

Ο ISPs-R-Us και ο Fly-By-Night-ISP δημοσιοποιούν blocks που περιέχουν τις διευθύνσεις του Organization 1

Ο ISPs-R-Us έχει μια πιο συγκεκριμένη* διαδρομή προς τον Organization 1



* -> μεγαλύτερου προθέματος (23>20). Άρα στέλνεται στον ISPs-R-US

Διευθυνσιοδότηση IP: τελευταίες λέξεις...

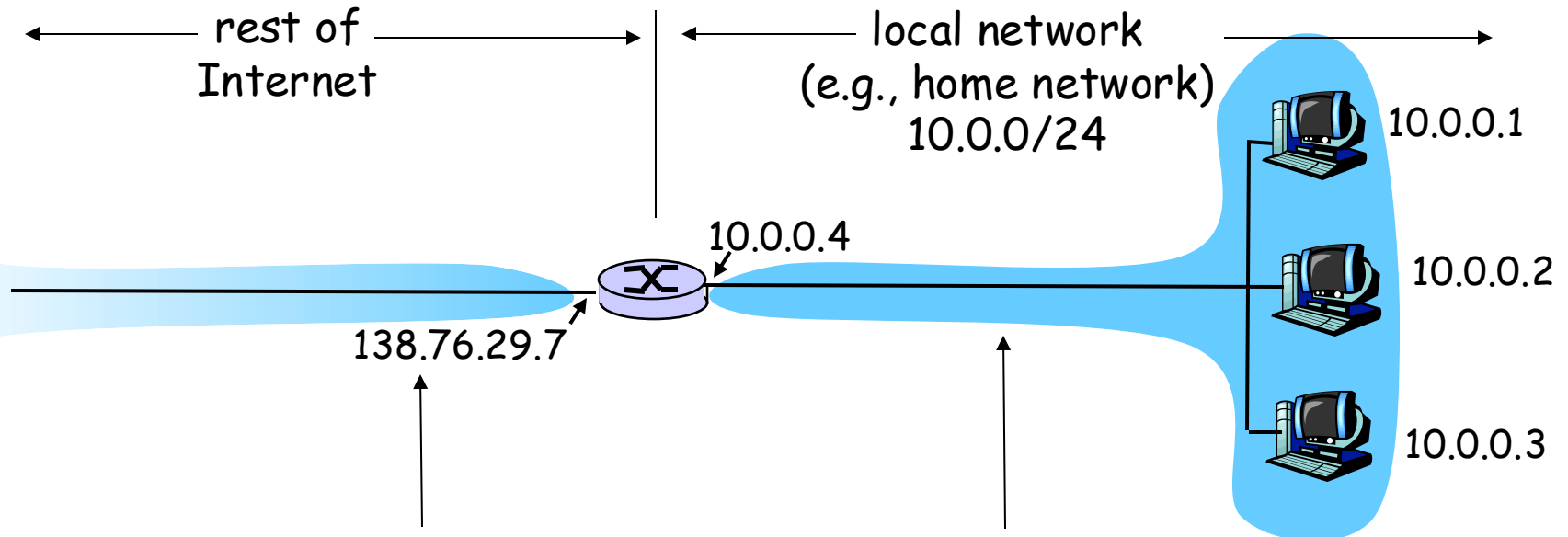
E: Πώς παίρνει ένας ISP ένα μπλοκ διευθύνσεων;

A: ICANN: Internet Corporation for Assigned Names and Numbers

- Κατανέμει διευθύνσεις
- Διαχειρίζεται το DNS
- Αποδίδει ονόματα τομέων (domain names), διευθετεί διαφορές

NAT: Network Address Translation

(Μετάφραση Διευθύνσεων Δικτύου)



Όλα τα datagrams που φεύγουν από το τοπικό δίκτυο έχουν την **ίδια** διεύθυνση IP προέλευσης NAT: 138.76.29.7, Διαφορετικούς αριθμούς θύρας προέλευσης

Τα datagrams με προέλευση ή προορισμό σε αυτό το δίκτυο έχουν διεύθυνση 10.0.0/24 ως προέλευση, προορισμό (όπως συνήθως)

NAT: Network Address Translation

- **Κίνητρο:** το τοπικό δίκτυο χρησιμοποιεί μόνο μία διεύθυνση IP όσο αφορά τον εξωτερικό κόσμο
 - Δεν απαιτείται σύνολο διευθύνσεων από τον ISP: μόνο μία διεύθυνση IP για όλες τις συσκευές
 - Μπορούν να αλλαχτούν οι διευθύνσεις των συσκευών στο τοπικό δίκτυο χωρίς να ειδοποιηθεί ο έξω κόσμος
 - Μπορεί να αλλάξει ο ISP χωρίς να αλλάξουν οι διευθύνσεις στο τοπικό δίκτυο
 - Μία συσκευή εντός του τοπικού δικτύου δεν είναι ορατή από τον έξω κόσμο ως σαφώς διευθυνσιοδοτημένη (ένα συν στην ασφάλεια).

NAT: Network Address Translation

Υλοποίηση από δρομολογητή NAT :

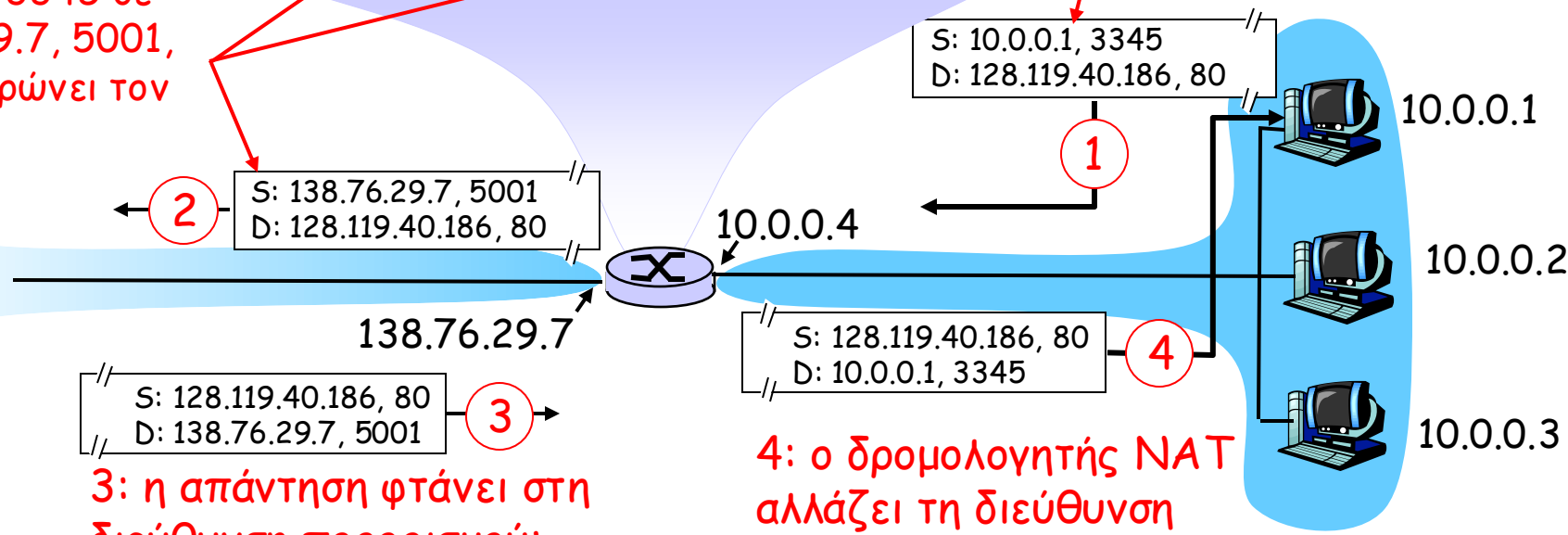
- *εξερχόμενα datagrams: αντικατάσταση* {διεύθυνση IP προέλευσης, # θύρας} κάθε εξερχόμενου datagram με {διεύθυνση IP NAT, νέος # θύρας}
... Οι απομακρυσμένοι πελάτες/εξυπηρέτες θα απαντήσουν χρησιμοποιώντας {διεύθυνση IP NAT, νέος # θύρας} ως διεύθυνση προορισμού
- *αποθήκευση (στον πίνακα μετάφρασης NAT (translation table))* κάθε ζεύγους μετάφρασης από {διεύθυνση IP προέλευσης, # θύρας} σε {διεύθυνση IP NAT, νέος # θύρας}
- *εισερχόμενα datagrams: αντικατάσταση* {διεύθυνση IP NAT, νέος # θύρας} στα πεδία προορισμού κάθε εισερχόμενου datagram με το αντίστοιχο {διεύθυνση IP προέλευσης, # θύρας} που είναι αποθηκευμένο στον πίνακα NAT

NAT: Network Address Translation

NAT translation table	
WAN side addr	LAN side addr
138.76.29.7, 5001	10.0.0.1, 3345
.....

2: ο δρομολογητής NAT αλλάζει τη διεύθυνση προέλευσης του datagram από 10.0.0.1, 3345 σε 138.76.29.7, 5001, και ενημερώνει τον πίνακα

1: ο υπολογιστής 10.0.0.1 στέλνει datagram στο 128.119.40.186, 80



3: η απάντηση φτάνει στη διεύθυνση προορισμού: 138.76.29.7, 5001

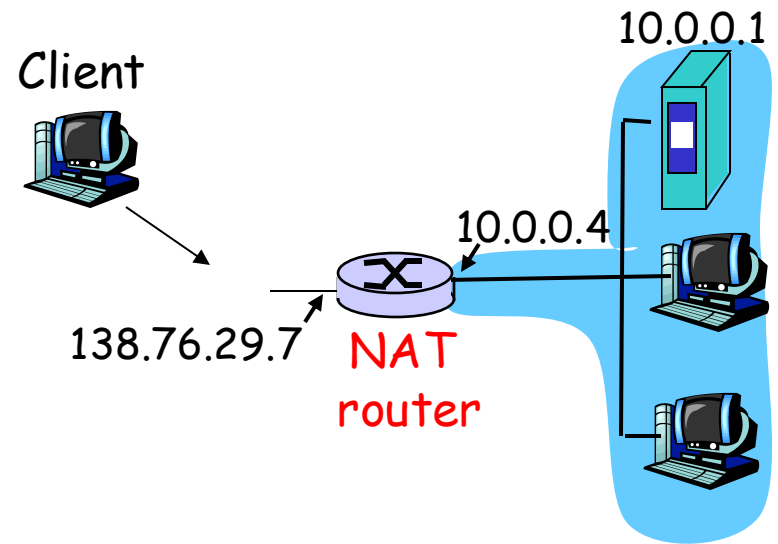
4: ο δρομολογητής NAT αλλάζει τη διεύθυνση προορισμού του datagram από 138.76.29.7, 5001 σε 10.0.0.1, 3345

NAT: Network Address Translation

- Πεδίο 16-bit αριθμού θύρας:
 - 60,000 ταυτόχρονες συνδέσεις με μία μόνο διεύθυνση στην πλευρά του LAN !
- Το NAT είναι αμφιλεγόμενο:
 - Οι δρομολογητές θα πρέπει να επεξεργάζονται πακέτα μόνο μέχρι το επίπεδο 3
 - Παραβιάζει την αρχή από-άκρο-σε-άκρο (end-to-end argument)
 - Η NAT δυνατότητα πρέπει να ληφθεί υπόψη από τους σχεδιαστές εφαρμογών, πχ, εφαρμογές P2P
 - Η έλλειψη διευθύνσεων θα πρέπει να επιλυθεί με το IPv6 (και όχι με το NAT)

Πρόβλημα εγκάρσιας διάβασης NAT

- ❑ Πελάτης θέλει να συνδεθεί σε εξυπηρετή με διεύθυνση 10.0.0.1
 - ❑ Η διεύθυνση του εξυπηρετή 10.0.0.1 είναι τοπική στο LAN (ο πελάτης δεν μπορεί να τη χρησιμοποιήσει ως διεύθυνση προορισμού)
 - ❑ Μόνο μία εξωτερικά ορατή διεύθυνση στην οποία εφαρμόζεται NAT: 138.76.29.7
- ❑ Λύση 1: στατική διαμόρφωση του NAT να προωθεί αιτήσεις εισερχόμενων συνδέσεων στη δοσμένη θύρα του εξυπηρετή
 - ❑ Π.χ., (138.76.29.7, θύρα 2500) προωθείται πάντα στο 10.0.0.1 θύρα 2500

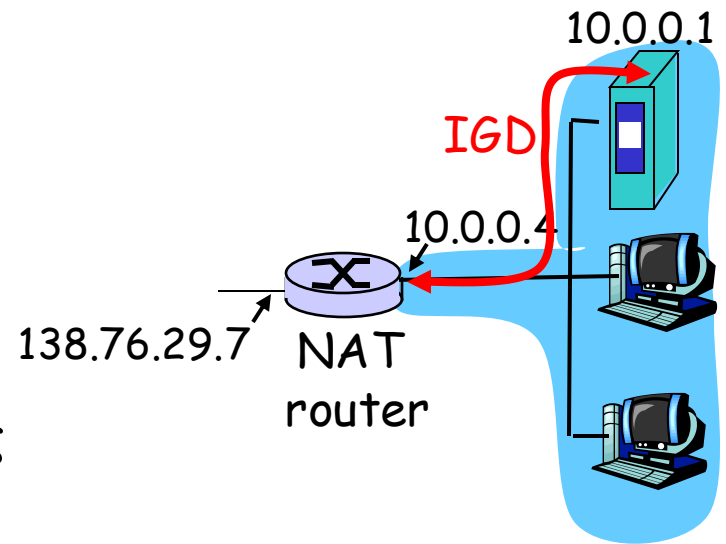


Πρόβλημα εγκάρσιας διάβασης NAT

- Λύση 2: Universal Plug and Play (UPnP) Internet Gateway Device (IGD) Protocol. Επιτρέπει σε υπολογιστή πίσω από NAT:

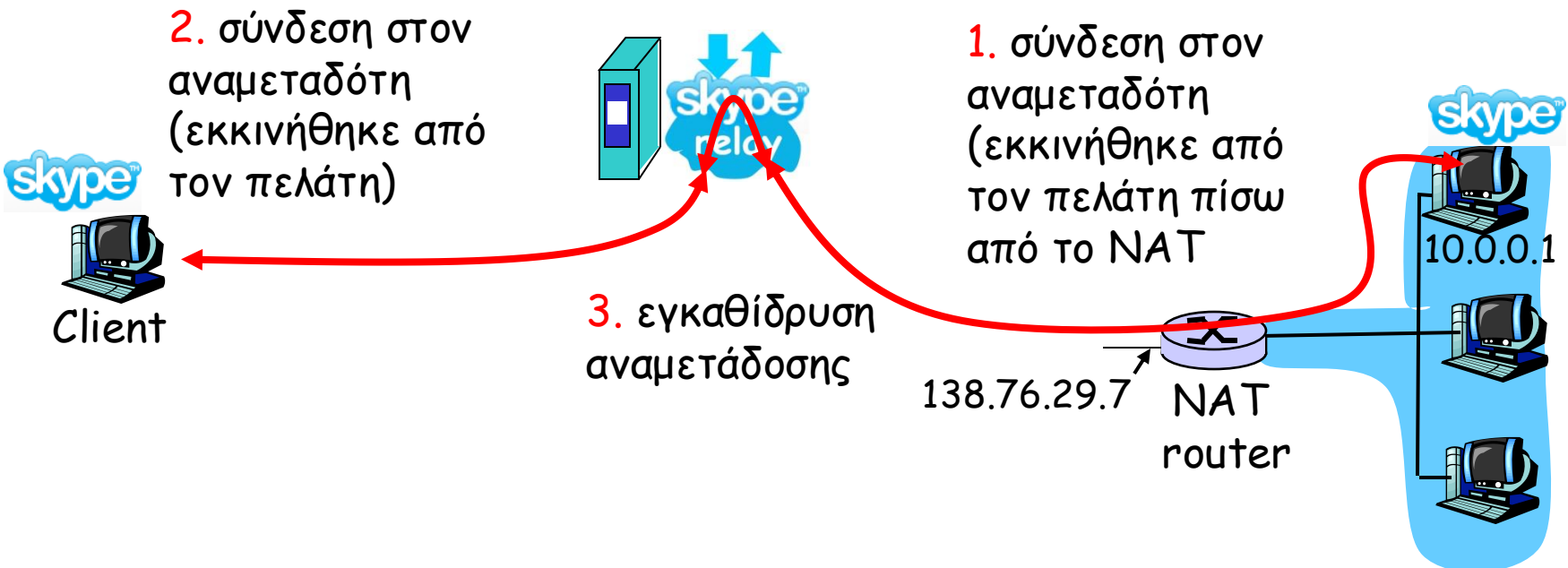
- ❖ Εκμάθηση δημόσιας διεύθυνσης IP (138.76.29.7)
- ❖ Προσθήκη/αφαίρεση αντιστοιχίσεων θυρών (με χρόνους μίσθωσης)

δηλ., αυτοματοποίηση της στατικής διαμόρφωσης αντιστοίχισης θυρών στο NAT



Πρόβλημα εγκάρσιας διάβασης NAT

- **Λύση 3: αναμετάδοση** (χρησιμοποιείται στο Skype)
 - Ο πελάτης πίσω από το NAT εγκαθιδρύει σύνδεση με αναμεταδότη
 - Ο εξωτερικός πελάτης συνδέεται με αναμεταδότη (relay)
 - Ο αναμεταδότης γεφυρώνει πακέτα μεταξύ των συνδέσεων

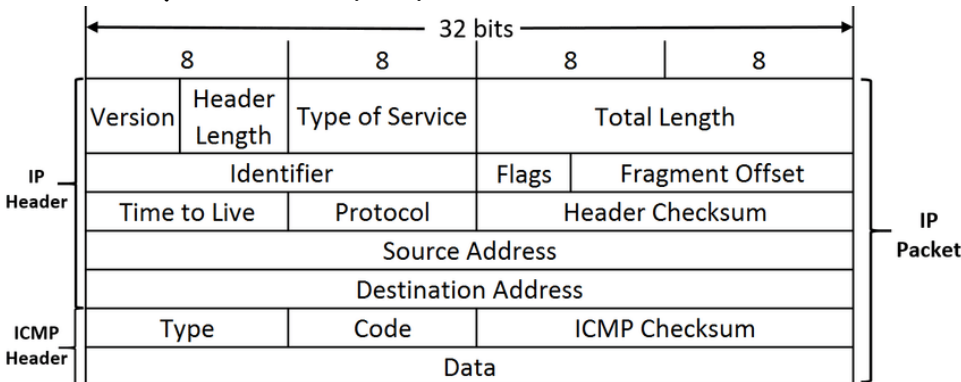


ICMP: Internet Control Message Protocol

(Πρωτόκολλο Ελέγχου Μηνυμάτων Διαδικτύου)

- ❑ Χρησιμοποιείται από υπολογιστές & δρομολογητές για ανταλλαγή πληροφορίας επιπέδου δικτύου
 - ❑ Αναφορά σφαλμάτων: μη προσπελάσιμος υπολογιστής, δίκτυο, θύρα, πρωτόκολλο
 - ❑ Αίτηση/απάντηση ηχούς (echo request/reply) (χρήση στο ping)
- ❑ Επίπεδο δικτύου "πάνω" από το IP:
 - ❑ Τα μηνύματα του ICMP μεταφέρονται σε IP datagrams
- ❑ **Μήνυμα ICMP:** τύπος, κωδικός και τα 8 πρώτα bytes του IP datagram που προκαλεί σφάλμα

<u>Type</u>	<u>Code</u>	<u>description</u>
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header



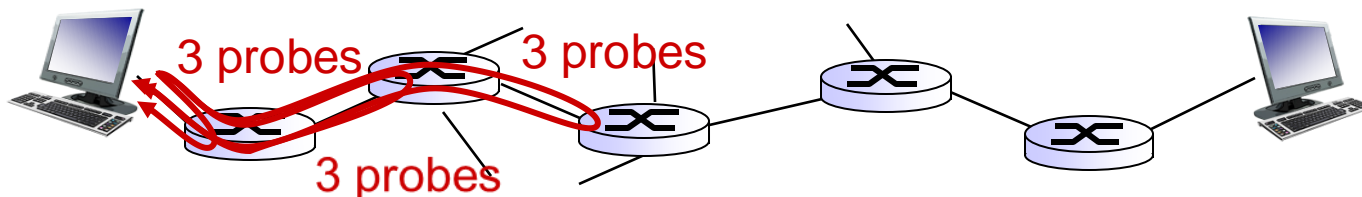
Traceroute και ICMP

- Η προέλευση στέλνει μια σειρά από τμήματα UDP στον προορισμό
 - Το πρώτο έχει TTL =1
 - Το δεύτερο έχει TTL=2, κτλ.
 - Ασυνήθιστος αριθμός θύρας
- Όταν το n-στο datagram φτάνει στο n-στο δρομολογητή:
 - Ο δρομολογητής απορρίπτει το datagram
 - Στέλνει στην πηγή ένα μήνυμα ICMP (τύπος 11, κωδικός 0)
 - Το μήνυμα περιέχει το όνομα και την IP διεύθυνση του δρομολογητή

- Όταν το μήνυμα ICMP φτάνει, η προέλευση υπολογίζει το RTT
- Το traceroute το κάνει αυτό 3 φορές.

Κριτήριο Λήξης

- Το τμήμα UDP τελικά φτάνει στον υπολογιστή προορισμό
- Ο προορισμός επιστρέφει πακέτο ICMP "host unreachable" (τύπος 3, κωδικός 3)
- Όταν η προέλευση παίρνει αυτό το πακέτο ICMP, σταματά.



IPv6: κίνητρο

- ❑ **Αρχικό κίνητρο:** Ο χώρος των 32-bit διευθύνσεων είχε δεσμευτεί σχεδόν πλήρως.
- ❑ **Επιπλέον κίνητρο:**
 - ❑ Η δομή της κεφαλίδας βοηθά στην ταχύτητα επεξεργασίας/ προώθησης
 - ❑ Η κεφαλίδα αλλάζει για να εξυπηρετήσει το QoS

Δομή IPv6 datagram:

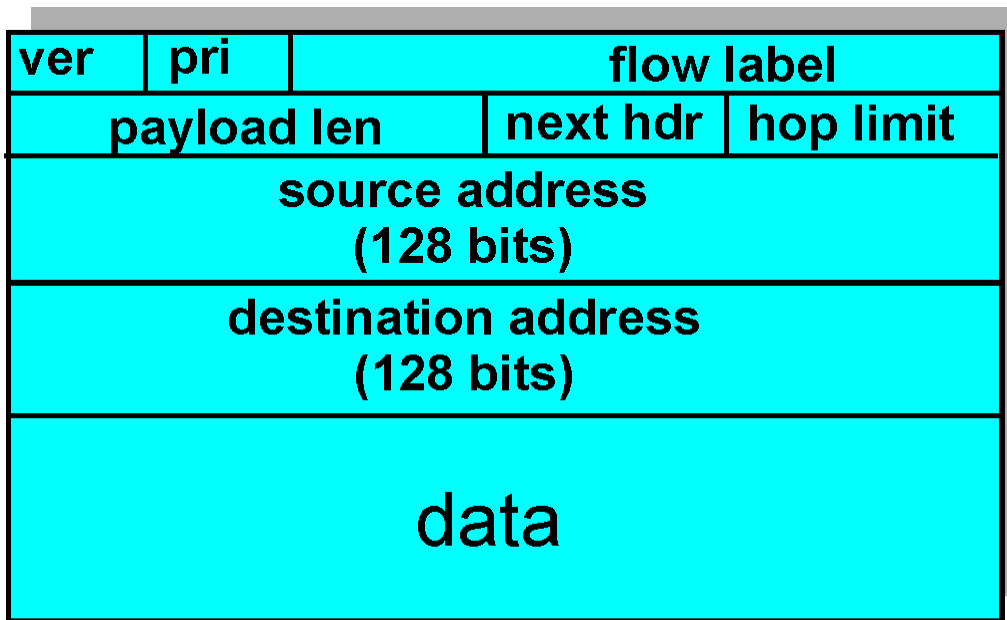
- ❑ Σταθερού μήκους κεφαλίδα 40 byte
- ❑ Δεν επιτρέπεται κατάτμηση

Δομή IPv6 Κεφαλίδας

Προτεραιότητα (Priority): καθορίζει την προτεραιότητα μεταξύ των datagrams στη ροή (flow)

Ετικέτα ροής (Flow Label): ταυτοποιεί datagrams στην ίδια "ροή" (η έννοια της "ροής" δεν είναι καλά ορισμένη)

Επόμενη κεφαλίδα (Next header): αναγνωρίζει το πρωτόκολλο ανώτερου επιπέδου για δεδομένα



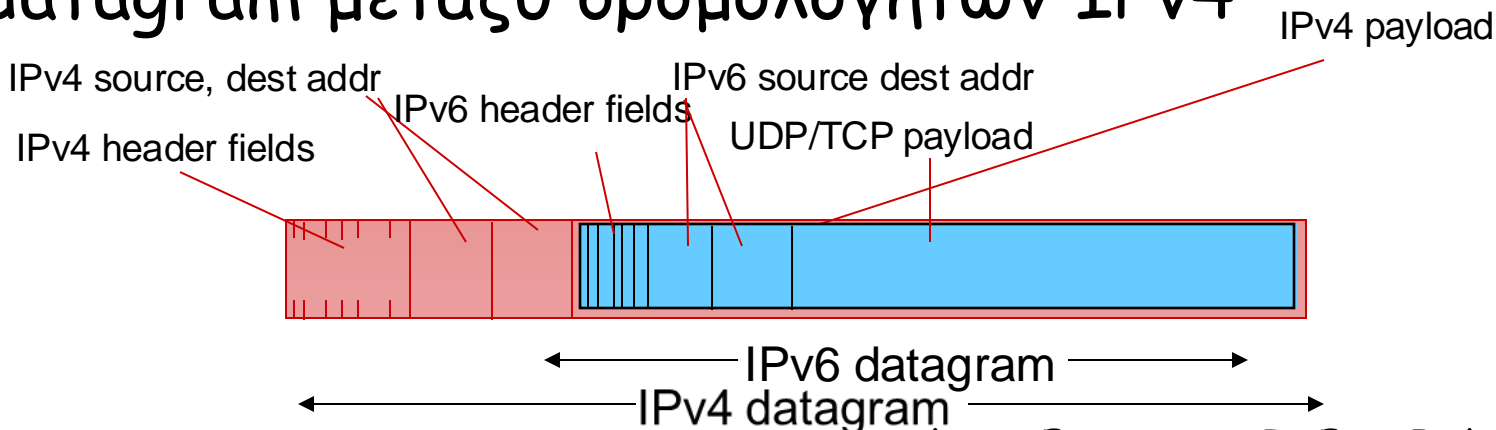
← 32 bits →

Άλλες αλλαγές σε σχέση με το IPv4

- ❑ **Άθροισμα ελέγχου (checksum)**: αφαιρέθηκε τελείως προκειμένου να μειωθεί ο χρόνος επεξεργασίας σε κάθε άλμα (hop)
- ❑ **Επιλογές**: επιτρέπονται, αλλά εκτός κεφαλίδας, καθορίζονται από το πεδίο "Next Header"
- ❑ **ICMPv6**: νέα έκδοση του ICMP
 - Πρόσθετοι τύποι μηνυμάτων, π.χ. "Πακέτο πολύ μεγάλο"
 - Συναρτήσεις διαχείρισης ομάδων πολυεκπομπής (multicast)

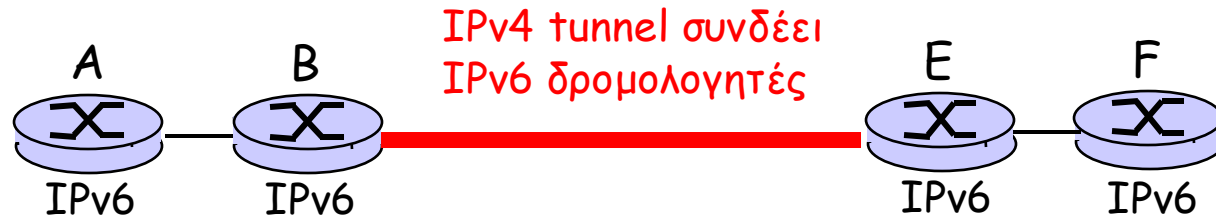
Μετάβαση από το IPv4 στο IPv6

- Δεν μπορούν να αναβαθμιστούν όλοι οι δρομολογητές ταυτόχρονα
 - χωρίς "ημερομηνία μετάβασης"
 - Πώς θα μπορέσει το δίκτυο να λειτουργήσει με μείγμα δρομολογητών IPv4 και IPv6;
- **Σηράγγωση (Tunneling):** Το IPv6 datagram μεταφέρεται ως ωφέλιμο φορτίο (payload) σε IPv4 datagram μεταξύ δρομολογητών IPv4

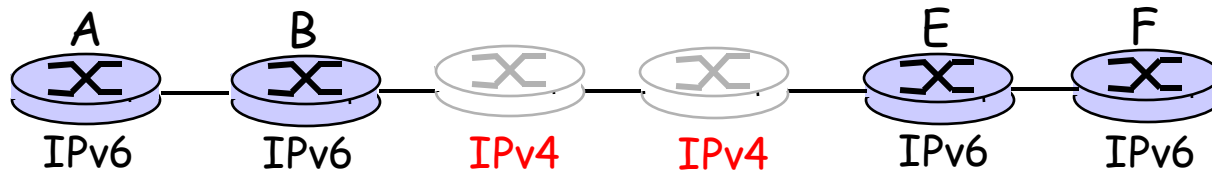


Σηράγγωση (Tunneling)

Λογική όψη:

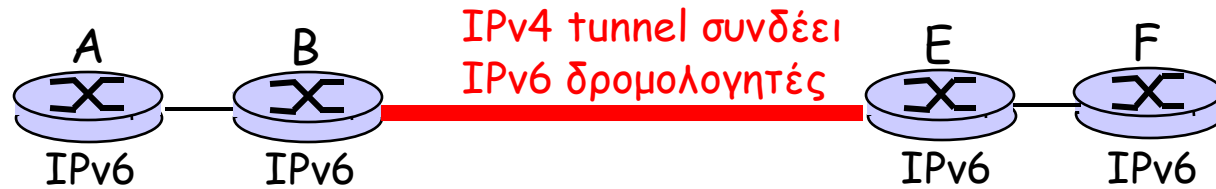


Φυσική όψη:

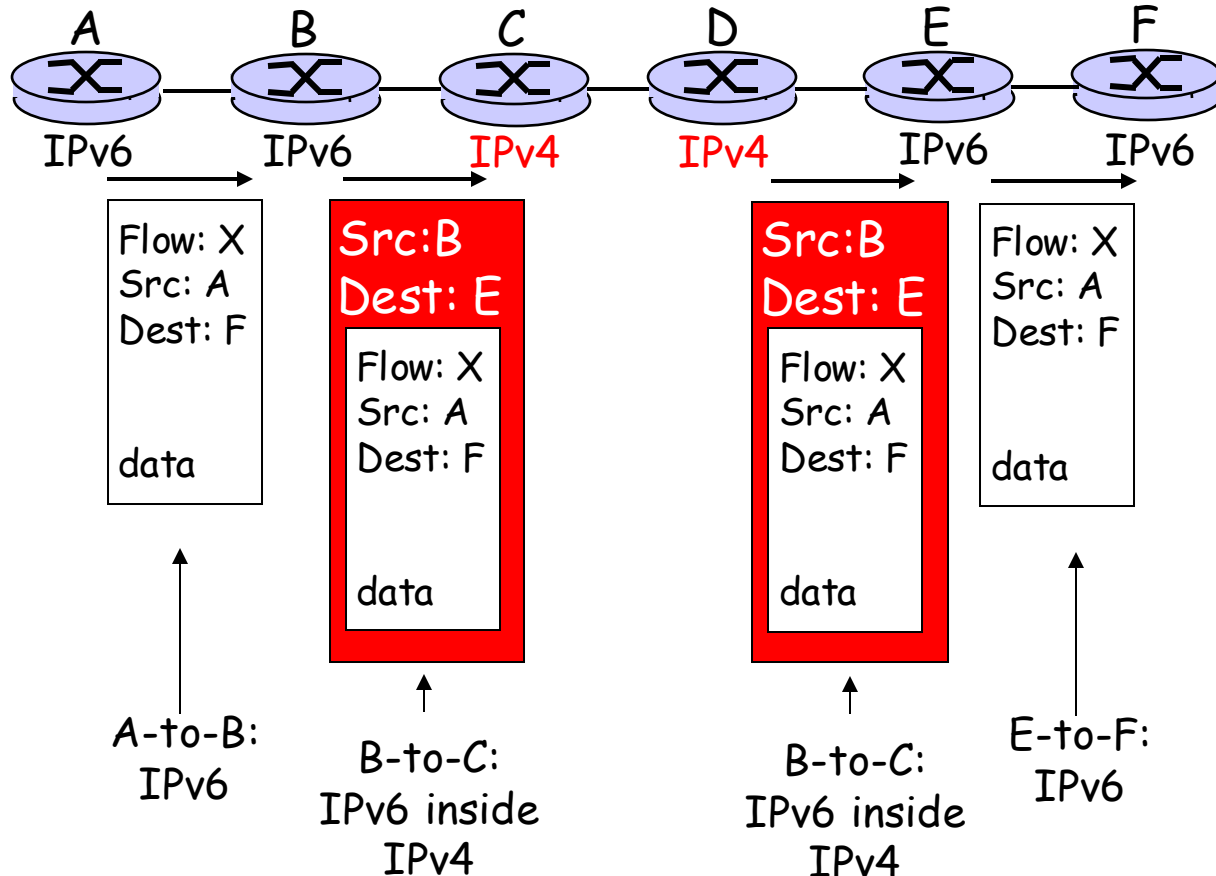


Σηράγγωση (Tunneling)

Λογική όψη:



Φυσική όψη:



Κεφάλαιο 4: Επίπεδο Δικτύου

4.1 Εισαγωγή

4.2 Δίκτυα εικονικού κυκλώματος και δεδομενογράμματος

4.3 Τι βρίσκεται μέσα σ' ένα δρομολογητή

4.4 IP: Πρωτόκολλο Διαδικτύου (Internet Protocol)

- ❑ Μορφή δεδομενογράμματος
- ❑ Διευθυνσιοδότηση IPv4
- ❑ ICMP
- ❑ IPv6

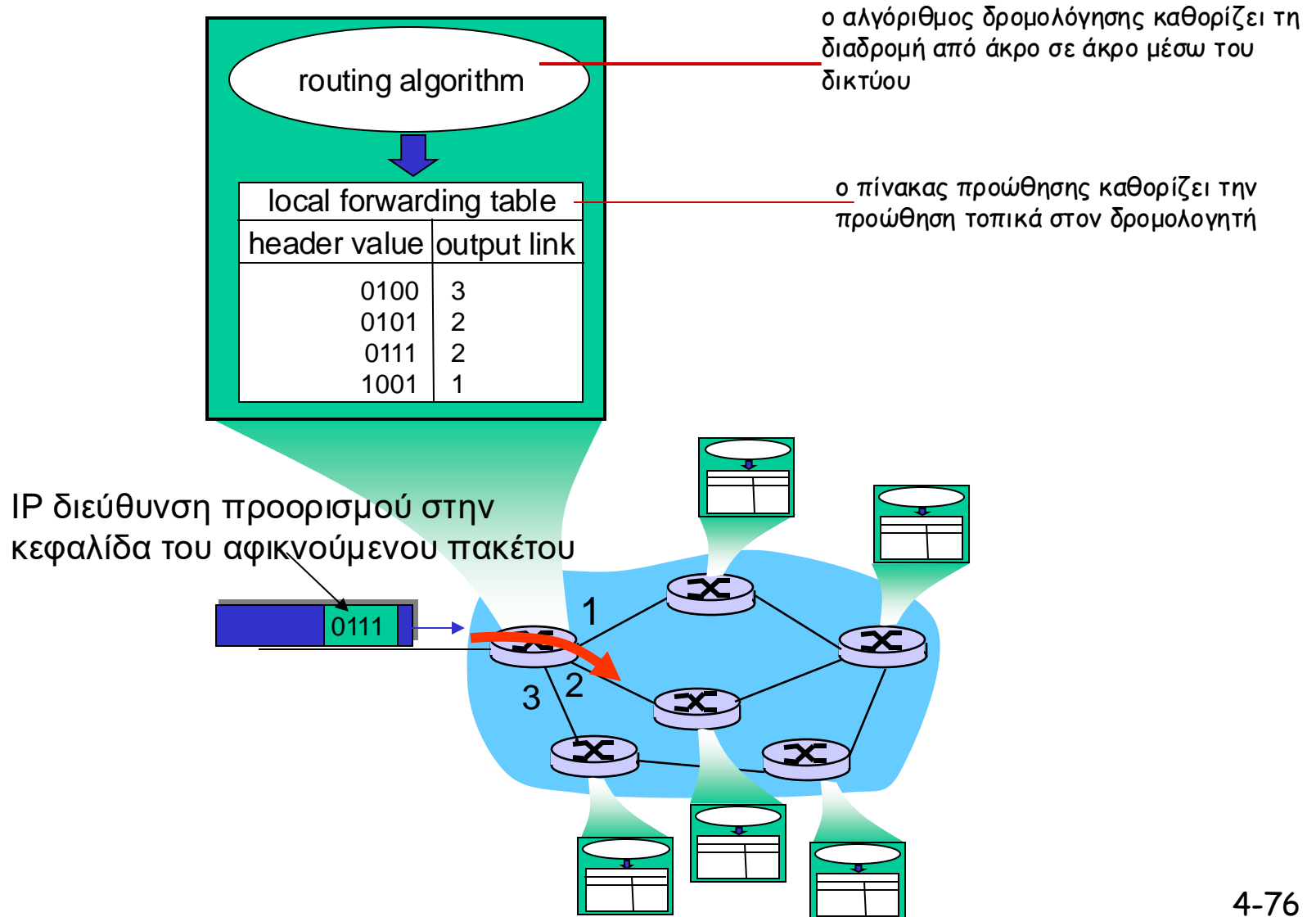
4.5 Αλγόριθμοι δρομολόγησης

- ❑ Κατάστασης ζεύξης (Link State)
- ❑ Διανύσματος απόστασης (Distance Vector)
- ❑ Ιεραρχική δρομολόγηση

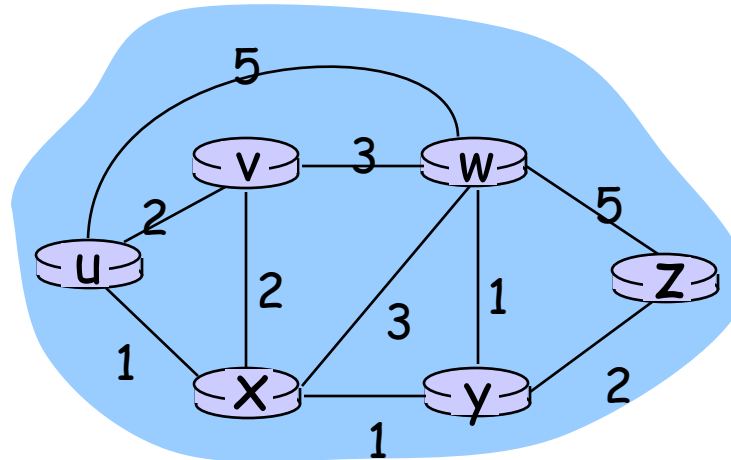
4.6 Δρομολόγηση στο Διαδίκτυο

- ❑ RIP
- ❑ OSPF
- ❑ BGP

Αλληλεπίδραση μεταξύ δρομολόγησης και προώθησης



Αφηρημένο μοντέλο γράφων



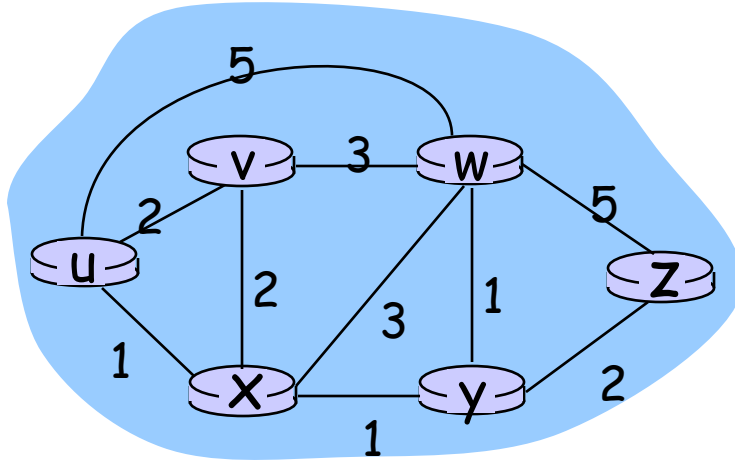
Γράφος: $G = (N,E)$

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: Το αφηρημένο μοντέλο γράφων είναι χρήσιμο και σε άλλα δικτυακά περιβάλλοντα, πχ στο P2P, όπου N είναι το σύνολο των ομότιμων και E είναι το σύνολο των TCP συνδέσεων

Αφηρημένο μοντέλο γράφων: κόστη



- ✓ $c(x,x')$ = κόστος της ζεύξης (x,x')
π.χ., $c(w,z) = 5$
- ✓ το κόστος θα μπορούσε να είναι πάντα 1, ή να σχετίζεται με το εύρος ζώνης, ή με τη συμφόρηση

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

Ερώτημα: Ποιά είναι η ελάχιστου κόστους διαδρομή μεταξύ των u και z ?

Αλγόριθμος δρομολόγησης (routing algorithm): αλγόριθμος που βρίσκει την ελάχιστου κόστους διαδρομή

Κατάταξη αλγορίθμων δρομολόγησης

Ε: Καθολική ή αποκεντρωμένη
πληροφορία;

Καθολική:

- όλοι οι δρομολογητές έχουν τη συνολική τοπολογία και πληροφορίες για το κόστος των ζεύξεων
- αλγόριθμοι κατάστασης ζεύξης ("link state")

Αποκεντρωμένη:

- ο δρομολογητής γνωρίζει τους φυσικά συνδεδεμένους γείτονες, κόστη ζεύξεως προς γείτονες
- επαναληπτική διαδικασία υπολογισμού, ανταλλαγής πληροφορίας με τους γείτονες
- αλγόριθμοι διανύσματος απόστασης ("distance vector")

Στατικός ή δυναμικός;

Στατικός:

- οι διαδρομές αλλάζουν αργά με το χρόνο

Δυναμικός:

- οι διαδρομές αλλάζουν πιο γρήγορα
 - περιοδική ενημέρωση
 - ως απόκριση σε αλλαγές του κόστους των ζεύξεων

Ένας αλγόριθμος κατάστασης ζεύξης

Αλγόριθμος του Dijkstra

- τοπολογία του δικτύου, κόστη ζεύξεων γνωστά σε όλους τους κόμβους
 - επιτυγχάνεται μέσω (ευρυ)εκπομπής (broadcast) κατάστασης ζεύξης
 - όλοι οι κόμβοι έχουν την ίδια πληροφορία
- υπολογίζει τις διαδρομές ελάχιστου κόστους από έναν κόμβο (προέλευση) προς όλους τους άλλους κόμβους
 - δίνει τον **πίνακα δρομολόγησης** γι' αυτόν τον κόμβο
- επαναληπτικός: μετά από k επαναλήψεις, είναι γνωστές οι ελάχιστου κόστους διαδρομές προς k προορισμούς

Συμβολισμοί:

- $c(i,j)$: κόστος ζεύξης από τον κόμβο i στον κόμβο $j = \infty$ αν δεν είναι άμεσοι γείτονες
- $D(v)$: τρέχουσα τιμή του κόστους της διαδρομής από την προέλευση στον προορισμό v
- $p(v)$: προηγούμενος από τον v κόμβος κατά μήκος της διαδρομής από την προέλευση στον v
- N' : σύνολο κόμβων για τους οποίους η ελάχιστου κόστους διαδρομή έχει σαφώς καθοριστεί

Αλγόριθμος του Dijkstra

1 **Initialization:**

2 $N' = \{u\}$

3 for all nodes v

4 if v adjacent to u

5 then $D(v) = c(u,v)$

6 else $D(v) = \infty$

7

8 **Loop**

9 find w not in N' such that $D(w)$ is a minimum

10 add w to N'

11 update $D(v)$ for all v adjacent to w and not in N' :

12 $D(v) = \min(D(v), D(w) + c(w,v))$

13 /* new cost to v is either old cost to v or known

14 shortest path cost to w plus cost from w to v */

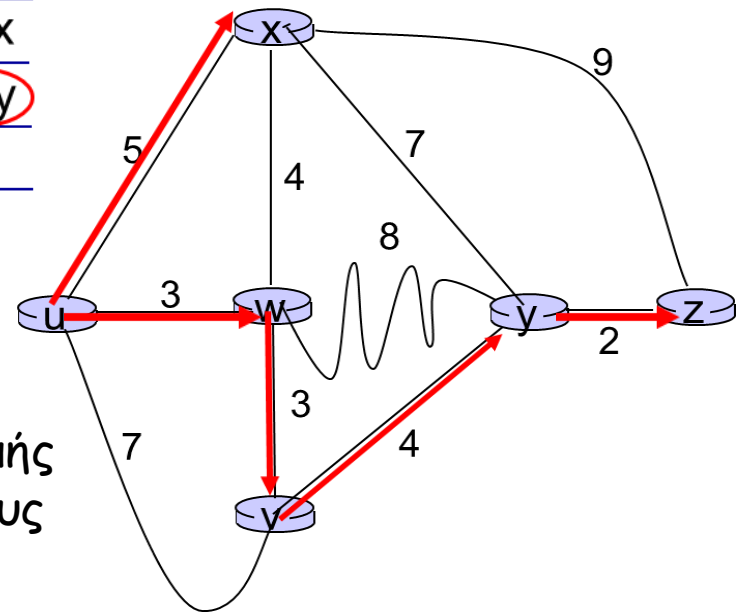
15 **until all nodes in N'**

Αλγόριθμος του Dijkstra: Παράδειγμα

Step	N'	D(v) p(v)	D(w) p(w)	D(x) p(x)	D(y) p(y)	D(z) p(z)
0	u	7,u	3,u	5,u	∞	∞
1	uw	6,w		5,u	11,w	∞
2	uwvx	6,w			11,w	14,x
3	uwxv				10,v	14,x
4	uwxvy				12,y	
5	uwxvyz					

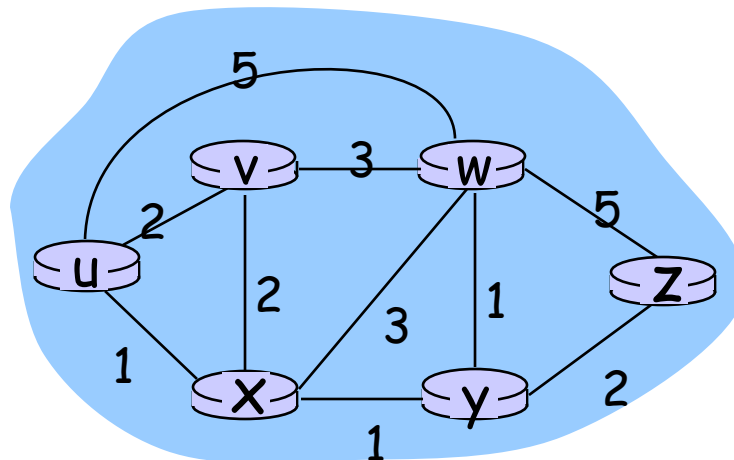
Παρατηρήσεις:

- ❑ Κατασκευή δέντρου συντομότερης διαδρομής εντοπίζοντας τους προηγούμενους κόμβους
- ❑ Μπορεί να υπάρχουν "ισοπαλίες" (σπάνε αυθαίρετα)



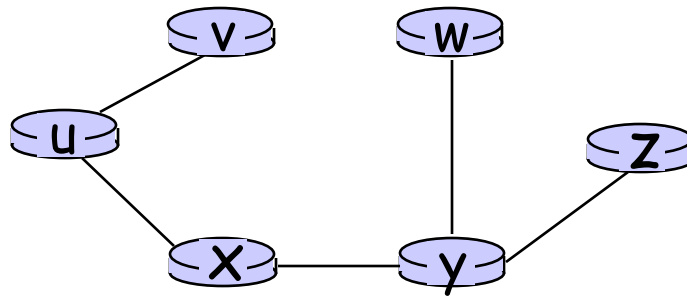
Αλγόριθμος του Dijkstra: άλλο παράδειγμα

Step	N'	D(v),p(v)	D(w),p(w)	D(x),p(x)	D(y),p(y)	D(z),p(z)
0	u	2,u	5,u	1,u	∞	∞
1	ux	2,u	4,x		2,x	∞
2	uxy	2,u	3,y			4,y
3	uxyv		3,y			4,y
4	uxynw					4,y
5	uxynwz					



Αλγόριθμος του Dijkstra: παράδειγμα (2)

Δέντρο ελάχιστων διαδρομών (shortest-path tree) από τον u:



Πίνακας δρομολόγησης στον u:

destination	link
v	(u,v)
x	(u,x)
y	(u,x)
w	(u,x)
z	(u,x)

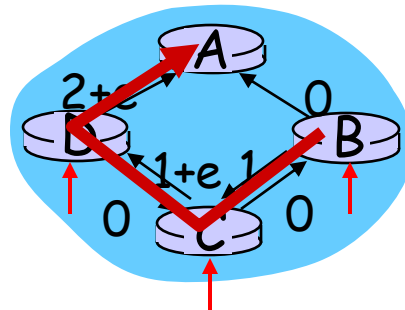
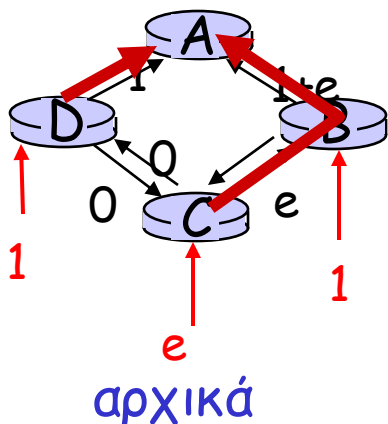
Αλγόριθμος του Dijkstra: συζήτηση

Πολυπλοκότητα του αλγορίθμου: n κόμβοι

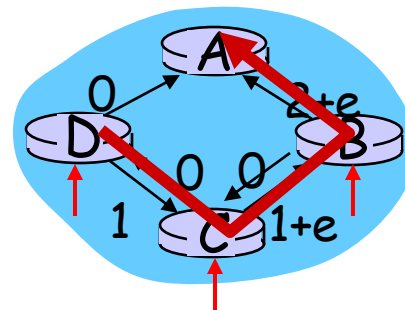
- ❑ κάθε επανάληψη: χρειάζεται να εξετάσει όλους τους κόμβους, w , που δεν ανήκουν στο N
- ❑ $n(n+1)/2$ συγκρίσεις: $O(n^2)$
- ❑ είναι δυνατές πιο αποδοτικές υλοποιήσεις: $O(n \log n)$

Είναι δυνατόν να εμφανιστούν ταλαντώσεις: (κίνηση προς A)

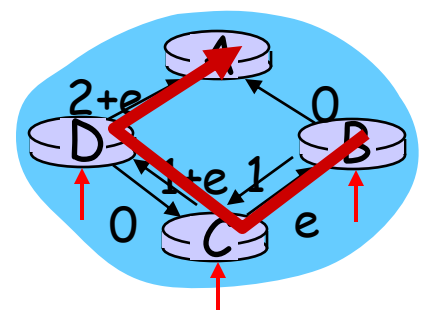
- ❑ π.χ., κόστος ζεύξης = ποσότητα κίνησης που μεταφέρεται



με δεδομένα αυτά τα κόστη, βρες νέα δρομολόγηση... προκύπτουν νέα κόστη



με δεδομένα αυτά τα κόστη, βρες νέα δρομολόγηση... προκύπτουν νέα κόστη



με δεδομένα αυτά τα κόστη, βρες νέα δρομολόγηση... προκύπτουν νέα κόστη

Αλγόριθμος Διανύσματος Απόστασης (Distance Vector)

Εξίσωση Bellman-Ford (δυναμικός προγραμματισμός)

Ορίζουμε

$d_x(y)$ = κόστος της ελάχιστου κόστους διαδρομής
από τον x στον y

Τότε

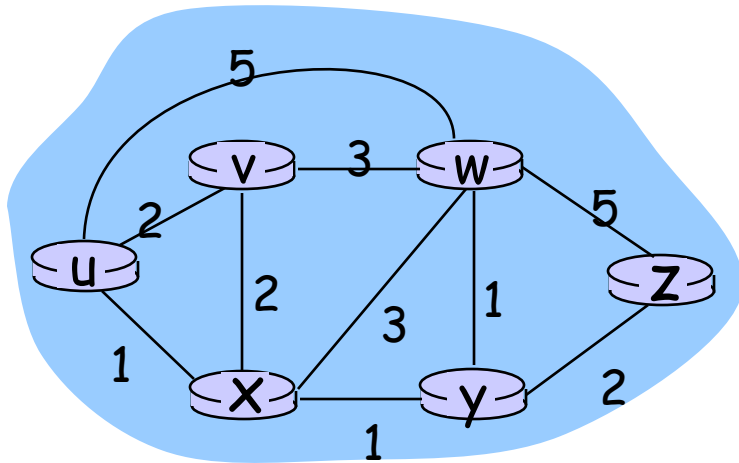
$$d_x(y) = \min_v \{ c(x,v) + d_v(y) \}$$

κόστος από γείτονα v στον προορισμό y

κόστος για τον γείτονα v

το \min λαμβάνεται πάνω σε όλους τους γείτονες v του x

Παράδειγμα Bellman-Ford



Σαφώς, $d_v(z) = 5$, $d_x(z) = 3$, $d_w(z) = 3$

Η εξίσωση B-F λέει:

$$\begin{aligned}d_u(z) &= \min \{ c(u,v) + d_v(z), \\ &\quad c(u,x) + d_x(z), \\ &\quad c(u,w) + d_w(z) \} \\ &= \min \{ 2 + 5, \\ &\quad 1 + 3, \\ &\quad 5 + 3 \} = 4\end{aligned}$$

Ο κόμβος που επιτυγχάνει το ελάχιστο είναι το επόμενο άλμα (hop) στη βραχύτερη διαδρομή, χρησιμοποιείται στον πίνακα προώθησης

Αλγόριθμος Διανύσματος Απόστασης

- $D_x(y)$ = εκτίμηση του ελάχιστου κόστους από τον x στον y
 - Ο κόμβος x διατηρεί το διάνυσμα απόστασης
 $D_x = [D_x(y): y \in N]$
- Ο κόμβος x :
 - γνωρίζει το κόστος προς κάθε γείτονα v : $c(x,v)$
 - Διατηρεί το διάνυσμα απόστασης των γειτόνων του. Για κάθε γείτονα v , ο κόμβος x διατηρεί
 $D_v = [D_v(y): y \in N]$

Αλγόριθμος διανύσματος απόστασης

Βασική ιδέα:

- ❑ Τακτικά, κάθε κόμβος στέλνει το δικό του διάνυσμα απόστασης στους γείτονές του
- ❑ Όταν ο κόμβος x λαμβάνει νέα εκτίμηση DV από γείτονα, ενημερώνει το δικό του DV με χρήση της B-F εξίσωσης:

$$D_x(y) \leftarrow \min_v \{c(x,v) + D_v(y)\} \quad \text{για κάθε κόμβο } y \in N$$

- ❑ Υπό φυσιολογικές συνθήκες, η εκτίμηση $D_x(y)$ συγκλίνει στο πραγματικό ελάχιστο κόστος $d_x(y)$

Αλγόριθμος Διανύσματος Απόστασης

Επαναληπτικός,
ασύγχρονος:

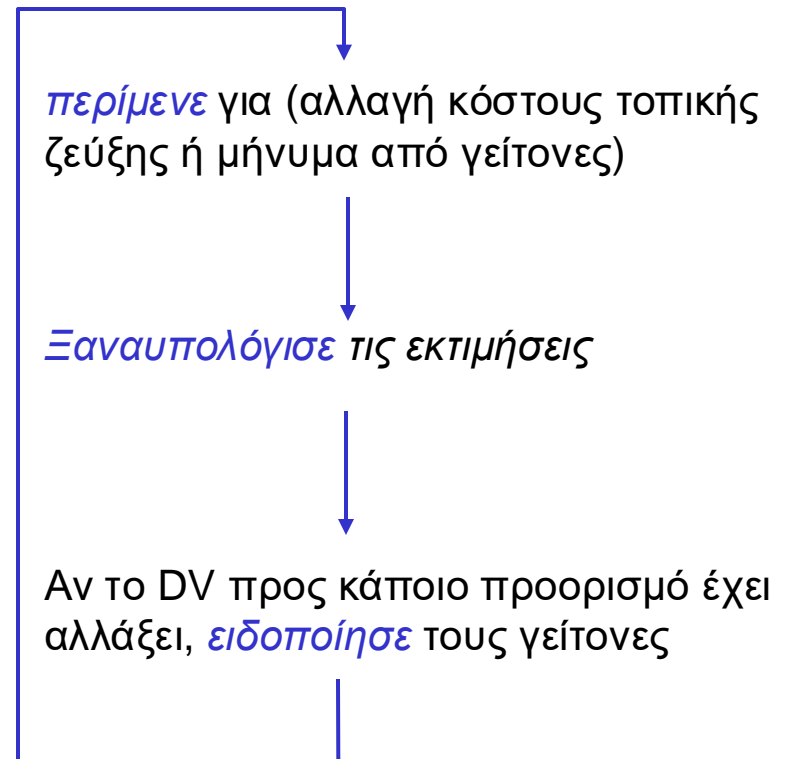
κάθε τοπική επανάληψη
προκαλείται από

- αλλαγή κόστους τοπικής ζεύξης
- μήνυμα ενημέρωσης DV από γείτονα

Κατανεμημένος:

- Κάθε κόμβος ειδοποιεί τους γείτονες μόνο όταν το DV του αλλάζει
 - Οι γείτονες τότε ειδοποιούν τους γείτονές τους αν χρειάζεται

Κάθε κόμβος:



$$D_x(y) = \min\{c(x,y) + D_y(y), c(x,z) + D_z(y)\} \\ = \min\{2+0, 7+1\} = 2$$

$$D_x(z) = \min\{c(x,y) + D_y(z), c(x,z) + D_z(z)\} \\ = \min\{2+1, 7+0\} = 3$$

node x table

		cost to		
		x	y	z
from	x	0	2	7
	y	∞	∞	∞
	z	∞	∞	∞

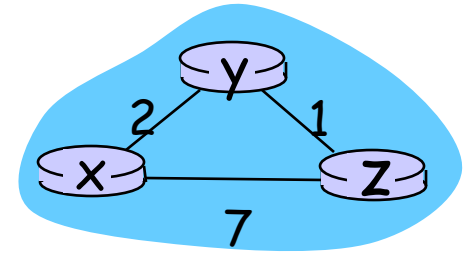
		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	7	1	0

node y table

		cost to		
		x	y	z
from	x	∞	∞	∞
	y	2	0	1
	z	∞	∞	∞

node z table

		cost to		
		x	y	z
from	x	∞	∞	∞
	y	∞	∞	∞
	z	7	1	0



► time

$$D_x(y) = \min\{c(x,y) + D_y(y), c(x,z) + D_z(y)\} \\ = \min\{2+0, 7+1\} = 2$$

$$D_x(z) = \min\{c(x,y) + D_y(z), c(x,z) + D_z(z)\} \\ = \min\{2+1, 7+0\} = 3$$

node x table

		cost to		
		x	y	z
from	x	0	2	7
	y	∞	∞	∞
	z	∞	∞	∞

		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	7	1	0

		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	3	1	0

node y table

		cost to		
		x	y	z
from	x	∞	∞	∞
	y	2	0	1
	z	∞	∞	∞

		cost to		
		x	y	z
from	x	0	2	7
	y	2	0	1
	z	7	1	0

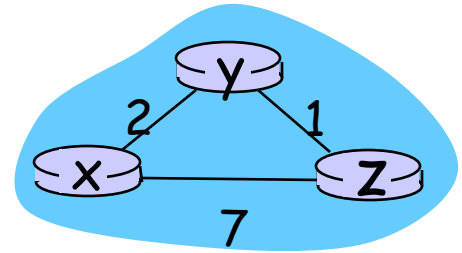
		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	3	1	0

node z table

		cost to		
		x	y	z
from	x	∞	∞	∞
	y	∞	∞	∞
	z	7	1	0

		cost to		
		x	y	z
from	x	0	2	7
	y	2	0	1
	z	3	1	0

		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	3	1	0



▶ time

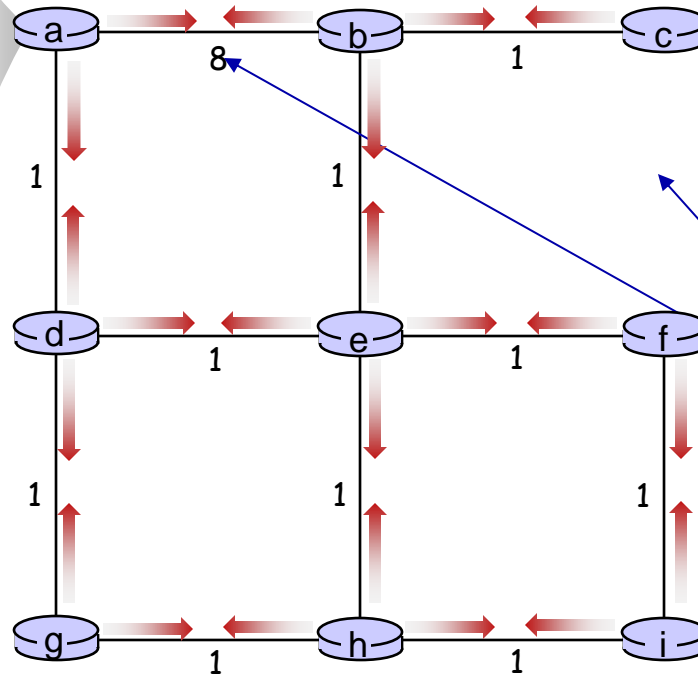
Distance vector: example



t=0

- All nodes have distance estimates to nearest neighbors (only)
- All nodes send their local distance vector to their neighbors

DV in
$D_a(a)=0$
$D_a(b) = 8$
$D_a(c) = \infty$
$D_a(d) = 1$
$D_a(e) = \infty$
$D_a(f) = \infty$
$D_a(g) = \infty$
$D_a(h) = \infty$
$D_a(i) = \infty$



A few asymmetries:

- missing link
- larger cost

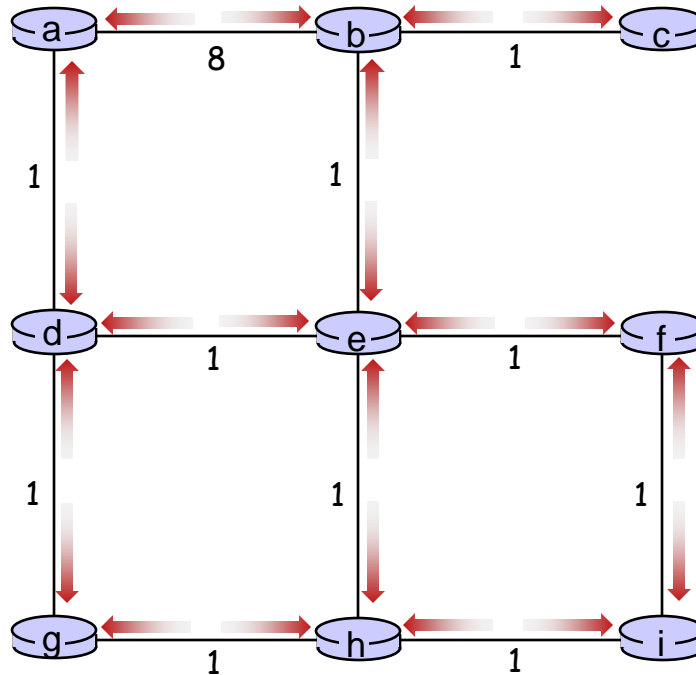
Distance vector example: iteration



$t=1$

All nodes:

- receive distance vectors from neighbors
- compute their new local distance vector
- send their new local distance vector to neighbors



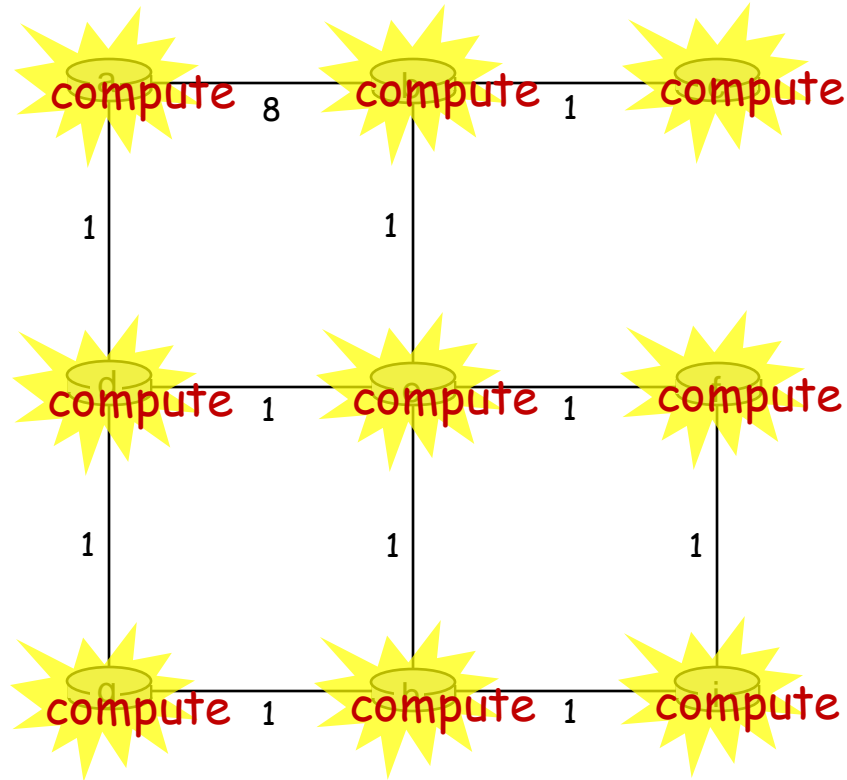
Distance vector example: iteration



$t=1$

All nodes:

- receive distance vectors from neighbors
- compute their new local distance vector
- send their new local distance vector to neighbors



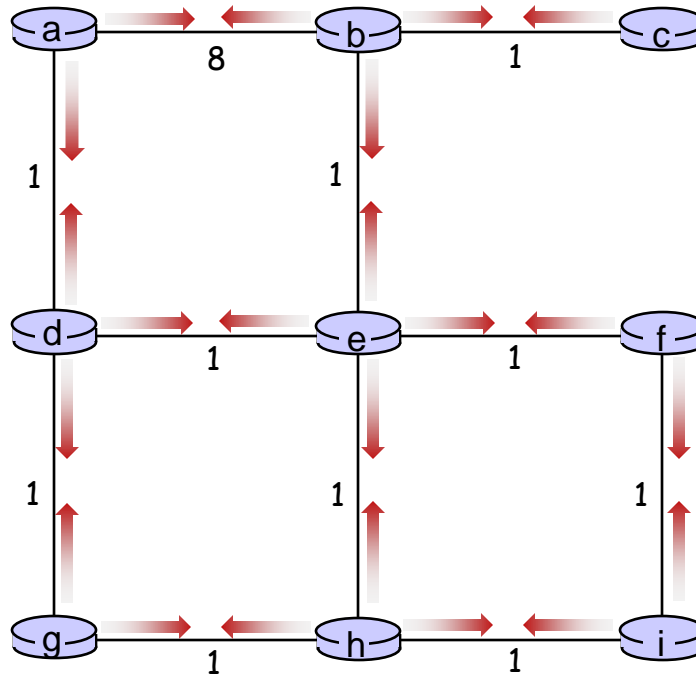
Distance vector example: iteration



$t=1$

All nodes:

- receive distance vectors from neighbors
- compute their new local distance vector
- send their new local distance vector to neighbors



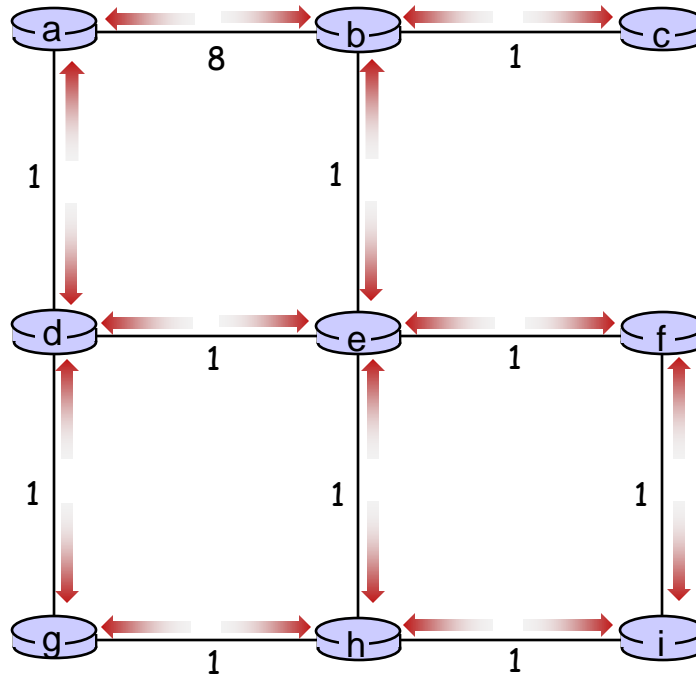
Distance vector example: iteration



t=2

All nodes:

- receive distance vectors from neighbors
- compute their new local distance vector
- send their new local distance vector to neighbors



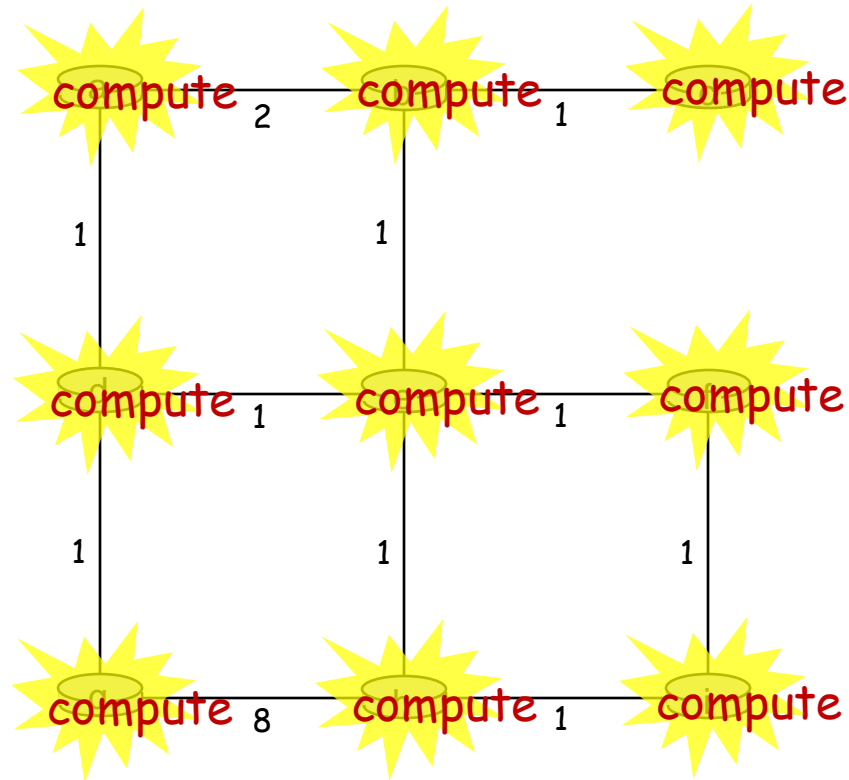
Distance vector example: iteration



t=2

All nodes:

- receive distance vectors from neighbors
- compute their new local distance vector
- send their new local distance vector to neighbors



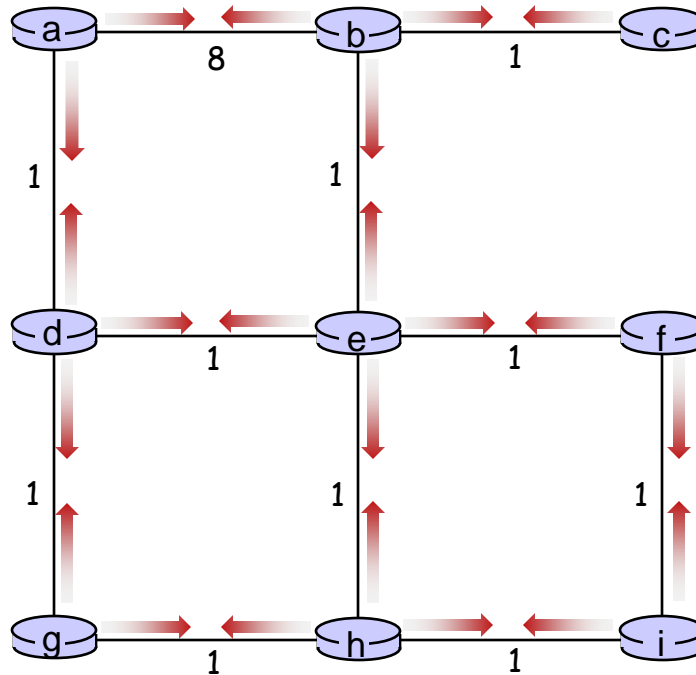
Distance vector example: iteration



t=2






All nodes:

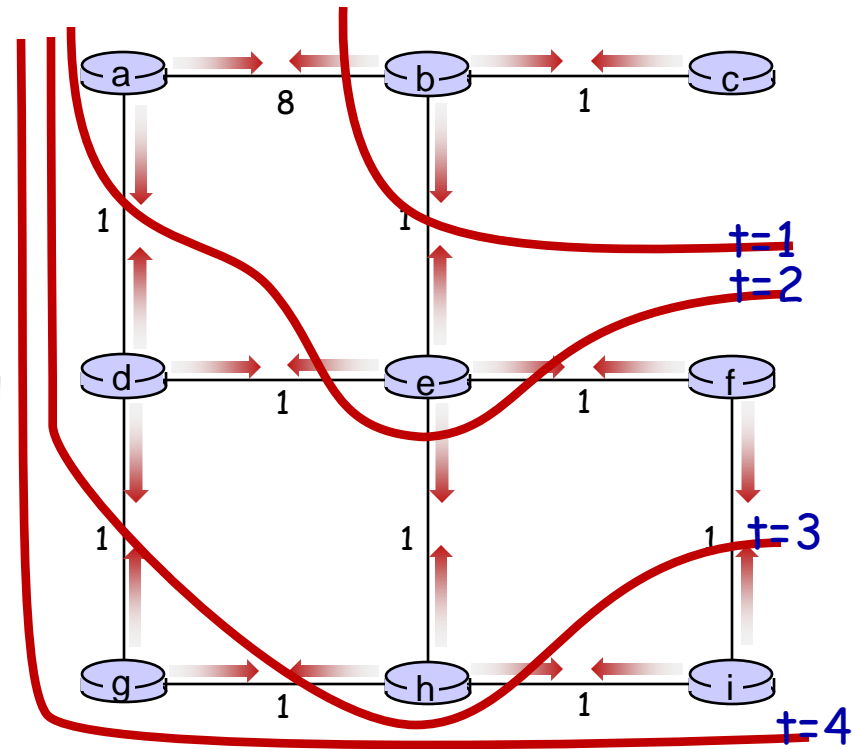
- receive distance vectors from neighbors
- compute their new local distance vector
- send their new local distance vector to neighbors



Distance vector: state information diffusion

Iterative communication, computation steps diffuses information through network:

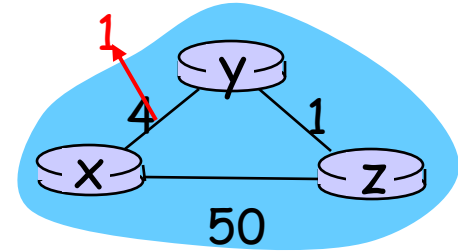
-  $t=0$ c's state at $t=0$ is at c only
-  $t=1$ c's state at $t=0$ has propagated to b, and may influence distance vector computations up to **1** hop away, i.e., at b
-  $t=2$ c's state at $t=0$ may now influence distance vector computations up to **2** hops away, i.e., at b and now at a, e as well
-  $t=3$ c's state at $t=0$ may influence distance vector computations up to **3** hops away, i.e., at b,a,e and now at d,f,h as well
-  $t=4$ c's state at $t=0$ may influence distance vector computations up to **4** hops away, i.e., at b,a,e, d, f, h and now at g,i as well



Διάνυσμα απόστασης: αλλαγές κόστους ζεύξης

Αλλαγές κόστους ζεύξης:

- Ο κόμβος ανιχνεύει αλλαγή κόστους τοπικής ζεύξης
- Ενημερώνει την πληροφορία δρομολόγησης, ξανα-υπολογίζει το διάνυσμα απόστασης
- Αν το DV αλλάζει, ειδοποιεί τους γείτονες



Για δρομολόγηση προς τον x (αφορά τους κόμβους y και z) :

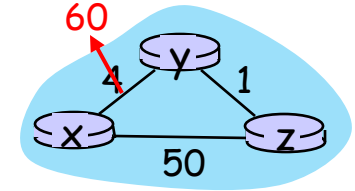
Σε χρόνο t_0 , ο y ανιχνεύει την αλλαγή στο κόστος της ζεύξης, ενημερώνει το DV του, και ενημερώνει τους γείτονές του

“τα καλά νέα ταξιδεύουν γρήγορα”

Σε χρόνο t_1 , ο z δέχεται την ενημέρωση από τον y και ενημερώνει τον πίνακά του. Υπολογίζει ένα νέο ελάχιστο κόστος προς τον x και στέλνει στους γείτονές του το DV του

Σε χρόνο t_2 , ο y δέχεται την ενημέρωση του z και ενημερώνει τον πίνακα απόστασης. Τα ελάχιστα κόστη του y δεν αλλάζουν και έτσι ο y δεν στέλνει κανένα μήνυμα στον z

Distance vector: link cost changes



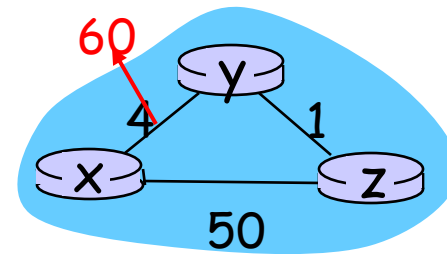
link cost changes:

- node detects local link cost change
- "bad news travels slow" - count-to-infinity problem:
 - y sees direct link to x has new cost 60, but z has said it has a path at cost of 5. So y computes "my new cost to x will be 6, via z); notifies z of new cost of 6 to x.
 - z learns that path to x via y has new cost 6, so z computes "my new cost to x will be 7 via y), notifies y of new cost of 7 to x.
 - y learns that path to x via z has new cost 7, so y computes "my new cost to x will be 8 via y), notifies z of new cost of 8 to x.
 - z learns that path to x via y has new cost 8, so z computes "my new cost to x will be 9 via y), notifies y of new cost of 9 to x.
 - ...
- see text for solutions. *Distributed algorithms are tricky!*

Διάνυσμα απόστασης: αλλαγές κόστους ζεύξης

Αλλαγές κόστους ζεύξης:

- ❑ Ο κόμβος ανιχνεύει αλλαγή κόστους τοπικής ζεύξης
- ❑ Τα κακά νέα ταξιδεύουν αργά - πρόβλημα "μέτρησης μέχρι το άπειρο"!
- ❑ 44 επαναλήψεις πριν σταθεροποιηθεί ο αλγόριθμος (γιατί;)



"poisoned reverse":

- ❑ Αν ο Z δρομολογεί μέσω του Y για να φτάσει στον X
 - Ο Z λέει στον Y ότι η απόστασή του από τον X είναι άπειρη (ώστε ο Y να μη δρομολογεί στον X μέσω του Z)

Σύγκριση των LS και DV αλγόριθμων

Πολυπλοκότητα μηνύματος

- ❑ **LS:** με n κόμβους, E ζεύξεις, στέλνονται $O(nE)$ μηνύματα
- ❑ **DV:** ανταλλαγή μόνο μεταξύ γειτόνων
 - ο χρόνος σύγκλισης ποικίλει

Ταχύτητα σύγκλισης

- ❑ **LS:** $O(n^2)$ ο αλγόριθμος απαιτεί $O(nE)$ μηνύματα
 - μπορεί να έχει ταλαντώσεις
- ❑ **DV:** ο χρόνος σύγκλισης ποικίλει
 - μπορεί να υπάρχουν βρόχοι δρομολόγησης
 - πρόβλημα μέτρησης μέχρι το άπειρο

Ευρωστία: τι συμβαίνει αν ένας δρομολογητής δυσλειτουργεί;

LS:

- ο κόμβος μπορεί να εκπέμψει λάθος κόστος **ζεύξης**
- κάθε κόμβος υπολογίζει μόνο το δικό του πίνακα

DV:

- ο DV κόμβος μπορεί να εκπέμψει εσφαλμένο κόστος διαδρομής
- ο πίνακας κάθε κόμβου χρησιμοποιείται από άλλους
- τα λάθη διαδίδονται μέσω του δικτύου

Ιεραρχική Δρομολόγηση

Η μελέτη της δρομολόγησης ως τώρα εξιδανικευμένη:

- όλοι οι δρομολογητές πανομοιότυποι
 - «επίπεδο» (flat) δίκτυο
- ... δεν ισχύει στην πράξη

Κλίμακα: με 600 εκατομμύρια προορισμούς

- ❑ δεν μπορεί να αποθηκευτούν όλοι οι προορισμοί στους πίνακες δρομολόγησης!
- ❑ η ανταλλαγή των πινάκων δρομολόγησης θα κατάκλυζε τις ζεύξεις

Διαχειριστική αυτονομία

- ❑ Διαδίκτυο = δίκτυο δικτύων
- ❑ κάθε διαχειριστής δικτύου ενδέχεται να θέλει να ελέγχει τη δρομολόγηση στο δικό του δίκτυο

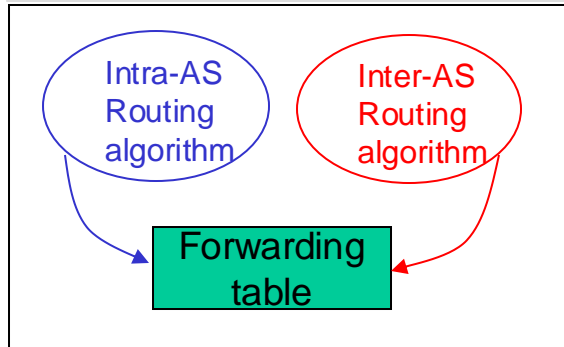
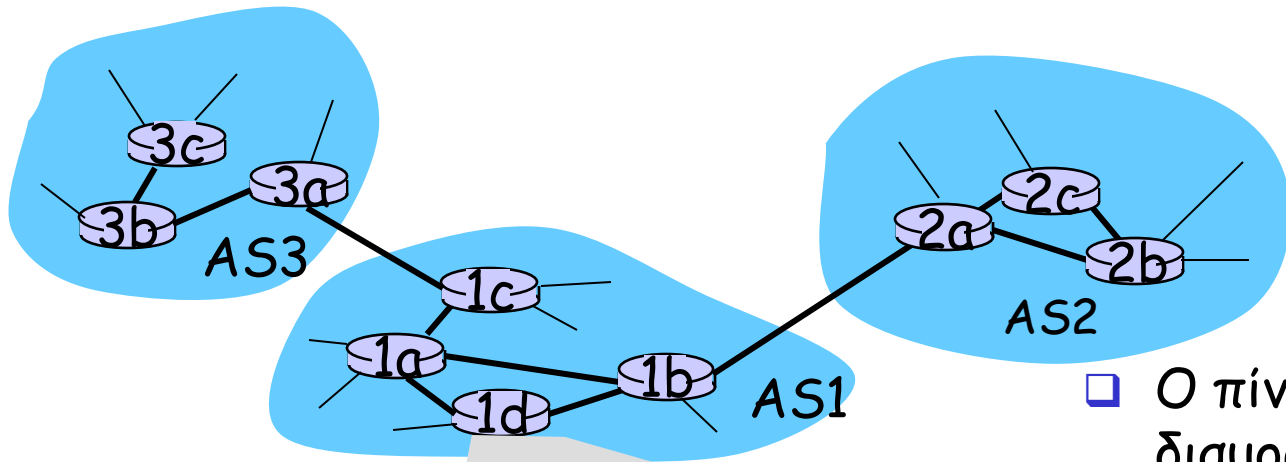
Ιεραρχική Δρομολόγηση

- ομαδοποίηση δρομολογητών σε περιοχές, «αυτόνομα συστήματα» ("autonomous systems" (AS))
- δρομολογητές του ίδιου AS τρέχουν το ίδιο πρωτόκολλο δρομολόγησης
 - πρωτόκολλο δρομολόγησης "intra-AS" [πρωτόκολλο δρομολόγησης ενδοαυτόνομου συστήματος]
 - δρομολογητές σε διαφορετικά AS μπορούν να τρέχουν διαφορετικά intra-AS πρωτόκολλα δρομολόγησης

Δρομολογητής πύλης (Gateway router)

- Στην "άκρη" του δικού του AS
- Έχει ζεύξη με δρομολογητή σε άλλο AS

Διασυνδεδεμένα AS



- Ο πίνακας προώθησης διαμορφώνεται από intra- (ενδο-) και inter- (δια-) AS αλγόριθμους δρομολόγησης
 - Οι intra-AS αλγόριθμοι ορίζουν καταχωρίσεις για εσωτερικούς προορισμούς
 - Τόσο οι inter-AS, όσο και οι intra-AS αλγόριθμοι ορίζουν καταχωρίσεις για εξωτερικούς προορισμούς

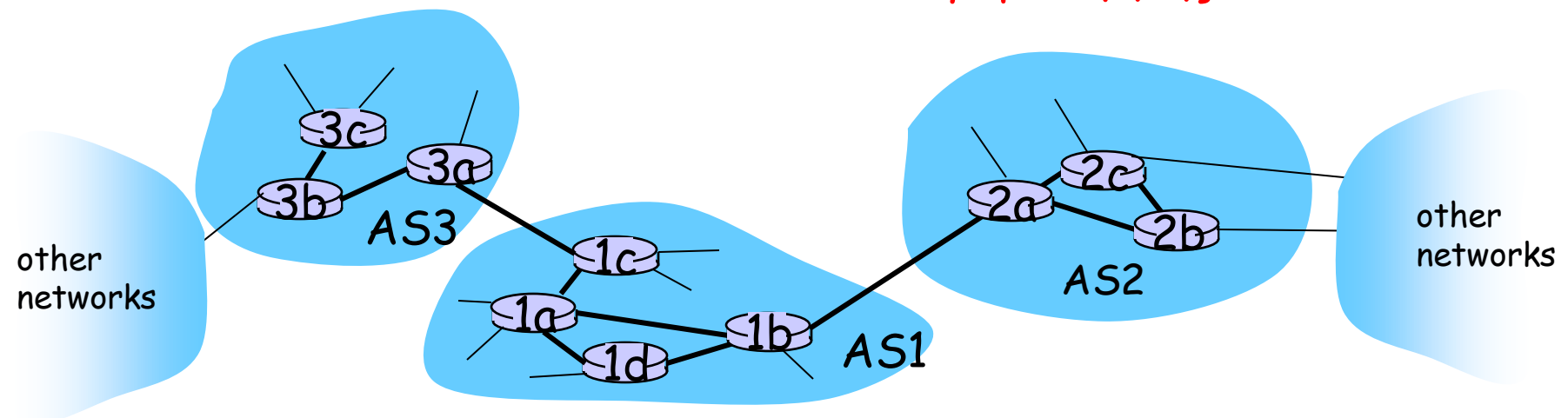
Εργασίες Inter-AS

- Έστω ότι δρομολογητής στο AS1 λαμβάνει datagram με προορισμό εκτός του AS1:
 - Ο δρομολογητής θα έπρεπε να προωθήσει το πακέτο σε δρομολογητή πύλης (gateway router), αλλά σε ποιόν;

Το AS1 πρέπει:

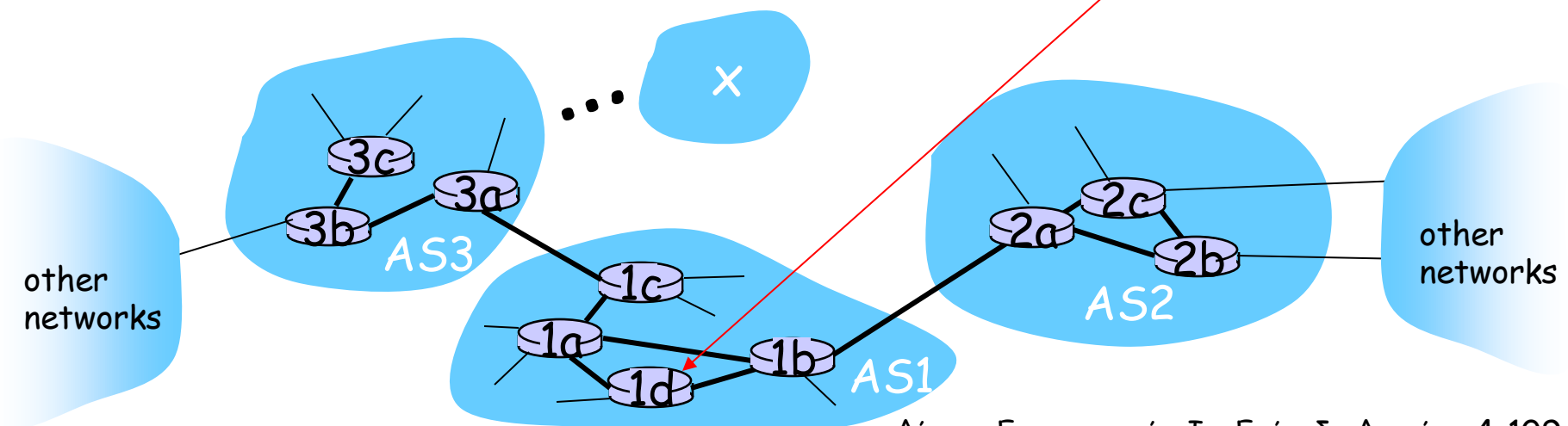
1. Να μάθει ποιοί προορισμοί είναι προσεγγίσιμοι μέσω του AS2 και ποιοί μέσω του AS3
2. Να διαδώσει την πληροφορία προσέγγισης σε όλους τους δρομολογητές στο AS1

Δουλειά της inter-AS δρομολόγησης!



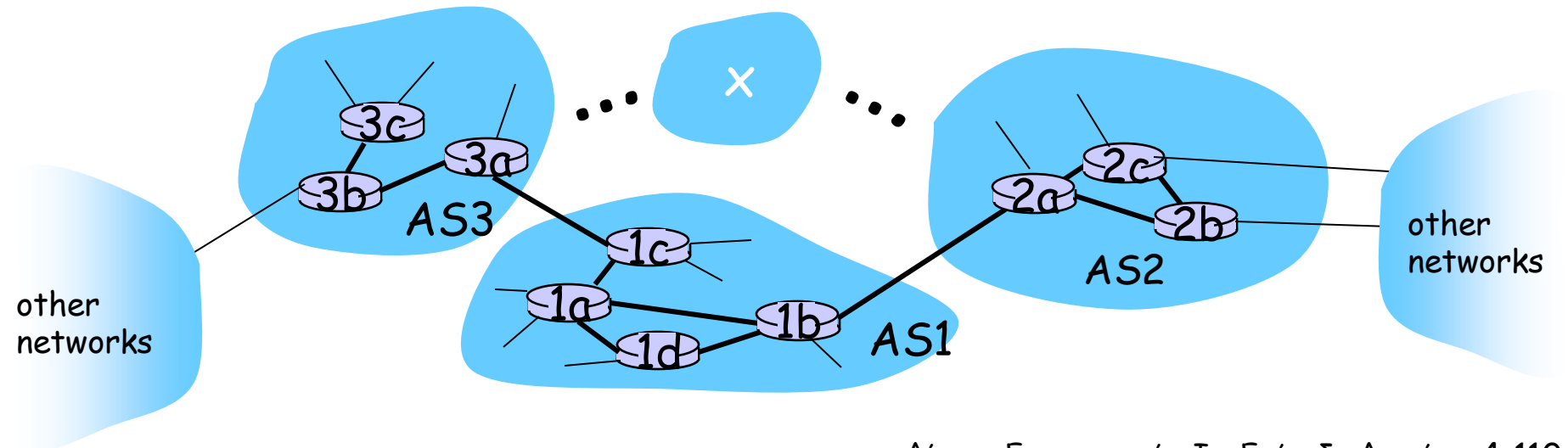
Παράδειγμα: Καθορισμός του πίνακα προώθησης στο δρομολογητή 1d

- Έστω ότι το AS1 μαθαίνει (μέσω πρωτοκόλλου inter-AS) ότι το υποδίκτυο x είναι προσεγγίσιμο μέσω του AS3 (δρομολογητής πύλης 1c) αλλά όχι μέσω του AS2.
 - Το πρωτόκολλο inter-AS διαδίδει την πληροφορία προσέγγισης σε όλους τους εσωτερικούς δρομολογητές
- Ο δρομολογητής 1d καθορίζει μέσω της intra-AS πληροφορίας δρομολόγησης ότι η διεπαφή του I είναι στη διαδρομή ελάχιστου κόστους προς το 1c.
 - Εισάγει καταχώριση στον πίνακα προώθησης (x,I)



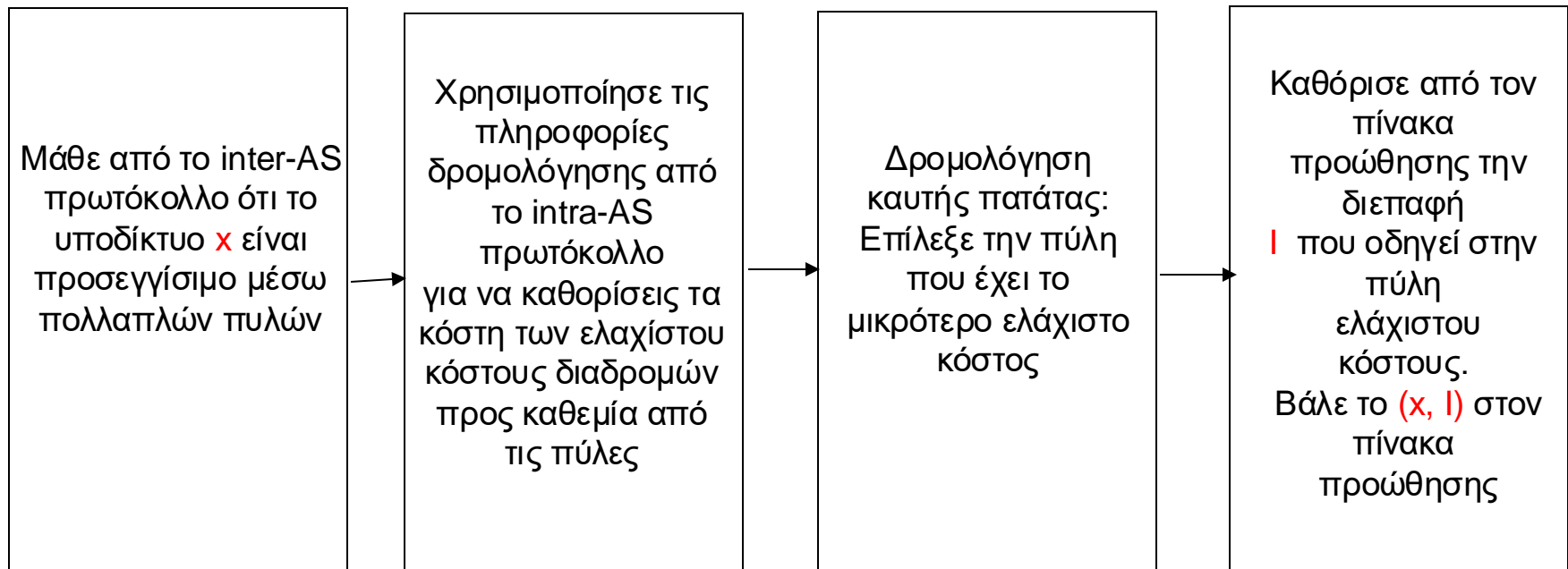
Παράδειγμα: Διαλέγοντας μεταξύ πολλαπλών AS

- Έστω τώρα ότι το AS1 μαθαίνει από το inter-AS πρωτόκολλο ότι το υποδίκτυο **X** είναι προσεγγίσιμο από το AS3 **και** από το AS2
- Για τη διαμόρφωση του πίνακα προώθησης, ο δρομολογητής 1d πρέπει να καθορίσει προς ποιόν δρομολογητή πύλης θα πρέπει να προωθεί τα πακέτα για τον προορισμό **X**
 - Αυτό είναι επίσης δουλειά του inter-AS πρωτοκόλλου δρομολόγησης!



Παράδειγμα: Διαλέγοντας μεταξύ πολλαπλών AS

- **Δρομολόγηση καυτής πατάτας (hot potato routing):** στείλε το πακέτο στον πιο κοντινό από τους δύο δρομολογητές



Κεφάλαιο 4: Επίπεδο Δικτύου

4.1 Εισαγωγή

4.2 Δίκτυα εικονικού κυκλώματος και δεδομενογράμματος

4.3 Τι βρίσκεται μέσα σ' ένα δρομολογητή

4.4 IP: Πρωτόκολλο Διαδικτύου (Internet Protocol)

- ❑ Μορφή δεδομενογράμματος
- ❑ Διευθυνσιοδότηση IPv4
- ❑ ICMP
- ❑ IPv6

4.5 Αλγόριθμοι δρομολόγησης

- ❑ Κατάστασης ζεύξης (Link State)
- ❑ Διανύσματος απόστασης (Distance Vector)
- ❑ Ιεραρχική δρομολόγηση

4.6 Δρομολόγηση στο Διαδίκτυο

- ❑ RIP
- ❑ OSPF
- ❑ BGP

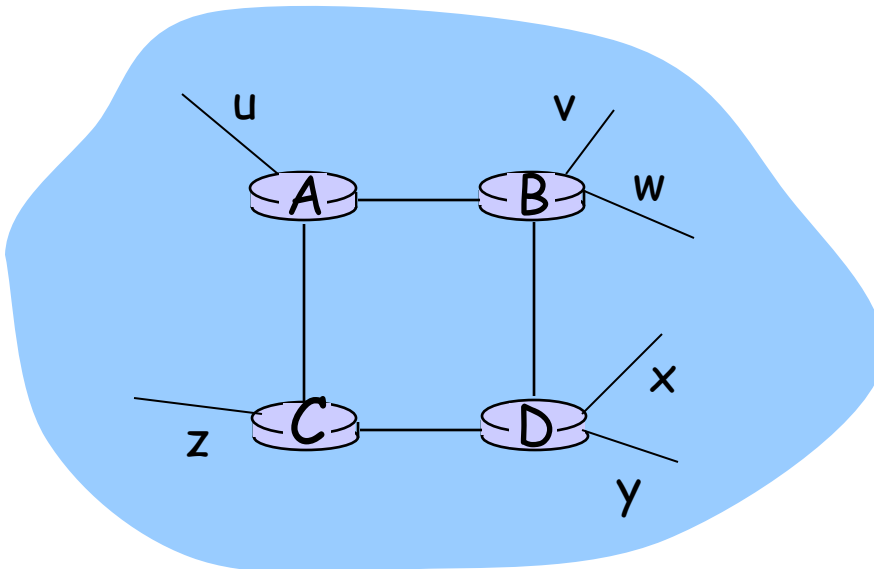
Δρομολόγηση intra-AS

- ❑ Γνωστά και ως **Interior Gateway Protocols (IGP)**
(πρωτόκολλα εσωτερικής πύλης)

- ❑ Τα πιο κοινά Intra-AS πρωτόκολλα δρομολόγησης :
 - RIP: Routing Information Protocol
 - OSPF: Open Shortest Path First
 - IGRP: Interior Gateway Routing Protocol
(ιδιοταγές της Cisco)

RIP (Routing Information Protocol)

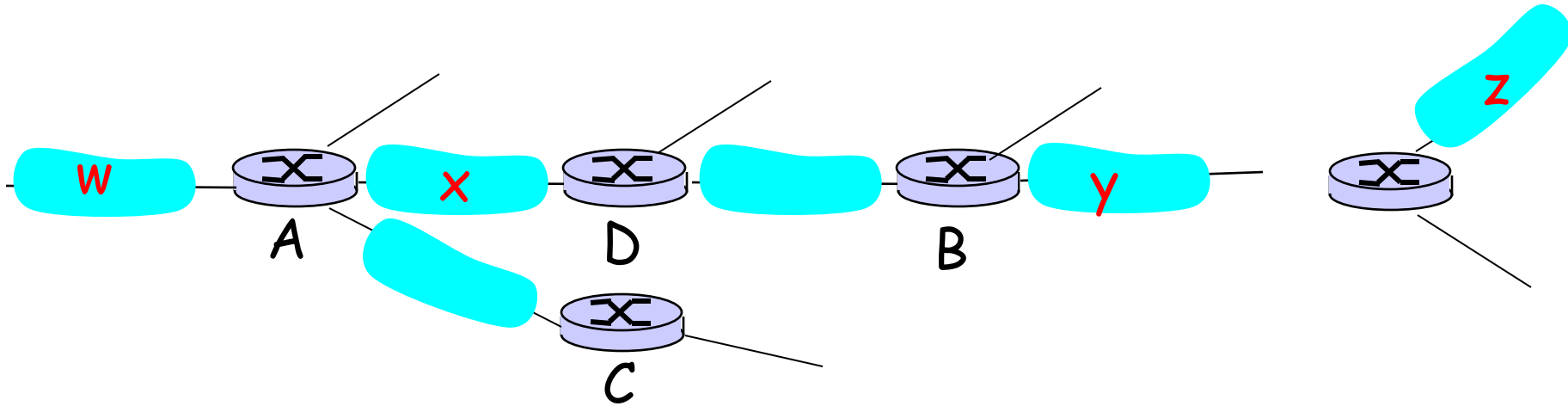
- ❑ Περιελήφθη στη διανομή BSD-UNIX το 1982
- ❑ Αλγόριθμος διανύσματος απόστασης
 - Μετρική απόστασης: # αλμάτων (max = 15 hops), κάθε ζεύξη έχει κόστος 1
 - Ανταλλάσσει DVs με τους γείτονες κάθε 30 sec σε απαντητικό μήνυμα (**δημοσιοποίηση**)
 - Κάθε δημοσιοποίηση: λίστα μέχρι και 25 υποδικτύα προορισμοί (υπό την έννοια της IP διευθυνσιοδότησης)



Από τον δρομολογητή A στα υποδίκτυα προορισμού:

<u>Προορισμός</u>	<u>Άλματα</u>
u	1
v	2
w	2
x	3
y	3
z	2

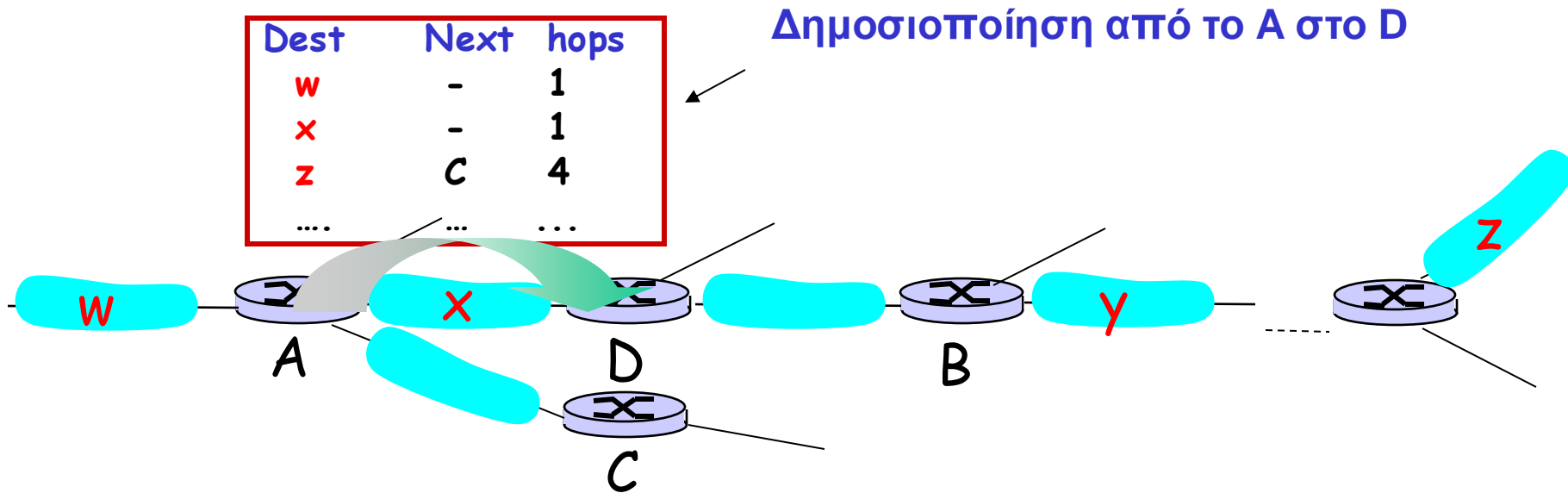
RIP: Παράδειγμα



Υποδίκτιο προορισμού	Επόμενος δρομολογητής	Πλήθος αλμάτων για προορισμό
w	A	2
y	B	2
z	B	7
x	--	1
....

Πίνακας δρομολόγησης στο δρομολογητή D

RIP: Παράδειγμα



Υποδίκτιο προορισμού	Επόμενος δρομολογητής	Πλήθος αλμάτων στον προορισμό
w	A	2
y	B	2
z	B A	7 5
x	--	1
....

Πίνακας δρομολόγησης στο δρομολογητή D

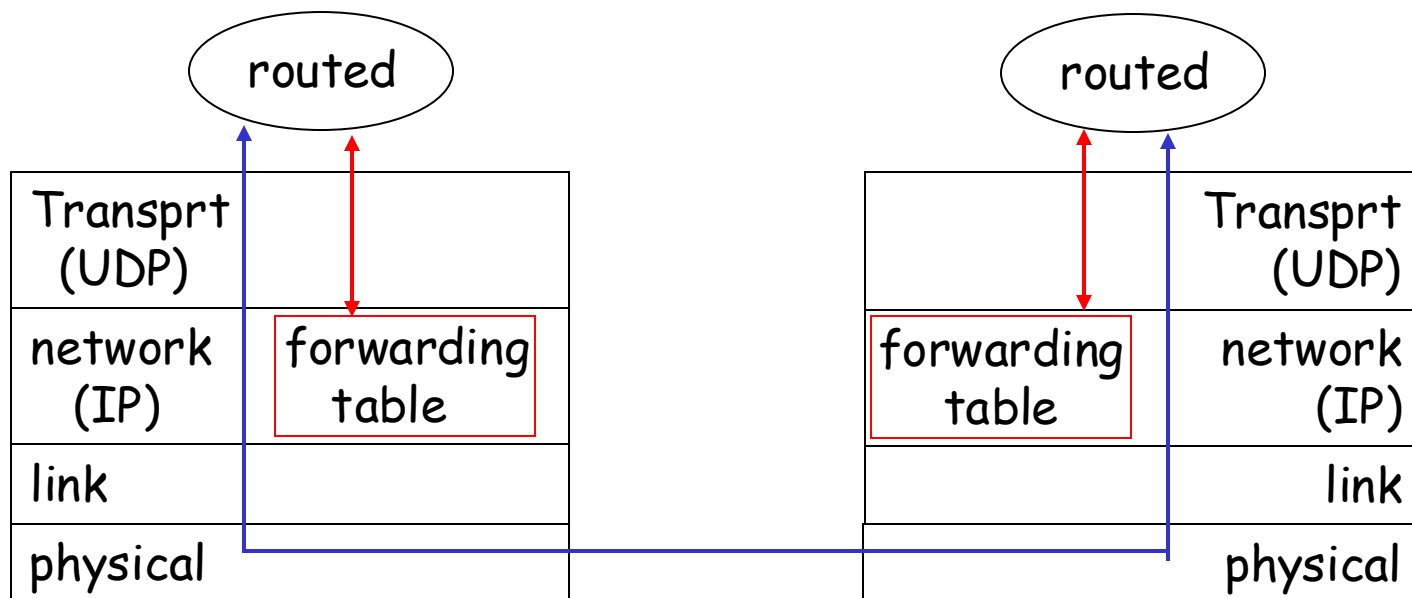
RIP: Αποτυχία ζεύξης και ανάνηψη

Αν δεν ακουστεί διαφήμιση μετά από 180 sec --> ο γείτονας/η ζεύξη θεωρείται νεκρός

- Οι διαδρομές μέσω του γείτονα παύουν να είναι έγκυρες
- Νέες διαφημίσεις στέλνονται στους γείτονες
- Οι γείτονες με τη σειρά τους στέλνουν νέες διαφημίσεις (αν άλλαξαν οι πίνακες)
- Η πληροφορία αποτυχίας της ζεύξης διαδίδεται γρήγορα (:) σε ολόκληρο το δίκτυο
- **Χρησιμοποιείται *poison reverse*** για την αποφυγή βρόχων ring-rong (άπειρη απόσταση = 16 άλματα)

Επεξεργασία πίνακα RIP

- Η διαχείριση των πινάκων δρομολόγησης του RIP γίνεται από διεργασία **επιπέδου εφαρμογής** που ονομάζεται route-d (daemon)
- Οι διαφημίσεις στέλνονται σε πακέτα UDP, επαναλαμβάνονται περιοδικά



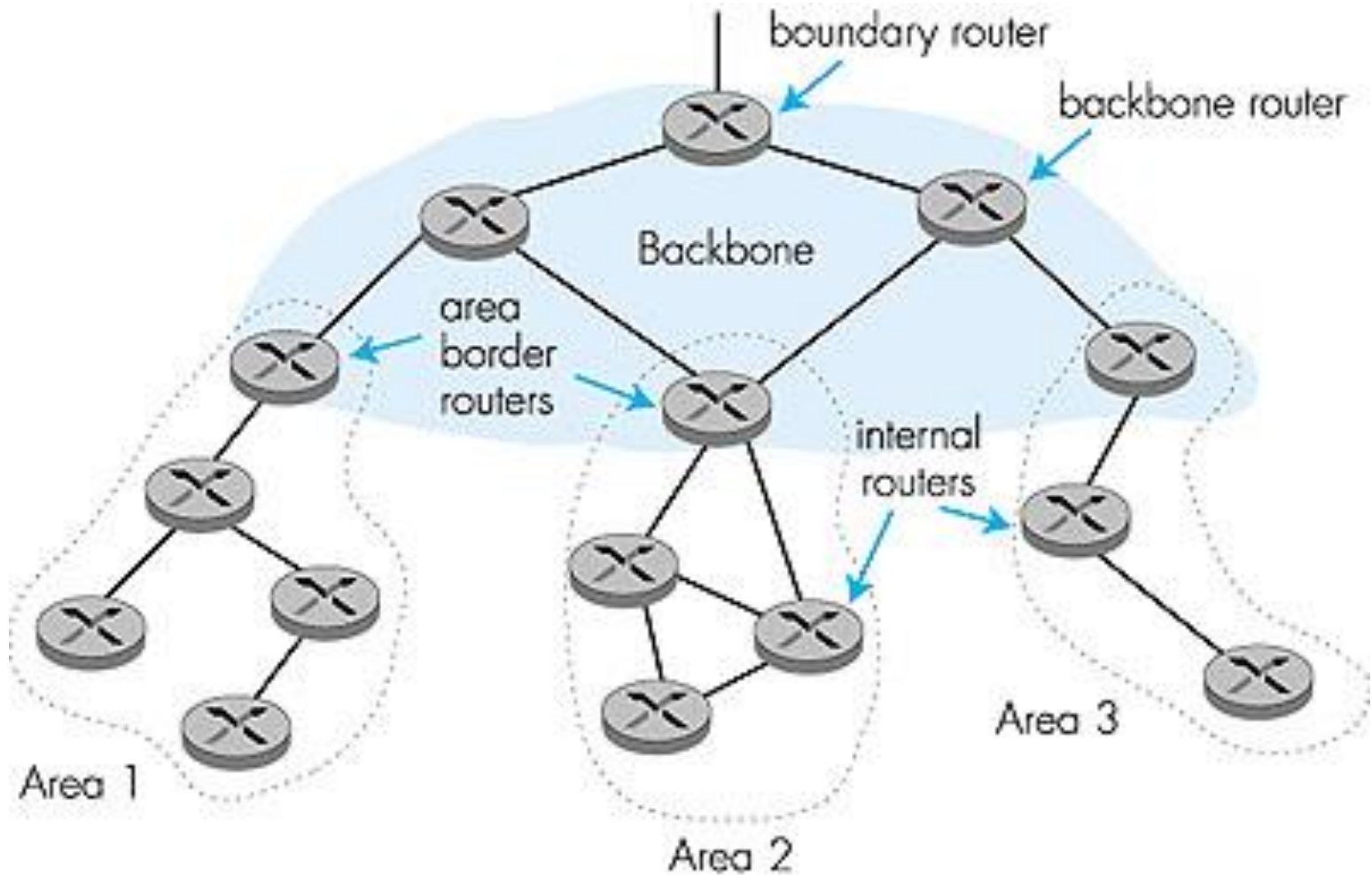
OSPF (Open Shortest Path First)

- ❑ “Ανοικτό”: δημόσια διαθέσιμη προδιαγραφή
- ❑ Χρησιμοποιεί αλγόριθμο κατάστασης ζεύξης (link state)
 - Διάδοση πακέτου LS
 - Χάρτης τοπολογίας σε κάθε κόμβο
 - Υπολογισμός διαδρομής με χρήση του αλγορίθμου του Dijkstra
- ❑ Η δημοσιοποίηση του OSPF μεταφέρει μόνο μία καταχώριση για κάθε γειτονικό δρομολογητή
- ❑ Οι δημοσιοποιήσεις διαδίδονται σε **ολόκληρο** το AS (μέσω πλημμύρας (flooding))
 - Μεταφέρονται σε μηνύματα του OSPF απευθείας επάνω στο IP (αντί για TCP ή UDP)
- ❑ **Πρωτόκολλο δρομολόγησης IS-IS: σχεδόν πανομοιότυπο με το OSPF**

«Προηγμένα» χαρακτηριστικά του OSPF (όχι στο RIP)

- ❑ **Ασφάλεια:** όλα τα μηνύματα του OSPF αυθεντικοποιούνται (για αποφυγή κακόβουλων εισβολών)
- ❑ Επιτρέπει **πολλαπλές διαδρομές** ίδιου κόστους (μόνο μία διαδρομή στο RIP)
- ❑ Για κάθε ζεύξη, πολλαπλές μετρικές κόστους για διαφορετικό **TOS** (π.χ., κόστος δορυφορικής ζεύξης θεωρείται «χαμηλό» για TOS βέλτιστης προσπάθειας, υψηλό για TOS πραγματικού χρόνου)
- ❑ Ενσωματωμένη υποστήριξη μονο- και πολυ-εκπομπής (unicast/multicast):
 - Το Multicast OSPF (MOSPF) (OSPF πολυ-εκπομπής) χρησιμοποιεί την ίδια βάση δεδομένων τοπολογίας με το OSPF
- ❑ **ιεραρχικό** OSPF σε μεγάλους τομείς (domains)

Ιεραρχικό OSPF



Ιεραρχικό OSPF

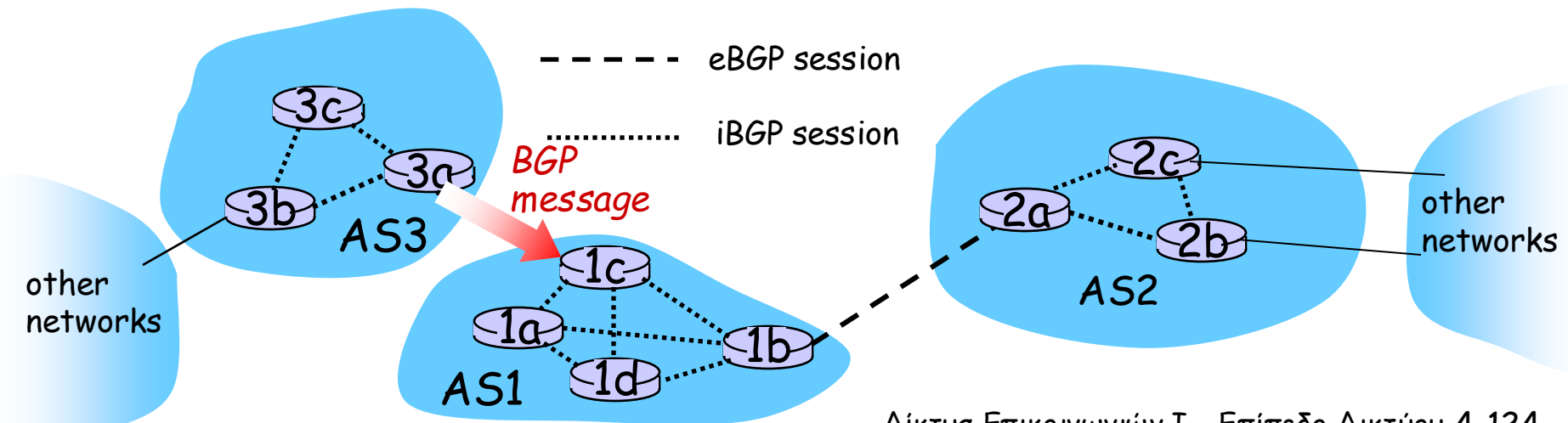
- **Δύο επίπεδα ιεραρχίας:** τοπική περιοχή, δικτυακός κορμός.
 - Δημοσιοποιήσεις κατάστασης ζεύξης μόνο στην περιοχή
 - Κάθε κόμβος έχει τη λεπτομερή τοπολογία; Γνωρίζει μόνο την κατεύθυνση (βραχύτερη διαδρομή) προς δίκτυα σε άλλες περιοχές
- **Παραμεθόριοι συνοριακοί δρομολογητές (area border routers):** "συνοψίζουν" τις αποστάσεις προς δίκτυα στη δική τους περιοχή, τις δημοσιοποιούν σε άλλους παραμεθόριους συνοριακούς δρομολογητές
- **Δρομολογητές δικτυακού κορμού (backbone routers):** τρέχουν δρομολόγηση OSPF περιορισμένη στον δικτυακό κορμό.
- **Συνοριακοί δρομολογητές (boundary routers):** συνδέουν με άλλα AS.

Δρομολόγηση inter-AS στο Διαδίκτυο: BGP

- ❑ **BGP (Border Gateway Protocol):** το de facto πρωτόκολλο δρομολόγησης μεταξύ τομέων (domains)
- ❑ Το BGP παρέχει σε κάθε AS ένα τρόπο για:
 - **eBGP:** να λαμβάνει πληροφορίες προσέγγισης υποδικτύου από γειτονικά AS.
 - **iBGP:** να διαδίδει τις πληροφορίες προσέγγισης σε όλους τους δρομολογητές που είναι εσωτερικοί στο AS
 - Να καθορίζει τις "καλές" διαδρομές προς άλλα δίκτυα με βάση τις πληροφορίες προσέγγισης και μια πολιτική
- ❑ Επιτρέπει σε κάθε υποδίκτυο να διαφημίζει την ύπαρξή του στο υπόλοιπο Διαδίκτυο: **"Είμαι εδώ"**

Τα βασικά του BGP

- **Σύνοδοι BGP (BGP sessions):** 2 BGP δρομολογητές (ομότιμοι) ανταλλάσσουν BGP μηνύματα:
 - Δημοσιοποιούν **διαδρομές** προς διαφορετικού προορισμού δικτυακά προθέματα (πρωτόκολλο "διανύσματος διαδρομής")
 - Ανταλλάσσονται πάνω από ημι-μόνιμες TCP συνδέσεις
- Όταν το AS3 διαφημίζει ένα πρόθεμα στον AS1:
 - Το AS3 **υπόσχεται** ότι θα προωθήσει datagrams προς αυτό το πρόθεμα.
 - Το AS3 μπορεί να ομαδοποιήσει τα προθέματα στη δημοσιοποίησή του

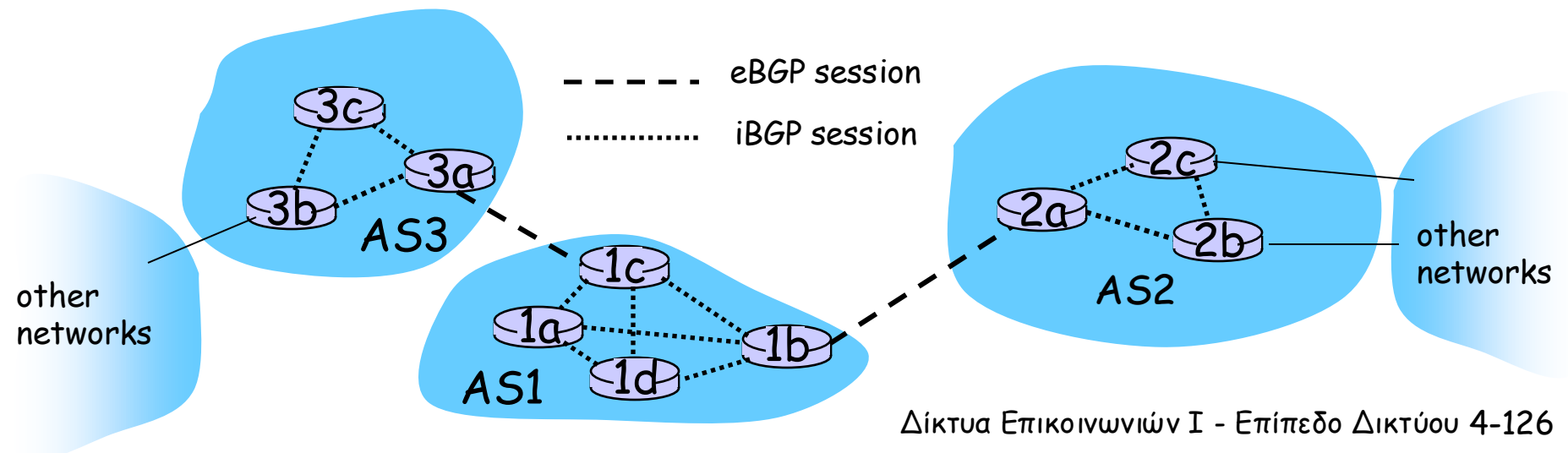


Μηνύματα BGP

- ❑ Τα μηνύματα BGP ανταλλάσσονται μεταξύ ομότιμων πάνω από TCP σύνδεση.
- ❑ Μηνύματα BGP:
 - **OPEN:** ανοίγει σύνδεση TCP προς τον ομότιμο και ταυτοποιεί τον αποστολέα
 - **UPDATE:** δημοσιοποιεί νέα διαδρομή (ή αποσύρει παλιά)
 - **KEEPALIVE:** διατηρεί τη σύνδεση ζωντανή ελλείψει UPDATES. Επίσης επιβεβαιώνει αίτηση OPEN
 - **NOTIFICATION:** αναφέρει σφάλματα σε προηγούμενο μήνυμα και χρησιμοποιείται για το κλείσιμο σύνδεσης

Διανομή πληροφορίας προσέγγισης

- Χρησιμοποιώντας σύνοδο eBGP μεταξύ των 3a και 1c, το AS3 στέλνει πληροφορίες δυνατότητας προσέγγισης προθέματος στο AS1.
 - Το 1c μπορεί τότε να χρησιμοποιήσει iBGP για να διανείμει νέες πληροφορίες προθέματος σε όλους τους δρομολογητές του AS1
 - Το 1b μπορεί τότε να διαφημίσει εκ νέου τις νέες πληροφορίες προσέγγισης στο AS2 μέσω της συνεδρίας eBGP από το 1b-στο-2a
- Όταν κάποιος δρομολογητής μαθαίνει για νέο πρόθεμα δημιουργεί καταχώριση για το πρόθεμα στον πίνακα προώθησής του



Ιδιοχαρακτηριστικά διαδρομής & διαδρομές

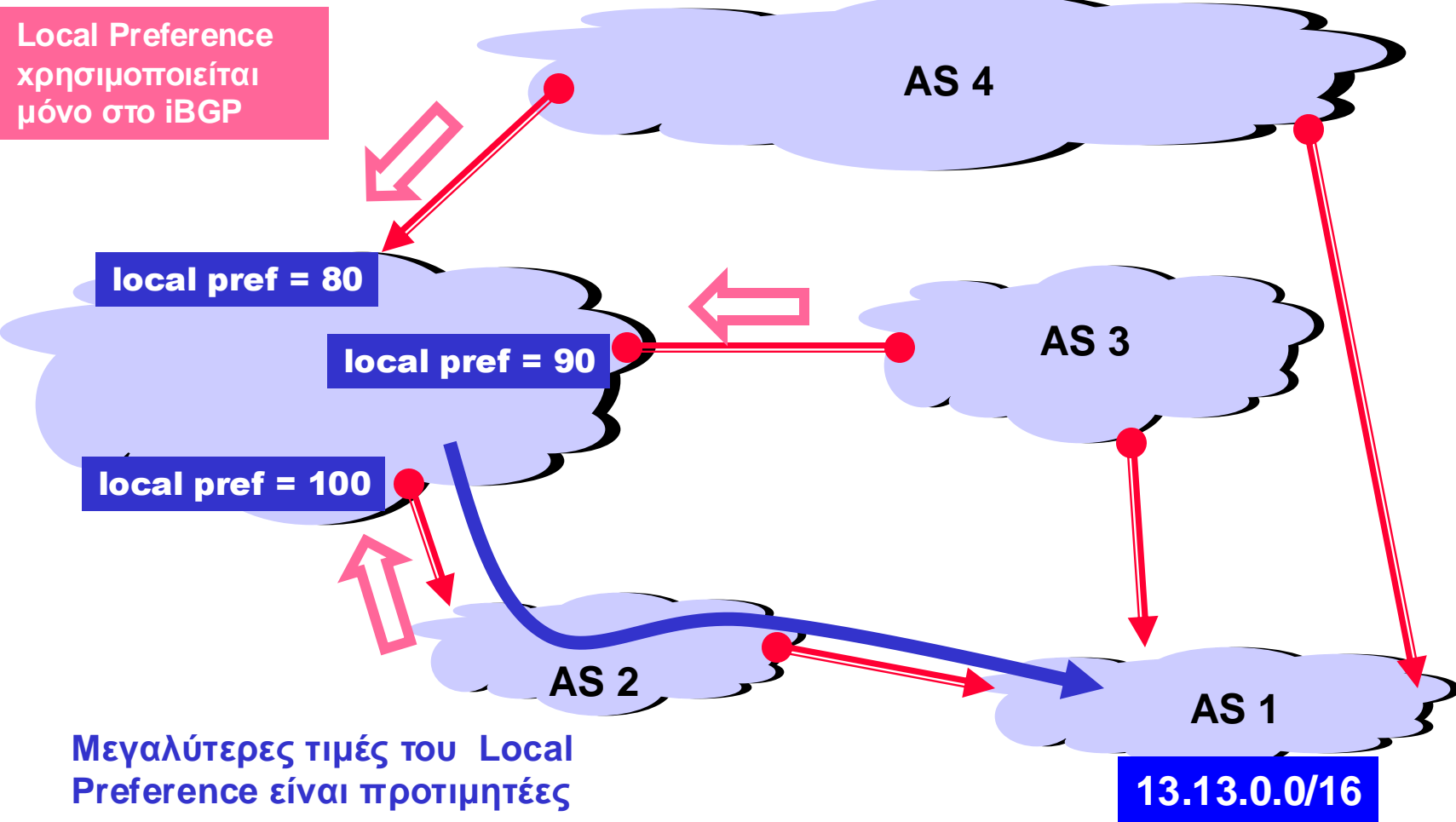
BGP

- Ένα δημοσιοποιούμενο πρόθεμα περιλαμβάνει χαρακτηριστικά του BGP
 - Πρόθεμα + ιδιοχαρακτηριστικά = "διαδρομή" (route)
- Δύο σημαντικά ιδιοχαρακτηριστικά:
 - **AS Path:** περιλαμβάνει τα AS μέσω των οποίων έχει περάσει η διαφήμιση για το πρόθεμα: π.χ., AS 67, AS 17
 - **Next Hop:** υποδεικνύει συγκεκριμένο εσωτερικό στο AS δρομολογητή προς το AS επόμενου άλματος (ενδέχεται να υπάρχουν πολλαπλές ζεύξεις από το τρέχον AS στο AS επόμενου άλματος)
- Όταν ένας δρομολογητής πύλης λαμβάνει τη διαφήμιση διαδρομής, χρησιμοποιεί την **πολιτική εισαγωγής (import policy)** για να δεχτεί/απορρίψει
 - π.χ., ποτέ μη δρομολογείς μέσω του AS x
 - δρομολόγηση βασισμένη σε κάποια πολιτική

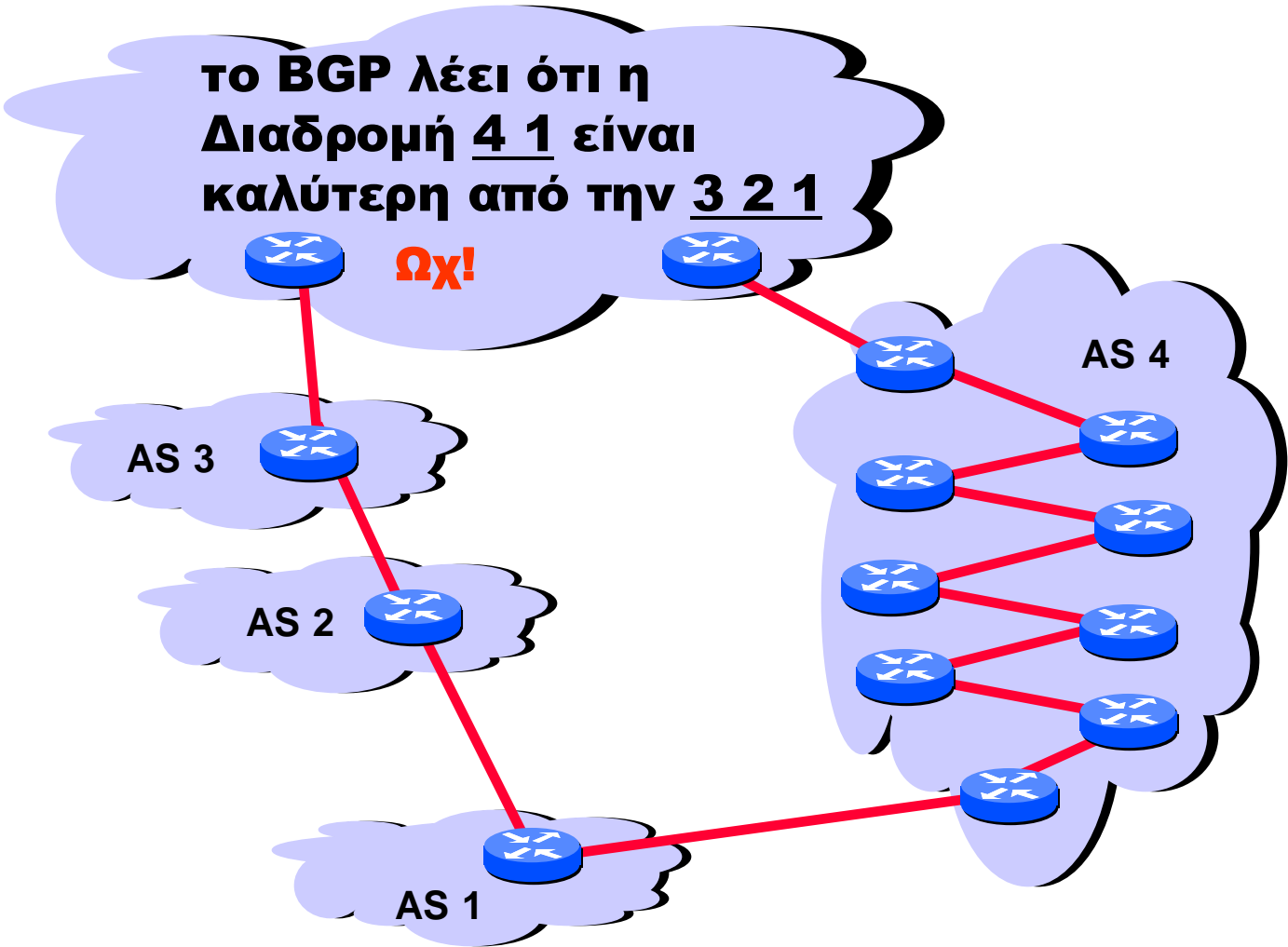
Επιλογή διαδρομής BGP

- Ο δρομολογητής ενδέχεται να μάθει για περισσότερες από μία διαδρομές για το AS προορισμού. Επιλέγει διαδρομή βασισμένος σε:
 1. Ιδιοχαρακτηριστικό τιμής τοπικής προτίμησης: απόφαση πολιτικής
 2. Βραχύτερο AS-PATH
 3. Πλησιέστερος δρομολογητής NEXT-HOP: δρομολόγηση καυτής πατάτας
 4. Πρόσθετα κριτήρια

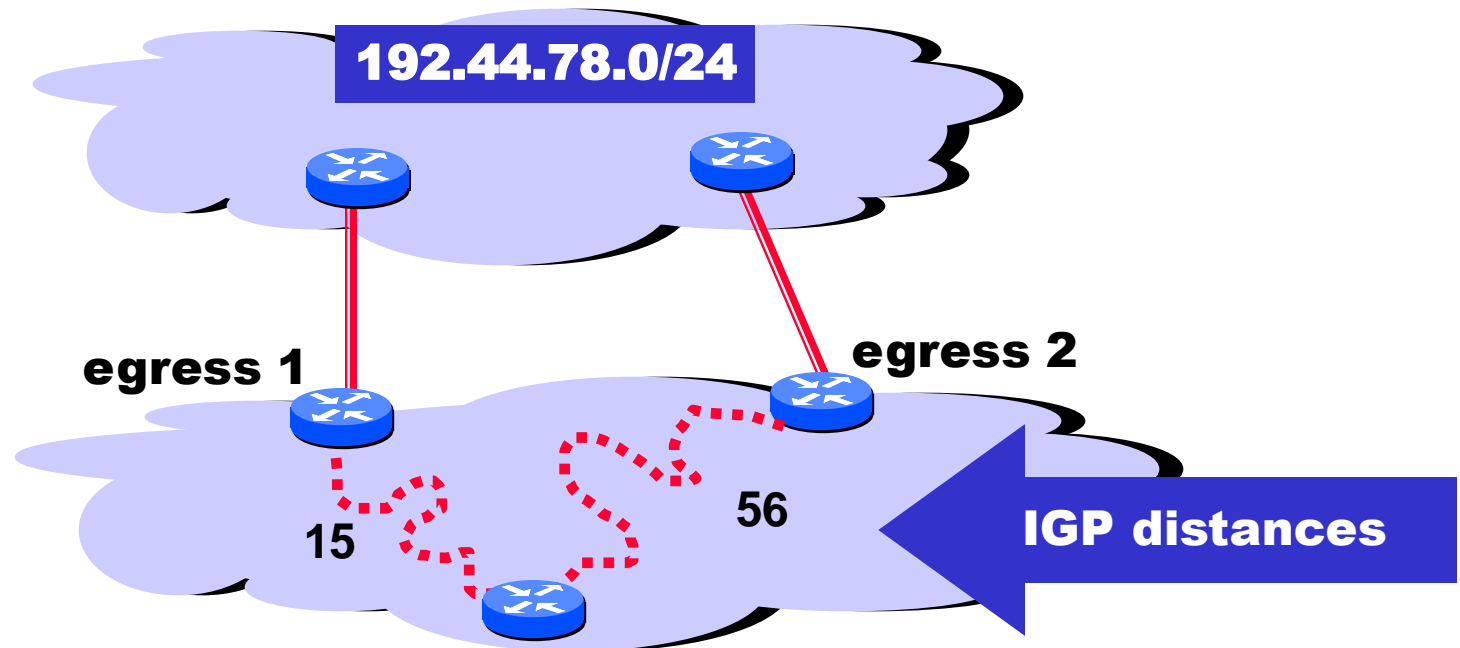
Τοπική Προτίμηση (Local Preference)



Συντομότερο δεν σημαίνει πάντα ...συντομότερο!



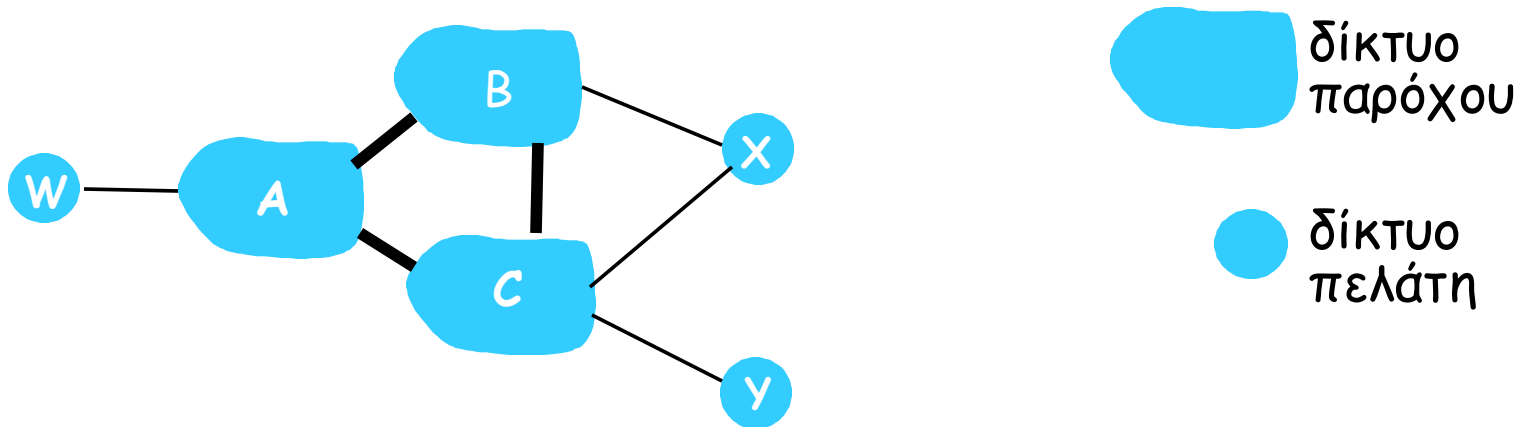
Hot Potato Routing: Επίλεξε το πλησιέστερο σημείο εξόδου



Αυτός ο δρομολογητής έχει δύο BGP routes προς 192.44.78.0/24.

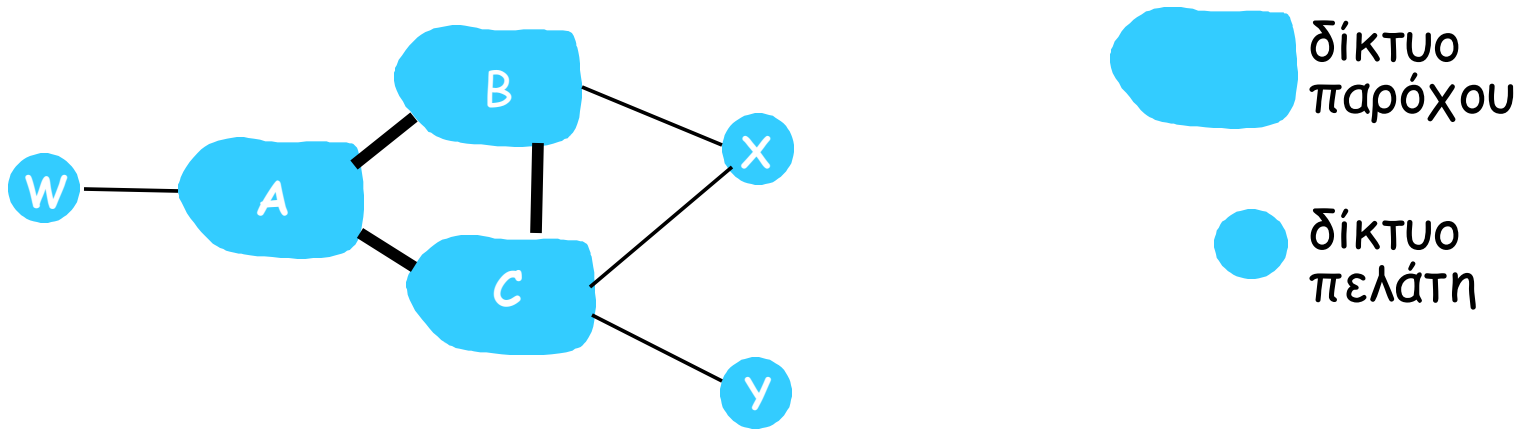
Hot potato: ξεφορτώσου την κίνηση από το δίκτυό σου το συντομότερο δυνατό. Επίλεξε egress 1!

Πολιτική δρομολόγησης BGP



- ❑ Τα A,B,C είναι **δίκτυα παρόχων (provider networks)**
- ❑ Τα X,W,Y είναι πελάτες (των δικτύων παρόχων)
- ❑ Το X είναι **διεστιακό (dual-homed)**: συνδέεται σε δύο δίκτυα
 - Το X δεν θέλει να δρομολογεί από το B μέσω του X προς το C
 - .. έτσι το X δεν θα δημοσιοποιήσει στο B μια διαδρομή προς το C

Πολιτική δρομολόγησης BGP (2)



- ❑ Το A διαφημίζει τη διαδρομή AW στο B
- ❑ Το B διαφημίζει τη διαδρομή BAW στο X
- ❑ Θα πρέπει το B να διαφημίσει τη διαδρομή BAW στο C;
 - Σε καμία περίπτωση! Το B δεν έχει όφελος από τη δρομολόγηση CBAW, καθώς ούτε το W ούτε το C είναι πελάτες του B
 - Το B θέλει να εξαναγκάσει το C να δρομολογεί προς το W μέσω του A
 - Το B θέλει να δρομολογεί **μόνο** προς/από τους πελάτες του!

Γιατί διαφορετική δρομολόγηση intra- και inter-AS ;

Πολιτική:

- ❑ inter-AS: ο διαχειριστής θέλει έλεγχο στο πώς δρομολογείται η κίνησή του, ποιός δρομολογεί μέσω του δικτύου του
- ❑ intra-AS: μοναδικός διαχειριστής, οπότε δεν χρειάζονται αποφάσεις πολιτικής

Κλίμακα:

- ❑ Η ιεραρχική δρομολόγηση εξοικονομεί μέγεθος πίνακα, μειωμένη κίνηση ενημερώσεων

Απόδοση:

- ❑ intra-AS: μπορεί να επικεντρώσει στην απόδοση
- ❑ inter-AS: η πολιτική μπορεί να κυριαρχήσει της απόδοσης

Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Άδεια Χρήσης

Σημείωμα Αναφοράς

Copyright Εθνικών και Καποδιστριακών Πανεπιστημίων Αθηνών,
Μεράκος Λάζαρος 2015. «Δίκτυα Επικοινωνιών Ι. Ενότητα 4:
Επίπεδο Δικτύου». Έκδοση: 1.01. Αθήνα 2015.

Διαθέσιμο από τη δικτυακή διεύθυνση:

<http://opencourses.uoa.gr/courses/DI19>

Χρηματοδότηση

- r Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- r Το έργο «**Ανοικτά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- r Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «**Εκπαίδευση και Δια Βίου Μάθηση**» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ