

Το Επίπεδο Δικτύου

Περίγραμμα

- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
- 4 Αρχές Δρομολόγησης
- 5 Δρομολόγηση στο Διαδίκτυο
- 6 IPv6
- 7 Ειδικά θέματα του στρώματος Δικτύου

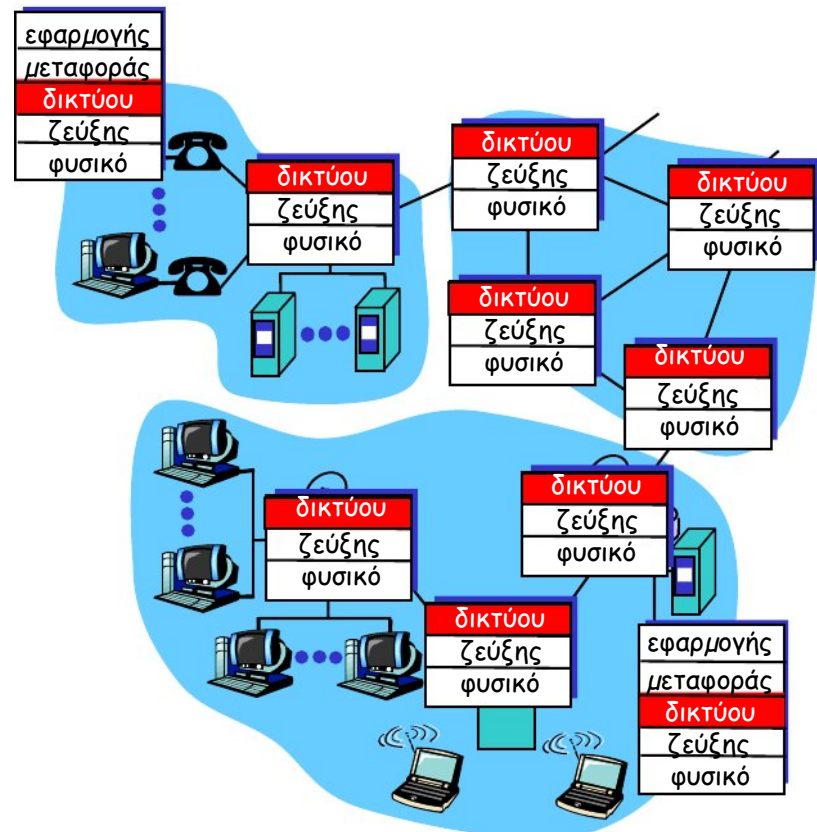
Λειτουργίες επιπέδου δικτύου

μεταφορά πακέτων από τον host-αποστολέα στον host-παραλήπτη
πρωτόκολλο επιπέδου δικτύου σε κάθε host, δρομολογητή

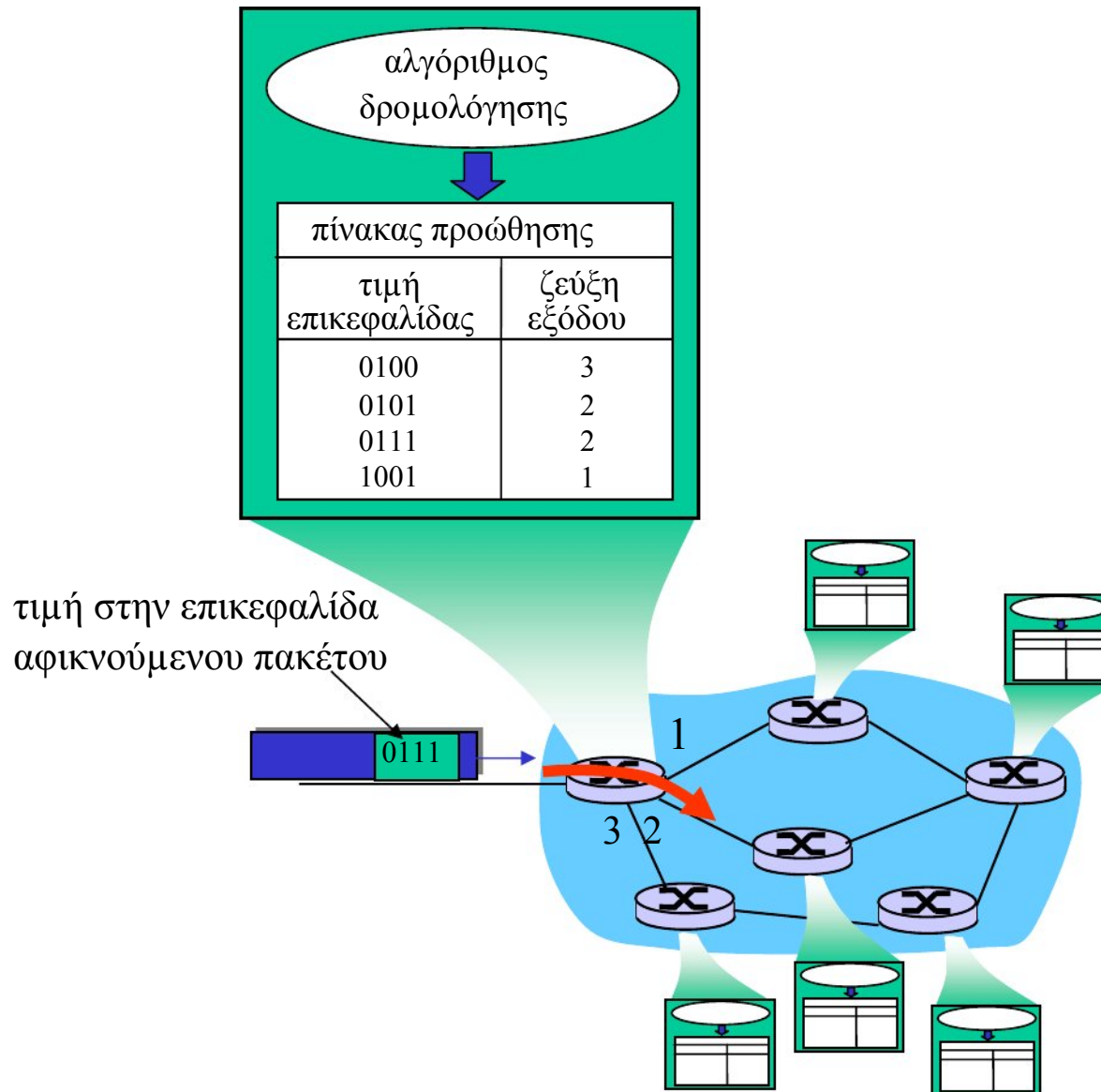
δύο σημαντικές λειτουργίες:

δρομολόγηση: προσδιορισμός διαδρομής που ακολουθούν τα πακέτα μεταξύ πηγής και προορισμού από αλγορίθμους δρομολόγησης

προώθηση: μεταφορά πακέτων από μία είσοδο στην κατάλληλη έξοδο ενός δρομολογητή



Δρομολόγηση και προώθηση



Εγκαθίδρυση σύνδεσης

Τρίτη σημαντική λειτουργία του επιπέδου δικτύου σε μερικές αρχιτεκτονικές δικτύων:
ATM, frame relay, X.25

Πριν από τη μεταφορά πακέτων, τα δύο τερματικά συστήματα και οι ενδιάμεσοι δρομολογητές εγκαθιδρύουν εικονική σύνδεση συμμετοχή των δρομολογητών

Υπηρεσία σύνδεσης σε επίπεδο δικτύου και μεταφοράς:

Επίπεδο δικτύου: μεταξύ δύο τερματικών συστημάτων

Επίπεδο μεταφοράς: μεταξύ δύο διεργασιών

Μοντέλο υπηρεσιών δικτύου

Το **μοντέλο υπηρεσιών** που παρέχεται από το επίπεδο δικτύου ορίζει τα χαρακτηριστικά της μεταφοράς δεδομένων μεταξύ τερματικών συστημάτων, όπως:

- εάν υπάρχουν απώλειες κατά τη μεταφορά
- εάν η παράδοση δεδομένων γίνεται με τη σωστή σειρά
- εάν διατηρείται ο χρονισμός των πακέτων
- εάν παρέχεται εγγυημένο bandwidth για τη μεταφορά
- εάν παρέχεται ένδειξη συμφόρησης στον αποστολέα από το δίκτυο

Μοντέλα υπηρεσιών στο επίπεδο δικτύου

Αρχιτεκτονική Δικτύου	Μοντέλο Υπηρεσιών	Εγγυήσεις				Ένδειξη
		Bandwidth	Απώλειες	Σειρά	Χρονισμός	Συμφόρησης
Διαδίκτυο	βέλτιστης προσπάθειας	όχι	όχι	όχι	όχι	όχι
ATM	CBR (constant bit rate)	σταθερός ρυθμός	ναι	ναι	ναι	δεν υφίσταται συμφόρηση
ATM	VBR (variable bit rate)	εγγυημένος ρυθμός	ναι	ναι	ναι	δεν υφίσταται συμφόρηση
ATM	ABR (available bit rate)	εγγυημένος ελάχιστος ρυθμός	όχι	ναι	όχι	ναι
ATM	UBR (unspecified bit rate)	όχι	όχι	ναι	όχι	όχι

Επέκταση του μοντέλου υπηρεσιών του Διαδικτύου: ενοποιημένες υπηρεσίες (Intserv), διαφοροποιημένες υπηρεσίες (Diffserv)

Επίπεδο δικτύου: υπηρεσία σύνδεσης και ασυνδειστρεφής υπηρεσία

Ένα επίπεδο δικτύου μπορεί να παρέχει **ασυνδειστρεφή υπηρεσία (connectionless service)** ή **υπηρεσία σύνδεσης (connection service)**
Υπηρεσίες ανάλογες με αυτές στο επίπεδο μεταφοράς, όμως με διαφορές:
επίπεδο δικτύου

υπηρεσία **από host σε host** που παρέχεται από το επίπεδο δικτύου στο επίπεδο μεταφοράς
το δίκτυο παρέχει τη **μία ή την άλλη υπηρεσία δίχως δυνατότητα επιλογής**
η υπηρεσία σύνδεσης υλοποιείται στον **πυρήνα του δικτύου** και στα **τερματικά συστήματα**

επίπεδο μεταφοράς

υπηρεσία **από διεργασία σε διεργασία** που παρέχεται από το επίπεδο μεταφοράς στο επίπεδο εφαρμογής
το επίπεδο εφαρμογής έχει τη **δυνατότητα επιλογής** μεταξύ ασυνδειστρεφούς και συνδειστρεφούς υπηρεσίας
η συνδειστρεφής υπηρεσία υλοποιείται στα **τερματικά συστήματα**

Τα δίκτυα που παρέχουν ασυνδειστρεφή υπηρεσία στο επίπεδο δικτύου καλούνται **δίκτυα αυτοδύναμων πακέτων**

Τα δίκτυα που παρέχουν υπηρεσία σύνδεσης στο επίπεδο δικτύου καλούνται **δίκτυα εικονικών κυκλωμάτων**

Εικονικά κυκλώματα (virtual circuits)

εγκαθίδρυση κλήσης πριν από τη μεταφορά δεδομένων και τερματισμός της μετά τη μεταφορά

κάθε πακέτο φέρει πεδίο που προσδιορίζει το εικονικό κύκλωμα στο οποίο ανήκει και που χρησιμοποιείται για την προώθησή του στο δίκτυο
κάθε δρομολογητής στη διαδρομή μεταξύ πηγής-προορισμού διατηρεί πληροφορία "κατάστασης" για κάθε διερχόμενη σύνδεση

η σύνδεση στο επίπεδο μεταφοράς ενέπλεκε μόνο δύο τερματικά συστήματα
είναι δυνατό να δεσμευθούν και να απονεμηθούν πόροι (bandwidth, buffers) σε ένα εικονικό κύκλωμα

ώστε να λάβουμε απόδοση όμοια με αυτή ενός "πραγματικού" κυκλώματος

"ένα εικονικό κύκλωμα συμπεριφέρεται παρόμοια με ένα τηλεφωνικό κύκλωμα"

ως προς την απόδοση

ως προς τις ενέργειες του δικτύου κατά μήκος του εικονικού κυκλώματος

Υλοποίηση εικονικών κυκλωμάτων

Ένα εικονικό κύκλωμα (Virtual Circuit, VC) αποτελείται από:

1. τη διαδρομή από πηγή σε προορισμό
2. αριθμούς VC, ένας αριθμός για κάθε ζεύξη κατά μήκος της διαδρομής
3. καταχωρήσεις στους πίνακες προώθησης των δρομολογητών κατά μήκος της διαδρομής

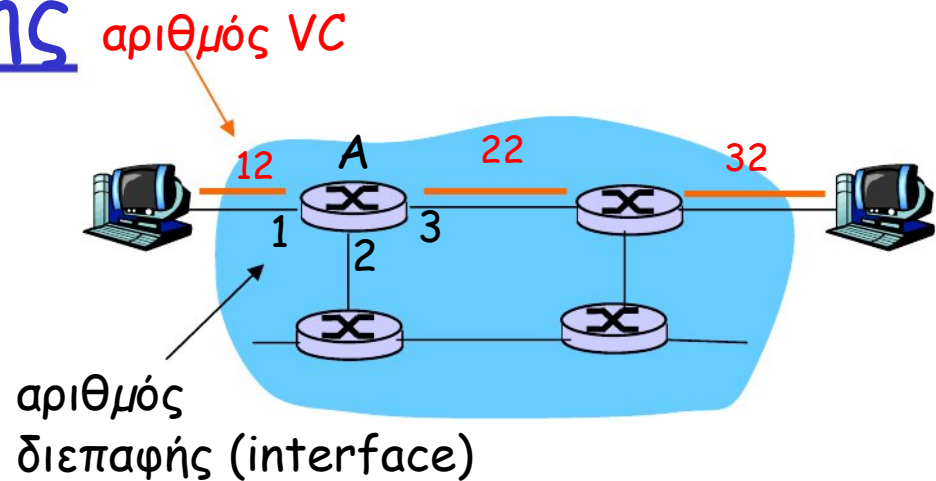
Κάθε πακέτο που ανήκει σε ένα VC φέρει στην επικεφαλίδα του έναν αριθμό VC

Ο αριθμός VC αλλάζει σε κάθε ζεύξη

ο νέος αριθμός VC καθορίζεται από τον πίνακα προώθησης

Πίνακας προώθησης

Πίνακας προώθησης στο δρομολογητή A:



Εισερχόμενη διεπαφή	Εισερχόμενος αριθμός VC	Εξερχόμενη διεπαφή	Εξερχόμενος αριθμός VC
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
...	...	...	...

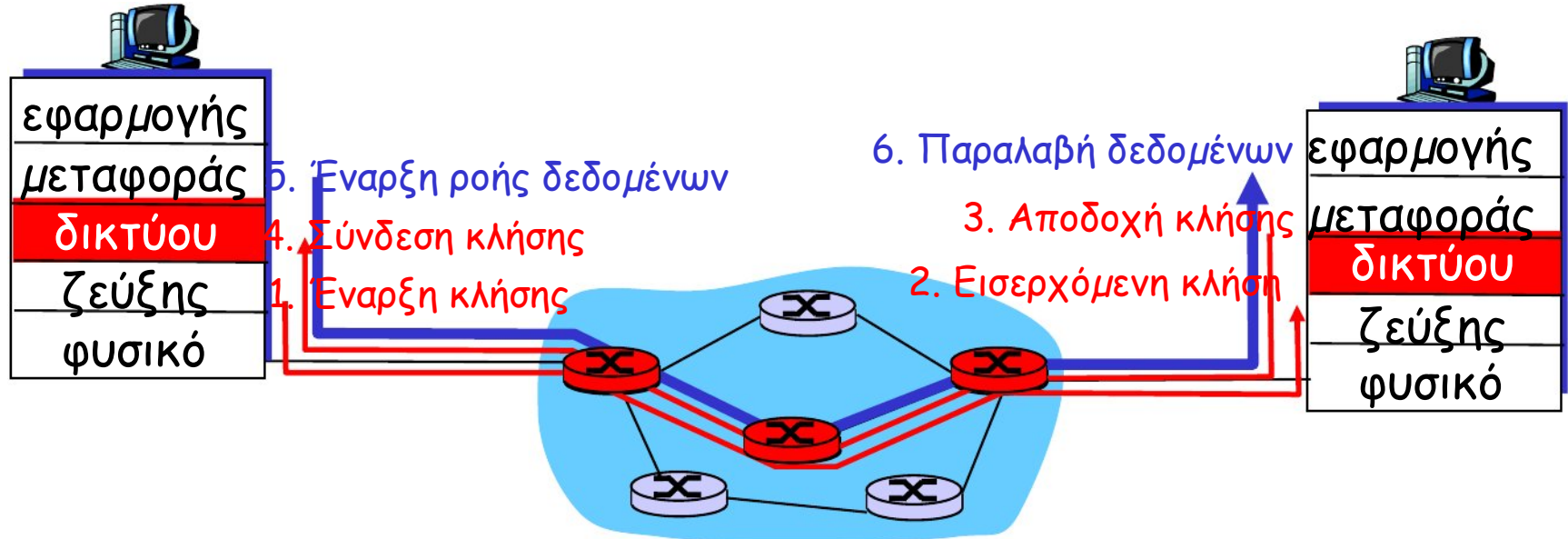
Οι δρομολογητές διατηρούν πληροφορία κατάστασης σύνδεσης

Εικονικά κυκλώματα: πρωτόκολλα σηματοδότησης

χρησιμοποιούνται για την εγκαθίδρυση, διατήρηση και τερματισμό εικονικών κυκλωμάτων

χρησιμοποιούνται στο ATM, frame-relay, X.25

δεν χρησιμοποιούνται στο Διαδίκτυο στην παρούσα του μορφή

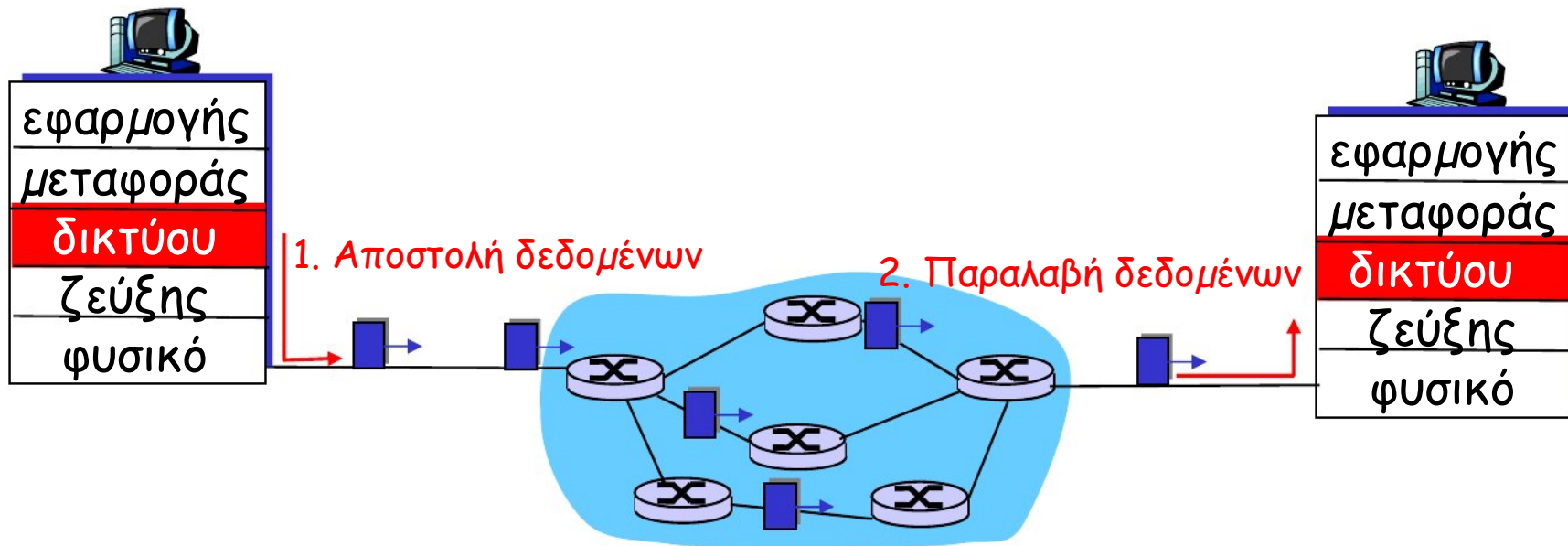


Δίκτυα αυτοδύναμων πακέτων: το μοντέλο του Διαδικτύου

δεν εφαρμόζεται εγκαθίδρυση κλήσης στο επίπεδο δικτύου
οι δρομολογητές δεν διατηρούν πληροφορία "κατάστασης"
για συνδέσεις από τερματικό σε τερματικό

δεν υπάρχει η έννοια της "σύνδεσης" στο επίπεδο δικτύου
τα πακέτα προωθούνται με βάση τη διεύθυνση του host
προορισμού

πακέτα μεταξύ του ίδιου ζεύγους πηγής-προορισμού ενδέχεται να
ακολουθήσουν διαφορετικές διαδρομές



Πίνακας προώθησης

4 δισεκατομμύρια
δυνατές καταχωρήσεις

Διευθύνσεις Προορισμού

11001000 00010111 00010000 00000000

ως

11001000 00010111 00010111 11111111

11001000 00010111 00011000 00000000

ως

11001000 00010111 00011000 11111111

11001000 00010111 00011001 00000000

ως

11001000 00010111 00011111 11111111

αλλιώς

Διεπαφή Ζεύξης

0

1

2

3

Longest Prefix Matching

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

Παραδείγματα

Διεύθυνση προορισμού:

11001000 00010111 00010110 10100001

διεπαφή 0

11001000 00010111 00011000 10101010

διεπαφή 1

Δίκτυα αυτοδύναμων πακέτων και εικονικών κυκλωμάτων: προέλευση

Διαδίκτυο

γεννήθηκε από την ανάγκη
διασύνδεσης υπολογιστών για την
ανταλλαγή δεδομένων

“ελαστική” υπηρεσία χωρίς
αυστηρές απαιτήσεις
χρονισμού

τερματικά συστήματα με ευφυία
(υπολογιστές)

μπορούν να εκτελούν έλεγχο
σφαλμάτων κλπ

απλότητα στο εσωτερικό του
δικτύου, πολυπλοκότητα στα
“άκρα”

πολλά είδη ζεύξεων με

διαφορετικά χαρακτηριστικά

ΑΤΜ

εξελίχθηκε από το τηλεφωνικό
δίκτυο

ανθρώπινη συνομιλία:

αυστηρές απαιτήσεις
χρονισμού και αξιοπιστίας
ανάγκη εγγυημένης
υπηρεσίας

τερματικά συστήματα χωρίς
ευφυία (τηλέφωνα)

πολυπλοκότητα στο
εσωτερικό του δικτύου

- εγκαθίδρυση κλήσης
- διατήρηση πληροφορίας
κατάστασης στους
μεταγωγείς πακέτων

Περιγραμμά

- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
- 4 Αρχές Δρομολόγησης
- 5 Δρομολόγηση στο Διαδίκτυο
- 6 IPv6
- 7 Ειδικά θέματα του στρώματος Δικτύου

Ιεραρχική Δρομολόγηση

Για τη μελέτη των αλγορίθμων δρομολόγησης θεωρήθηκε το εξιδανικευμένο μοντέλο ενός ομοιογενούς δικτύου: όλοι οι δρομολογητές εκτελούν τον ίδιο αλγόριθμο δρομολόγησης οι δρομολογητές υπολογίζουν τις διαδρομές δρομολόγησης για ολόκληρο το δίκτυο

Προβλήματα εφαρμογής στην πράξη:

μεγάλη κλίμακα

(200 εκατομμύρια προορισμοί):
αδυναμία αποθήκευσης όλων των προορισμών σε πίνακες δρομολόγησης (περιορισμοί μνήμης)
αδυναμία ανταλλαγής πληροφοριών πινάκων δρομολόγησης (περιορισμοί χωρητικότητας)

διαχειριστική αυτονομία

διαδίκτυο = δίκτυο από δίκτυα
κάθε διαχειριστής δικτύου ενδέχεται να θέλει να ελέγχει τη δρομολόγηση στο δικό του δίκτυο

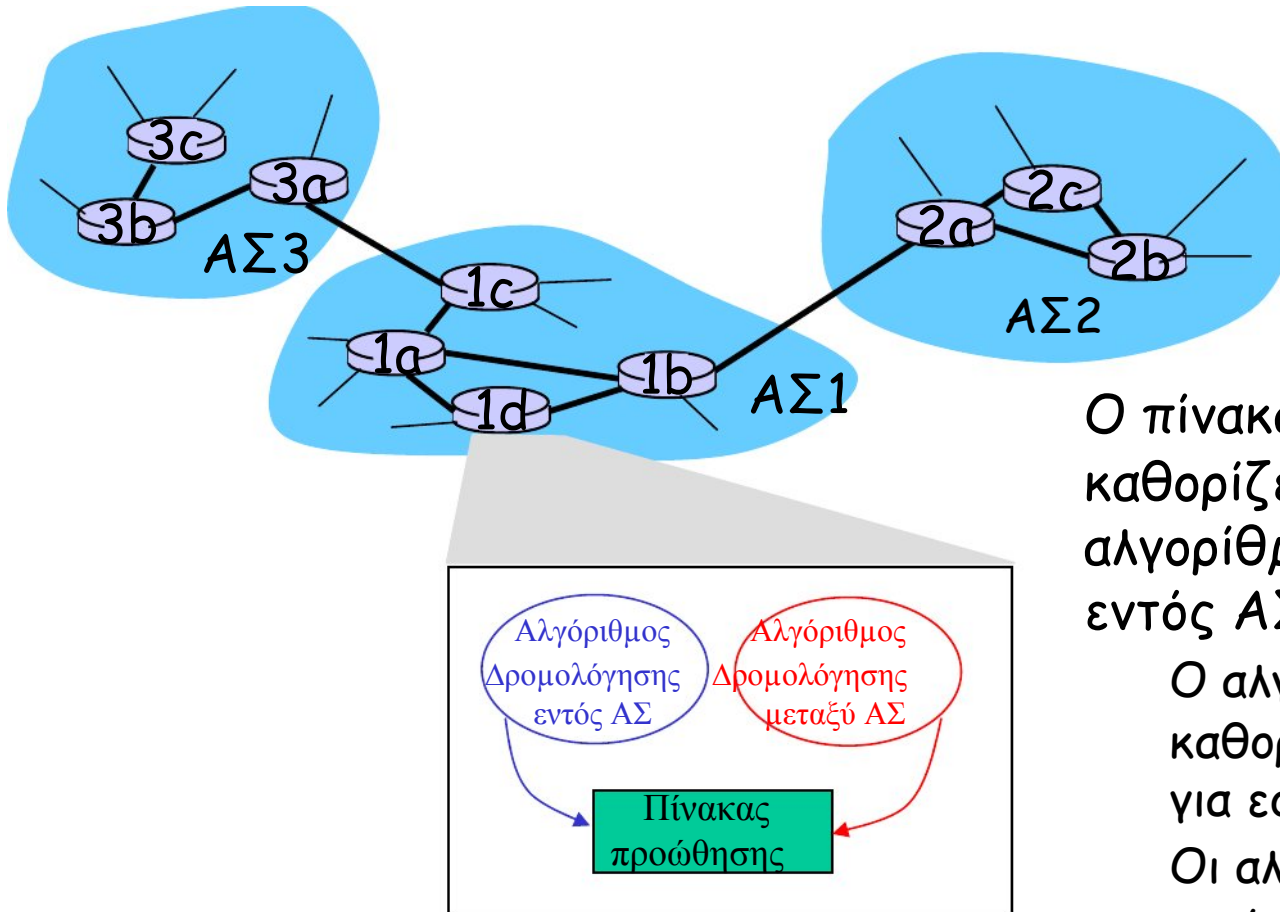
Ιεραρχική Δρομολόγηση

οργάνωση δρομολογητών
κατά περιοχές σε
αυτόνομα συστήματα (ΑΣ)
δρομολογητές στο ίδιο ΑΣ
εκτελούν το ίδιο
πρωτόκολλο δρομολόγησης
πρωτόκολλο δρομολόγησης
"εντός ΑΣ" (Intra-AS)
δρομολογητές σε
διαφορετικά ΑΣ μπορούν
να εκτελούν διαφορετικά
πρωτόκολλα δρομολόγησης

δρομολογητές πύλης

ειδικοί δρομολογητές στο
ΑΣ
εκτελούν το πρωτόκολλο
δρομολόγησης **"εντός
ΑΣ"** με τους υπόλοιπους
δρομολογητές του ΑΣ
επίσης υπεύθυνοι για τη
δρομολόγηση σε
προορισμούς εκτός ΑΣ
εκτελούν το πρωτόκολλο
δρομολόγησης **"μεταξύ ΑΣ"**
(Inter-AS) με άλλους
δρομολογητές πύλης

Διασυνδεδεμένα Αυτόνομα Συστήματα



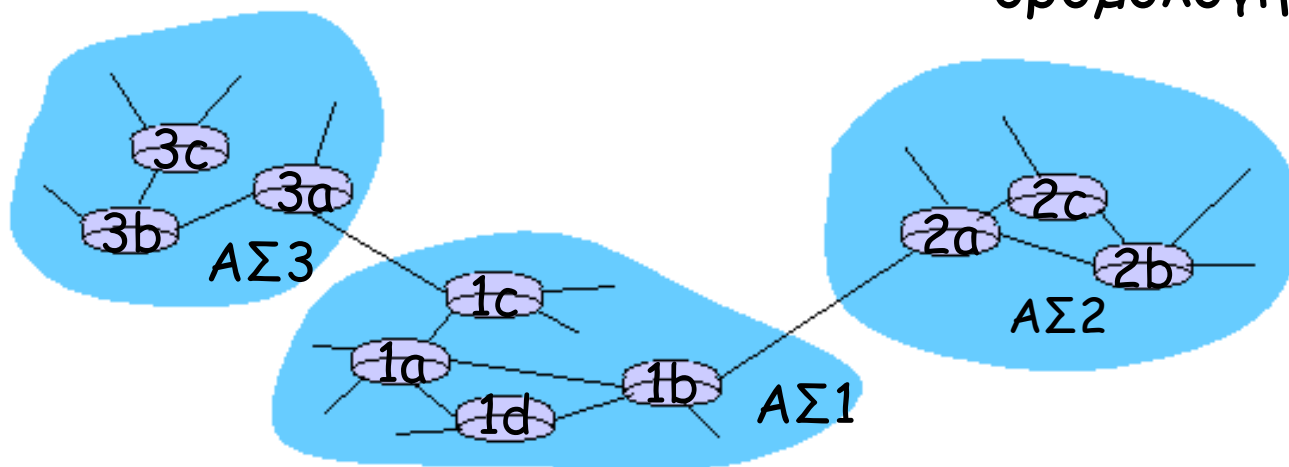
Ο πίνακας προώθησης καθορίζεται και από τους δύο αλγορίθμους δρομολόγησης: εντός ΑΣ και μεταξύ ΑΣ

Ο αλγόριθμος εντός ΑΣ καθορίζει τις καταχωρήσεις για εσωτερικούς προορισμούς
Οι αλγόριθμοι μεταξύ ΑΣ και εντός ΑΣ καθορίζουν τις καταχωρήσεις για εξωτερικούς προορισμούς

Έργο αλγορίθμου δρομολόγησης μεταξύ ΑΣ

Έστω ότι ένας δρομολογητής στο ΑΣ1 λαμβάνει ένα πακέτο με προορισμό στο εξωτερικό του ΑΣ1

Ο δρομολογητής πρέπει να προωθήσει το πακέτο προς έναν από τους δρομολογητές πύλης, αλλά ποιον;



Το ΑΣ1 χρειάζεται:

1. να μάθει ποιοι προορισμοί είναι προσπελάσιμοι μέσω του ΑΣ2 και ποιοι είναι προσπελάσιμοι μέσω του ΑΣ3
2. να διαδώσει αυτήν την πληροφορία προσπελασιμότητας σε όλους τους δρομολογητές στο ΑΣ1

Έργο της δρομολόγησης μεταξύ ΑΣ

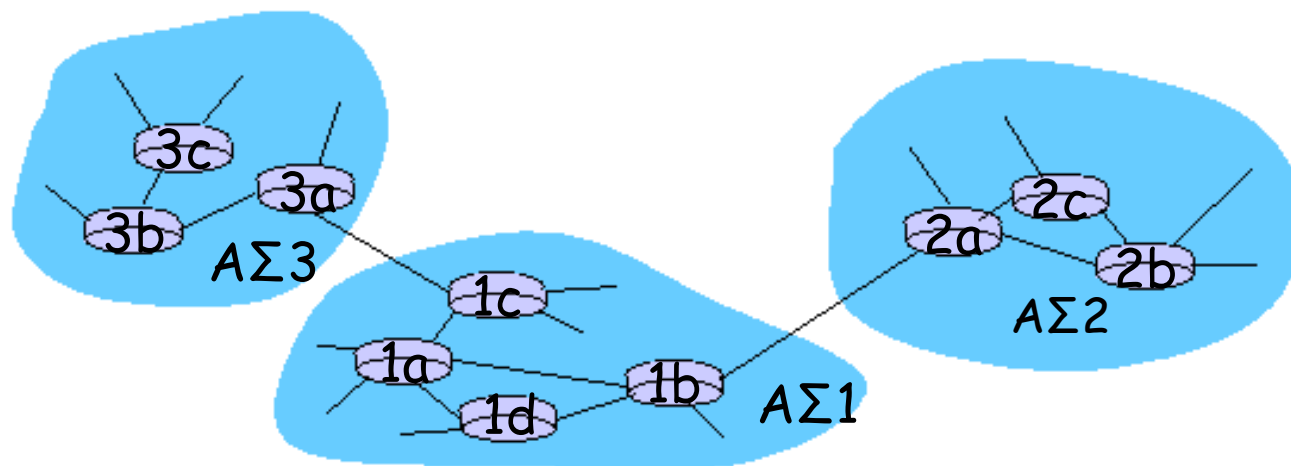
Παράδειγμα: Καθορισμός πίνακα προώθησης στο δρομολογητή 1d

Έστω ότι το ΑΣ1 μαθαίνει από το πρωτόκολλο μεταξύ ΑΣ ότι το υποδίκτυο **x** είναι προσπελάσιμο μέσω του ΑΣ3 (πύλη 1c) όχι όμως μέσω του ΑΣ2

Το πρωτόκολλο μεταξύ ΑΣ διαδίδει αυτήν την πληροφορία προσπελασιμότητας σε όλους τους εσωτερικούς δρομολογητές του ΑΣ1

Ο δρομολογητής 1d προσδιορίζει από την πληροφορία δρομολόγησης εντός ΑΣ ότι η διεπαφή του **I** βρίσκεται στην διαδρομή ελαχίστου κόστους προς την πύλη 1c

Εισάγει στον πίνακα προώθησής του την καταχώρηση (**x, I**)



Παράδειγμα: Επιλογή μεταξύ πολλαπλών ΑΣ

Έστω ότι το ΑΣ1 μαθαίνει από το πρωτόκολλο μεταξύ ΑΣ ότι το υποδίκτυο **×** είναι προσπελάσιμο μέσω του ΑΣ3 (πύλη 1c) και μέσω του ΑΣ2 (πύλη 1b)

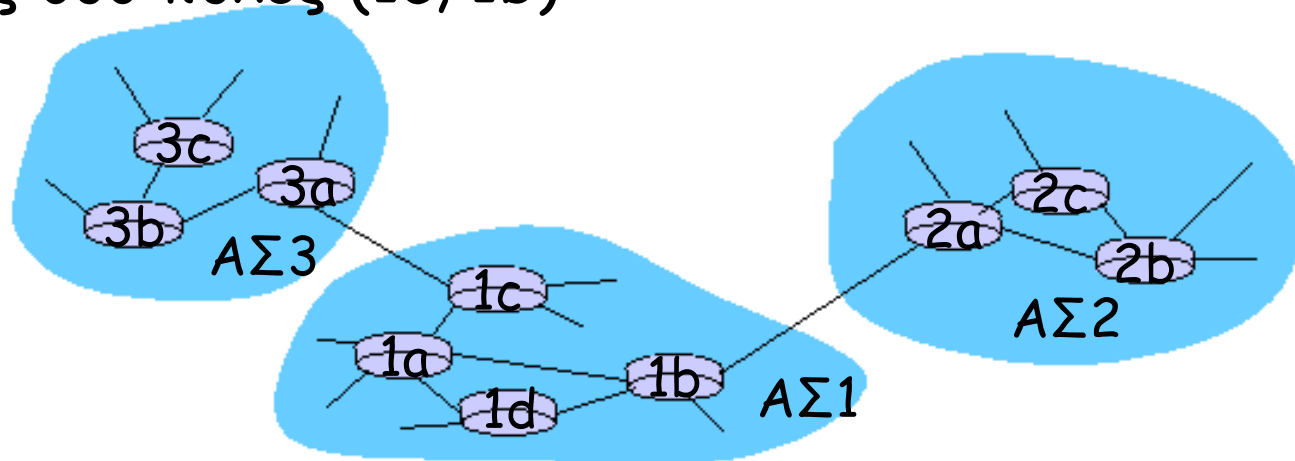
Ο δρομολογητής 1d πρέπει να προσδιορίσει σε ποια από τις δύο πύλες (1c, 1b)

θα προωθήσει τα πακέτα με προορισμό το υποδίκτυο **×**

Αυτό είναι επίσης έργο του πρωτοκόλλου μεταξύ ΑΣ

Δρομολόγηση καυτής πατάτας (hot potato routing):

αποστολή πακέτου προς τον "πλησιέστερο" δρομολογητή



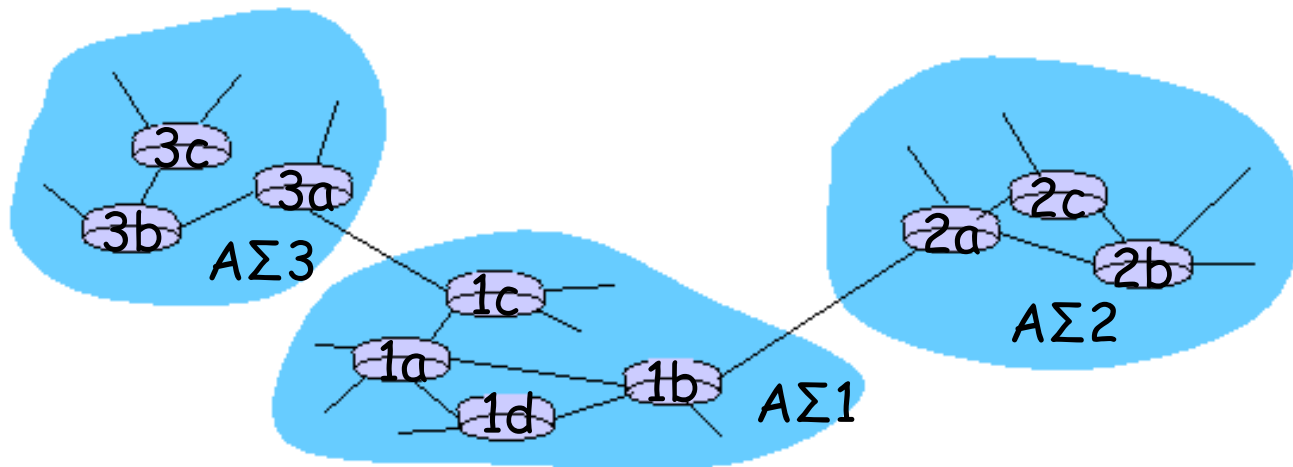
Παράδειγμα: Επιλογή μεταξύ πολλαπλών ΑΣ

Ένα ΑΣ μαθαίνει από το πρωτόκολλο μεταξύ ΑΣ ότι το υποδίκτυο x είναι προσπελάσιμο μέσω πολλαπλών πυλών

Χρησιμοποιεί την πληροφορία δρομολόγησης από το πρωτόκολλο εντός ΑΣ για να προσδιορίσει το κόστος των διαδρομών ελαχίστου κόστους προς καθεμία από τις πύλες

Δρομολόγηση καυτής πατάτας:
Επιλέγει την πύλη που έχει το ελάχιστο κόστος

Προσδιορίζει από τον πίνακα προώθησης την διεπαφή I που οδηγεί στην πύλη ελαχίστου κόστους
Εισάγει (x, I) στον πίνακα προώθησης

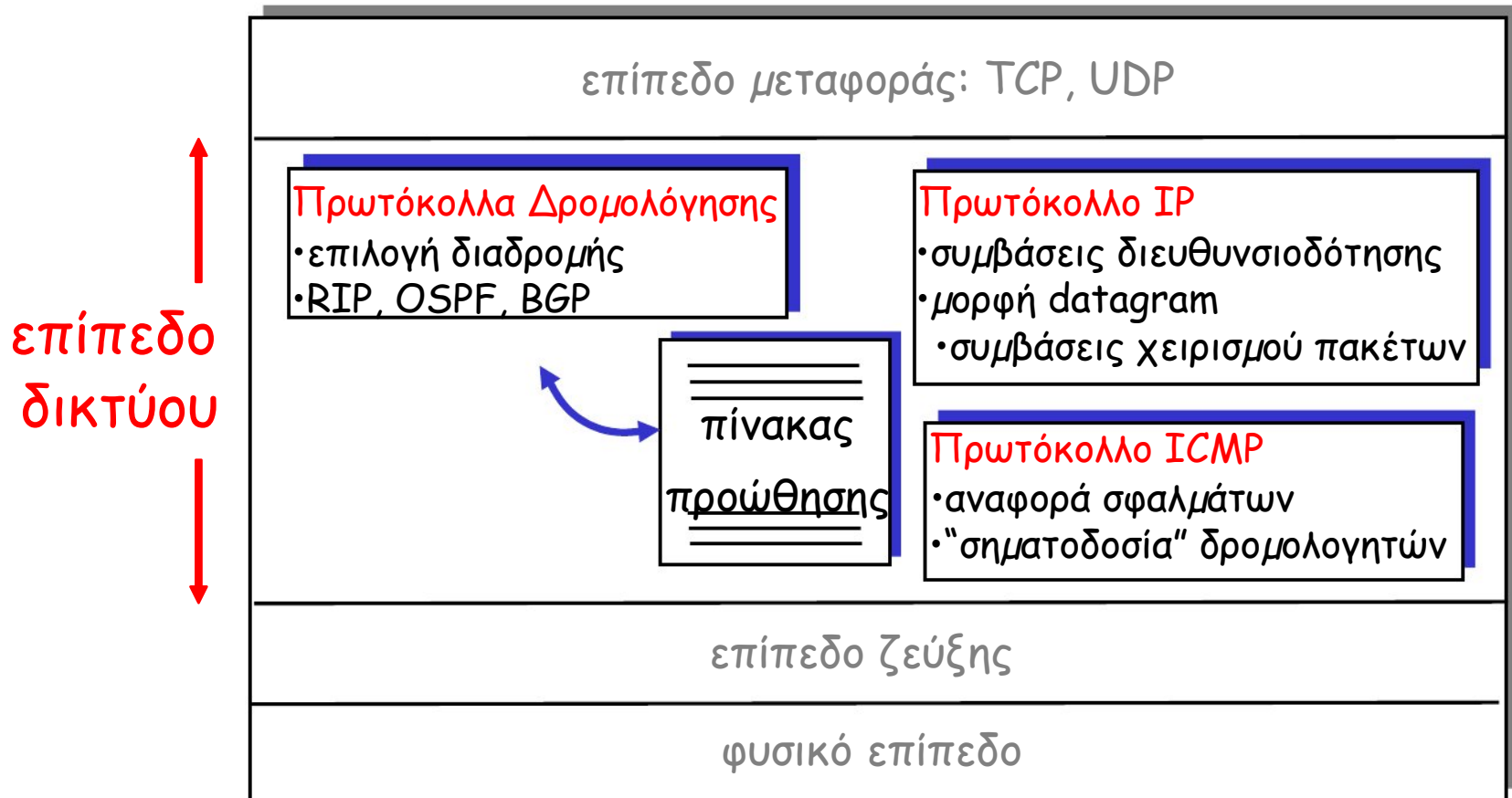


Περίγραμμα

- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
 - 3.1 Διευθυνσιοδότηση στο IPv4
 - 3.2 Μεταφορά αυτοδύναμου πακέτου από πηγή σε προορισμό
 - 3.3 Δομή αυτοδύναμου πακέτου (datagram)
 - 3.4 Κατάτμηση στο IP
 - 3.5 ICMP: Internet Control Message Protocol
 - 3.6 DHCP: Dynamic Host Configuration Protocol
 - 3.7 NAT: Network Address Translation
- 4 Αρχές Δρομολόγησης
- 5 Δρομολόγηση στο Διαδίκτυο
- 6 IPv6
- 7 Ειδικά θέματα του στρώματος Δικτύου

Το επίπεδο δικτύου στο Διαδίκτυο

Λειτουργίες του επιπέδου δικτύου σε τερματικά συστήματα και δρομολογητές:

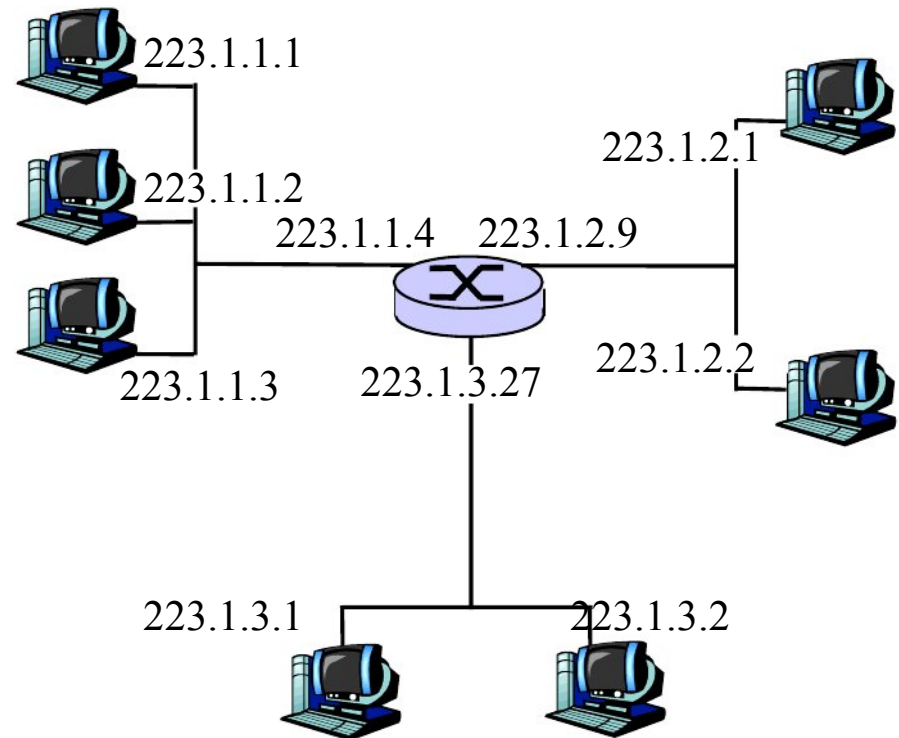


Διευθυνσιοδότηση IP: εισαγωγή

διεπαφή (interface): όριο μεταξύ τερματικού συστήματος ή δρομολογητή και φυσικής ζεύξης

οι δρομολογητές συνήθως έχουν πολλαπλές διεπαφές
τα τερματικά συστήματα συνήθως έχουν μία διεπαφή
κάθε διεπαφή σχετίζεται με μία διεύθυνση IP

Διεύθυνση IP: πεδίο από 32 bits που προσδιορίζει τη διεπαφή ενός τερματικού συστήματος ή ενός δρομολογητή - παγκοσμίως μοναδική



223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

δεκαδικός συμβολισμός με τελείες
(dotted decimal notation)

Υποδίκτυα

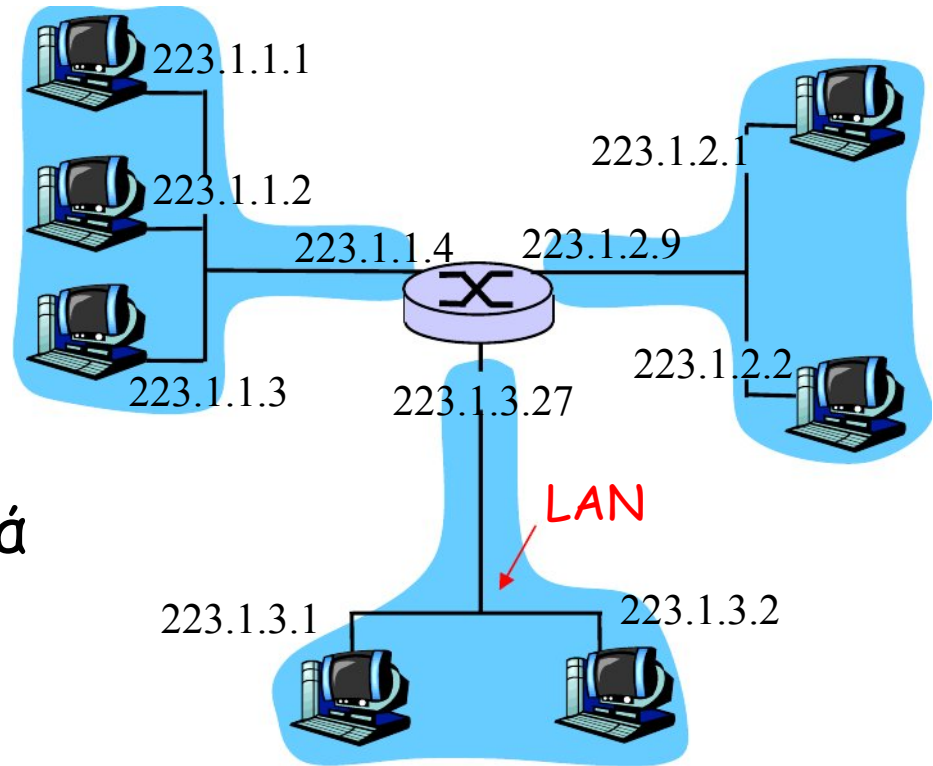
Διεύθυνση IP:

τμήμα δικτύου (bits ανώτερης τάξης)

τμήμα host (bits κατώτερης τάξης)

Από την άποψη των διευθύνσεων IP, **υποδίκτυο (subnet)** ή **δίκτυο IP** ή απλά **δίκτυο** είναι το σύνολο των διεπαφών

μεταξύ των οποίων υπάρχει φυσική σύνδεση χωρίς τη μεσολάβηση ενός δρομολογητή των οποίων οι διευθύνσεις IP έχουν το ίδιο τμήμα δικτύου



δίκτυο αποτελούμενο από 3 υποδίκτυα
Διευθύνσεις υποδικτύων: 223.1.1.0/24,
223.1.2.0/24, 223.1.3.0/24
όπου /24: **μάσκα ή πρόθεμα υποδικτύου**

Υποδίκτυα

Προσδιορισμός υποδικτύων
σε ένα σύστημα από hosts
και δρομολογητές:

Αποσύνδεση διεπαφών από
δρομολογητές, hosts

Δημιουργία "νησίδων" από
απομονωμένα δίκτυα

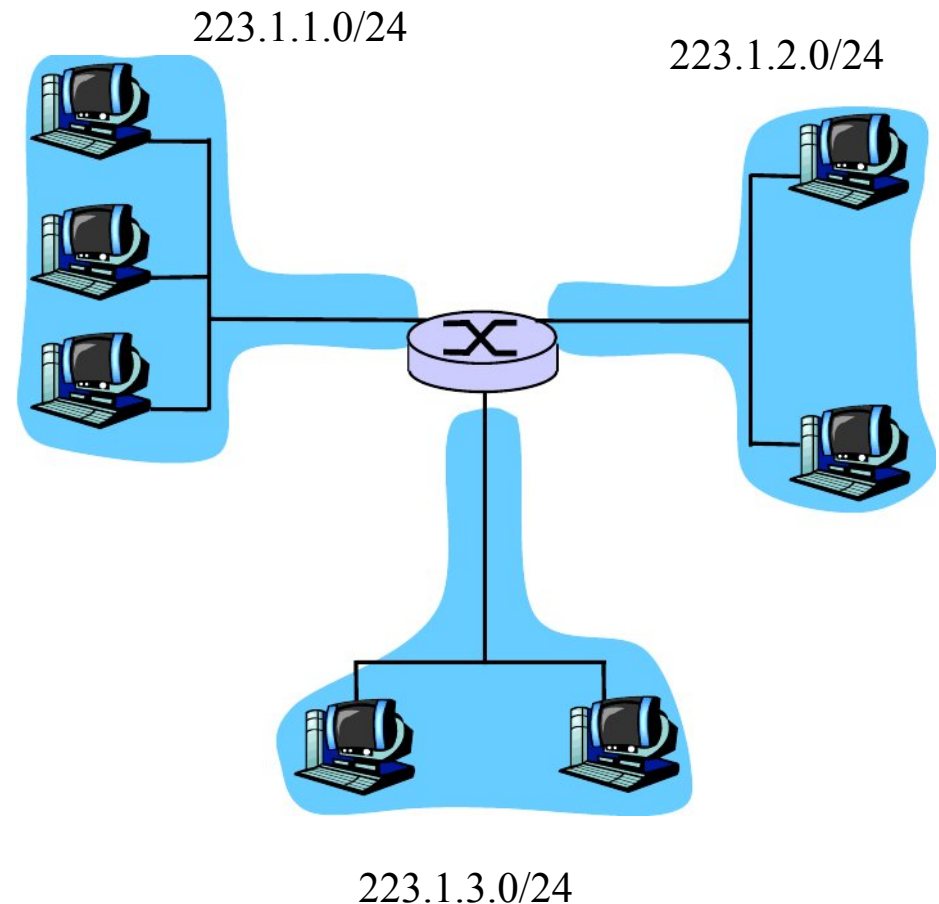
Κάθε απομονωμένο δίκτυο
καλείται **υποδίκτυο**
(subnet)

3 υποδίκτυα με διευθύνσεις:

223.1.1.0/24

223.1.2.0/24

223.1.3.0/24



μάσκα υποδικτύου: /24

Υποδίκτυα

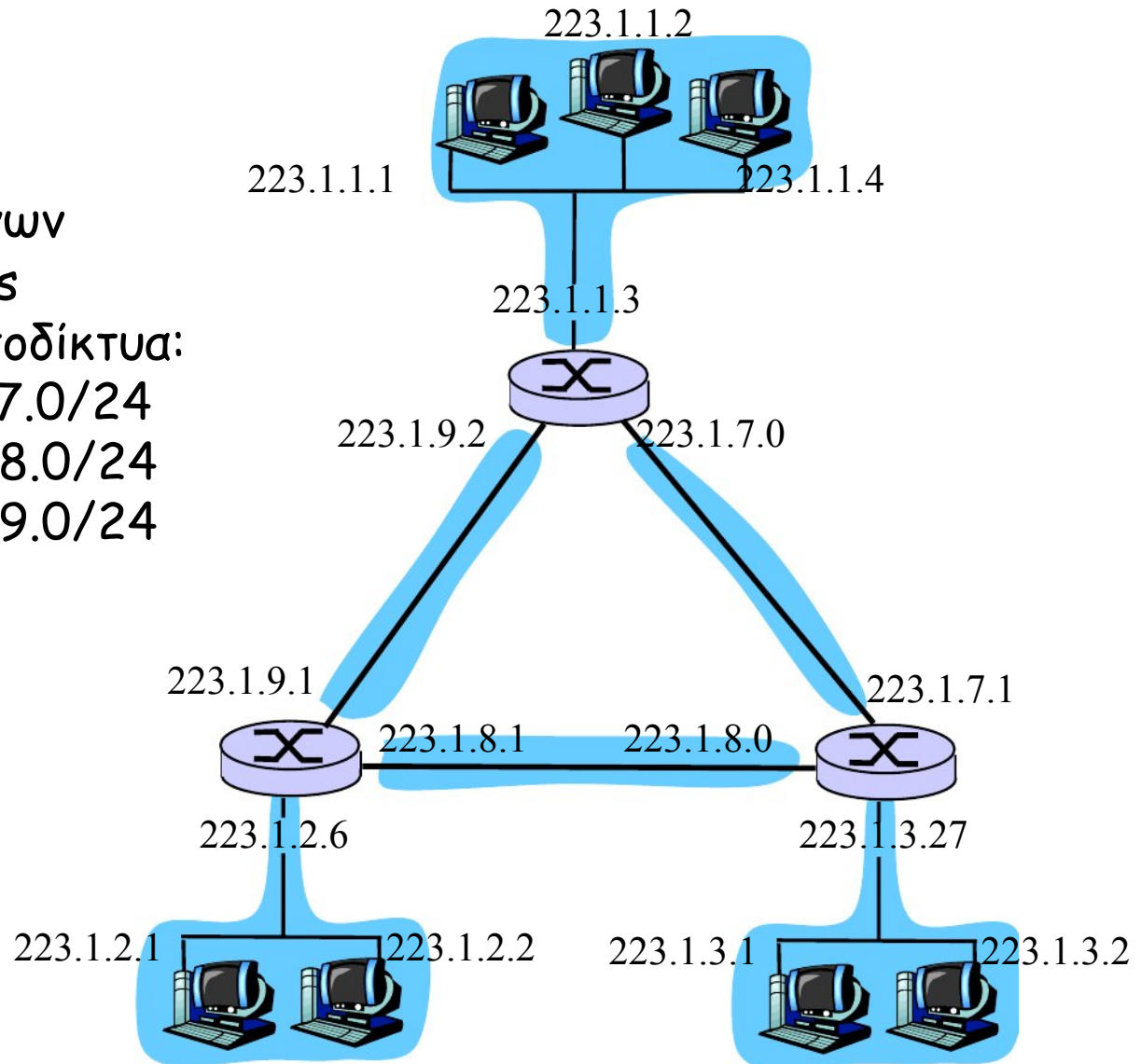
Πόσα υποδίκτυα;

Σύστημα διασυνδεδεμένων
δρομολογητών και hosts
αποτελούμενο από 6 υποδίκτυα:

223.1.1.0/24 223.1.7.0/24

223.1.2.0/24 223.1.8.0/24

223.1.3.0/24 223.1.9.0/24



Διευθύνσεις IP

αρχική αρχιτεκτονική διευθυνσιοδότησης Διαδικτύου

διευθυνσιοδότηση με βάση κλάσεις ("class-full"):

κλάση

A	0	δίκτυο		host		1.0.0.0 έως 127.255.255.255
B	10		δίκτυο		host	128.0.0.0 έως 191.255.255.255
C	110		δίκτυο		host	192.0.0.0 έως 223.255.255.255
D	1110		διεύθυνση multicast			224.0.0.0 έως 239.255.255.255

← 32 bits →

Διευθυνσιοδότηση IP: CIDR

Διευθυνσιοδότηση με βάση κλάσεις:

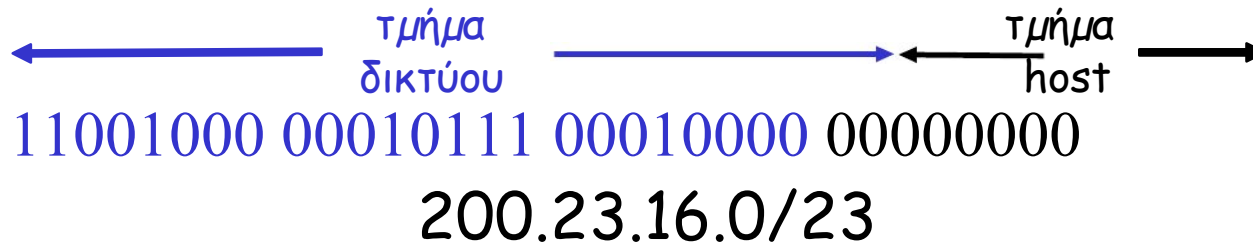
αναποτελεσματική εκχώρηση διευθύνσεων, γρήγορη εξάντληση διευθύνσεων κλάσης B

π.χ. σε δίκτυο κλάσης B εκχωρείται αριθμός διευθύνσεων για 65K hosts, ακόμη και αν στο δίκτυο περιλαμβάνονται μόνο 2K hosts

CIDR: Classless InterDomain Routing

τμήμα δικτύου διεύθυνσης έχει οποιοδήποτε μήκος

μορφή διεύθυνσης: **a.b.c.d/x**, όπου x ο αριθμός των bits στο τμήμα δικτύου της διεύθυνσης



π.χ. οργανισμός με 2000 hosts λαμβάνει μπλοκ διευθύνσεων με $2_{11}=2048$ διευθύνσεις hosts: a.b.c.d/21

Εκχώρηση διευθύνσεων σε ISPs

Ο οργανισμός **ICANN** (Internet **C**orporation for **A**ssigned **N**ames and **N**umbers)

διαχειρίζεται τις διευθύνσεις IP σε παγκόσμιο επίπεδο και εκχωρεί μπλοκ διευθύνσεων σε ISPs και οργανισμούς (μέσω RIPE, ARIN, APNIC)

διαχειρίζεται τους DNS root servers

κάνει ανάθεση ονομάτων τομέων (domain names) και επιλύει σχετικές διενέξεις

Εκχώρηση διεύθυνσης σε δίκτυο

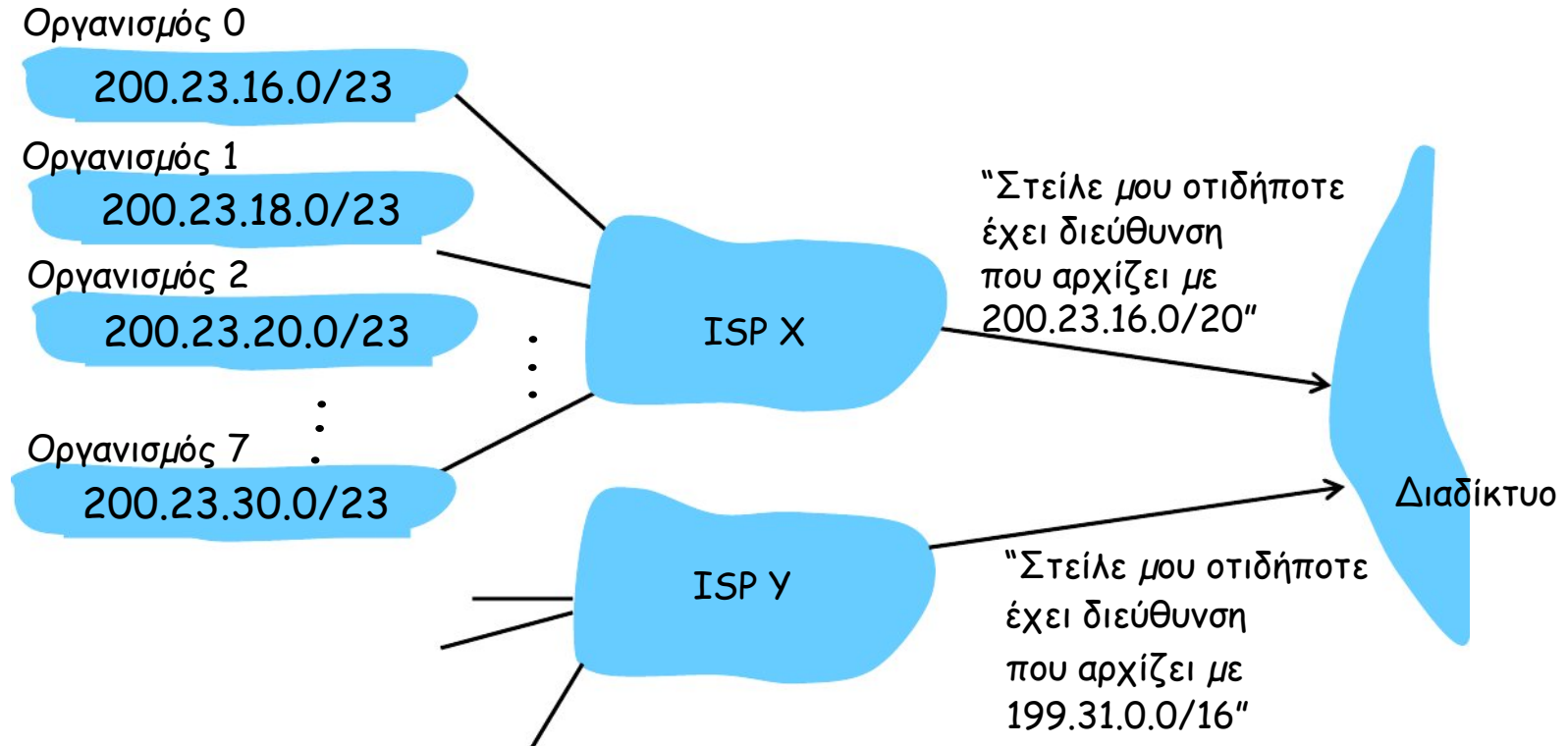
Ο διαχειριστής δικτύου ενός οργανισμού αποκτά
ένα μπλοκ διευθύνσεων από τον ISP του

μπλοκ διευθύνσεων του ISP:

	<u>11001000 00010111 00010000 00000000</u>	200.23.16.0/20
Οργανισμός 0	<u>11001000 00010111 00010000 00000000</u>	200.23.16.0/23
Οργανισμός 1	<u>11001000 00010111 00010010 00000000</u>	200.23.18.0/23
Οργανισμός 2	<u>11001000 00010111 00010100 00000000</u>	200.23.20.0/23
...		
Οργανισμός 7	<u>11001000 00010111 00011110 00000000</u>	200.23.30.0/23

Ιεραρχική διευθυνσιοδότηση: συνάθροιση διαδρομών

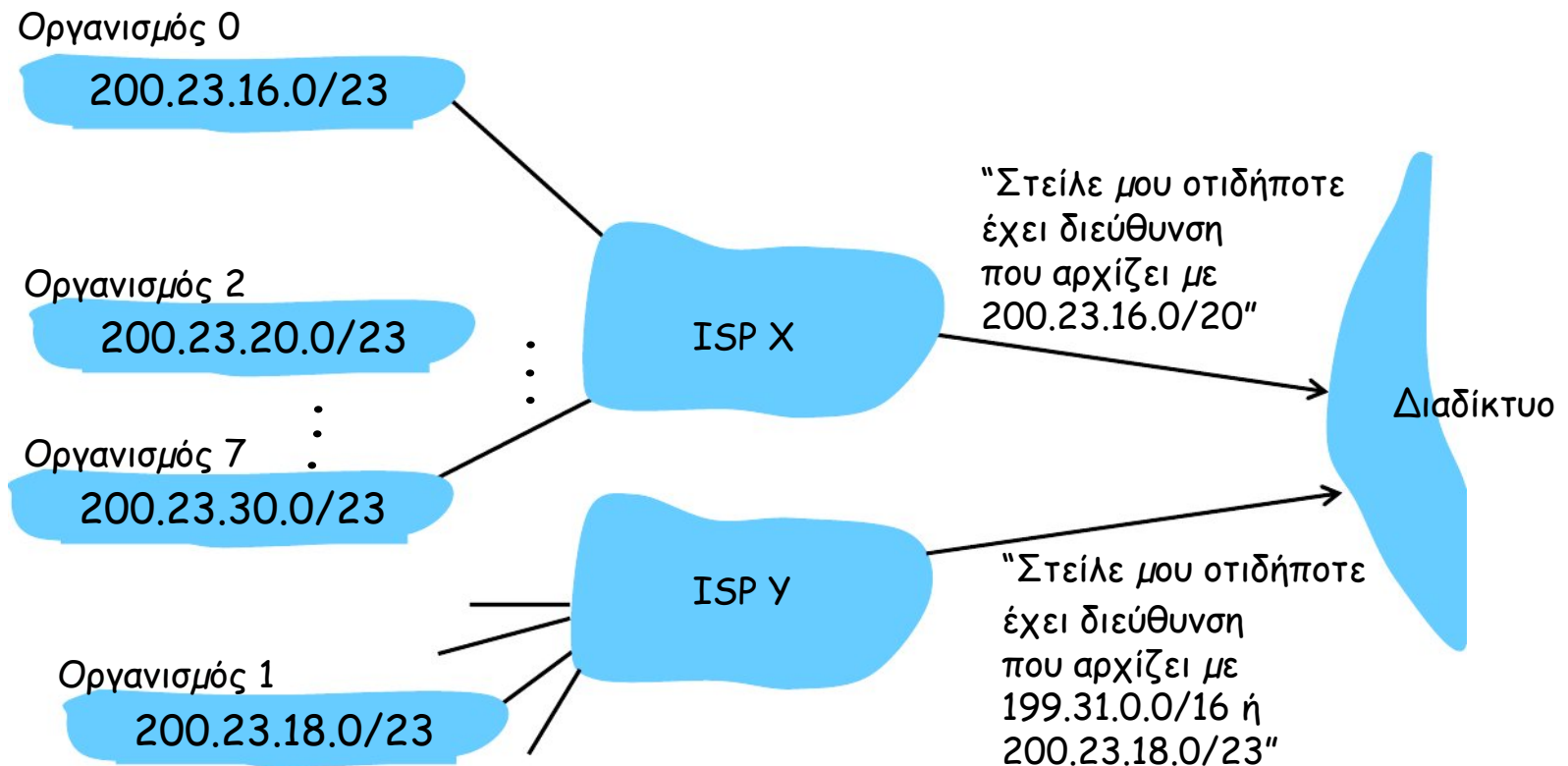
Η ιεραρχική διευθυνσιοδότηση επιτρέπει την αποτελεσματική ανακοίνωση (advertisement) πληροφορίας δρομολόγησης και διευκολύνει την ιεραρχική δρομολόγηση:



συνάθροιση διευθύνσεων ή διαδρομών (address or route aggregation): ικανότητα ανακοίνωσης πληροφορίας δρομολόγησης για πολλαπλά δίκτυα χρησιμοποιώντας ένα μόνο πρόθεμα δικτύου

Ιεραρχική διευθυνσιοδότηση: πιο συγκεκριμένες διαδρομές

Ο ISP Y έχει μία πιο συγκεκριμένη διαδρομή για τον Οργανισμό 1 από ότι ο ISP X



Εκχώρηση διεύθυνσης σε δρομολογητή ή host

δρομολογητής:

με παρέμβαση του διαχειριστή δικτύου στον δρομολογητή ή εξ αποστάσεως με λογισμικό διαχείρισης δικτύου

host:

με παρέμβαση του διαχειριστή δικτύου στο κατάλληλο αρχείο διευθέτησης:

Wintel: control-panel → network → configuration →

tcp/ip → properties

UNIX: /etc/rc.config

χρησιμοποιώντας το πρωτόκολλο **DHCP** (Dynamic **Host Configuration Protocol**)

αυτόματη απόκτηση διεύθυνσης από DHCP server - σύνδεση και άμεση λειτουργία ("plug-and-play")

(περισσότερα σε λίγο)

Μεταφορά datagram από πηγή σε προορισμό

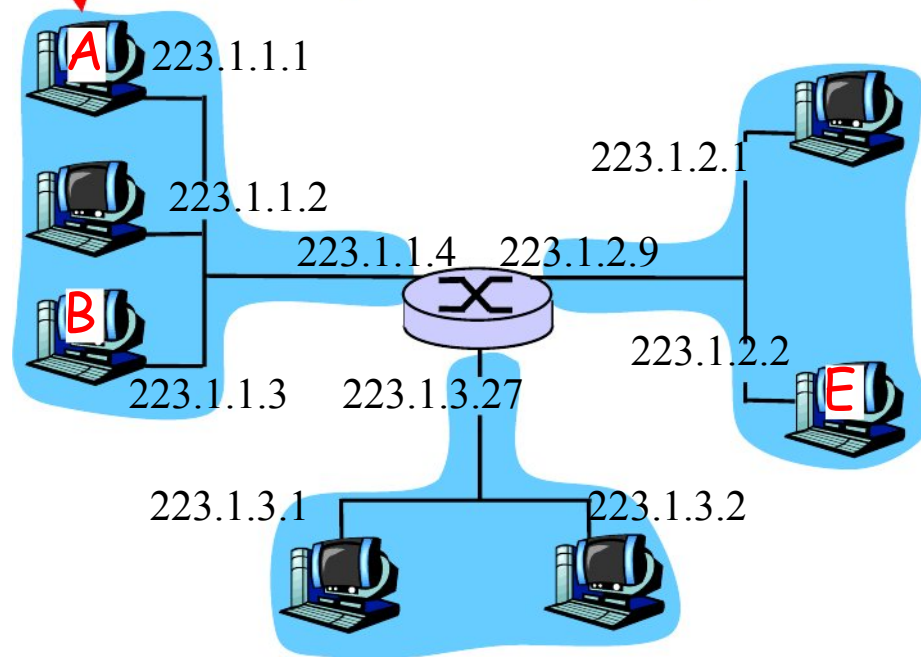
IP datagram:

διάφορα πεδία	διεύθυνση IP πηγής	διεύθυνση IP προορ.	δεδομένα
---------------	--------------------	---------------------	----------

τα τρία παραπάνω πεδία του datagram παραμένουν **αμετάβλητα** καθώς αυτό μεταφέρεται από την πηγή στον προορισμό

πίνακας προώθησης στον A

Δίκτυο προορισμού	Επόμενος Δρομολογητής	Αριθμός Αλμάτων
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2



Μεταφορά datagram από πηγή σε προορισμό

διάφορα πεδία	223.1.1.1	223.1.1.3	δεδομένα
---------------	-----------	-----------	----------

Ο host A στέλνει IP datagram στον host B:

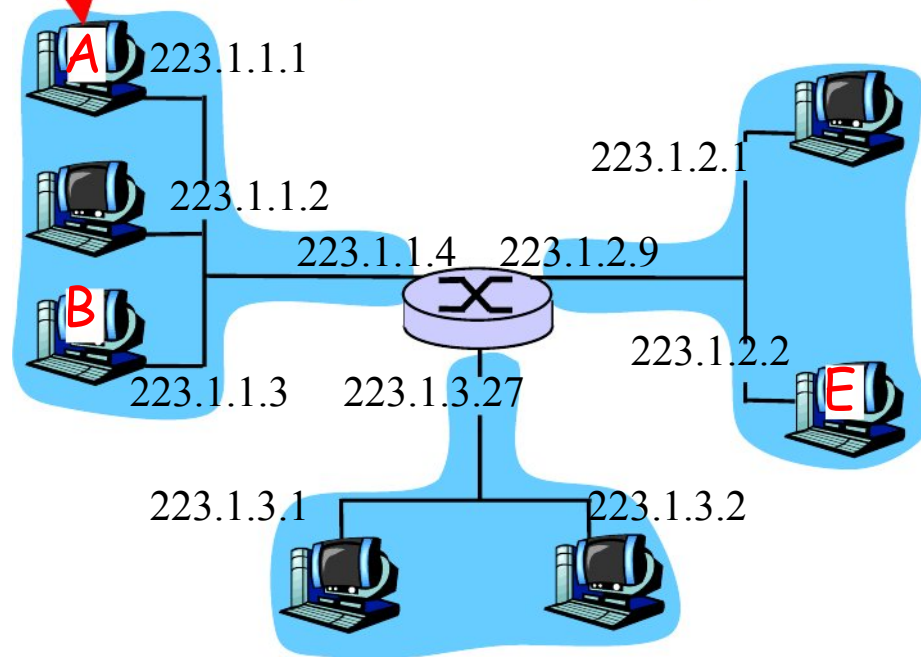
αναζήτηση διεύθυνσης δικτύου του B στον πίνακα προώθησης διαπιστώνεται ότι ο B βρίσκεται στο ίδιο δίκτυο με τον A

οι B και A συνδέονται απευθείας

το επίπεδο ζεύξης στέλνει το datagram απευθείας στον B μέσα σε ένα frame

πίνακας προώθησης στον A

Δίκτυο προορισμού	Επόμενος Δρομολογητής	Αριθμός Αλμάτων
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2



Μεταφορά datagram από πηγή σε προορισμό

διάφορα πεδία	223.1.1.1	223.1.2.2	δεδομένα
---------------	-----------	-----------	----------

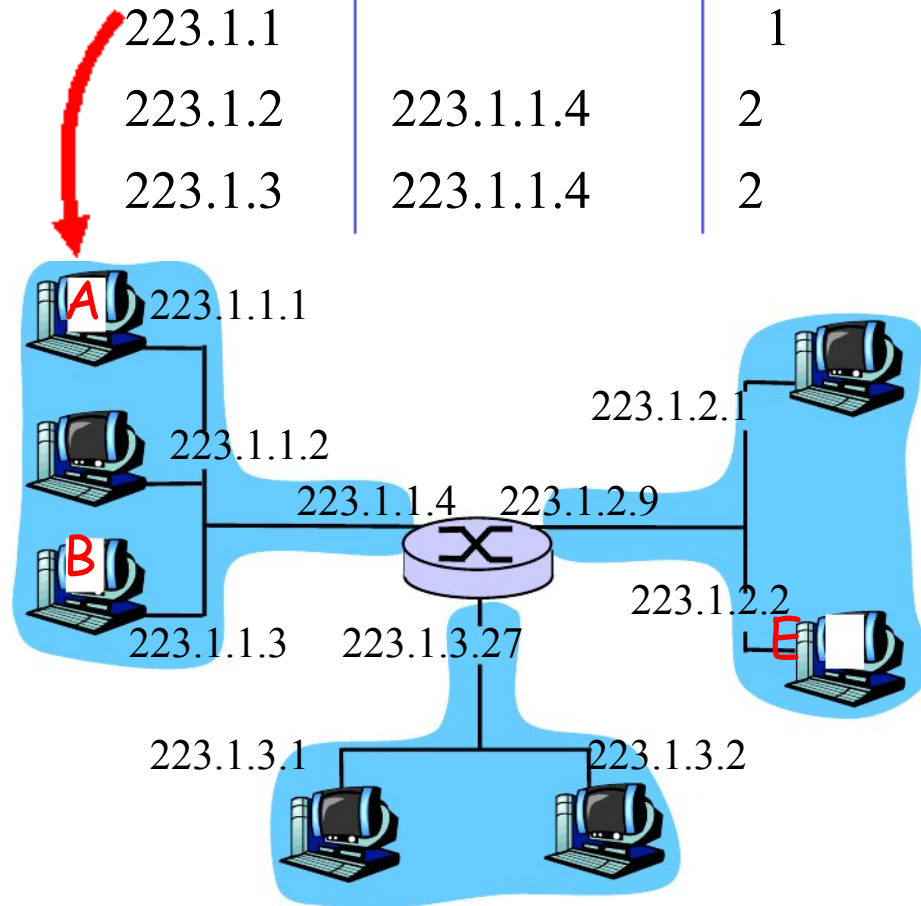
Πηγή Α, προορισμός Ε:

αναζήτηση διεύθυνσης δικτύου του Ε στον πίνακα προώθησης
Ε βρίσκεται σε διαφορετικό δίκτυο
οι Α και Ε δεν είναι απευθείας συνδεδεμένοι

πίνακας προώθησης: το επόμενο άλμα προς τον Ε είναι ο δρομολογητής 223.1.1.4
το επίπεδο ζεύξης στέλνει το datagram στον δρομολογητή 223.1.1.4 μέσα σε frame
το datagram φθάνει στο 223.1.1.4 συνεχίζεται ...

πίνακας προώθησης στον Α

Δίκτυο προορισμού	Επόμενος Δρομολογητής	Αριθμός Αλμάτων
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2



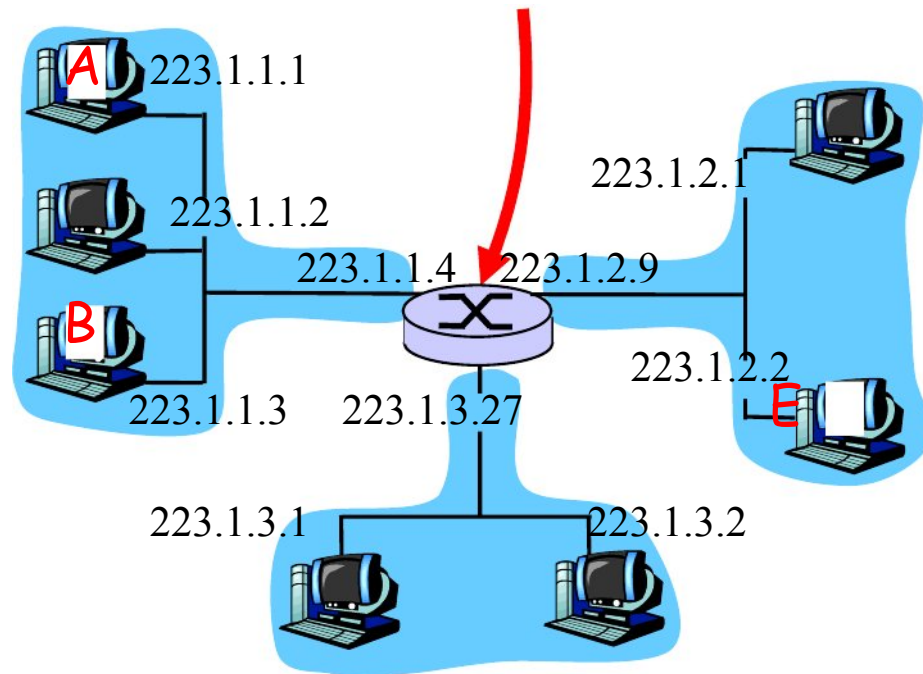
Μεταφορά datagram από πηγή σε προορισμό

διάφορα πεδία	223.1.1.1	223.1.2.2	δεδομένα
---------------	-----------	-----------	----------

πίνακας προώθησης στο δρομολογητή

Δίκτυο προορισμού	Επόμενος Δρομολογητής	Αριθμός Αλμάτων	Διεπαφή
223.1.1	-	1	223.1.1.4
223.1.2	-	1	223.1.2.9
223.1.3	-	1	223.1.3.27

datagram φθάνει στο 223.1.1.4,
προορίζεται για 223.1.2.2
αναζήτηση διεύθυνσης δικτύου του
Ε στον πίνακα προώθησης του
δρομολογητή
Ε βρίσκεται στο ίδιο δίκτυο με τη
διεπαφή 223.1.2.9 του δρομολογητή
δρομολογητής, Ε απευθείας
συνδεδεμένοι
επίπεδο ζεύξης στέλνει το
datagram στο 223.1.2.2 μέσα σε
frame μέσω διεπαφής 223.1.2.9
το datagram φθάνει στο 223.1.2.2



Δομή IP datagram (IPv4)

Πεδία:

version #: έκδοση πρωτοκόλλου IP

header length: μήκος επικεφαλίδας σε λέξεις των 32 bits

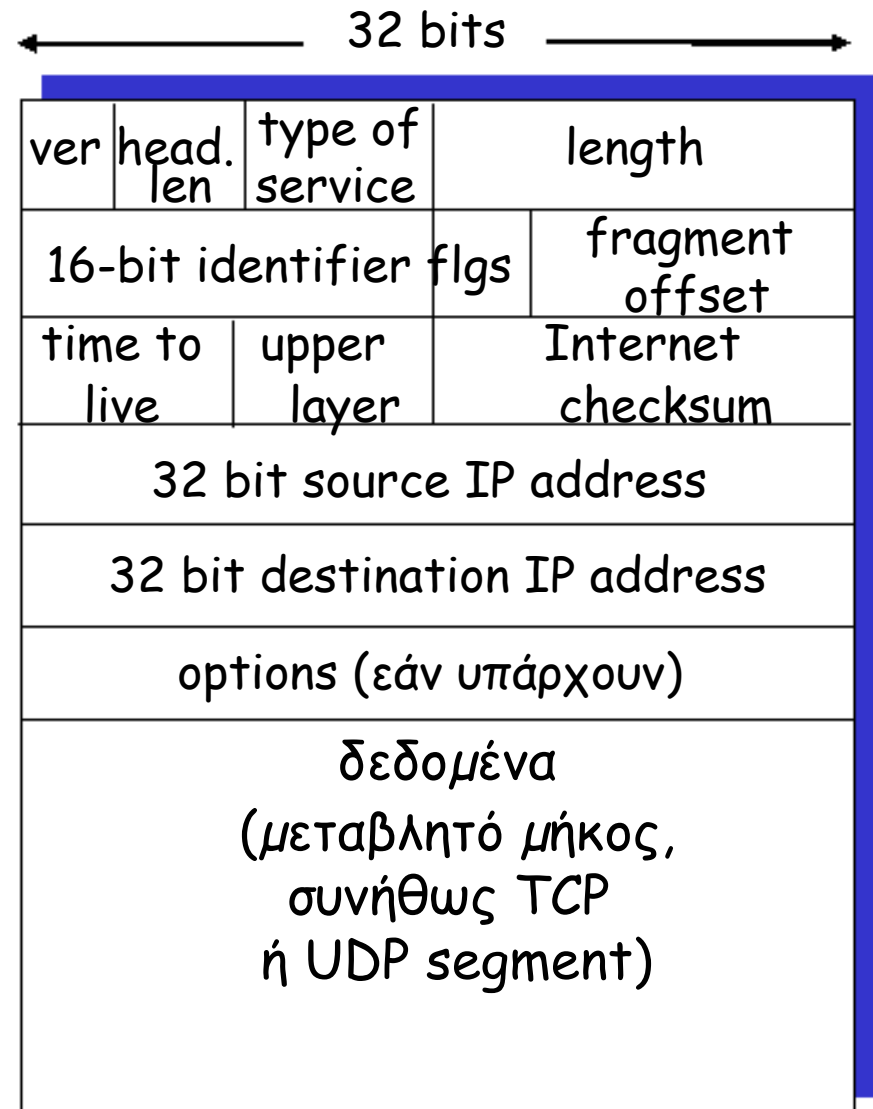
type of service: επιτρέπει να διακρίνουμε διαφορετικά είδη datagrams ως προς προτεραιότητα και απαιτήσεις της εφαρμογής

datagram length: συνολικό μήκος datagram σε bytes ≤ 65535 bytes

identifier, flags, IP fragmentation offset: αφορούν κατάτμηση και συναρμολόγηση

time to live: διασφαλίζει ότι ένα datagram δεν θα περιφέρεται άένα στο δίκτυο

μειώνεται κατά ένα σε κάθε άλμα



Δομή IP datagram (συνέχεια)

upper layer protocol:

υποδεικνύει το πρωτόκολλο επιπέδου μεταφοράς για το οποίο προορίζεται το datagram, π.χ.

6 □ TCP, 17 □ UDP, 1 □ ICMP

Internet checksum: ανίχνευση σφαλμάτων στην επικεφαλίδα του datagram

υπολογίζεται εκ νέου σε κάθε δρομολογητή

source/destination IP address:

διευθύνσεις IP πηγής, προορισμού
255.255.255.255: broadcast διεύθυνση IP

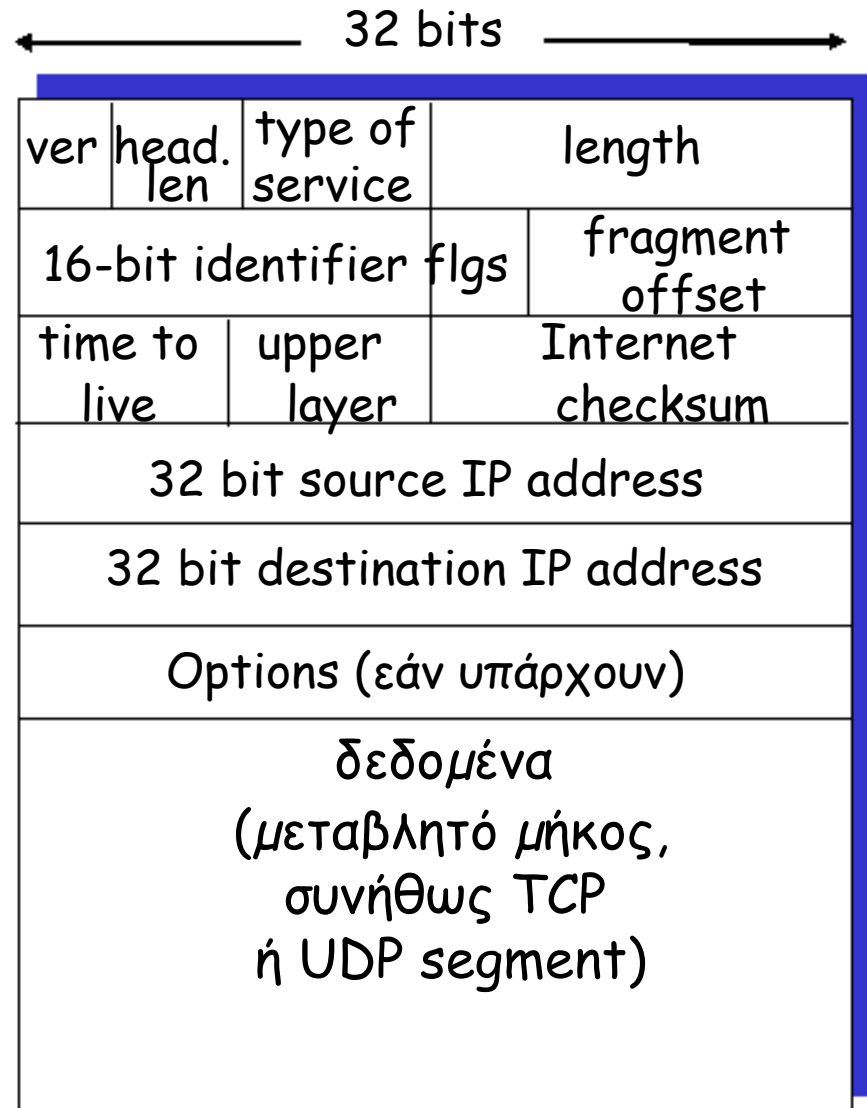
options: προαιρετικές επιλογές
π.χ. record route, timestamp

data: δεδομένα

π.χ. TCP ή UDP segment, μήνυμα ICMP

Επικεφαλίδα IP ε 20 bytes

Επικεφαλίδες IP και TCP ε 40 bytes



Κατάτμηση και συναρμολόγηση στο IP

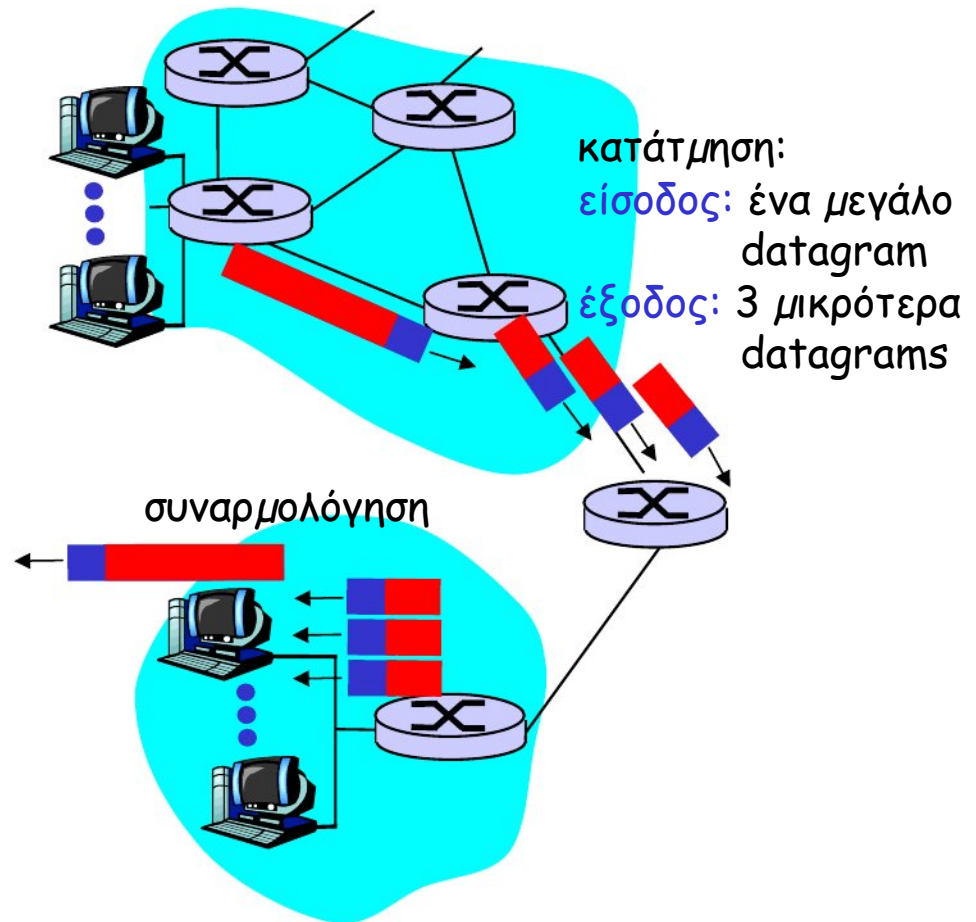
το πρωτόκολλο επιπέδου ζεύξης χαρακτηρίζεται από το μέγεθος MTU: μέγιστο δυνατό ωφέλιμο φορτίο στο frame (π.χ. Ethernet MTU = 1500 bytes)

διαφορετικά είδη ζεύξεων,
διαφορετικά μεγέθη MTU
μεγάλο IP datagram τεμαχίζεται
μέσα στο δίκτυο

από ένα datagram
προκύπτουν πολλαπλά
datagrams

“συναρμολόγηση” στον τελικό
προορισμό μόνο

τα πεδία της επικεφαλίδας
IP χρησιμοποιούνται για την
συναρμολόγηση των
τεμαχίων



Κατάτμηση και συναρμολόγηση στο IP

Παράδειγμα

μήκος αρχικού
datagram = 4000
bytes
MTU = 1500 bytes

length	ID	fragflag	offset	3980 bytes
=4000	=x	=0	=0	

Ένα μεγάλο datagram τεμαχίζεται
σε μερικά μικρότερα datagrams

length	ID	fragflag	offset	1480 bytes
=1500	=x	=1	=0	

length	ID	fragflag	offset	1480 bytes
=1500	=x	=1	=1480	

length	ID	fragflag	offset	1020 bytes
=1040	=x	=0	=2960	

Αποφυγή κατάτμησης με κατάλληλη επιλογή του μεγέθους των TCP/UDP segments:

MTU ε 576 bytes MSS = 576 - 20 - 20 = 536 bytes

ICMP: Internet Control Message Protocol

χρησιμοποιείται από hosts και δρομολογητές για να ανταλλάξουν μεταξύ τους πληροφορίες επιπέδου δικτύου

αναφορά σφαλμάτων: μη προσπελάσιμος host, δίκτυο, θύρα, πρωτόκολλο
echo request/reply
(χρησιμοποιείται από το ping)

επίπεδο δικτύου "επάνω" από IP:

τα μηνύματα ICMP μεταφέρονται μέσα σε IP datagrams

μήνυμα ICMP: περιέχει πεδία type, code, επικεφαλίδα IP και πρώτα 8 bytes του IP datagram που προκαλεί τη δημιουργία του μηνύματος ICMP

<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 πηγή στέλνει μία σειρά από UDP segments στον προορισμό

Το πρώτο έχει TTL = 1

Το δεύτερο έχει TTL= 2, κλπ.

με απίθανο αριθμό θύρας

Όταν το n-οστό datagram φθάσει στον n-οστό δρομολογητή:

Ο δρομολογητής απορρίπτει το datagram

Και στέλνει στην πηγή ένα μήνυμα ICMP (type 11, code 0 - TTL expired)

Το μήνυμα περιλαμβάνει το όνομα και τη διεύθυνση IP του δρομολογητή

Όταν λάβει το μήνυμα ICMP, η πηγή υπολογίζει το RTT

Το Traceroute επαναλαμβάνει αυτή τη διαδικασία 3 φορές

Κριτήριο τερματισμού

Το UDP segment φθάνει τελικά στον host προορισμού

Ο προορισμός επιστρέφει ένα μήνυμα ICMP "port unreachable" (type 3, code 3)

Όταν η πηγή λάβει αυτό το μήνυμα ICMP, σταματά να στέλνει datagrams

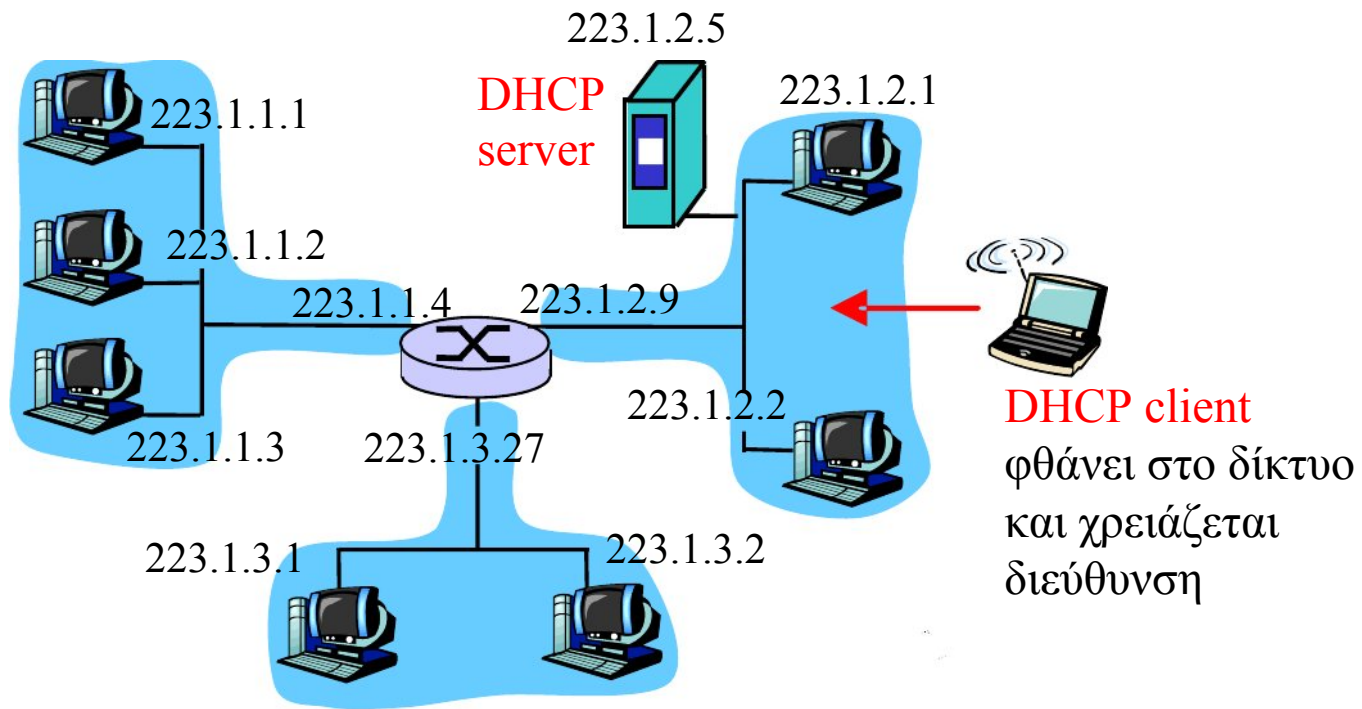
DHCP: Dynamic Host Configuration Protocol

Στόχος: να επιτρέψει στον host να αποκτήσει τη διεύθυνση IP του δυναμικά κατά τη σύνδεσή του με το δίκτυο
Μπορεί να ανανεώσει το χρόνο χρήσης της διεύθυνσης
Επιτρέπει ανακύκλωση των διευθύνσεων
Υποστηρίζει κινούμενους χρήστες που θέλουν να συνδεθούν με το δίκτυο

Επισκόπηση DHCP:

Ο host μεταδίδει με broadcast μήνυμα "DHCP discover"
Ο DHCP server αποκρίνεται με μήνυμα "DHCP offer"
Ο host ζητά διεύθυνση IP: μήνυμα "DHCP request"
Ο DHCP server στέλνει διεύθυνση: μήνυμα "DHCP ack"

Σενάριο DHCP client-server



Σενάριο DHCP client-server

DHCP server: 223.1.2.5

DHCP discover

src : 0.0.0.0, 68
dest.: 255.255.255.255, 67
yiaddr: 0.0.0.0
transaction ID: 654

αφικνούμενος
client



DHCP offer

src: 223.1.2.5, 67
dest: 255.255.255.255, 68
yiaddr: 223.1.2.4
transaction ID: 654
Lifetime: 3600 secs

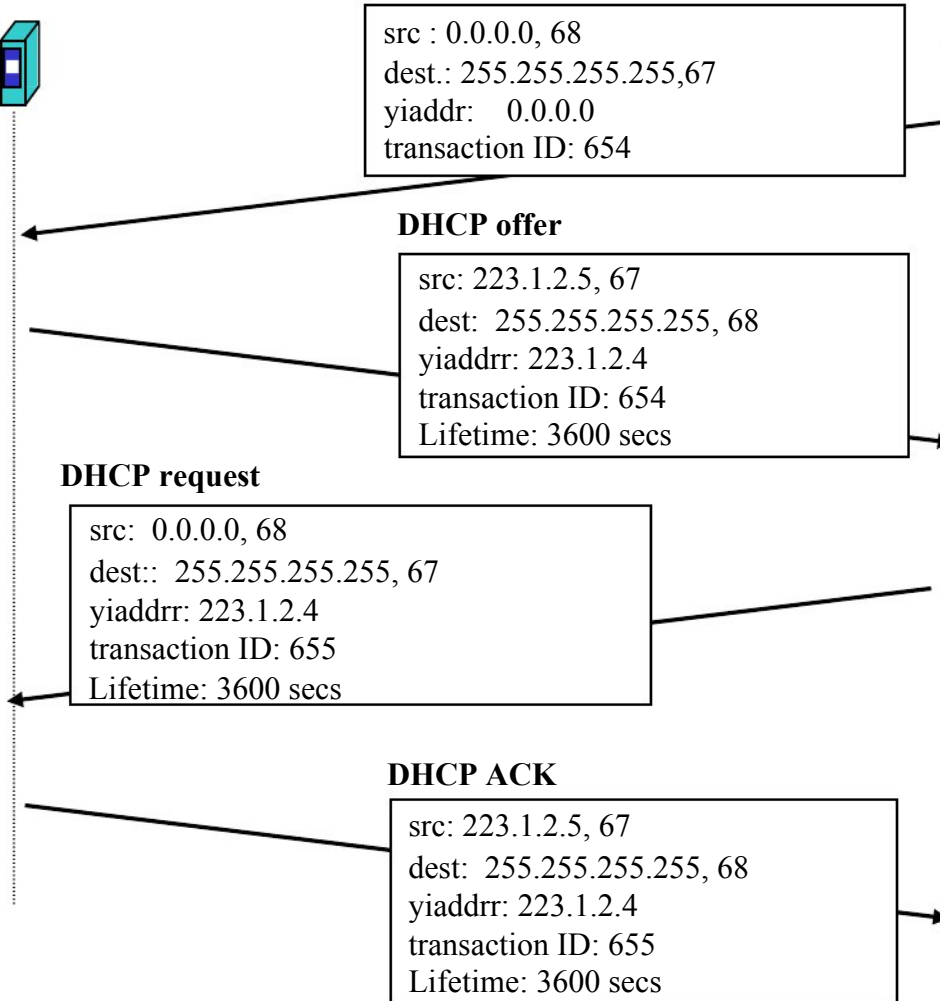
DHCP request

src: 0.0.0.0, 68
dest.: 255.255.255.255, 67
yiaddr: 223.1.2.4
transaction ID: 655
Lifetime: 3600 secs

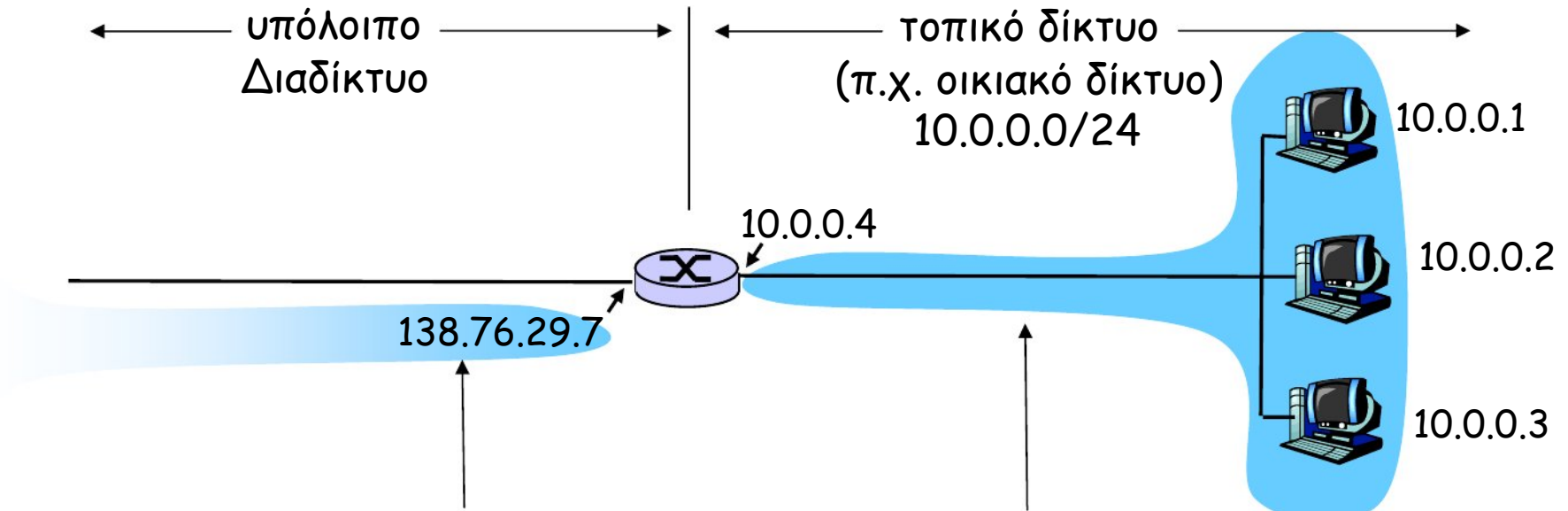
DHCP ACK

src: 223.1.2.5, 67
dest: 255.255.255.255, 68
yiaddr: 223.1.2.4
transaction ID: 655
Lifetime: 3600 secs

χρόνος



NAT: Network Address Translation



Όλα τα datagrams που φεύγουν από το τοπικό δίκτυο έχουν την ίδια μοναδική διεύθυνση IP: 138.76.29.7

Datagrams με πηγή ή προορισμό στο τοπικό δίκτυο έχουν διεύθυνση πηγής ή προορισμού 10.0.0.x

Ο δρομολογητής NAT "κρύβει" τις λεπτομέρειες του οικιακού δικτύου από τον έξω κόσμο

NAT: Network Address Translation

Κίνητρο: το τοπικό δίκτυο χρησιμοποιεί μία μόνο διεύθυνση IP όσον αφορά τον έξω κόσμο:

δεν χρειάζεται η εκχώρηση μπλοκ διευθύνσεων από τον ISP: χρησιμοποιείται μία μόνο διεύθυνση IP για όλες τις συσκευές

αλλαγή των διευθύνσεων των συσκευών του τοπικού δικτύου χωρίς να ενημερωθεί ο έξω κόσμος

αλλαγή του ISP χωρίς να αλλάξουν οι διευθύνσεις των συσκευών του τοπικού δικτύου

οι συσκευές στο εσωτερικό του τοπικού δικτύου δεν είναι άμεσα διευθυνσιοδοτήσιμες, "ορατές" από τον έξω κόσμο (πλεονέκτημα ασφάλειας)

NAT: Network Address Translation

Υλοποίηση

Ο δρομολογητής NAT πρέπει να:

αντικαθιστά το ζεύγος (διεύθυνση IP πηγής, # θύρας) σε κάθε εξερχόμενο datagram με το ζεύγος (διεύθυνση IP NAT, νέος # θύρας)

... ένας απομακρυσμένος client ή server θα αποκριθεί χρησιμοποιώντας ως διεύθυνση, αριθμό θύρας προορισμού το ζεύγος (διεύθυνση IP NAT, νέος # θύρας)

αποθηκεύει (στον πίνακα μετάφρασης NAT) κάθε αντιστοιχία ζευγών (διεύθυνση IP πηγής, # θύρας) - (διεύθυνση IP NAT, νέος # θύρας)

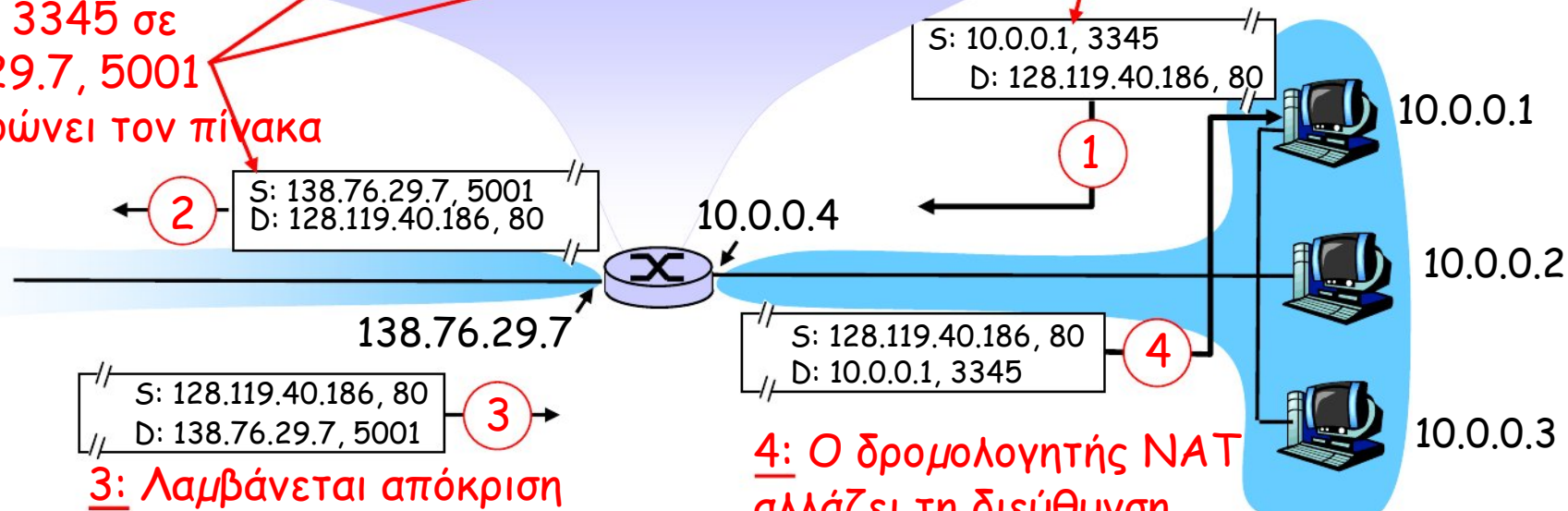
αντικαθιστά το ζεύγος (διεύθυνση IP NAT, νέος # θύρας) των πεδίων προορισμού σε κάθε εισερχόμενο datagram με το αντίστοιχο ζεύγος (διεύθυνση IP πηγής, # θύρας) που βρίσκεται αποθηκευμένο στον πίνακα NAT

NAT: Network Address Translation

2: Ο δρομολογητής NAT αλλάζει τη διεύθυνση πηγής του datagram από 10.0.0.1, 3345 σε 138.76.29.7, 5001 & ενημερώνει τον πίνακα

Πίνακα μετάφρασης NAT	
Διεύθυνση (WAN)	Διεύθυνση (LAN)
138.76.29.7, 5001	10.0.0.1, 3345
.....	

1: Ο host 10.0.0.1 στέλνει datagram με 128.119.40, 80



3: Λαμβάνεται απόκριση με 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

Ο αριθμός θύρας είναι πεδίο με 16 bits:

δυνατότητα για 60000 ταυτόχρονες συνδέσεις με μία μόνο διεύθυνση στην πλευρά του WAN

Η τεχνική NAT δέχεται επικρίσεις:

παραβιάζει το αρχιτεκτονικό μοντέλο του IP όπου η διεύθυνση που αντιστοιχεί σε κάθε συσκευή είναι παγκοσμίως μοναδική

- το λογισμικό του Διαδικτύου στηρίζεται στο μοντέλο αυτό

μεταβάλλει τον ασυνδεσμικό (connectionless) χαρακτήρα του Διαδικτύου εισάγοντας συνδεσμικά στοιχεία

- το NAT διατηρεί πληροφορία κατάστασης για τις συνδέσεις (αστοχία στο NAT σημαίνει καταστροφή όλων των συνδέσεων)

παραβιάζει τον θεμελιώδη κανόνα της απομόνωσης και ανεξαρτησίας των επιπέδων

το έλλειμμα διευθύνσεων πρέπει να αντιμετωπισθεί χρησιμοποιώντας IPv6

Περίγραμμα

- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
- 4 Αρχές Δρομολόγησης
 - Δρομολόγηση Κατάστασης Ζεύξης
 - Δρομολόγηση Ανύσματος Απόστασης
- 5 Δρομολόγηση στο Διαδίκτυο
- 6 IPv6

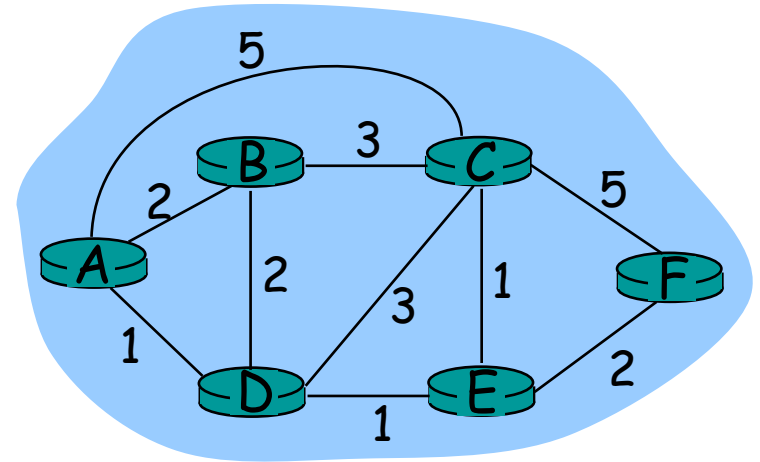
Routing

Routing protocol

Goal: determine "good" path (sequence of routers) thru network from source to dest.

Graph abstraction for routing algorithms:

- graph nodes are routers
- graph edges are physical links
 - link cost: delay, \$ cost, or congestion level



- "good" path:
 - typically means minimum cost path
 - other def's possible

Routing Algorithm classification

Global or decentralized information?

Global:

- all routers have complete topology, link cost info
- "link state" algorithms

Decentralized:

- router knows physically-connected neighbors, link costs to neighbors
- iterative process of computation, exchange of info with neighbors
- "distance vector" algorithms

Static or dynamic?

Static:

- routes change slowly over time

Dynamic:

- routes change more quickly
 - periodic update
 - in response to link cost changes

A Link-State Routing Algorithm

Dijkstra's algorithm

- net topology, link costs known to all nodes
 - accomplished via "link state broadcast"
 - all nodes have same info
- computes least cost paths from one node ('source') to all other nodes
 - gives **routing table** for that node
- iterative: after k iterations, know least cost path to k dest.'s

Notation:

- $c(i,j)$: link cost from node i to j . cost infinite if not direct neighbors
- $D(v)$: current value of cost of path from source to dest. V
- $p(v)$: predecessor node along path from source to v , that is next v
- N : set of nodes whose least cost path definitively known

Dijkstra's Algorithm

1 **Initialization:**

2 $N = \{A\}$

3 for all nodes v

4 if v adjacent to A

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

6 else $D(v) = \text{infinity}$

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

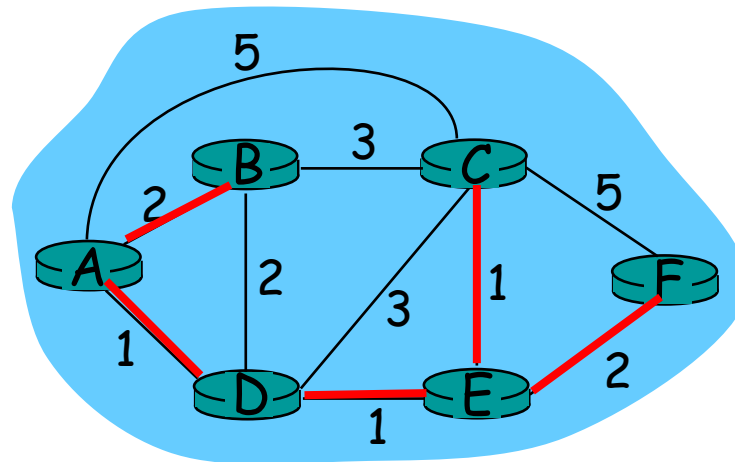
14 shortest path cost to w plus cost from w to v */

15 **until all nodes in N**



Dijkstra's algorithm: example

Step	start N	D(B),p(B)	D(C),p(C)	D(D),p(D)	D(E),p(E)	D(F),p(F)
→ 0	A	2,A	5,A	1,A	infinity	infinity
→ 1	AD	2,A	4,D		2,D	infinity
→ 2	ADE	2,A	3,E			4,E
→ 3	ADEB		3,E			4,E
→ 4	ADEBC					4,E
5	ADEBCF					



Dijkstra's algorithm, discussion

Algorithm complexity: n nodes

- each iteration: need to check all nodes, w , not in N
- $n*(n+1)/2$ comparisons: $O(n^2)$

Distance Vector Routing Algorithm

iterative:

- continues until no nodes exchange info.
- *self-terminating*: no "signal" to stop

asynchronous:

- nodes need *not* exchange info/iterate in lock step!

distributed:

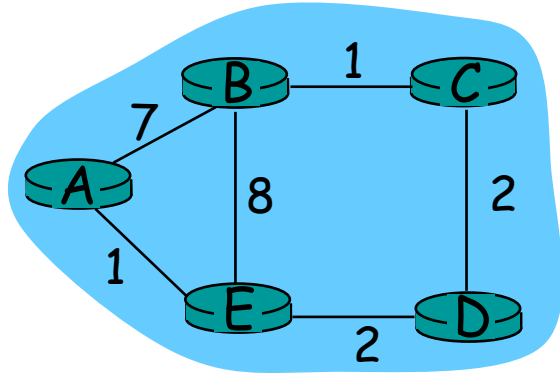
- each node communicates *only* with directly-attached neighbors

Distance Table data structure

- each node has its own
- row for each possible destination
- column for each directly-attached neighbor to node
- example: in node X, for dest. Y via neighbor Z:

$$\begin{aligned} D^X(Y, Z) &= \text{distance from X to Y, via Z as next hop} \\ &= c(X, Z) + \min_w \{D^Z(Y, w)\} \end{aligned}$$

Distance Table: example



$$D^E(C,D) = c(E,D) + \min_w \{D^D(C,w)\} \\ = 2+2 = 4$$

$$D^E(A,D) = c(E,D) + \min_w \{D^D(A,w)\} \\ = 2+3 = 5 \text{ loop!}$$

$$D^E(A,B) = c(E,B) + \min_w \{D^B(A,w)\} \\ = 8+6 = 14 \text{ loop!}$$

cost to destination via

$D^E()$	A	B	D
A	1	14	5
B	7	8	5
C	6	9	4
D	4	11	2

destination

Distance table gives routing table

		cost to destination via		
destination	$D^E()$	A	B	D
	A	1	14	5
	B	7	8	5
	C	6	9	4
	D	4	11	2

		Outgoing link to use, cost	
destination	A	A	1
	B	D	5
	C	D	4
	D	D	2

Distance table $\longrightarrow$ Routing table

Distance Vector Routing: overview

Iterative, asynchronous:

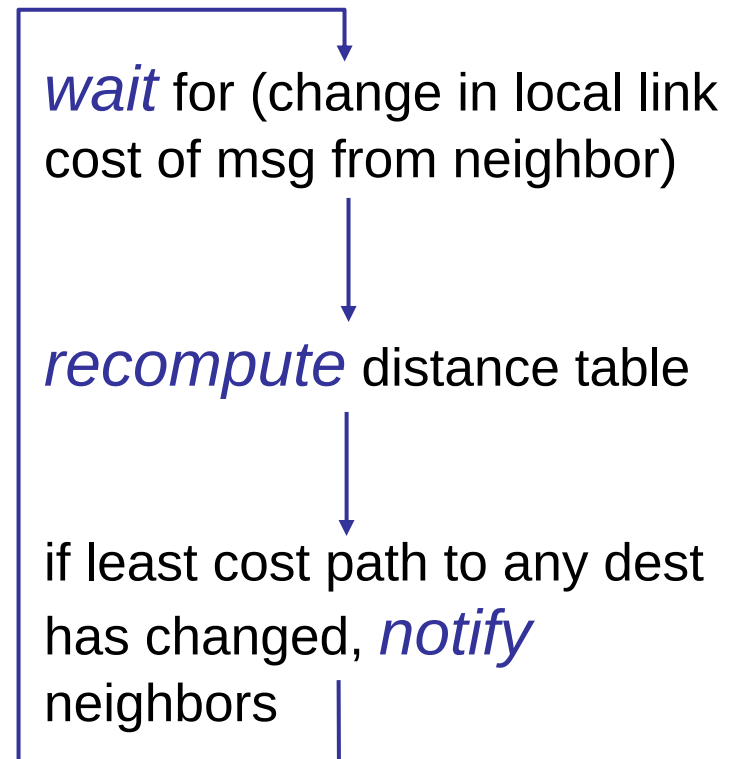
each local iteration caused by:

- local link cost change
- message from neighbor: its least cost path change from neighbor

Distributed:

- each node notifies neighbors *only* when its least cost path to any destination changes
 - neighbors then notify their neighbors if necessary

Each node:



Distance Vector Algorithm:

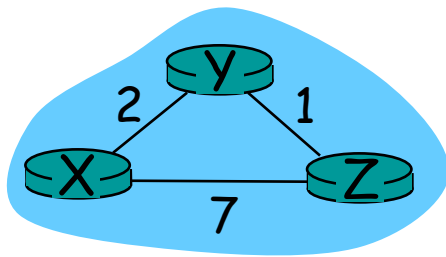
At all nodes, X:

- 1 Initialization:
- 2 for all adjacent nodes v:
- 3 $D^X(*,v) = \text{infinity}$ /* the * operator means "for all rows" */
- 4 $D^X(v,v) = c(X,v)$
- 5 for all destinations, y
- 6 send $\min_w D^X(y,w)$ to each neighbor /* w over all X's neighbors */

Distance Vector Algorithm (cont.):

```
8 loop
9   wait (until I see a link cost change to neighbor V
10      or until I receive update from neighbor V)
11
12   if (c(X,V) changes by d)
13     /* change cost to all dest's via neighbor v by d */
14     /* note: d could be positive or negative */
15     for all destinations y:  $D^X(y,V) = D^X(y,V) + d$ 
16
17   else if (update received from V wrt destination Y)
18     /* shortest path from V to some Y has changed */
19     /* V has sent a new value for its  $\min_w DV(Y,w)$  */
20     /* call this received new value is "newval" */
21     for the single destination y:  $D^X(Y,V) = c(X,V) + \text{newval}$ 
22
23   if we have a new  $\min_w D^X(Y,w)$  for any destination Y
24     send new value of  $\min_w D^X(Y,w)$  to all neighbors
25
26 forever
```

Distance Vector Algorithm: example



		cost via	
		Y	Z
destination	X		
	Y	2	∞
	Z	∞	7

		cost via	
		Y	Z
destination	X		
	Y	2	8
	Z	3	7

		cost via	
		Y	Z
destination	X		
	Y		
	Z		

		cost via	
		X	Z
destination	Y		
	X	2	∞
	Z	∞	1

		cost via	
		X	Z
destination	Y		
	X	2	8
	Z	9	1

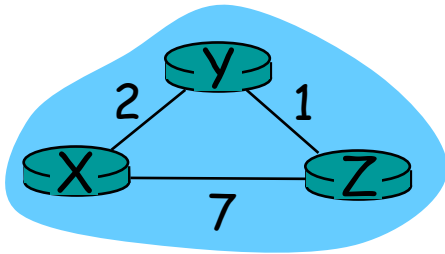
		cost via	
		X	Z
destination	Y		
	X		
	Z		

		cost via	
		X	Y
destination	Z		
	X	7	∞
	Y	∞	1

		cost via	
		X	Y
destination	Z		
	X	7	3
	Y	9	1

		cost via	
		X	Y
destination	Z		
	X		
	Y		

Distance Vector Algorithm: example



		cost via	
		Y	Z
d e s t	X		
	Y	2	∞
	Z	∞	7

		cost via	
		X	Z
d e s t	Y		
	X	2	∞
	Z	∞	1

		cost via	
		X	Y
d e s t	Z		
	X	7	∞
	Y	∞	1

		cost via	
		Y	Z
d e s t	X		
	Y	2	8
	Z	3	7

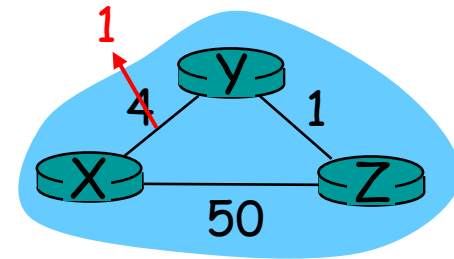
$$D^X(Y, Z) = c(X, Z) + \min_w \{D^Z(Y, w)\} \\ = 7 + 1 = 8$$

$$D^X(Z, Y) = c(X, Y) + \min_w \{D^Y(Z, w)\} \\ = 2 + 1 = 3$$

Distance Vector: link cost changes

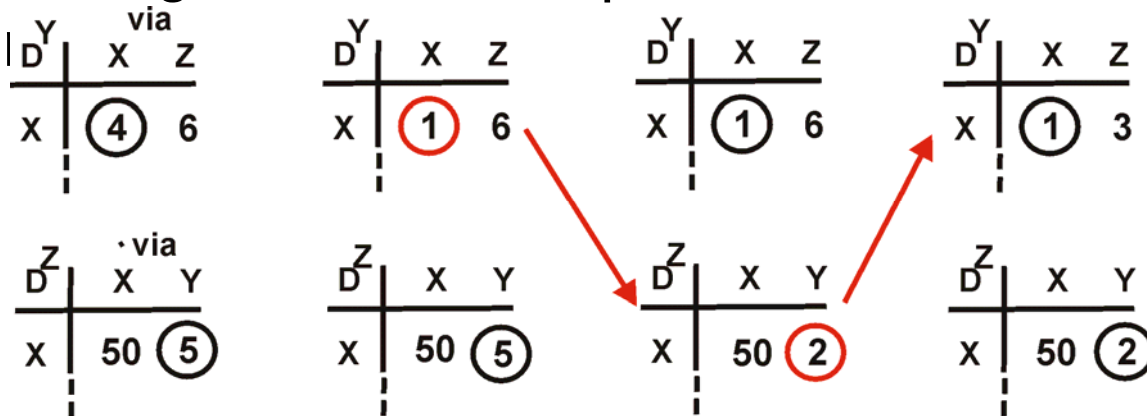
Link cost changes:

- node detects local link cost change
- updates distance table (line 15)
- if cost change in least cost path, notify

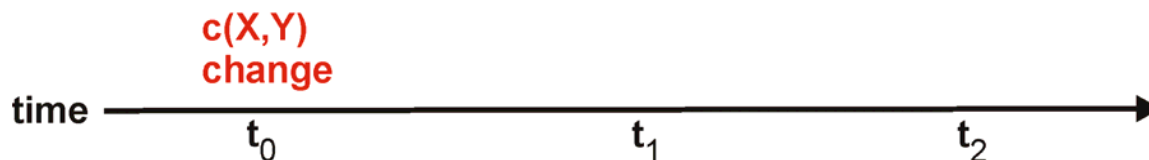


“good news travels fast”

notify



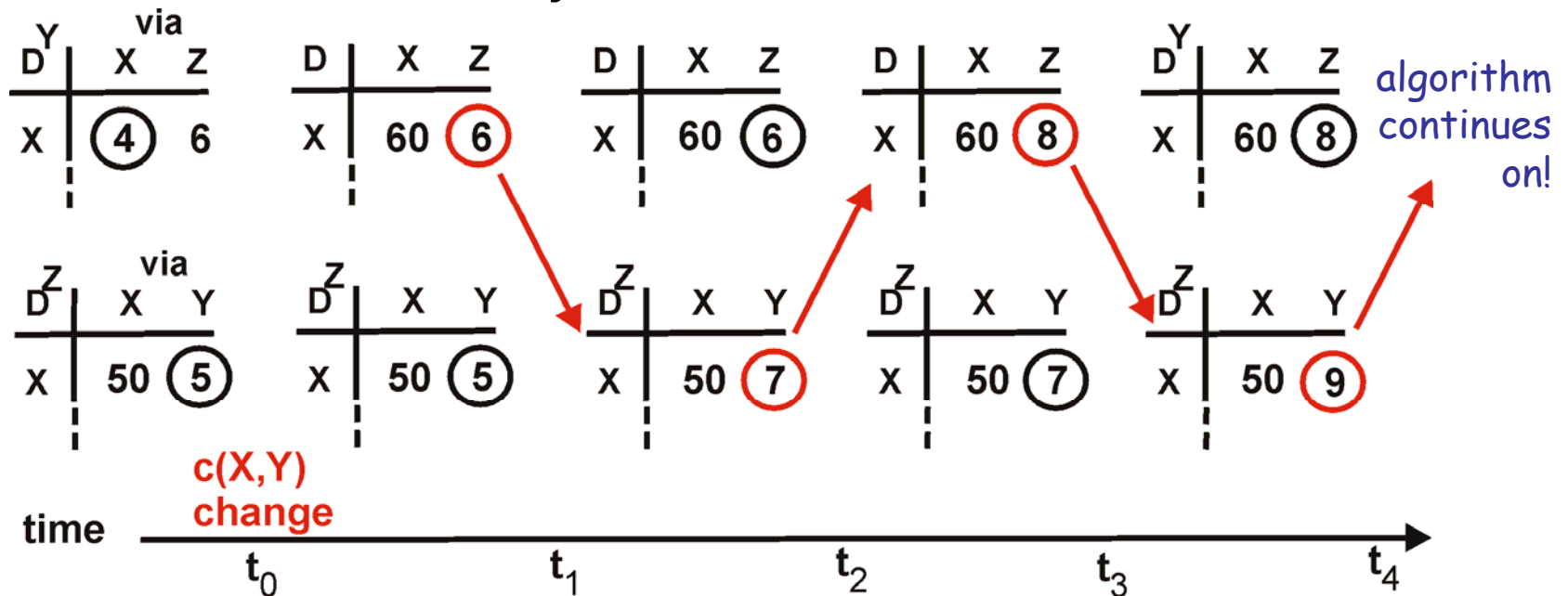
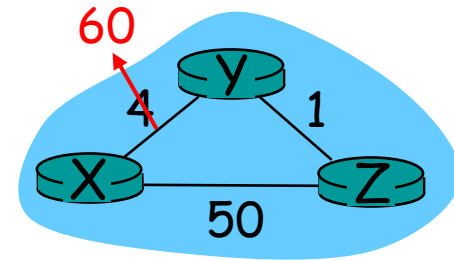
algorithm terminates



Distance Vector: link cost changes

Link cost changes:

- good news travels fast
- bad news travels slow - "count to infinity"

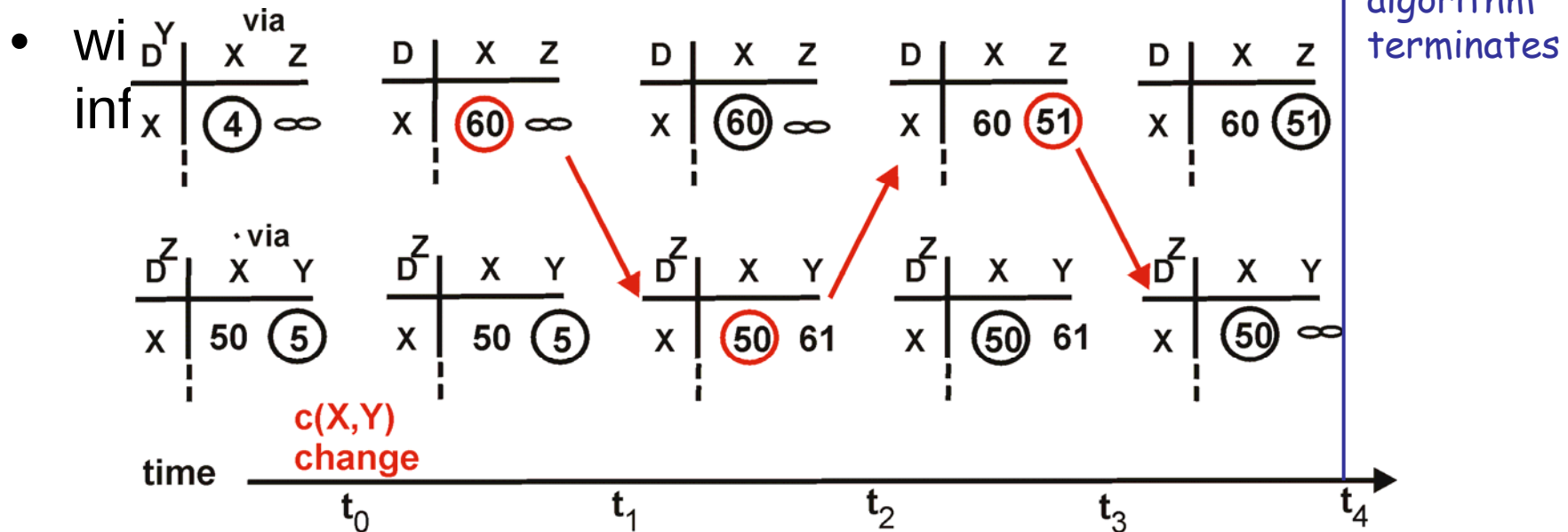
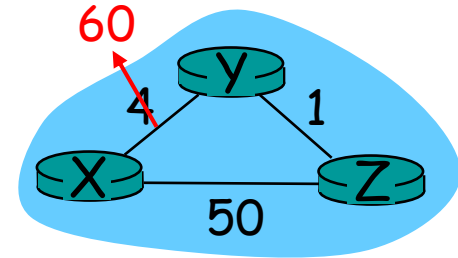


Distance Vector: poisoned reverse

If Z routes through Y to get to X

:

- Z tells Y its (Z's) distance to X is infinite (so Y won't route to X via Z)



Comparison of LS and DV algorithms

Message complexity

- LS: with n nodes, E links, $O(nE)$ msgs sent each
- DV: exchange between neighbors only
 - convergence time varies

Speed of Convergence

- LS: $O(n^2)$ algorithm requires $O(nE)$ msgs
 - may have oscillations
- DV: convergence time varies
 - may be routing loops
 - count-to-infinity problem

Robustness: what happens if router malfunctions?

LS:

- node can advertise incorrect *link* cost
- each node computes only its *own* table

DV:

- DV node can advertise incorrect *path* cost
- each node's table used by others
 - error propagation through network

Περίγραμμα

- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
- 4 Αρχές Δρομολόγησης
- 5 Δρομολόγηση στο Διαδίκτυο
 - 5.1 Δρομολόγηση εντός ΑΣ: RIP και OSPF
 - 5.2 Δρομολόγηση μεταξύ ΑΣ: BGP
- 6 IPv6
- 7 Ειδικά θέματα του στρώματος Δικτύου

Δρομολόγηση στο Διαδίκτυο

Το Παγκόσμιο Διαδίκτυο αποτελείται από **Αυτόνομα Συστήματα (ΑΣ)** που διασυνδέονται μεταξύ τους

Τα αυτόνομα συστήματα διακρίνονται σε:

Stub: μία μόνο σύνδεση με ένα άλλο ΑΣ - μεταφέρει τοπική κίνηση που προέρχεται από ή προορίζεται για το ίδιο

Multihomed: πολλαπλές συνδέσεις με περισσότερα από ένα άλλα ΑΣ - μεταφέρει τοπική κίνηση, δεν επιτρέπει διερχόμενη κίνηση

Transit: πάροχος που συνδέει πολλά ΑΣ μεταξύ τους - μεταφέρει τοπική και διερχόμενη κίνηση

Δρομολόγηση σε δύο επίπεδα:

Εντός ΑΣ: διαχειριστής υπεύθυνος για την επιλογή του αλγορίθμου δρομολόγησης μέσα στο δίκτυο

Μεταξύ ΑΣ: ένα μόνο πρότυπο δρομολόγησης μεταξύ ΑΣ: BGP

Πρωτόκολλα δρομολόγησης εντός ΑΣ

Γνωστά επίσης ως Πρωτόκολλα Εσωτερικής Πύλης -
Interior Gateway Protocols (IGP)

Πιο διαδεδομένα πρωτόκολλα δρομολόγησης εντός ΑΣ:

RIP: Routing Information Protocol

- αρχικό πρωτόκολλο δρομολόγησης στο ARPANET

OSPF: Open Shortest Path First

- συνιστώμενο πρωτόκολλο δρομολόγησης στο Internet

IGRP: Interior Gateway Routing Protocol

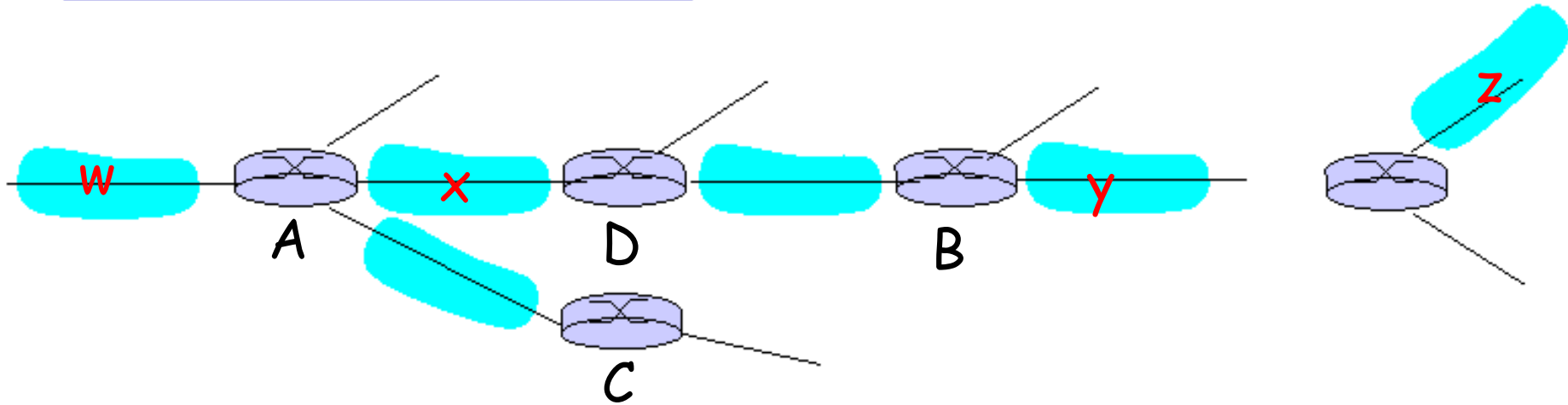
- ιδιοταγές πρωτόκολλο της Cisco

RIP (Routing Information Protocol)

συμπεριλήφθηκε στην έκδοση BSD του UNIX το 1982
αλγόριθμος διανύσματος αποστάσεων (distance vector)
μετρική απόστασης: αριθμός αλμάτων (μέγιστο = 15 άλματα)
τα διανύσματα αποστάσεων ανταλλάσσονται μεταξύ
γειτόνων κάθε 30 sec μέσω μηνυμάτων απόκρισης RIP
(RIP response messages), ή αλλιώς, **αγγελίες RIP (RIP
advertisements)**

Κάθε αγγελία RIP μπορεί να περιέχει μία λίστα με καταχωρήσεις
για έως 25 δίκτυα προορισμού μέσα στο AS

RIP: Παράδειγμα



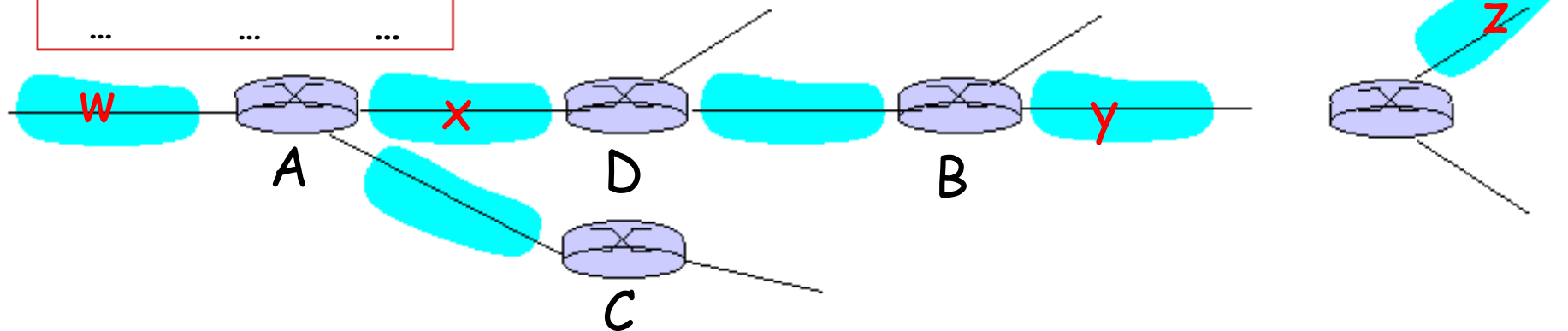
Δίκτυο Προορισμού	Επόμενος Δρομολογητής	Αριθμός αλμάτων ως τον προορισμό
W	A	2
Y	B	2
Z	B	7
x	-	1
...	...	...

Πίνακας δρομολόγησης στον D

RIP: Παράδειγμα

Προορ	Επομ	Άλμ
w	-	1
x	-	1
z	C	4
...	...	...

Αγγελία από
τον A στον D



Δίκτυο Προορισμού	Επόμενος Δρομολογητής	Αριθμός αλμάτων ως τον προορισμό
w	A	2
y	B	2
z	B A	7 5
x	-	1
...	...	...

Πίνακας δρομολόγησης στον D

RIP: Αστοχία ζεύξεων και ανάκαμψη

Εάν ένας δρομολογητής δεν λάβει αγγελία RIP από ένα γείτονα μέσα σε 180 sec, θεωρεί ότι ο γείτονας δεν είναι πλέον προσεγγίσιμος

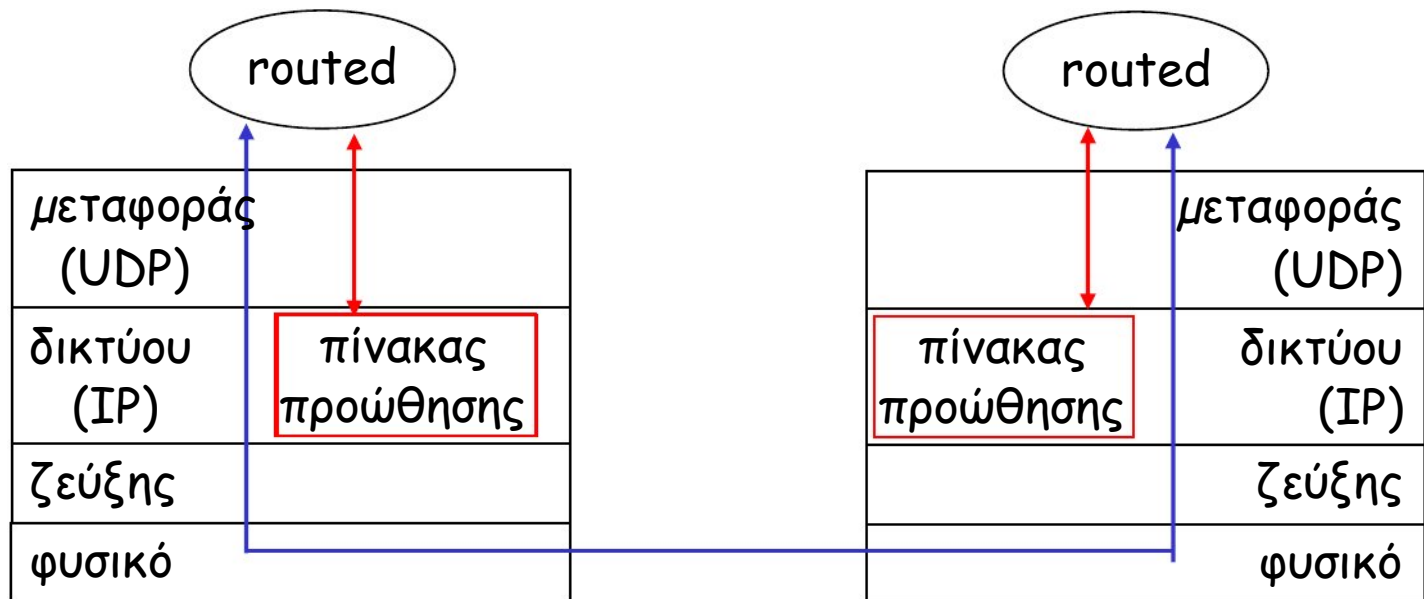
ακυρώνει τις διαδρομές που διέρχονται από το γείτονα
στέλνει νέες αγγελίες στους γείτονες

οι γείτονες με τη σειρά τους στέλνουν νέες αγγελίες
(εφόσον μεταβληθούν οι πίνακες δρομολόγησής τους)

η πληροφορία σχετικά με την αστοχία διαδίδεται σε
ολόκληρο το δίκτυο

Υλοποίηση RIP στο UNIX

Η διεργασία route-d στο επίπεδο εφαρμογής διαχειρίζεται τους πίνακες δρομολόγησης RIP ανταλλάσσει αγγελίες με τις αντίστοιχες διεργασίες στους γείτονες δρομολογητές μέσω UDP (αριθμός θύρας 520), οι αγγελίες επαναλαμβάνονται περιοδικά



Παράδειγμα πίνακα RIP

netstat -rn στο δρομολογητή giroflee.eurocom.fr

Destination	Gateway	Flags Ref	Use	Interface
-----	-----	-----		
127.0.0.1	127.0.0.1	UH	0 26492	lo0
192.168.2.	192.168.2.5	U	2 13	fa0
193.55.114.	193.55.114.6	U	3 58503	le0
192.168.3.	192.168.3.5	U	2 25	qaa0
224.0.0.0	193.55.114.6	U	3 0	le0
default	193.55.114.129	UG	0 143454	

Τρία συνδεδεμένα δίκτυα κλάσης C (LANs)

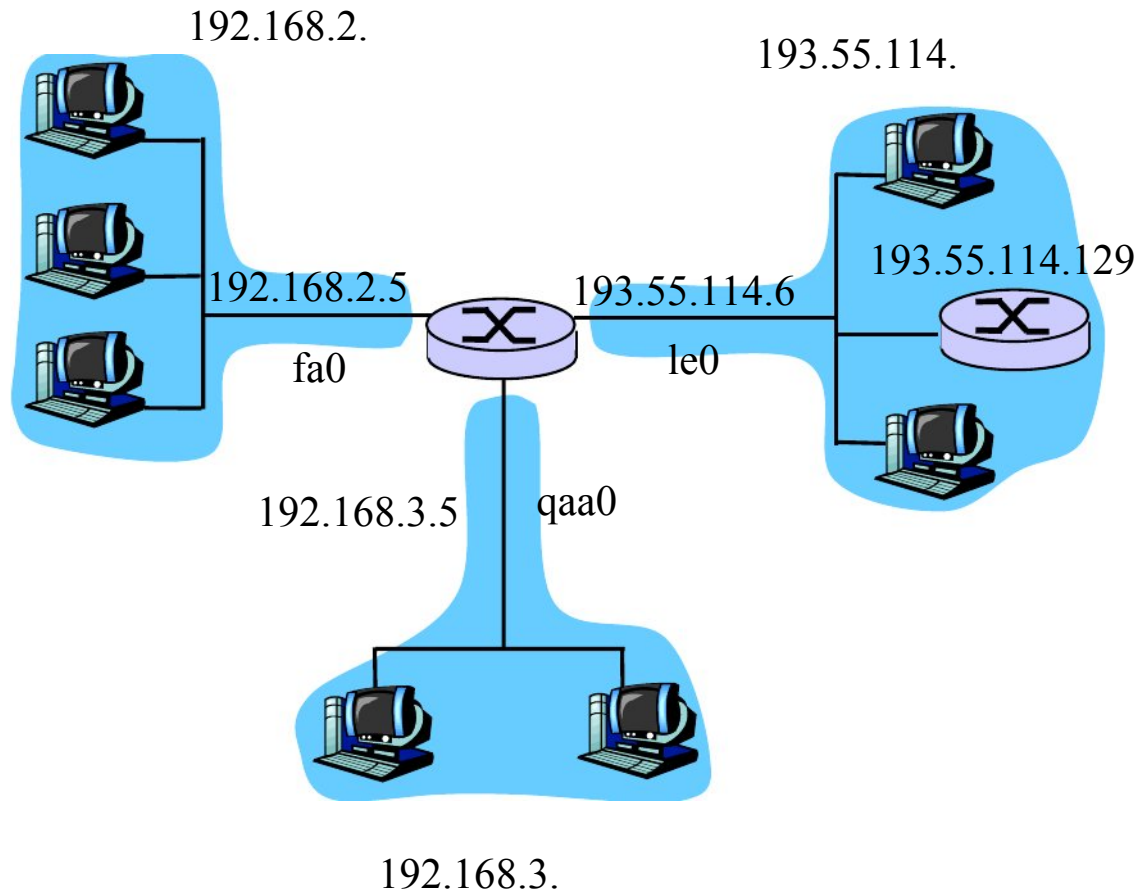
Ο δρομολογητής γνωρίζει τις διαδρομές προς τα συνδεδεμένα δίκτυα μόνο

Στον προεπιλεγμένο (default) δρομολογητή αποστέλλονται όλα τα datagrams με προορισμούς που δεν περιέχονται στον πίνακα

Δρομολόγηση διεύθυνσης multicast: 224.0.0.0

Διεπαφή βρόχου επαναφοράς (loopback interface) (για debugging)

Παράδειγμα πίνακα RIP (συνέχεια)



OSPF (Open Shortest Path First)

"open": δημόσια διαθέσιμο

Χρησιμοποιεί έναν αλγόριθμο κατάστασης ζεύξεων (Link State)

Διανομή πακέτων LS

Χάρτης τοπολογίας σε κάθε κόμβο

Υπολογισμός διαδρομών με τον αλγόριθμο του Dijkstra

- Μετρική κόστους επιλέγεται από τον διαχειριστή

Κάθε αγγελία OSPF μεταφέρει μία καταχώρηση για κάθε γείτονα δρομολογητή

Μηνύματα OSPF HELLO

Οι αγγελίες διαδίδονται σε **ολόκληρο** το AS (μέθοδος πλημμύρας)

Μεταφέρονται μέσα σε μηνύματα OSPF απευθείας πάνω από IP (χωρίς μεσολάβηση TCP ή UDP)

Χαρακτηριστικά του OSPF (που δεν υπάρχουν στο RIP)

Ασφάλεια: πιστοποίηση ταυτότητας για όλα τα μηνύματα OSPF

Επιτρέπει **πολλαπλές διαδρομές** ιδίου κόστους (μόνο μία διαδρομή στο RIP)

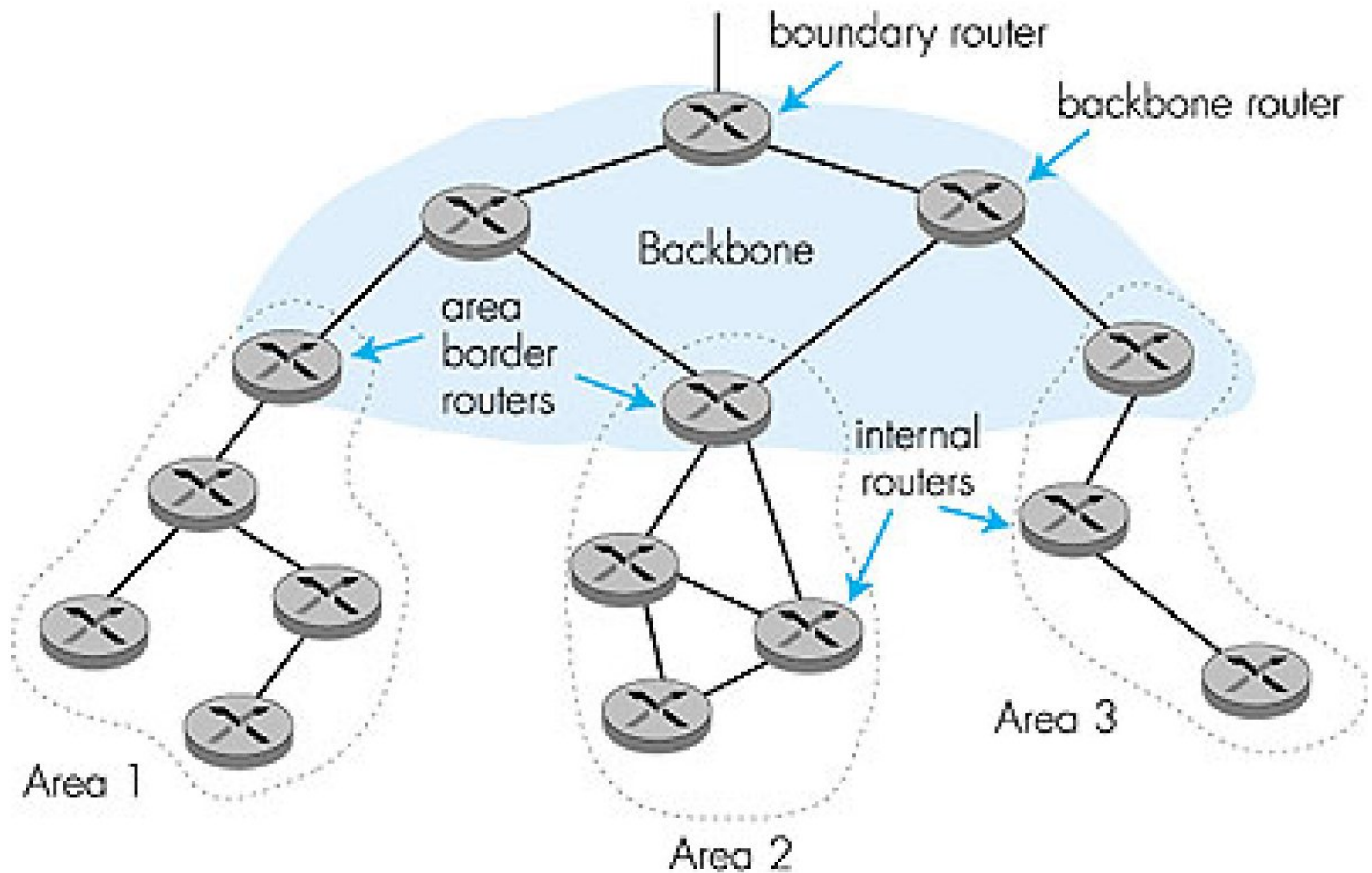
Επιτρέπει πολλαπλές μετρικές κόστους σε κάθε ζεύξη για κάθε είδος υπηρεσίας **TOS** (π.χ. το κόστος δορυφορικής ζεύξης "μικρό" για υπηρεσία βέλτιστης προσπάθειας, υψηλό για υπηρεσία πραγματικού χρόνου)

Ενοποιημένη υποστήριξη unicast και **multicast**:

Το Multicast OSPF (MOSPF) χρησιμοποιεί την ίδια βάση δεδομένων για την τοπολογία με το OSPF

Ιεραρχικό OSPF σε μεγάλα αυτόνομα συστήματα

Ιεραρχικό OSPF



Ιεραρχία OSPF

Ιεραρχία δύο επιπέδων: τοπική περιοχή, backbone

Είδη δρομολογητών:

Εσωτερικοί (Internal)

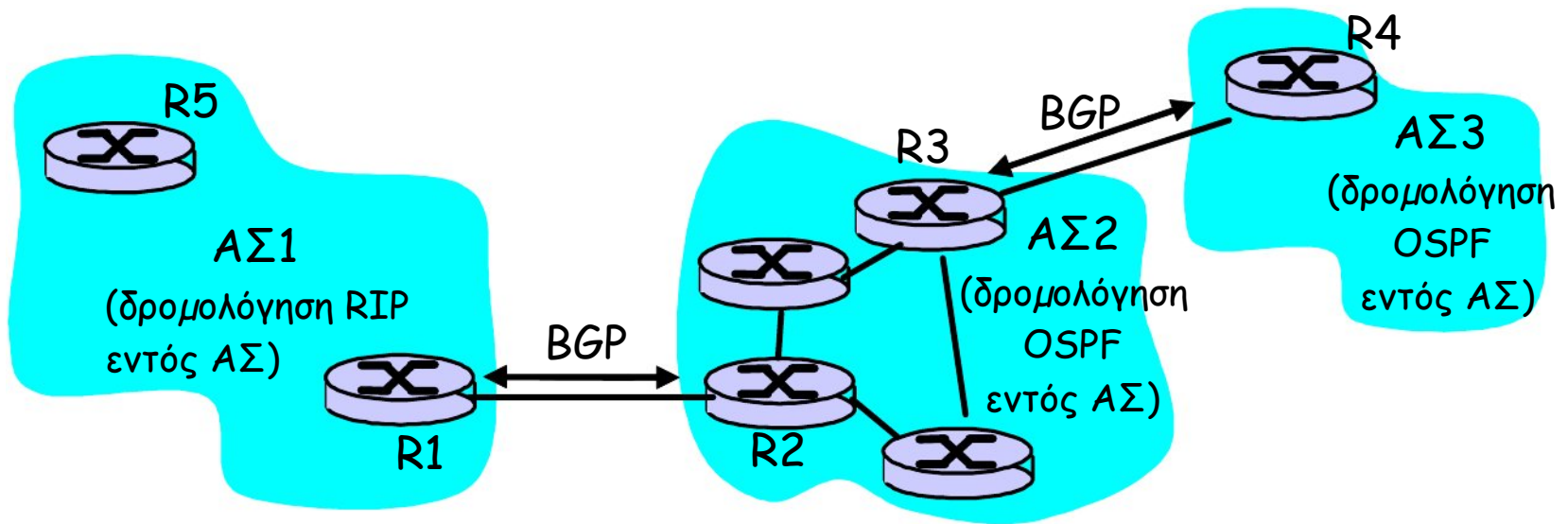
Αγγελίες LS στην τοπική περιοχή μόνο
κάθε κόμβος έχει λεπτομερή τοπολογία της περιοχής
- γνωρίζει μόνο κατεύθυνση (συντομότερη διαδρομή)
προς δίκτυα σε άλλες περιοχές

Ορίου περιοχής (Area border): "συνοψίζουν" τις
αποστάσεις προς δίκτυα στη δική τους περιοχή,
αναγγέλλουν σε άλλους δρομολογητές ορίου περιοχής

Κορμού (Backbone): εκτελούν δρομολόγηση OSPF στο
backbone

Συνοριακοί (Boundary): συνδέονται με άλλα ΑΣ

Δρομολόγηση μεταξύ ΑΣ στο Διαδίκτυο: BGP



Δρομολόγηση μεταξύ ΑΣ στο Διαδίκτυο: BGP

BGP (Border Gateway Protocol - Πρωτόκολλο Συνοριακών Τυλών): η χρήση του έχει καθιερωθεί για τη δρομολόγηση μεταξύ ΑΣ στο Διαδίκτυο

Το BGP παρέχει σε κάθε ΑΣ τα μέσα ώστε:

1. να αποκτά πληροφορία προσεγγισιμότητας υποδικτύων από γειτονικά ΑΣ
2. να διαδίδει την πληροφορία προσεγγισιμότητας σε όλους τους εσωτερικούς δρομολογητές του ΑΣ
3. να προσδιορίζει "καλές" διαδρομές προς τα υποδίκτυα με βάση την πληροφορία προσεγγισιμότητας και την πολιτική δρομολόγησης

Επιτρέπει σε ένα υποδίκτυο να αναγγείλει την ύπαρξη του στο υπόλοιπο Διαδίκτυο: **"Βρίσκομαι εδώ"**

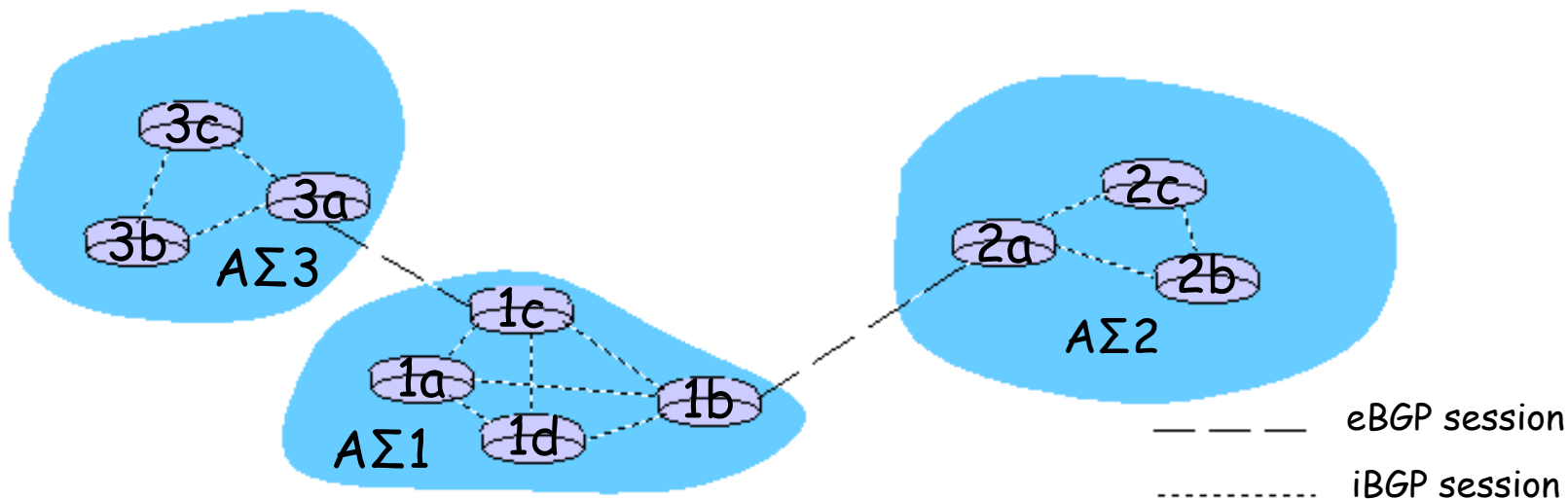
BGP: Βασικά χαρακτηριστικά

Ζεύγη δρομολογητών (BGP peers) ανταλλάσσουν πληροφορία δρομολόγησης μέσω ημι-μόνιμων συνδέσεων TCP: **BGP sessions**

τα BGP sessions δεν αντιστοιχούν σε φυσικές ζεύξεις

Όταν το ΑΣ2 αναγγέλλει ένα πρόθεμα στο ΑΣ1, το ΑΣ2 **υπόσχεται** ότι θα προωθήσει οποιοδήποτε πακέτο με προορισμό αυτό το πρόθεμα προς το πρόθεμα

το ΑΣ2 μπορεί να συναθροίσει τα προθέματα στην αγγελία του
π.χ. συνάθροιση των προθεμάτων 138.16.64/24, 138.16.65/24, 138.16.66/24, 138.16.67/24 σε ένα μόνο πρόθεμα: 138.16.64/22



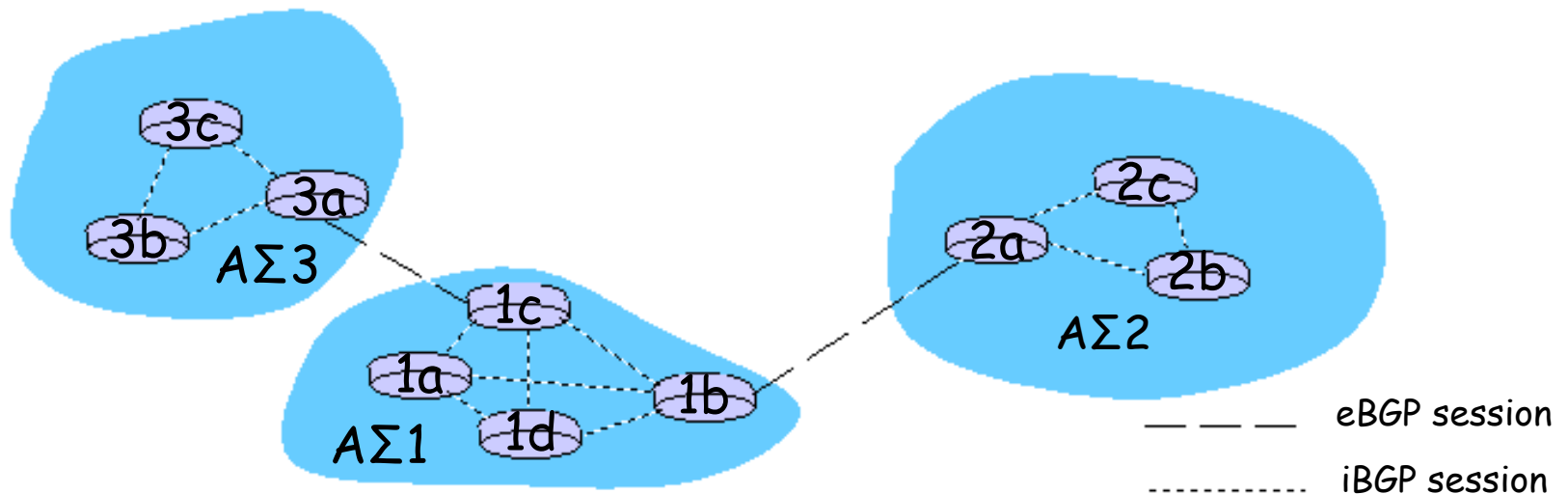
Διανομή πληροφορίας προσεγγισιμότητας

Το ΑΣ3 στέλνει πληροφορία προσεγγισιμότητας προθέματος στο ΑΣ1 με το eBGP session μεταξύ των πυλών 3a και 1c

Στη συνέχεια, η πύλη 1c μπορεί τότε να χρησιμοποιήσει το iBGP για να διανείμει αυτήν τη νέα πληροφορία προσεγγισιμότητας προθέματος σε όλους τους δρομολογητές στο ΑΣ1

Μετά, η πύλη 1b μπορεί να αναγγείλει τη νέα πληροφορία προσεγγισιμότητας στο ΑΣ2 μέσω του eBGP session από την πύλη 1b στην πύλη 2a

Όταν ένας δρομολογητής μάθει σχετικά με ένα νέο πρόθεμα, δημιουργεί μία καταχώρηση για το πρόθεμα αυτό στο πίνακα προώθησής του



Ιδιότητες διαδρομής (Path attributes) & Διαδρομές BGP (BGP routes)

Η αγγελία ενός προθέματος περιλαμβάνει τις ιδιότητες (attributes) της διαδρομής

πρόθεμα + ιδιότητες = "διαδρομή"

Δύο σημαντικές ιδιότητες:

AS-PATH: περιέχει τα ΑΣ από τα οποία έχει διέλθει η αγγελία για το πρόθεμα: ΑΣ67 ΑΣ17

NEXT-HOP: υποδεικνύει τον συγκεκριμένο δρομολογητή (στο εσωτερικό του ΑΣ) προς το επόμενο ΑΣ (Ενδέχεται να υπάρχουν πολλαπλές ζεύξεις από ένα ΑΣ στο επόμενο)

Όταν ο δρομολογητής πύλης λάβει την αγγελία μίας διαδρομής, χρησιμοποιεί την **πολιτική εισαγωγής (import policy)** για να την κάνει αποδεκτή ή για να την απορρίψει π.χ. η πολιτική μπορεί να απορρίψει μία διαδρομή επειδή διέρχεται μέσα από ένα ανεπιθύμητο ΑΣ, για λόγους κόστους, προς αποφυγή βρόχων κλπ.

Επιλογή διαδρομής στο BGP

Ένας δρομολογητής ενδέχεται να πληροφορηθεί περισσότερες από μία διαδρομές προς κάποιο πρόθεμα □ ο δρομολογητής πρέπει να επιλέξει μία από αυτές

Κανόνες επιλογής διαδρομής (ή αποκλεισμού διαδρομών):

1. Μέγιστη τιμή ιδιότητας τοπικής προτίμησης (LOCAL_PREF): απόφαση πολιτικής
2. Συντομότερο AS-PATH
3. Πλησιέστερος δρομολογητής NEXT-HOP: δρομολόγηση hot potato
4. Πρόσθετα κριτήρια

Μηνύματα BGP

Για την ανταλλαγή μηνυμάτων BGP χρησιμοποιείται το TCP (αριθμός θύρας 179)

Είδη μηνυμάτων BGP:

OPEN

ανοίγει μία σύνδεση TCP με τον γείτονα (peer) και πιστοποιεί την ταυτότητα του αποστολέα

UPDATE

αναγγέλλει νέα διαδρομή (ή αποσύρει παλαιά διαδρομή που δεν ισχύει πλέον)

χρησιμοποιείται επίσης ως επιβεβαίωση σε μία αίτηση OPEN

KEEPALIVE:

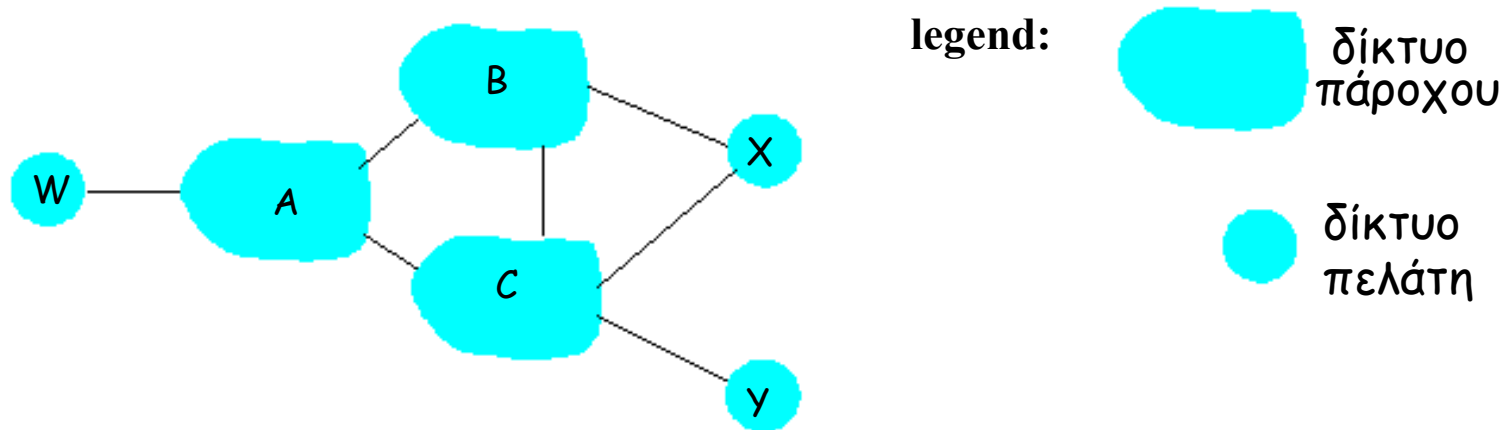
διατηρεί τη σύνδεση ενεργή όσο δεν ανταλλάσσονται μηνύματα UPDATE

NOTIFICATION:

αναφέρει σφάλματα στο προηγούμενο μήνυμα

χρησιμοποιείται επίσης για τον τερματισμό της σύνδεσης

BGP: πολιτική δρομολόγησης



A,B,C: **δίκτυα παρόχων**

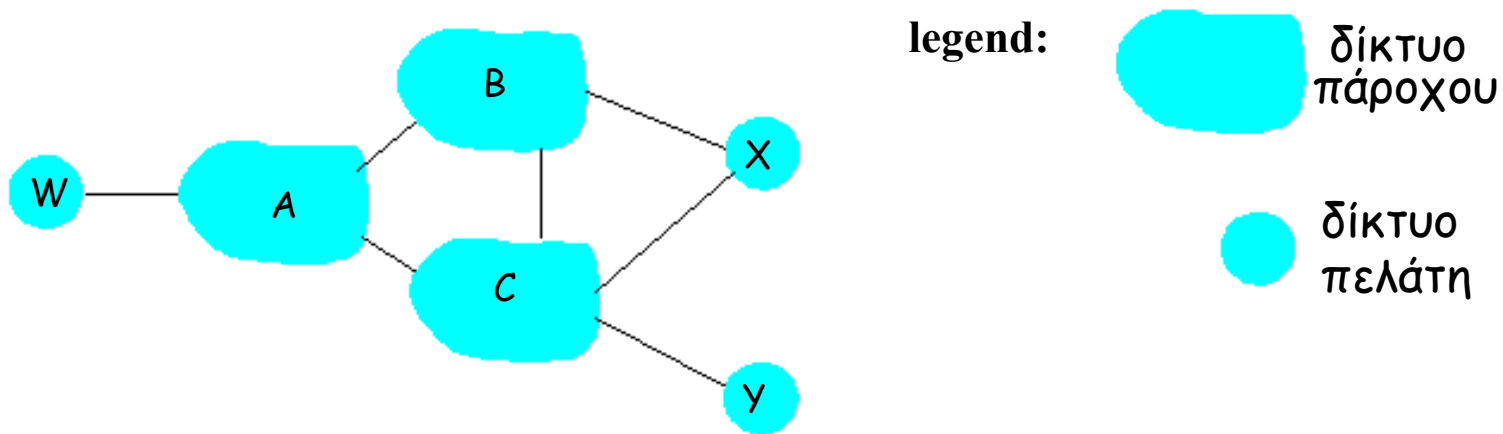
X,W,Y: δίκτυα πελατών (των παρόχων)

το δίκτυο X καλείται **dual-homed**: είναι συνδεδεμένο σε δύο δίκτυα

το X δεν θέλει να δρομολογήσει κίνηση από το B στο C μέσω X

οπότε το X δεν αναγγέλλει στο B τη διαδρομή προς το C

BGP: πολιτική δρομολόγησης



το δίκτυο A αναγγέλλει στο B τη διαδρομή AW

το δίκτυο B αναγγέλλει στο X τη διαδρομή BAW

το δίκτυο B δεν αναγγέλλει τη διαδρομή BAW στο C

B δεν έχει κέρδος από τη δρομολόγηση CBAW αφού ούτε το W ούτε το C είναι πελάτες του B

το B θέλει να εξαναγκάσει το C να δρομολογεί στο w μέσω του A

το B θέλει να δρομολογεί **μόνο** προς/από τους πελάτες του

Γιατί διαφορετικά πρωτόκολλα δρομολόγησης εντός ΑΣ και μεταξύ ΑΣ;

Πολιτική:

Μεταξύ ΑΣ: ο διαχειριστής επιθυμεί έλεγχο του τρόπου με τον οποίο δρομολογείται η δική του κίνηση καθώς και της κίνησης που διέρχεται από το δίκτυό του

Εντός ΑΣ: δεν εφαρμόζονται αποφάσεις πολιτικής (ένας διαχειριστής)

Κλίμακα:

με την ιεραρχική δρομολόγηση μικραίνει το μέγεθος των πινάκων, μειώνεται η κίνηση λόγω πληροφoρίας δρομολόγησης

Απόδοση:

Εντός ΑΣ: εστίαση στην απόδοση

Μεταξύ ΑΣ: η πολιτική μπορεί να κυριαρχεί έναντι της απόδοσης

Περίγραμμα

- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
- 4 Αρχές Δρομολόγησης
- 5 Δρομολόγηση στο Διαδίκτυο
- 6 IPv6
- 7 Ειδικά θέματα του στρώματος Δικτύου

IPv6

Αρχικό κίνητρο: Οι διευθύνσεις IP των 32 bits κινδύνευαν να εξαντληθούν μέχρι το 2008

το μέγεθος της διεύθυνσης IP αυξάνεται από 32 σε 128 bits

Πρόσθετες αλλαγές:

απλοποίηση δομής επικεφαλίδας ώστε να επιταχύνεται η επεξεργασία και προώθηση πακέτων

- βασική επικεφαλίδα σταθερού μήκους 40 bytes - οι επιλογές (options) υλοποιούνται εκτός βασικής επικεφαλίδας με το πεδίο next header
- δεν επιτρέπεται η κατάτμηση στους δρομολογητές
- δεν εφαρμόζεται έλεγχος σφαλμάτων επικεφαλίδας IP

αλλαγές στην επικεφαλίδα ώστε να διευκολύνεται η παροχή Ποιότητας Υπηρεσίας (Quality of Service, QoS)

- εισαγωγή πεδίου flow label

νέα διεύθυνση "anycast" προσδιορίζει ένα σύνολο διεπαφών hosts: ένα πακέτο με διεύθυνση "anycast" δρομολογείται στον "καλύτερο" (π.χ. πλησιέστερο) από τους host

Δομή αυτοδύναμου πακέτου IPv6

Πεδία:

version #: έκδοση πρωτοκόλλου IP

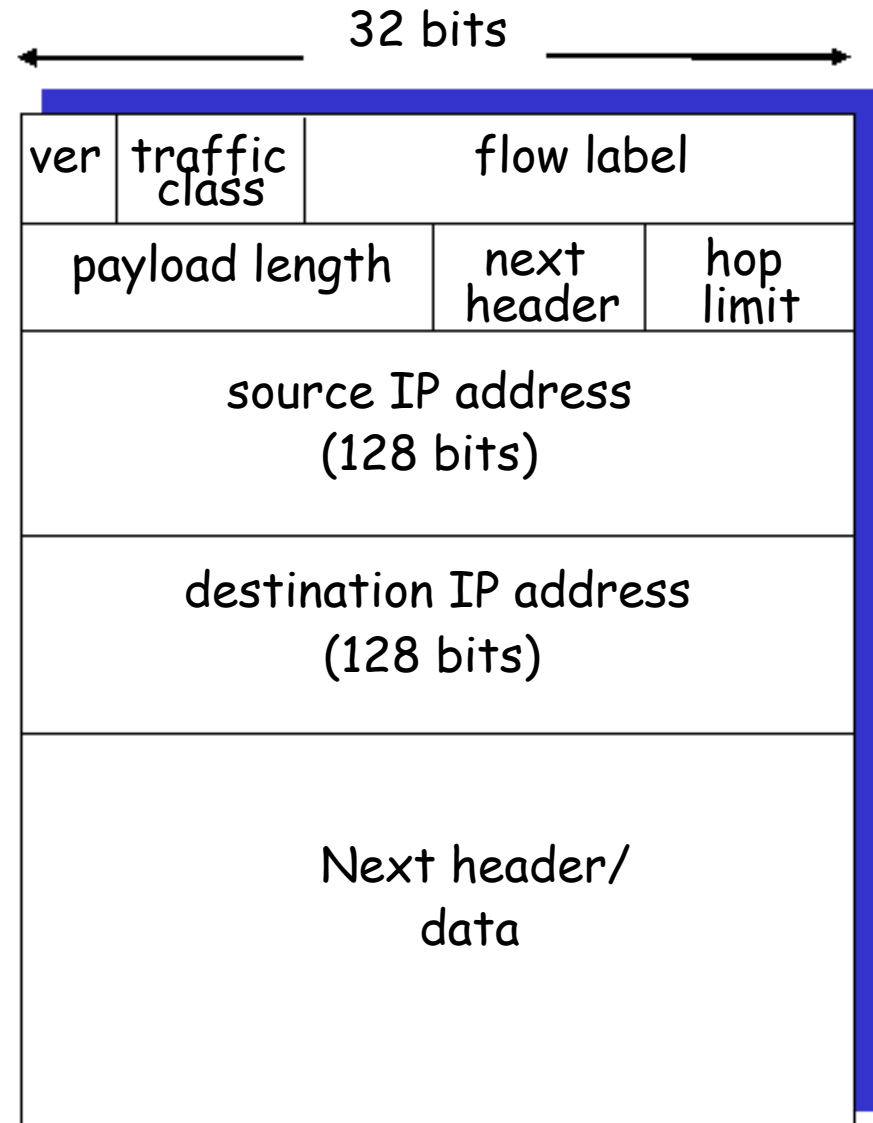
traffic class: αντίστοιχο με πεδίο Type of Service στο IPv4

flow label: προσδιορίζει πακέτα που ανήκουν στην ίδια "ροή (flow)"

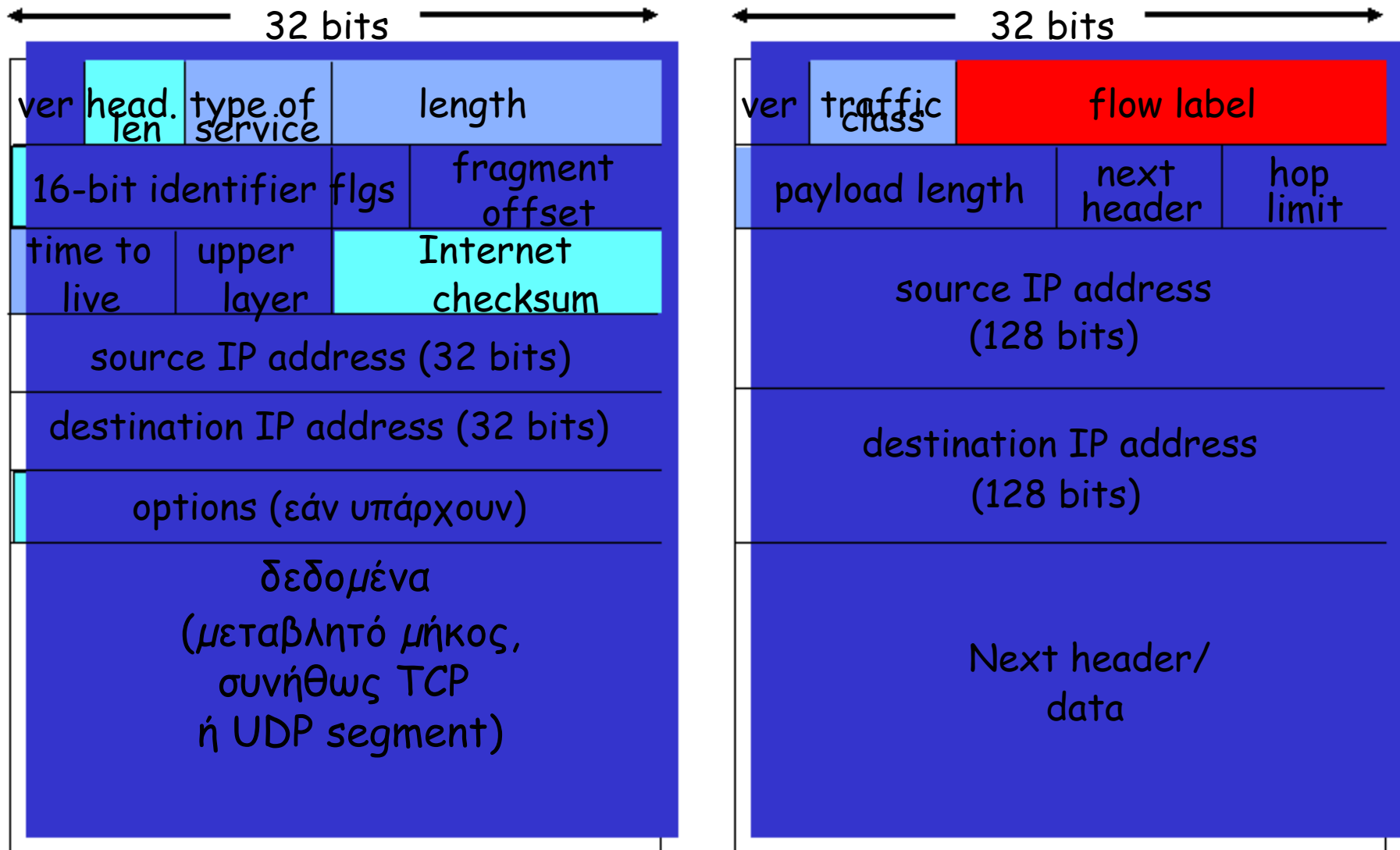
payload length: συνολικό μήκος ωφέλιμου φορτίου (payload) σε bytes ≤ 65535 bytes

next header: προσδιορίζει το πρωτόκολλο ανωτέρου επιπέδου για το οποίο προορίζονται τα δεδομένα (αντίστοιχο με πεδίο upper layer protocol στο IPv4) ή την επόμενη επικεφαλίδα επέκτασης που ακολουθεί την επικεφαλίδα

hop limit: αντίστοιχο με πεδίο Time to live στο IPv4



Διαφορές επικεφαλίδων IPv4 και IPv6



■ πεδία που αφαιρούνται
■ πεδία που αντικαθίστανται

■ πεδία που εισάγονται

Διευθύνσεις IPv6 πηγής/προορισμού

Αριθμός διευθύνσεων IPv6:

$$2_{128} \approx 3 \times 10_{38}$$

Συμβολισμός (colon hexadecimal notation)

8 ομάδες των τεσσάρων δεκαεξαδικών ψηφίων που χωρίζονται με άνω και κάτω τελείες:

8000:0000:0000:0000:0123:4567:89AB:DEF

Συμπίεση μηδενικών

8000:0:0:0:123:4567:89AB:DEF

8000::123:4567:89AB:DEF

Δομή διεύθυνσης unicast (aggregatable global unicast address)

	m bits	n bits	o bits	p bits	125-m-n-o-p bits
010	Registry ID	Provider ID	Subscriber ID	Subnet ID	Interface ID

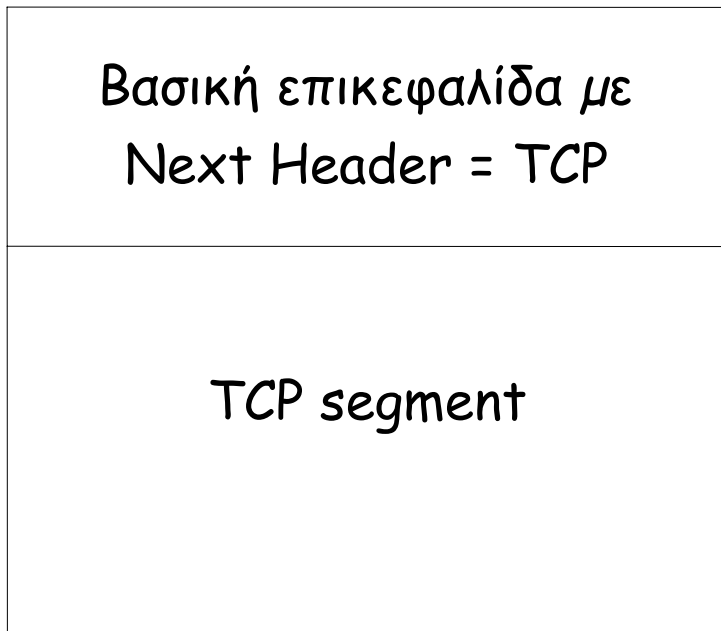
Επικεφαλίδες επέκτασης

τοποθετούνται μεταξύ της βασικής επικεφαλίδας και του ωφέλιμου φορτίου και υποστηρίζουν λειτουργίες που δεν παρέχονται από τη βασική επικεφαλίδα (αντίστοιχα με πεδίο options στο IPv4)

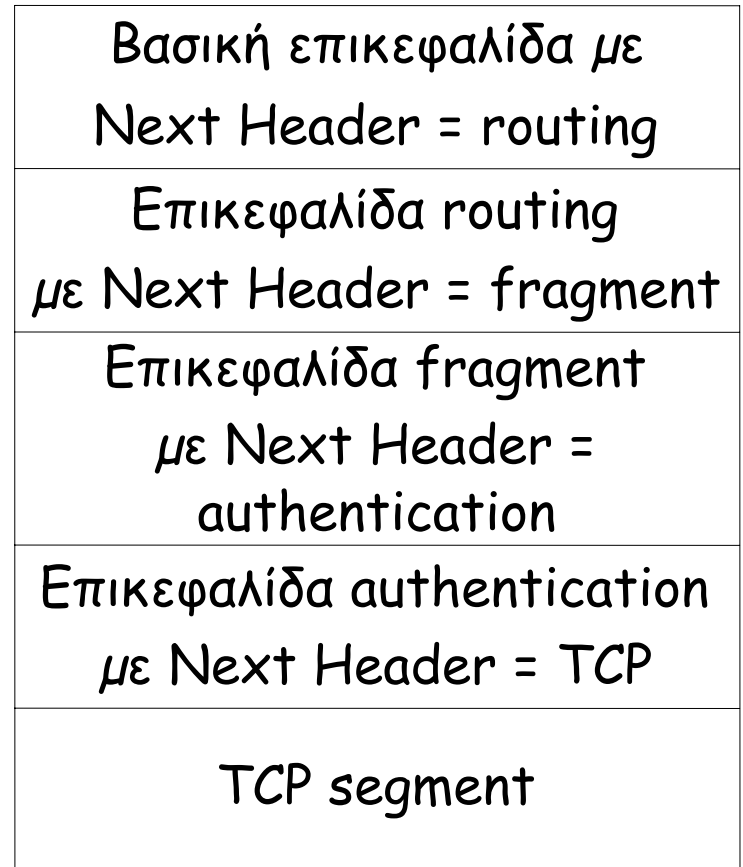
Επικεφαλίδα επέκτασης	Περιγραφή
Hop-by-hop options	Διάφορες πληροφορίες για τους δρομολογητές
Destination options	Πρόσθετες πληροφορίες για τον προορισμό
Routing	Κατάλογος δρομολογητών που πρέπει να ακολουθήσει το πακέτο
Fragmentation	Διαχείριση τεμαχισμένων πακέτων
Authentication	Πιστοποίηση ταυτότητας αποστολέα
Encrypted security payload	Πληροφορίες σχετικά με κρυπτογραφημένο περιεχόμενο πακέτου

Δομή αυτοδύναμου πακέτου IPv6 (συνέχεια)

πακέτο με βασική
επικεφαλίδα μόνο:



πακέτο με επικεφαλίδες
επέκτασης:



Μετάβαση από το IPv4 στο IPv6

Δεν είναι δυνατή η ταυτόχρονη αναβάθμιση όλων των δρομολογητών

όπως έγινε με τη μετάβαση από NCP σε TCP στις 1/1/1983

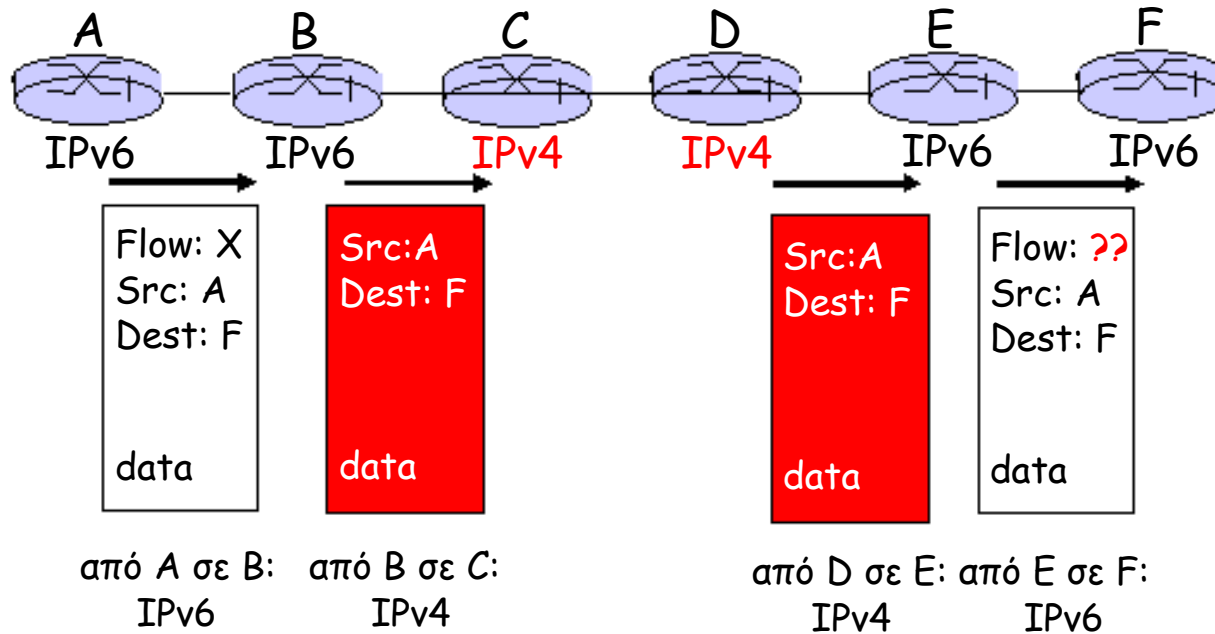
Αναγκαία η λειτουργία του δικτύου με δύο ειδών δρομολογητές: IPv4 και IPv6

Δύο προτεινόμενες προσεγγίσεις:

Διπλός σωρός (dual stack): ορισμένοι δρομολογητές με διπλό σωρό (v6, v4) μπορούν να "μεταφράζουν" πακέτα IPv4 σε πακέτα IPv6 και αντίστροφα

Χρήση σήραγγας (tunneling): ένα πακέτο IPv6 μεταφέρεται ως ωφέλιμο φορτίο εντός πακέτου IPv4 μεταξύ δρομολογητών IPv4

Διπλός σωρός (dual stack)

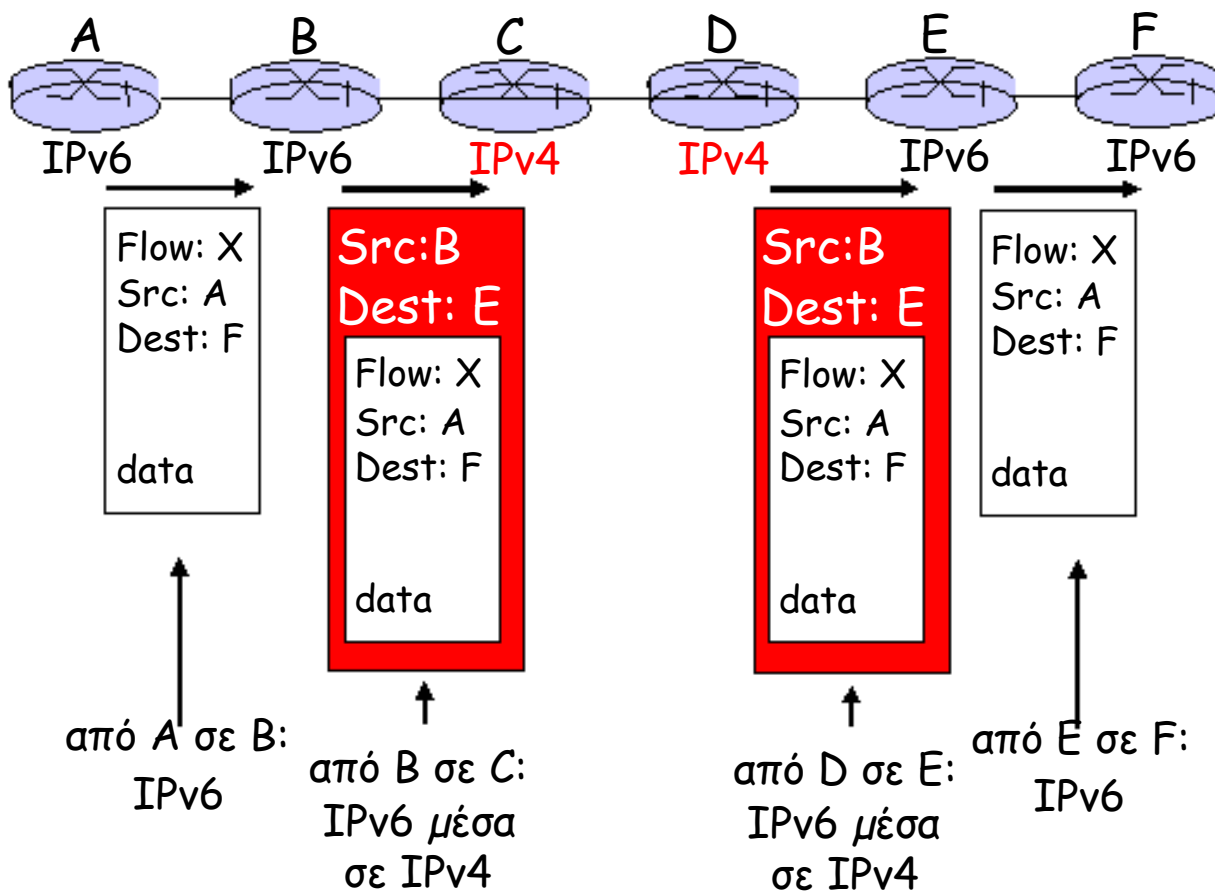


Χρήση σήραγγας (tunneling)

Λογική
Θεώρηση:



Φυσική
Θεώρηση:



Περίγραμμα

- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
- 4 Αρχές Δρομολόγησης
- 5 Δρομολόγηση στο Διαδίκτυο
- 6 IPv6
- 7 Ειδικά θέματα του στρώματος Δικτύου
 - 7.1 Broadcast/Multicast
 - 7.2 Mobility
 - 7.3 Quality of Service (QoS)

Ευρυεκπομπή (Broadcast)

- Μόνο στο IPv4
- Χρησιμεύει για τη μετάδοση πακέτων σε όλους τους κόμβους ενός δικτύου
- Οι δρομολογητές εκτελούν αντιγραφή των πακέτων
- Η Broadcast διεύθυνση ενός δικτύου έχει όλο άσους στο host τμήμα της αντίστοιχης IP διεύθυνσης, π.χ.
172.31.255.255
 - Η διεύθυνση 255.255.255.255 είναι η broadcast διεύθυνση του δικτύου 0.0.0.0 που υποδηλώνει το local network

Πολυεκπομπή (Multicast)

multicast: αποστολή πακέτων από έναν ή περισσότερους αποστολείς σε πολλαπλούς παραλήπτες

παραδείγματα εφαρμογών

μετάδοση ραδιοφωνικών ή τηλεοπτικών εκπομπών

τηλεκπαίδευση, τηλεδιασκέψεις

ενημέρωση βάσεων δεδομένων (π.χ. web caches)

κατανεμημένα υπολογιστικά συστήματα

συνδεδεισμενής (connection-oriented) υπηρεσία

σε αντίθεση με την υπηρεσία μονοεκπομπής (unicast), η υπηρεσία multicast που παρέχεται από το επίπεδο δικτύου είναι συνδεδεισμενής (connection-oriented)

απαιτεί την εγκαθίδρυση και διατήρηση πληροφορίας

κατάστασης στους δρομολογητές που προωθούν πακέτα

multicast τα οποία ανταλλάσσονται μεταξύ των hosts που

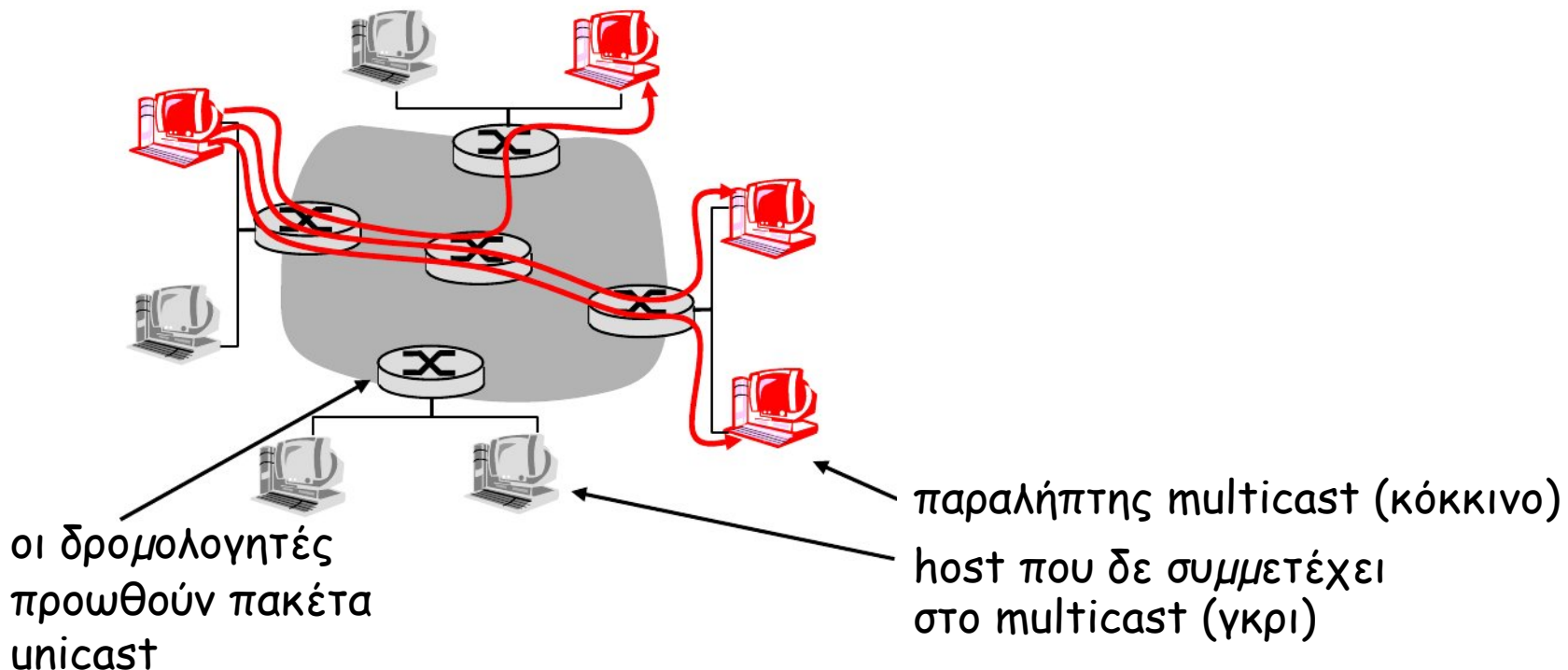
ανήκουν στην ίδια ομάδα πολυεκπομπής (multicast group)

Multicast: από έναν αποστολέα σε πολλούς παραλήπτες

Τρόποι επίτευξης πολυεκπομπής

Πολυεκπομπή μέσω πολλαπλών μονοεκπομπών

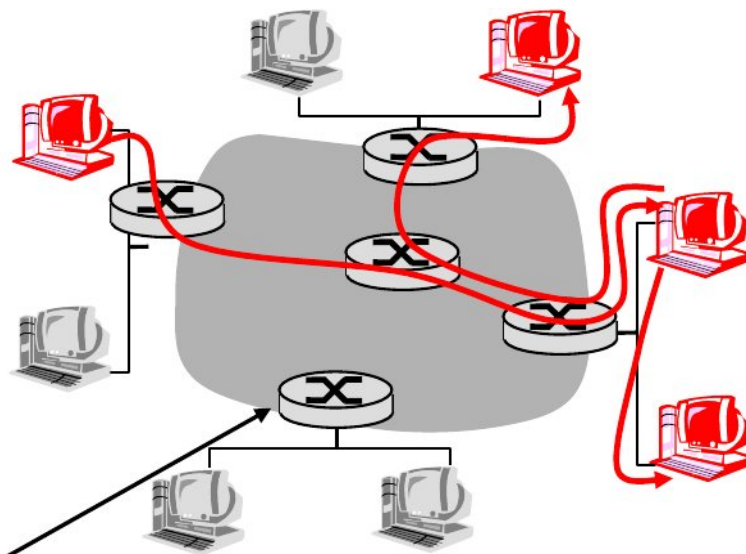
ο αποστολέας στέλνει N πακέτα unicast, καθένα με διεύθυνση τη διεύθυνση IP ενός από N παραλήπτες



Multicast: από έναν αποστολέα σε πολλούς παραλήπτες

πολυεκπομπή στο επίπεδο εφαρμογής

τα τερματικά συστήματα συμμετέχουν στην πολυεκπομπή:
αντιγράφουν και προωθούν πακέτα unicast μεταξύ τους

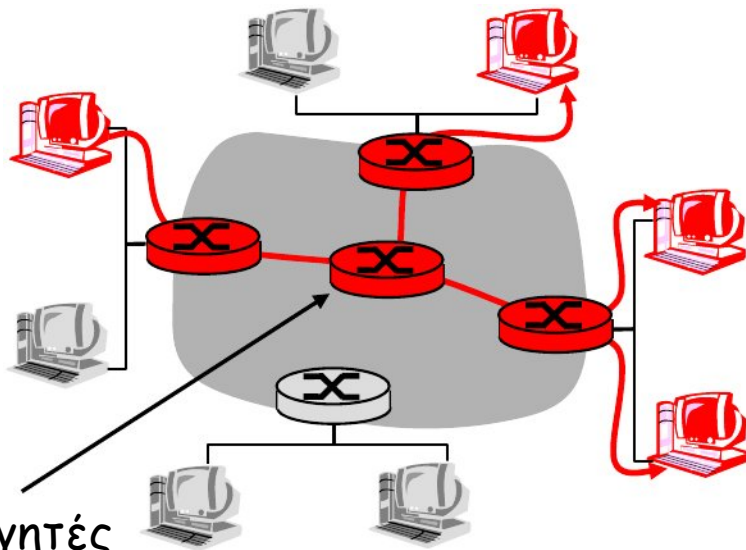


οι δρομολογητές
προωθούν πακέτα
unicast

Multicast: από έναν αποστολέα σε πολλούς παραλήπτες

πολυεκπομπή στο επίπεδο δικτύου

οι δρομολογητές συμμετέχουν ενεργά στην πολυεκπομπή:
δημιουργούν τα απαιτούμενα αντίγραφα των πακέτων τα
οποία προωθούν προς τους παραλήπτες της πολυεκπομπής



Δρομολογητές
multicast (κόκκινο) επαναλαμβάνουν
και προωθούν πακέτα multicast

Μοντέλο υπηρεσιών multicast στο Διαδίκτυο

Ζητήματα προς επίλυση

Με ποιο τρόπο αναγνωρίζονται οι παραλήπτες ενός πακέτου multicast;

Ποια η διεύθυνση του πακέτου multicast που προορίζεται για τους παραλήπτες αυτούς;

Διαδικασία έμμεσης διευθυνσιοδότησης (address **indirection**) στο Διαδίκτυο

Η ομάδα των παραληπτών αναγνωρίζεται από μία multicast διεύθυνση IP (κλάση D)

Η ομάδα των αποστολέων και παραληπτών που σχετίζονται με την ίδια διεύθυνση multicast αποτελούν μία **ομάδα πολυεκπομπής (multicast group)**

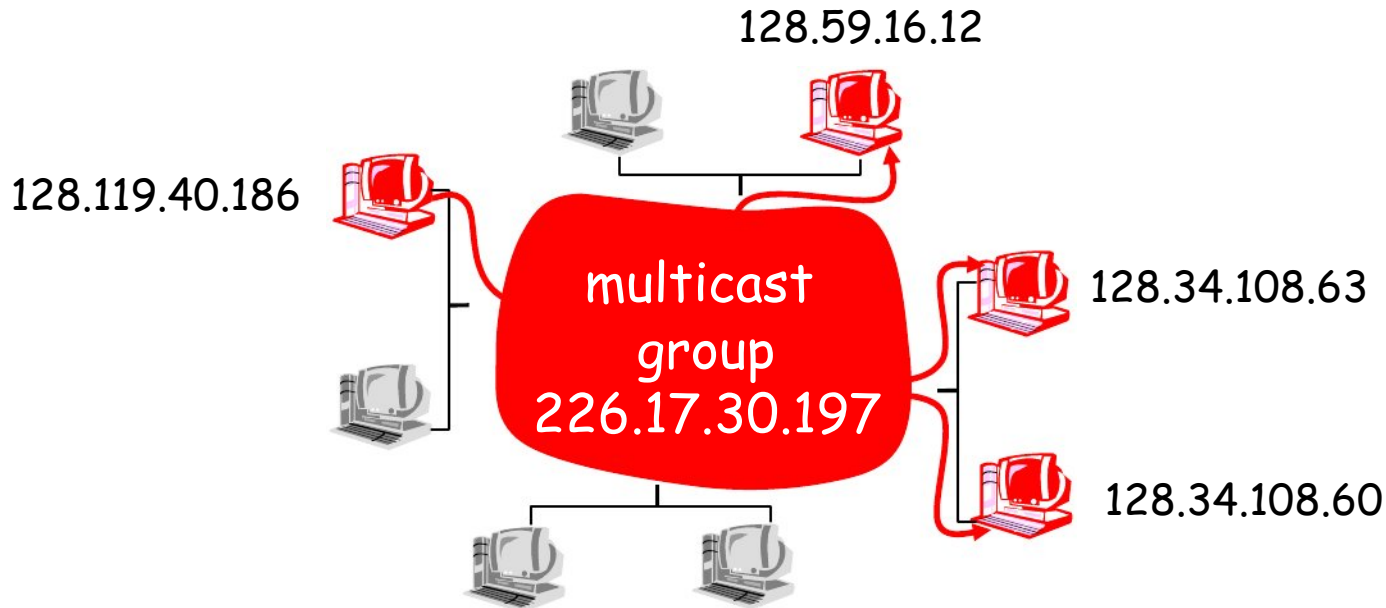


← 28 bits →

Πακέτα που προορίζονται για μία ομάδα multicast έχουν διεύθυνση τη διεύθυνση multicast της ομάδας

οι δρομολογητές προωθούν τα πακέτα multicast στους hosts που συμμετέχουν στη συγκεκριμένη ομάδα multicast

Μοντέλο υπηρεσιών multicast στο Διαδίκτυο



Οι διευθύνσεις unicast των μελών μίας ομάδας multicast είναι ανεξάρτητες από τη διεύθυνση multicast της ομάδας

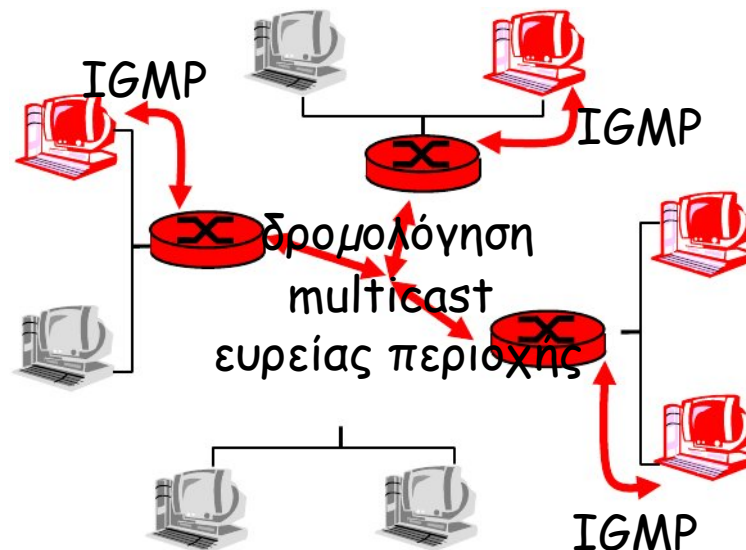
απαιτείται: υποδομή που να παραδίδει τα πακέτα με μία δεδομένη διεύθυνση multicast σε όλα τα μέλη της αντίστοιχης ομάδας multicast

Προσχώρηση σε μία ομάδα multicast

Διαδικασία με δύο βήματα:

τοπικά: ο host εκφράζει στον τοπικό δρομολογητή multicast την επιθυμία του να συμμετέχει στην ομάδα: IGMP (Internet Group Management Protocol)

ευρεία περιοχή: ο τοπικός δρομολογητής αλληλεπιδρά με άλλους δρομολογητές ώστε να λάβει τα πακέτα multicast πρωτόκολλα δρομολόγησης multicast (π.χ. DVMRP, MOSPF, PIM)



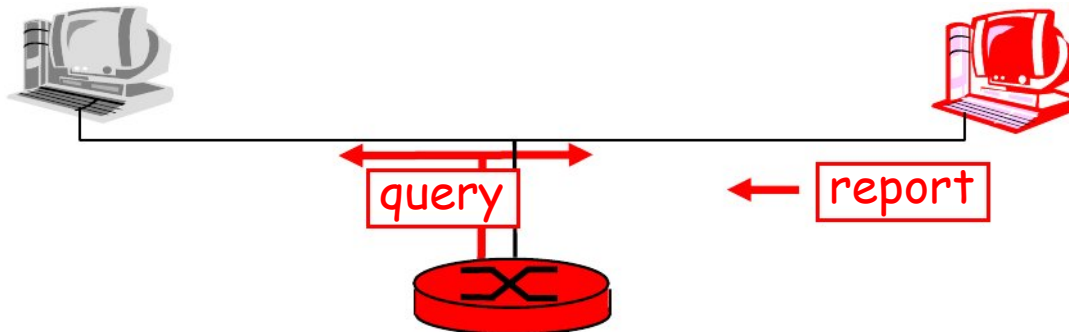
IGMP: Internet Group Management Protocol

Ο **host** στέλνει μήνυμα IGMP report όταν μία εφαρμογή προσχωρήσει σε μία ομάδα multicast

Ο host δεν υποχρεούται να αναφέρει την αποχώρησή του από μία ομάδα multicast

Ο **δρομολογητής** στέλνει μηνύματα IGMP query σε τακτά χρονικά διαστήματα (multicast σε όλους τους hosts στο LAN στη διεύθυνση IP 224.0.0.1)

οι hosts που είναι μέλη μίας ομάδας multicast πρέπει να αποκριθούν στο μήνυμα IGMP query με μήνυμα IGMP report
αρκεί ο δρομολογητής να λάβει απόκριση από έναν μόνο host:
"τυχαιοποίηση" της καθυστέρησης απόκρισης και αναστολή απόκρισης εάν προηγηθεί απόκριση από άλλο μέλος της ομάδας



IGMP

IGMP έκδοση 1 (RFC 1112)

Τύπος μηνύματος	Αποστολέας	Σκοπός
Membership query: general	δρομολογητής	Ο δρομολογητής ρωτά να μάθει τις ομάδες multicast στις οποίες συμμετέχουν οι συνδεδεμένοι hosts
Membership report	host	Ο host που στέλνει το μήνυμα είναι μέλος ή επιθυμεί να γίνει μέλος σε δεδομένη ομάδα multicast

IGMP (συνέχεια)

IGMP έκδοση 2 (RFC 2236) - προσθήκες

Τύπος μηνύματος	Αποστολέας	Σκοπός
Membership query: specific	δρομολογητής	Ο δρομολογητής ρωτά να μάθει εάν υπάρχουν hosts οι οποίοι συμμετέχουν σε δεδομένη ομάδα multicast (ελέγχει εάν έχουν απομείνει μέλη της δεδομένης ομάδας)
Leave group	host	Ο host που στέλνει το μήνυμα αναφέρει αποχώρηση από δεδομένη ομάδα multicast

IGMP (συνέχεια)

Δομή μηνύματος IGMP (v2)

Τύπος Μηνύματος	Μέγιστος χρόνος απόκρισης	Άθροισμα ελέγχου
Διεύθυνση ομάδας multicast		

Τα μηνύματα IGMP ενθυλακώνονται μέσα σε IP datagrams (αριθμός πρωτοκόλλου ανώτερου επιπέδου = 2)

Στο IPv6 το IGMP έχει ενσωματωθεί στο ICMPv6

Το IGMPv3 (RFC 3376) υποστηρίζει Source-Specific Multicast

ένα μέλος μπορεί να ζητήσει να λαμβάνει πακέτα μόνο από συγκεκριμένες διευθύνσεις πηγής

"Soft State" στο IGMP

Παράδειγμα "soft state" στο IGMP:

Εάν ο τοπικός δρομολογητής δεν λάβει απόκριση σε ένα μήνυμα membership query μέχρι το timeout, συνάγει ότι όλα τα μέλη της ομάδας έχουν αποχωρήσει και η κατάσταση διαγράφεται αυτόματα από τη μνήμη χωρίς να απαιτείται κάθε μέλος που αποχωρεί να στέλνει ρητό μήνυμα αποχώρησης ("hard state")

Πρωτόκολλα με "soft state" είναι απλούστερα με αντάλλαγμα το bandwidth:

- διαγράφουν την κατάσταση αυτόματα

- δεν απαιτούν μηχανισμούς ανάκτησης κατάστασης για την ανάκαμψη από αστοχίες

Μοντέλο υπηρεσιών multicast (συμπεράσματα)

Στο επίπεδο δικτύου:

Η προσχώρηση σε μία ομάδα multicast οδηγείται από τον παραλήπτη (receiver-driven)

οποιοσδήποτε host μπορεί να προσχωρήσει σε μία ομάδα

ο αποστολέας δεν έχει έλεγχο ως προς το ποιος συμμετέχει στην ομάδα και λαμβάνει τα πακέτα multicast

Δεν υπάρχει έλεγχος ως προς το ποιος στέλνει πακέτα στην ομάδα

Οι hosts δεν γνωρίζουν την ταυτότητα των μελών της ομάδας

Δεν υπάρχει συντονισμός ως προς τη χρήση διευθύνσεων multicast

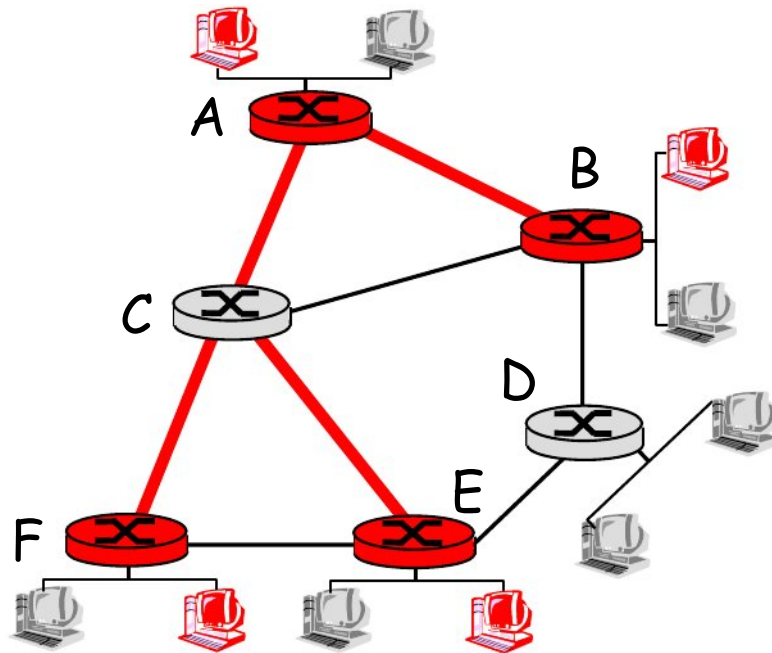
Είναι δυνατό δύο διαφορετικές ομάδες multicast να επιλέξουν την ίδια διεύθυνση multicast

Μηχανισμοί για την αντιμετώπιση των παραπάνω αδυναμιών μπορούν να υλοποιηθούν στο επίπεδο εφαρμογής

Δρομολόγηση multicast: διατύπωση προβλήματος

Στόχος: ανεύρεση δένδρου multicast

δένδρο multicast: δένδρο ζεύξεων που συνδέει τους δρομολογητές στους οποίους είναι προσαρτημένα τα μέλη της ομάδας multicast

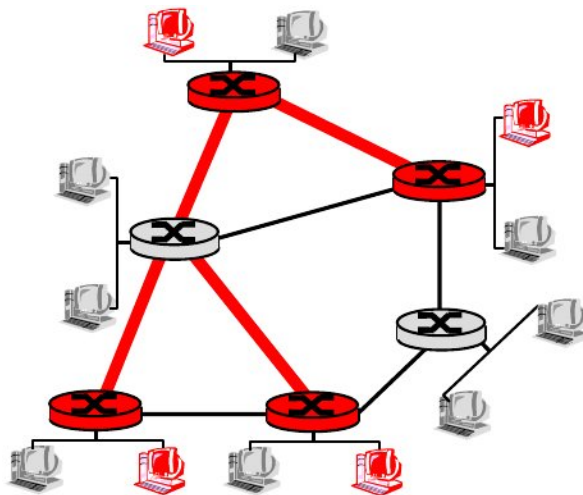


Δημιουργία δένδρου multicast

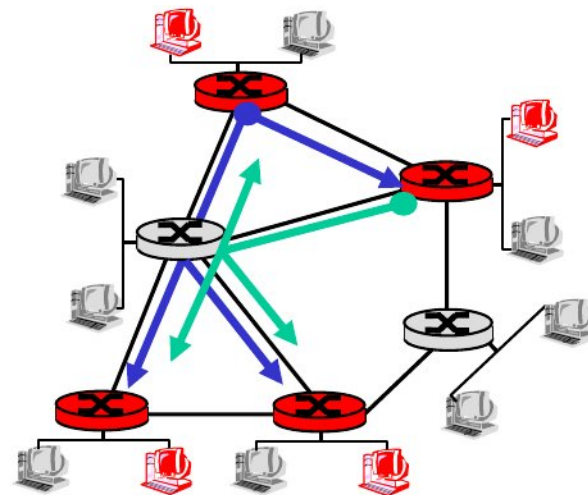
Δύο προσεγγίσεις προσδιορισμού του δένδρου multicast:

κοινού δένδρου (shared-tree): όλα τα μέλη της ομάδας χρησιμοποιούν το ίδιο δένδρο (ένα δένδρο για την ομάδα)
ελάχιστο εκτεινόμενο δένδρο (minimum spanning tree) (δένδρο Steiner)
δένδρο βασισμένο στο κέντρο (center-based tree)

δένδρου βασισμένου στην πηγή (source-based): διαφορετικό δένδρο από κάθε αποστολέα προς τους παραλήπτες (ένα δένδρο ανά πηγή)
δένδρο συντομότερων διαδρομών (shortest path tree)
προώθηση ανάστροφης διαδρομής (reverse path forwarding)



κοινό δένδρο



δένδρο βασισμένο στην πηγή

Κοινά δένδρα: το δένδρο Steiner

Δένδρο Steiner: δένδρο ελαχίστου κόστους που συνδέει όλους τους δρομολογητές στους οποίους είναι προσαρτημένα μέλη της ομάδας

η εύρεση του δένδρου Steiner είναι πρόβλημα NP-complete

υπάρχουν εξαιρετικά καλές προσεγγιστικές μέθοδοι
δεν χρησιμοποιείται στην πράξη:

- υπολογιστική πολυπλοκότητα

- απαιτείται πληροφορία σχετικά με όλες τις ζεύξεις του δικτύου

- μονολιθικό: ο αλγόριθμος εκτελείται ξανά κάθε φορά που ένας προστίθεται ή εξαιρείται ένας δρομολογητής

- δεν αξιοποιεί τους πίνακες δρομολόγησης που έχουν υπολογισθεί για τη δρομολόγηση unicast

Κοινά δένδρα: δένδρα βασισμένα στο κέντρο

Ένας δρομολογητής επιλέγεται ως **"κέντρο"** του δένδρου

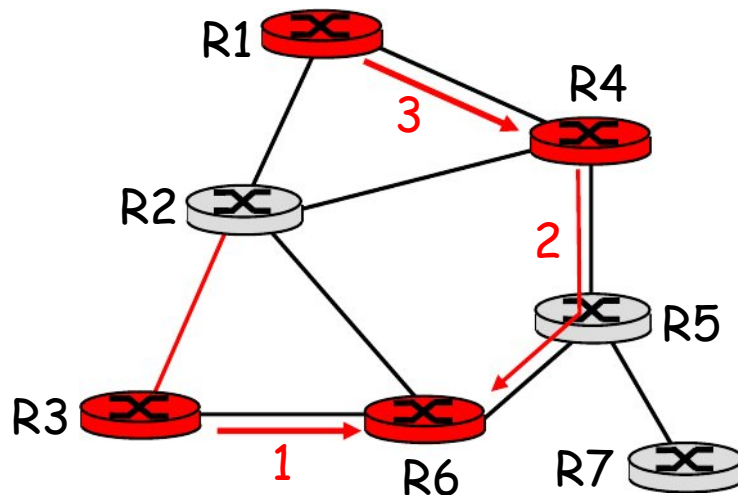
Για την συνένωση ενός ακραίου δρομολογητή με το κέντρο:

ο δρομολογητής στέλνει (unicast) ένα μήνυμα συνένωσης (join-msg) με προορισμό τον κεντρικό δρομολογητή
οι ενδιαμέσοι δρομολογητές "επεξεργάζονται" το μήνυμα συνένωσης και το προωθούν προς το κέντρο

το μήνυμα συνένωσης είτε φθάνει σε ένα δρομολογητή που ανήκει ήδη στο δένδρο multicast είτε φθάνει στο κέντρο
η διαδρομή που ακολουθείται από το μήνυμα συνένωσης αποτελεί ένα νέο κλάδο του δένδρου για αυτόν τον δρομολογητή

Δένδρα βασισμένα στο κέντρο: παράδειγμα

Έστω ότι ως κέντρο επιλέγεται ο δρομολογητής R6:



υπόμνημα



δρομολογητής στον οποίο είναι προσαρτημένα μέλη της ομάδας



δρομολογητής χωρίς προσαρτημένα μέλη ομάδας

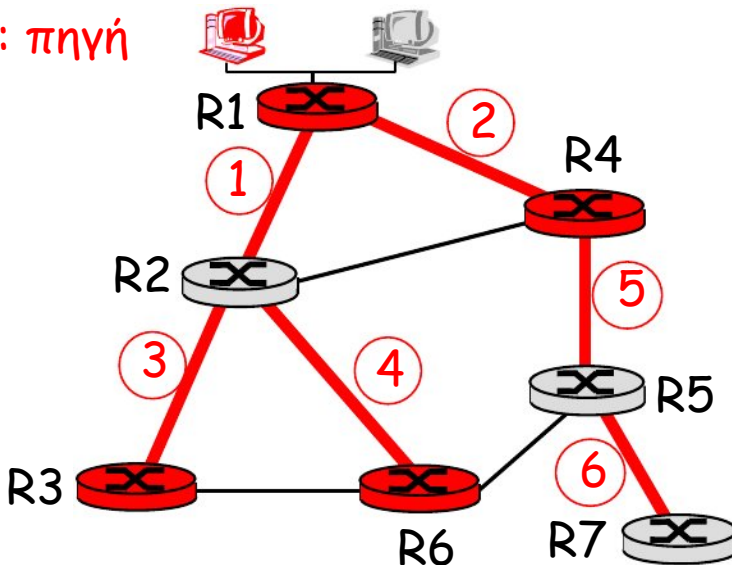


διαδρομή/σειρά με την οποία παράγονται τα μηνύματα συνένωσης

Δένδρο συντομότερων διαδρομών

δένδρο multicast: το δένδρο των συντομότερων διαδρομών από την πηγή προς όλους τους παραλήπτες
αλγόριθμος του Dijkstra

S: πηγή



υπόμνημα



δρομολογητής στον οποίο είναι προσαρτημένα μέλη της ομάδας



δρομολογητής χωρίς προσαρτημένα μέλη ομάδας



ζεύξη που χρησιμοποιείται για προώθηση,
ο δείκτης i υποδεικνύει τη σειρά με την οποία ο αλγόριθμος προσέθεσε τη ζεύξη

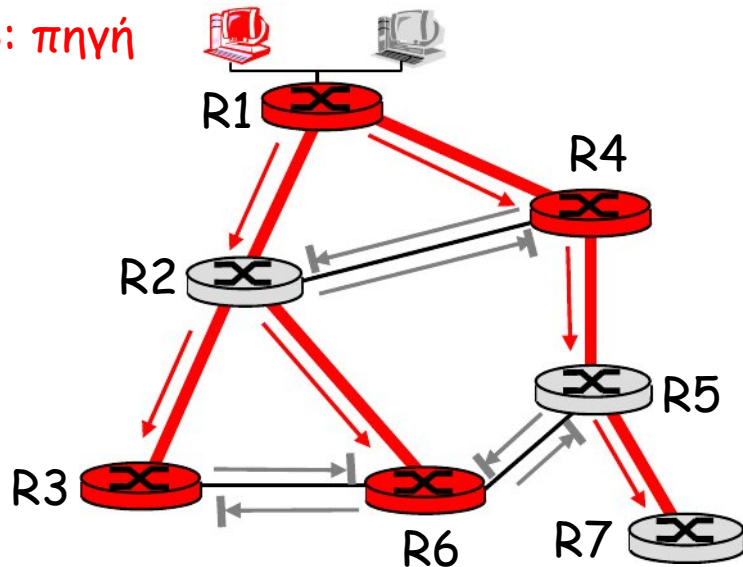
Πρωώθηση Ανάστροφης Διαδρομής (Reverse Path Forwarding, RPF)

βασίζεται στη γνώση που έχει ο κάθε δρομολογητής της συντομότερης διαδρομής unicast από τον ίδιο στον αποστολέα
κάθε δρομολογητής προωθεί ένα πακέτο multicast το οποίο λαμβάνει στην εισερχόμενη ζεύξη ως εξής:

εάν το πακέτο ληφθεί μέσω της συντομότερης διαδρομής προς τον αποστολέα, το πακέτο προωθείται σε καθεμία από τις εξερχόμενες ζεύξεις αλλιώς το πακέτο απορρίπτεται

Πρώθηση ανάστροφης διαδρομής: παράδειγμα

S: πηγή



υπόμνημα

 δρομολογητής στον οποίο είναι προσαρτημένα μέλη της ομάδας

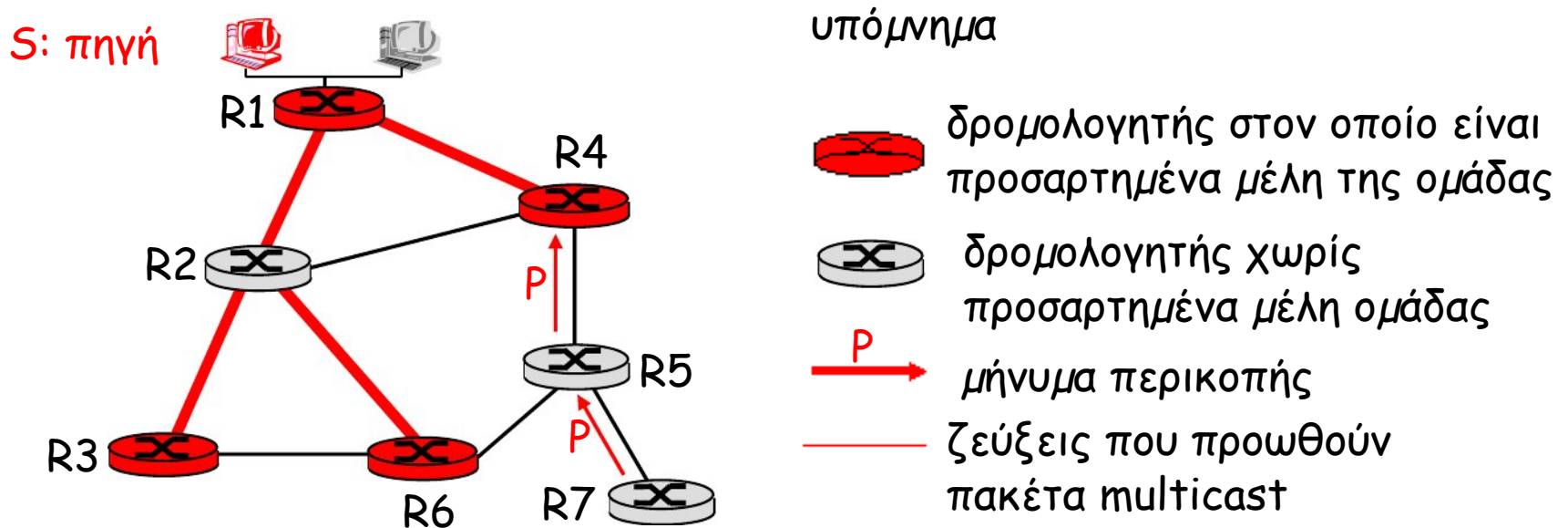
 δρομολογητής χωρίς προσαρτημένα μέλη ομάδας

 το πακέτο θα προωθηθεί

 το πακέτο δεν θα προωθηθεί

Πρώθηση ανάστροφης διαδρομής: περικοπή

το δένδρο κατά μήκος του οποίου προωθούνται τα πακέτα περιέχει υπο-δένδρα που δεν περιλαμβάνουν μέλη της ομάδας multicast
άσκοπη η πρώθηση πακέτων κατά μήκος αυτών των υπο-δένδρων
μηνύματα "περικοπής" στέλνονται "upstream" από τους δρομολογητές που δεν έχουν μέλη ομάδας "downstream"



Δρομολόγηση multicast στο Διαδίκτυο : DVMRP

DVMRP: distance vector multicast routing protocol (RFC 1075)

το DVMRP υλοποιεί δένδρα βασισμένα στην πηγή με **προώθηση ανάστροφης διαδρομής** και **περικοπή**

τα δένδρα που χρησιμοποιούνται στην προώθηση ανάστροφης διαδρομής δημιουργούνται με βάση τους πίνακες δρομολόγησης του ίδιου του DVMRP το οποίο χρησιμοποιεί έναν αλγόριθμο διανυσμάτων απόστασης

το αρχικό πακέτο που προορίζεται για την ομάδα multicast στέλνεται σε όλους τους δρομολογητές (πλημμύρα) με RPF οι δρομολογητές που δεν επιθυμούν να λάβουν τα πακέτα multicast στέλνουν "upstream" μηνύματα περικοπής

DVMRP (συνέχεια)

soft state: ο δρομολογητής DVMRP επαναφέρει περιοδικά (κάθε 1 min) τα κλαδιά του δένδρου που έχουν περικοπεί:

- πακέτα multicast διανύουν και πάλι τα κλαδιά που είχαν περικοπεί προηγουμένως

- οι δρομολογητές μπορούν να στείλουν νέο μήνυμα περικοπής ή να συνεχίσουν να λαμβάνουν τα πακέτα

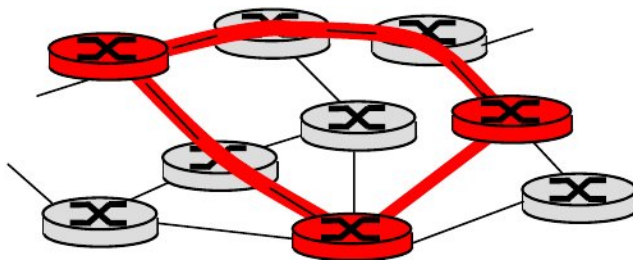
- οι δρομολογητές μπορούν να επανασυνδεθούν γρήγορα με το δένδρο

- στέλνοντας upstream μηνύματα "graft"

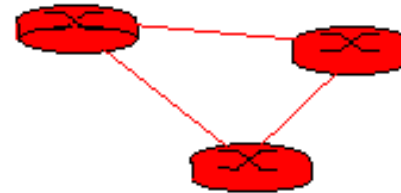
το DVMRP χρησιμοποιείται για τη δρομολόγηση στο Mbone

Χρήση σήραγγας

Στο Διαδίκτυο "νησίδες" δρομολογητών multicast συνδέονται μεταξύ τους με τη μέθοδο της σήραγγας



φυσική τοπολογία



λογική τοπολογία

τα πακέτα multicast ενθυλακώνονται μέσα σε πακέτα unicast
τα πακέτα unicast στέλνονται στον παραλήπτη δρομολογητή multicast μέσω "σήραγγας"

ο παραλήπτης δρομολογητής multicast αποθυλακώνει το πακέτο multicast από το πακέτο unicast

PIM: Protocol Independent Multicast

ανεξάρτητο από το πρωτόκολλο δρομολόγησης που χρησιμοποιείται για unicast

δύο διαφορετικά σενάρια multicast ανάλογα με τον που κατανέμονται τα μέλη της ομάδας:

Πυκνός τρόπος (dense mode):

τα μέλη της ομάδας πυκνά κατανεμημένα, οι περισσότεροι δρομολογητές της περιοχής συμμετέχουν στην πολυεκπομπή άφθονο bandwidth

Αραιός τρόπος (sparse mode):

αριθμός δικτύων με μέλη της ομάδας μικρός σε σχέση με τον αριθμό των διασυνδεδεμένων δικτύων μέλη της ομάδας "ευρέως διασκορπισμένα" περιορισμένο bandwidth

Συνέπειες της διχοτόμησης σε αραιό/πυκνό τρόπο:

Πυκνός τρόπος:

Θεωρείται ότι όλοι οι δρομολογητές συμμετέχουν στην πολυεκπομπή έως ότου οι δρομολογητές στείλουν μήνυμα περικοπής η δημιουργία του δένδρου multicast καθοδηγείται από τα δεδομένα (data-driven) (π.χ. RPF) σπατάλη bandwidth και άσκοπη επεξεργασία από δρομολογητές που δεν έχουν μέλη της ομάδας

Αραιός τρόπος:

οι δρομολογητές πρέπει να ζητήσουν συνένωση με το δένδρο multicast η δημιουργία του δένδρου multicast καθοδηγείται από τους παραλήπτες (receiver-driven) (π.χ. δένδρα βασισμένα στο κέντρο) συντηρητική χρήση του bandwidth και της επεξεργασίας από δρομολογητές που δεν έχουν μέλη της ομάδας

ΡΙΜ- Πυκνός τρόπος

μέθοδος προώθησης ανάστροφης διαδρομής με πλημμύρα και περικοπή, παρόμοια με το πρωτόκολλο DVMRP με τις ακόλουθες διαφορές:

λαμβάνει πληροφορία για την προώθηση ανάστροφης διαδρομής των πακέτων που εισέρχονται σε ένα δρομολογητή από το πρωτόκολλο unicast

επειδή βασίζεται στο υποκείμενο πρωτόκολλο unicast, η μέθοδος της πλημμύρας που εφαρμόζει είναι λιγότερο πολύπλοκη και αποδοτική από αυτή του DVMRP

PIM - Αραιός τρόπος

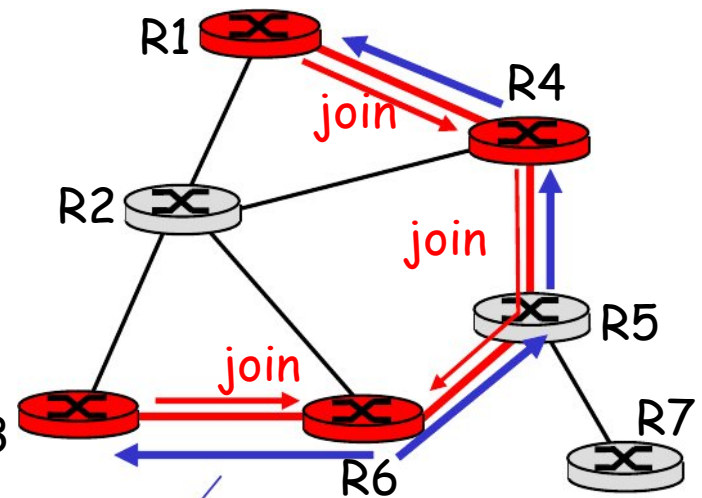
μέθοδος δένδρου βασισμένου
στο κέντρο

ο δρομολογητής στέλνει
μήνυμα συνένωσης στο
κέντρο (rendezvous point,
RP)

οι ενδιαμέσοι δρομολογητές
ενημερώνουν την "κατάσταση"
και προωθούν το μήνυμα
συνένωσης

μετά τη συνένωση μέσω RP, ο R3
δρομολογητής μπορεί να
χρησιμοποιήσει ένα δένδρο
βασισμένο στην πηγή

βελτίωση απόδοσης: μικρότερη
συγκέντρωση κίνησης,
συντομότερες διαδρομές



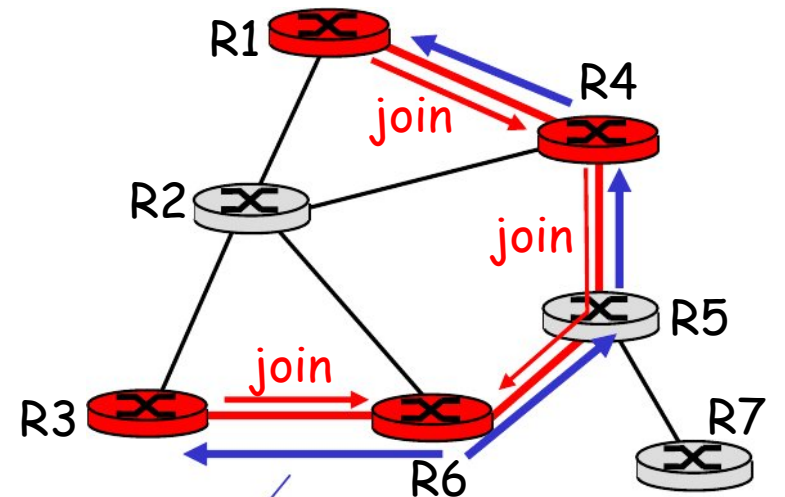
πολυεκπομπή όλων
των δεδομένων από
το κέντρο (rendezvous
point)

rendezvous
point RP

PIM - Αραιός τρόπος

αποστολέας:

στέλνει δεδομένα με unicast στο κέντρο RP, το οποίο τα διανέμει κατά μήκος του δένδρου με ρίζα στο RP
το κέντρο RP μπορεί να στείλει μήνυμα stop στον αποστολέα εφόσον δεν υπάρχουν προσαρτημένοι παραλήπτες



πολυεκπομπή όλων των δεδομένων από το κέντρο (rendezvous point)

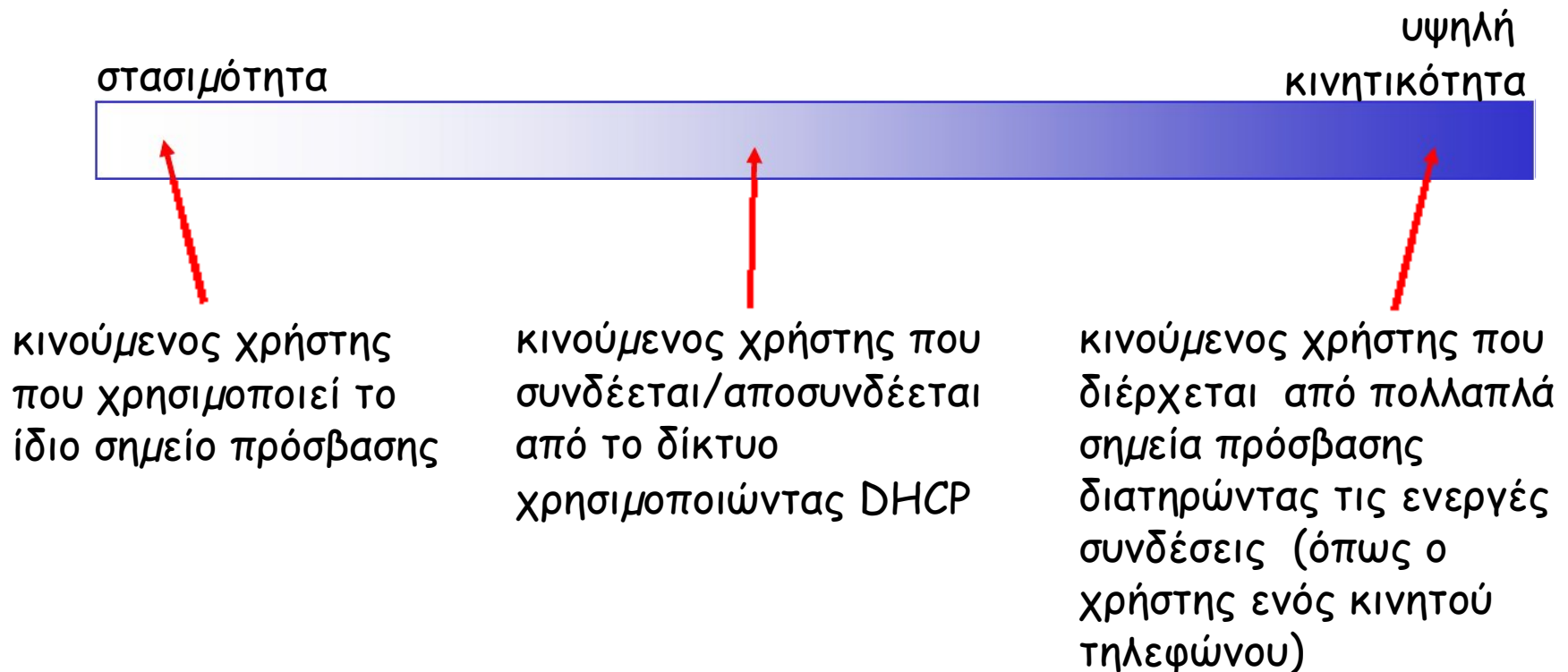
rendezvous point

Περίγραμμα

- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
- 4 Αρχές Δρομολόγησης
- 5 Δρομολόγηση στο Διαδίκτυο
- 6 IPv6
- 7 Ειδικά θέματα του στρώματος Δικτύου
 - 7.1 Broadcast/Multicast
 - 7.2 Mobility
 - 7.3 Quality of Service (QoS)

Τι είναι κινητικότητα (mobility);

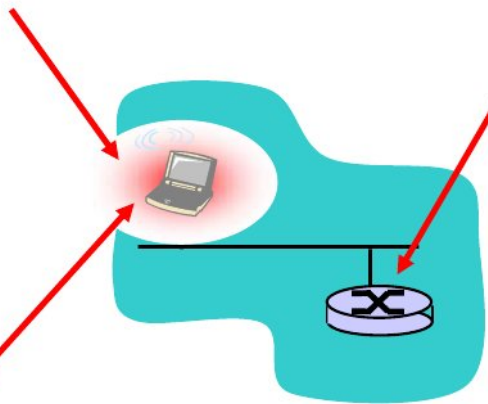
φάσμα κινητικότητας από την άποψη του δικτύου:



Κινητικότητα: ορολογία

οικιακό δίκτυο (home network):

“μόνιμη κατοικία” του κινητού
(π.χ. 128.119.40/24)

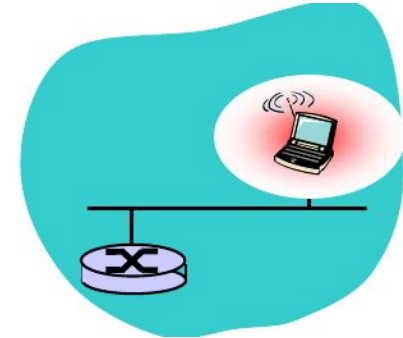


**μόνιμη διεύθυνση
(permanent address):**

διεύθυνση του κινητού
στο οικιακό δίκτυο -
μπορεί να χρησιμοποιηθεί
πάντοτε για να
προσπελάσουμε το κινητό
(π.χ. 128.119.40.186)

οικιακός πράκτορας (home agent):

οντότητα που εκτελεί λειτουργίες
διαχείρισης κινητικότητας εκ
μέρους του κινητού, όταν το κινητό
βρίσκεται εκτός οικιακού δικτύου



Κινητικότητα: ορολογία

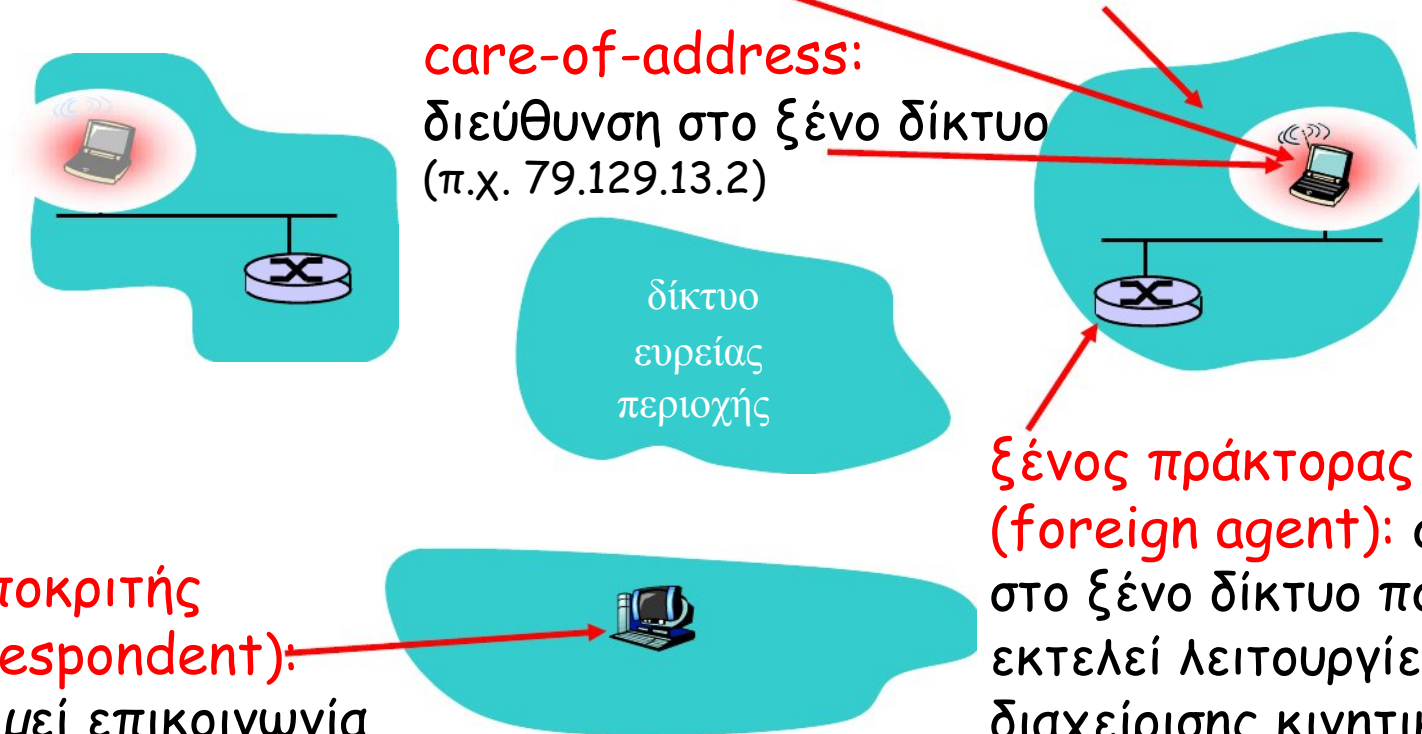
μόνιμη διεύθυνση: παραμένει σταθερή (π.χ. 128.119.40.186)

ξένο δίκτυο (foreign network) ή δίκτυο επίσκεψης (visited network): δίκτυο στο οποίο είναι προσαρτημένο το κινητό όταν δεν βρίσκεται στο οικιακό δίκτυο (π.χ. 79.129.13/24)

care-of-address: διεύθυνση στο ξένο δίκτυο (π.χ. 79.129.13.2)

ανταποκριτής (correspondent): επιθυμεί επικοινωνία με το κινητό

ξένος πράκτορας (foreign agent): οντότητα στο ξένο δίκτυο που εκτελεί λειτουργίες διαχείρισης κινητικότητας εκ μέρους του κινητού



Κινητικότητα: προσεγγίσεις

Ανάθεση της διαχείρισης της κινητικότητας στους δρομολογητές:

οι δρομολογητές αναγγέλλουν τις μόνιμες διευθύνσεις των κινητών κόμβων που επισκέπτονται το δίκτυό τους μέσω της συνήθους ανταλλαγής πληροφορίας δρομολόγησης

οι πίνακες δρομολόγησης υποδεικνύουν το δίκτυο στο οποίο είναι προσαρτημένο κάθε κινητό

καμία αλλαγή στα τερματικά συστήματα

Ανάθεση της διαχείρισης της κινητικότητας στα ακραία συστήματα του δικτύου:

ο home agent στο home network παρακολουθεί τις μετακινήσεις του κινητού και γνωρίζει το foreign network στο οποίο είναι προσαρτημένο

ο home agent μεσολαβεί ώστε τα πακέτα που προορίζονται για το κινητό να δρομολογούνται στο foreign network

το κινητό έχει δύο διευθύνσεις: permanent address στο home network και care-of-address στο foreign network

Κινητικότητα: προσεγγίσεις

Ανάθεση της διαχείρισης της κινητικότητας στους δρομολογητές:
οι δρομολογητές αναγγέλλουν τις μόνιμες διευθύνσεις των
κινητών κόμβων που επισκέπτονται το δίκτυό τους μέσω της
συνήθους ανταλλαγής πληροφορίας δρομολόγησης
οι πίνακες δρομολόγησης υποδεικνύουν το δίκτυο στο οποίο
είναι προσαρτημένο κάθε κινητό
καμία αλλαγή στα τερματικά συστήματα

δεν
εφαρμόζεται
σε εκατομμύρια
κινητών

Ανάθεση της διαχείρισης της κινητικότητας στα ακραία συστήματα του δικτύου:

ο home agent στο home network παρακολουθεί τις μετακινήσεις του κινητού και γνωρίζει το foreign network στο οποίο είναι προσαρτημένο

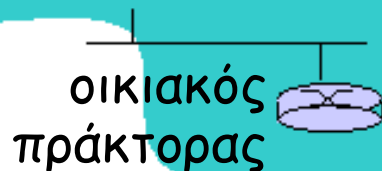
ο home agent μεσολαβεί ώστε τα πακέτα που προορίζονται για το κινητό να δρομολογούνται στο foreign network

το κινητό έχει δύο διευθύνσεις: permanent address στο home network και care-of-address στο foreign network

Κινητικότητα: διεύθυνση επιμέλειας

οικιακό δίκτυο
128.119.40/24

μόνιμη διεύθυνση
128.119.40.186



δίκτυο
ευρείας
περιοχής

ξένο δίκτυο
79.129.13/24

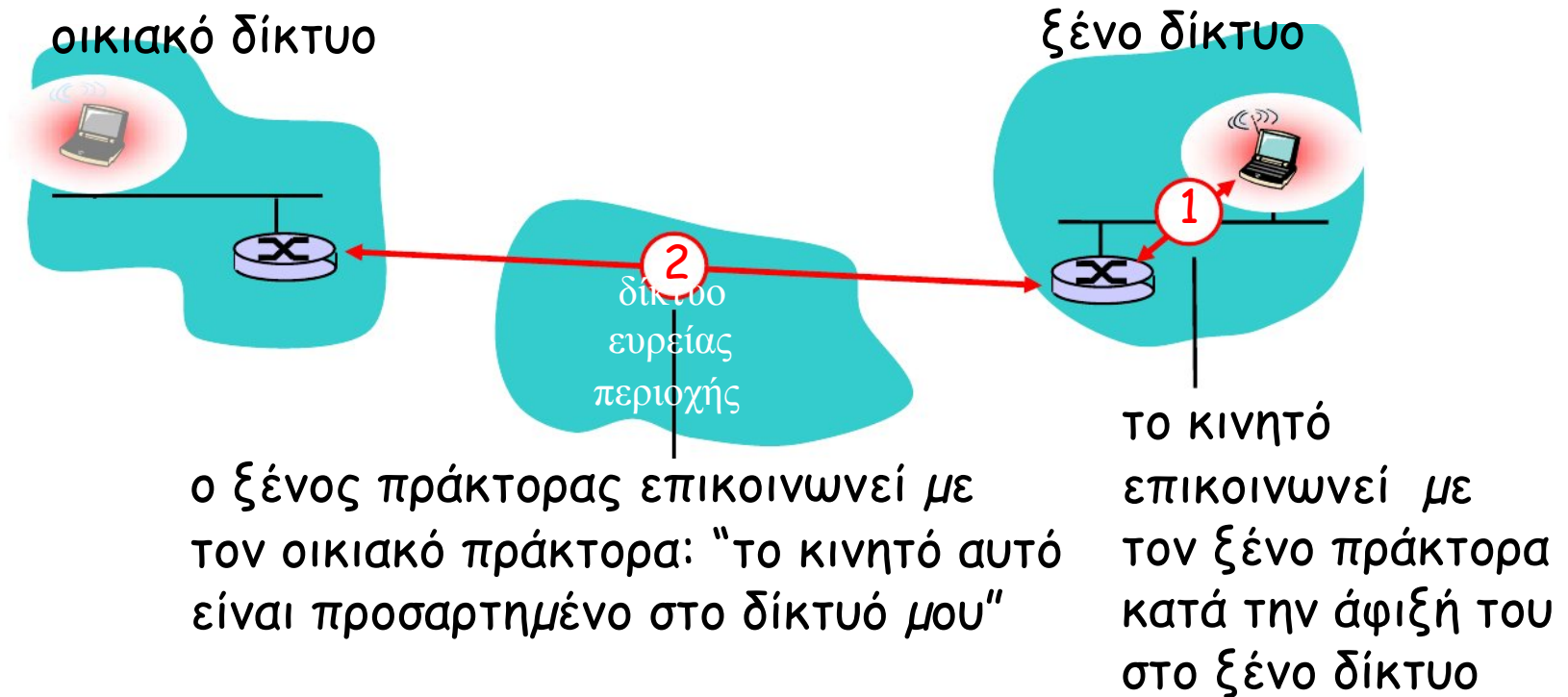
διεύθυνση επιμέλειας
79.129.13.2

κινητός
κόμβος



correspondent

Κινητικότητα: εγγραφή (registration)



Αποτέλεσμα:

Ο ξένος πράκτορας γνωρίζει σχετικά με το κινητό

Ο οικιακός πράκτορας γνωρίζει που βρίσκεται το κινητό

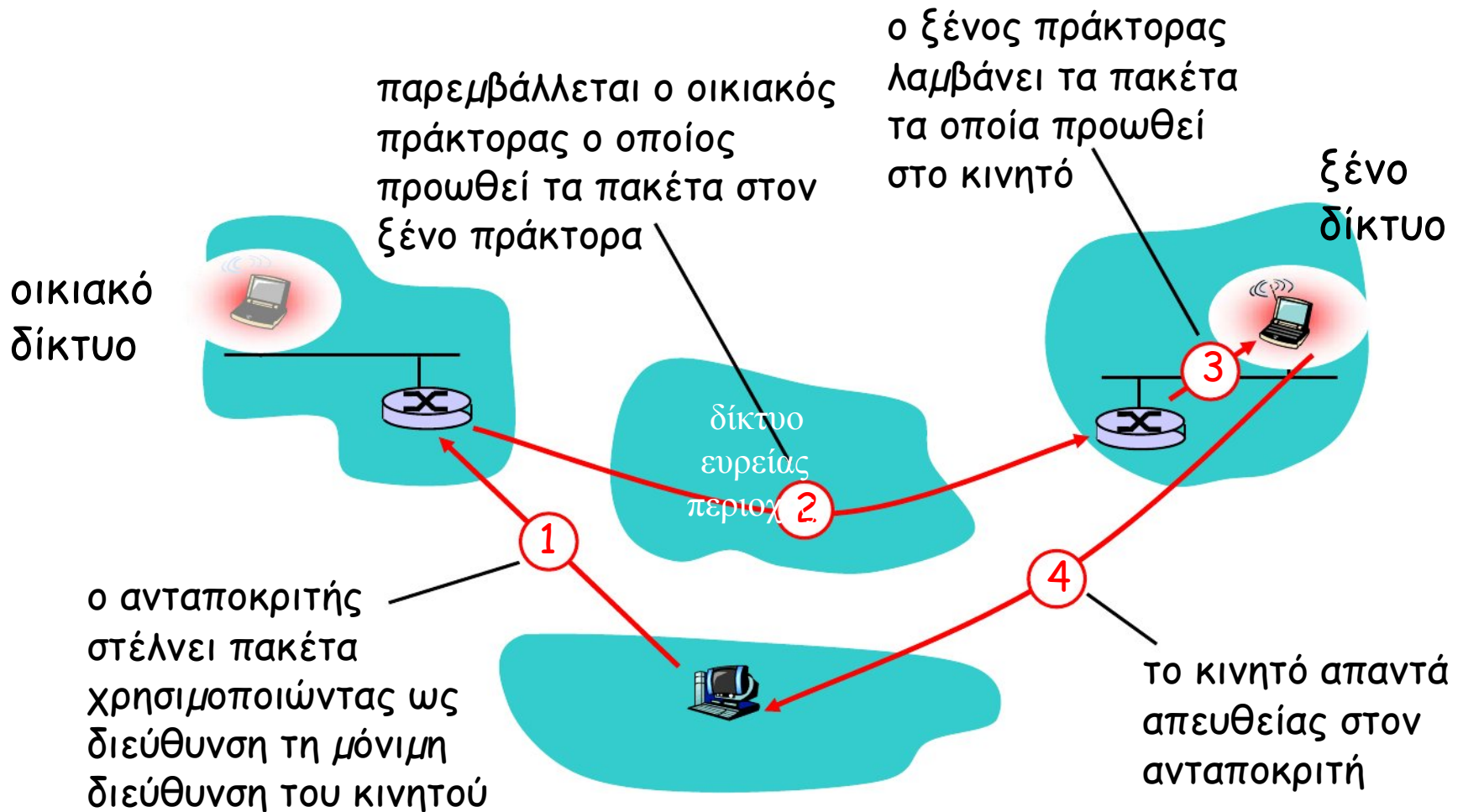
Κινητικότητα: προώθηση πακέτων

Δύο μέθοδοι προώθησης πακέτων από τον ανταποκριτή στο κινητό:

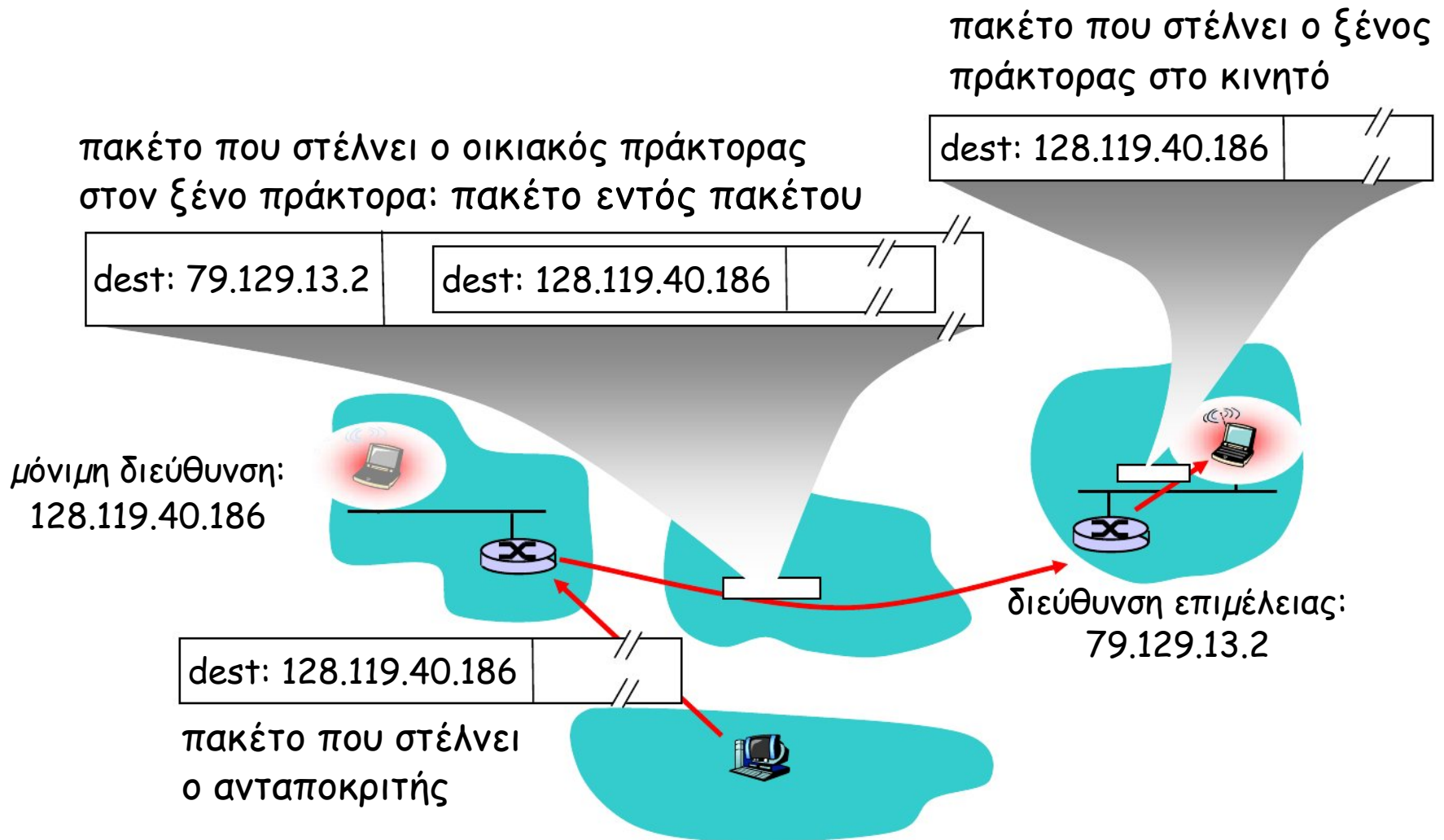
έμμεση δρομολόγηση: πακέτα από τον ανταποκριτή για το κινητό διέρχονται από τον οικιακό πράκτορα, στη συνέχεια προωθούνται στο ξένο δίκτυο

άμεση δρομολόγηση: ο ανταποκριτής αφού λάβει τη διεύθυνση του κινητού στο ξένο δίκτυο, στέλνει πακέτα απευθείας στο κινητό

Έμμεση δρομολόγηση



Πρώθηση πακέτων στο απομακρυσμένο κινητό



Έμμεση δρομολόγηση: σχόλια

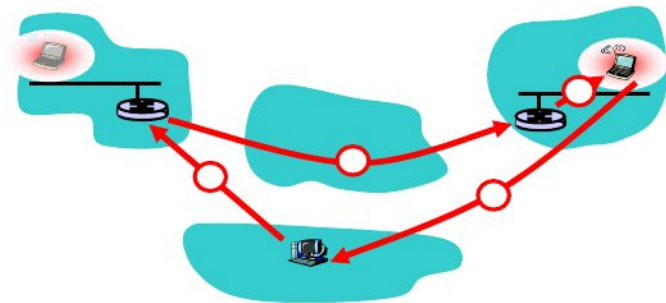
το κινητό χρησιμοποιεί δύο διευθύνσεις:

μόνιμη διεύθυνση: χρησιμοποιείται από τον ανταποκριτή (η θέση του κινητού είναι **διαφανής** στον ανταποκριτή)

διεύθυνση επιμέλειας: χρησιμοποιείται από τον οικιακό πράκτορα για να προωθήσει τα πακέτα στο κινητό

οι λειτουργίες του ξένου πράκτορα μπορούν εναλλακτικά να εκτελεσθούν από το ίδιο το κινητό

δρομολόγηση τριγώνου: ανταποκριτής - οικιακό δίκτυο - κινητό
αναποτελεσματική όταν
ανταποκριτής και κινητό
βρίσκονται στο ίδιο δίκτυο



Έμμεση δρομολόγηση: μετακίνηση μεταξύ δικτύων

έστω ότι ο κινητός χρήστης μετακινείται σε ένα άλλο δίκτυο

εγγραφή του κινητού στο νέο ξένο πράκτορα

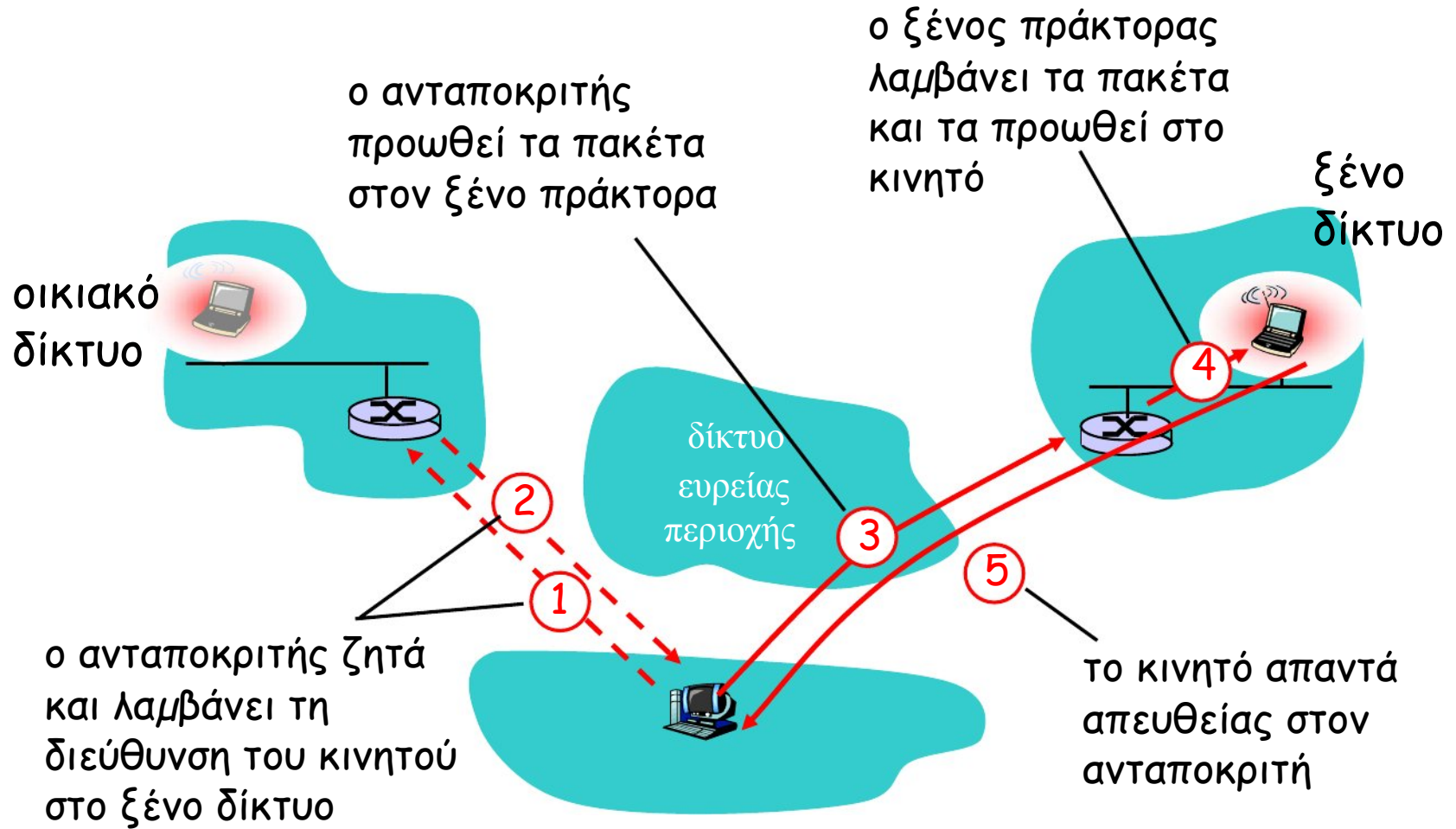
εγγραφή του νέου ξένου πράκτορα στον οικιακό πράκτορα

ο οικιακός πράκτορας ενημερώνει την διεύθυνση επιμέλειας του κινητού

τα πακέτα συνεχίζουν να προωθούνται στο κινητό (με νέα διεύθυνση επιμέλειας)

κινητικότητα, μετακίνηση μεταξύ ξένων δικτύων
διαφανείς: **οι ενεργές συνδέσεις διατηρούνται
κατά τις μετακινήσεις**

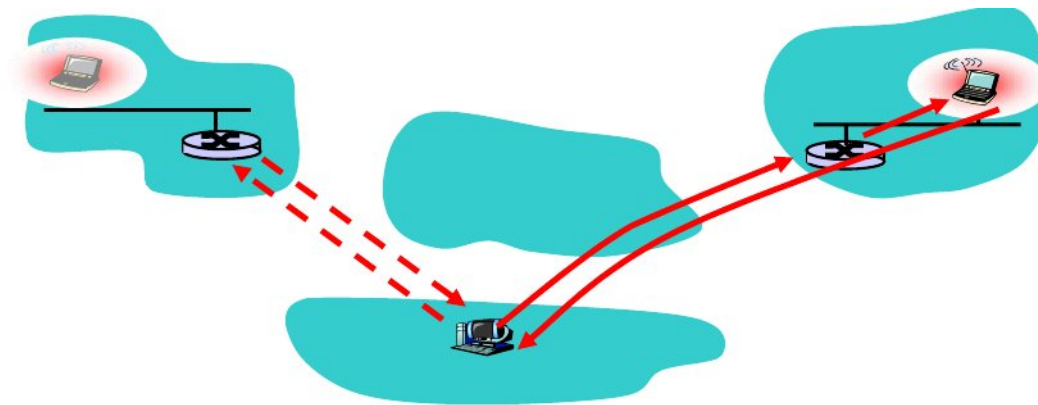
Άμεση δρομολόγηση



Άμεση δρομολόγηση: σχόλια

δεν αντιμετωπίζει το πρόβλημα της
δρομολόγησης τριγώνου

η κινητικότητα δεν είναι διαφανής στον
ανταποκριτή: ο ανταποκριτής πρέπει να λάβει τη
διεύθυνση επιμέλειας από τον οικιακό πράκτορα
Τι συμβαίνει όταν το κινητό αλλάζει δίκτυο;



Mobile IP

RFC 3344

περιλαμβάνει πολλά από τα χαρακτηριστικά που αναφέρθηκαν:

home agent, foreign agent, εγγραφή στον foreign-agent, διεύθυνση επιμέλειας, ενθυλάκωση

τρεις κύριες ενότητες στο πρότυπο:

διαδικασία ανακάλυψης πράκτορα (agent discovery)

διαδικασία εγγραφής κινητού στον οικιακό πράκτορα

έμμεση δρομολόγηση πακέτων

Περίγραμμα

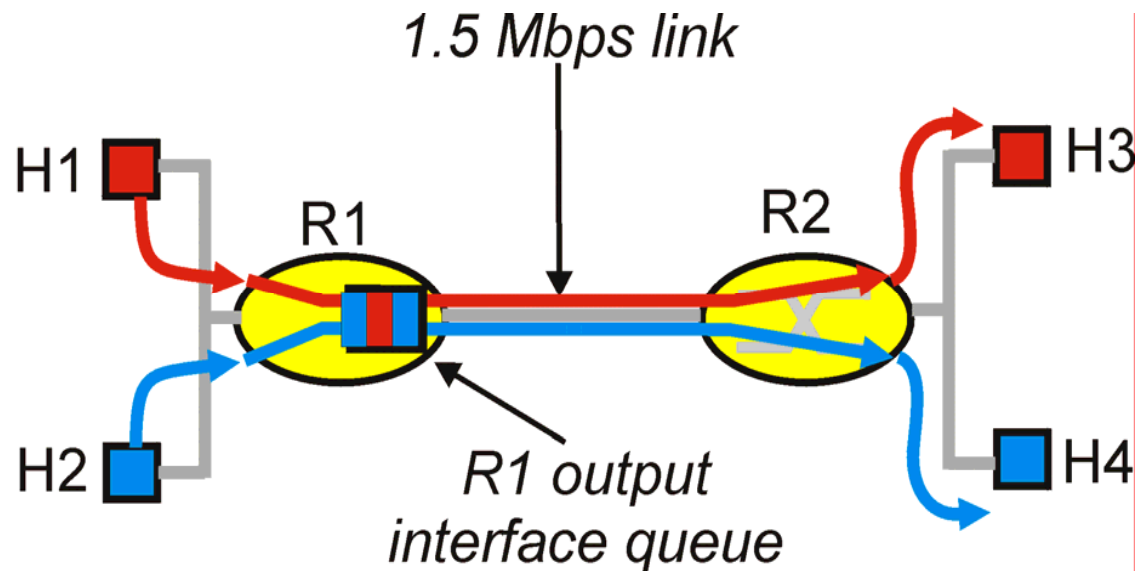
- 1 Εισαγωγή και Μοντέλα Υπηρεσιών Δικτύου
- 2 Ιεραρχική Δρομολόγηση
- 3 Το Πρωτόκολλο IP (Internet Protocol)
- 4 Αρχές Δρομολόγησης
- 5 Δρομολόγηση στο Διαδίκτυο
- 6 IPv6
- 7 Ειδικά θέματα του στρώματος Δικτύου
 - 7.1 Broadcast/Multicast
 - 7.2 Mobility
 - 7.3 Quality of Service (QoS)

Improving QoS in IP Networks

Thus far: “making the best of best effort”

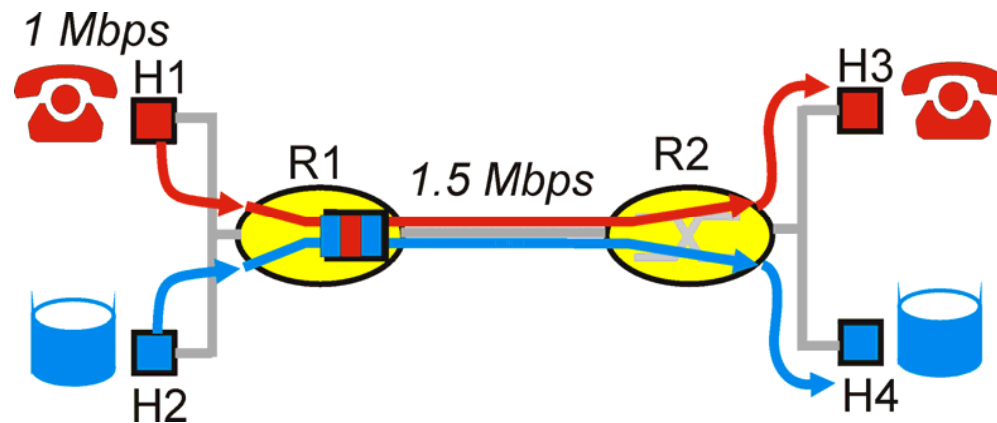
Future: next generation Internet with QoS guarantees

- **RSVP:** signaling for resource reservations
 - **Differentiated Services:** differential guarantees
 - **Integrated Services:** firm guarantees
- simple model for sharing and congestion studies:



Principles for QOS Guarantees

- Example: 1Mbps P phone, FTP share 1.5 Mbps link.
 - bursts of FTP can congest router, cause audio loss
 - want to give priority to audio over FTP

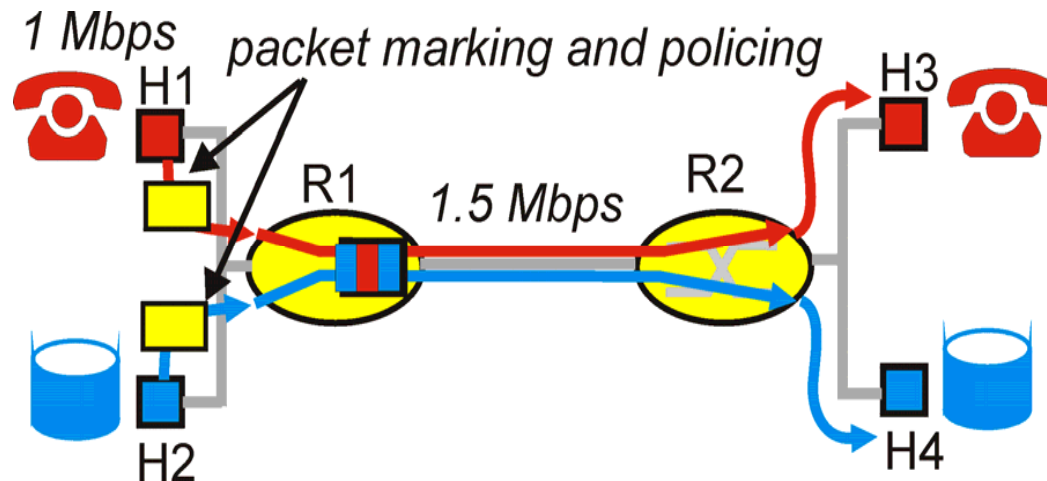


Principle 1

packet marking needed for router to distinguish between different classes; and new router policy to treat packets accordingly

Principles for QOS Guarantees

- what if applications misbehave (audio sends higher than declared rate)
 - policing: force source adherence to bandwidth allocations
- marking and policing at network edge:
 - similar to ATM UNI (User Network Interface)

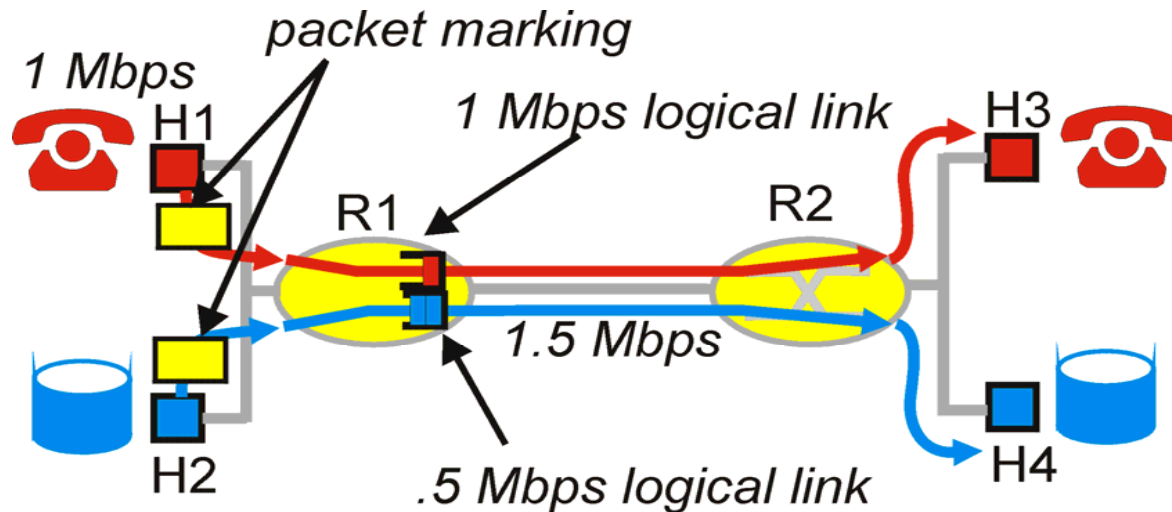


Principle 2

provide protection (*isolation*) for one class from others

Principles for QOS Guarantees

- Allocating *fixed* (non-sharable) bandwidth to flow: *inefficient* use of bandwidth if flows doesn't use its allocation

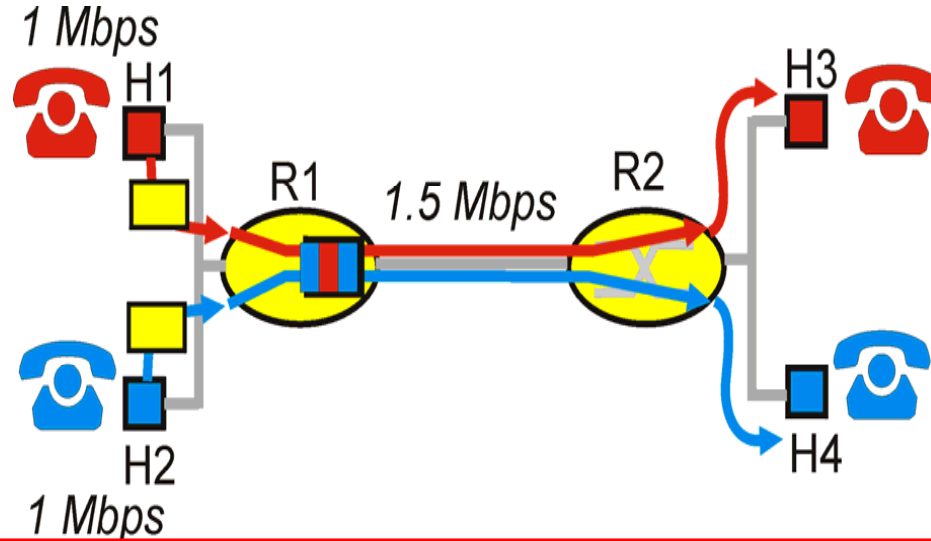


Principle 3

While providing isolation, it is desirable to use resources as efficiently as possible

Principles for QOS Guarantees (more)

- *Basic fact of life:* can not support traffic demands beyond link capacity

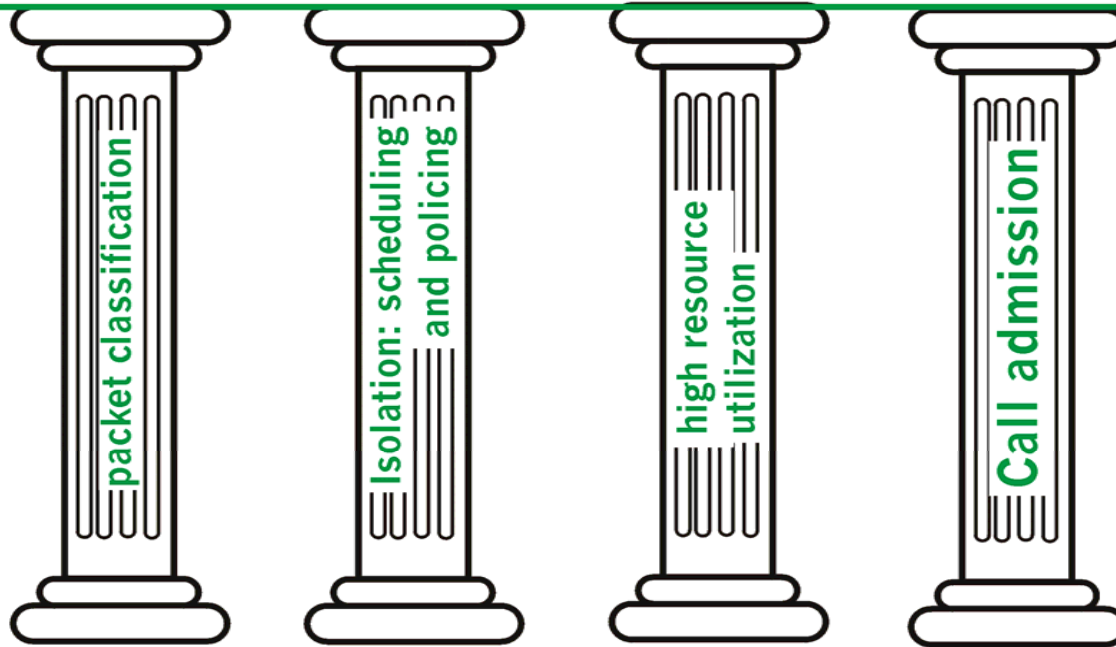


Principle 4

Call Admission: flow declares its needs, network may block call (e.g., busy signal) if it cannot meet needs

Summary of QoS Principles

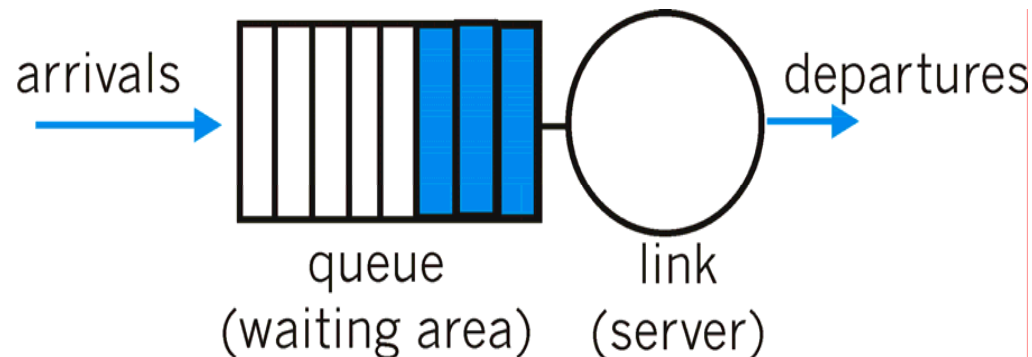
QoS for networked applications



Let's next look at mechanisms for achieving this

Scheduling And Policing Mechanisms

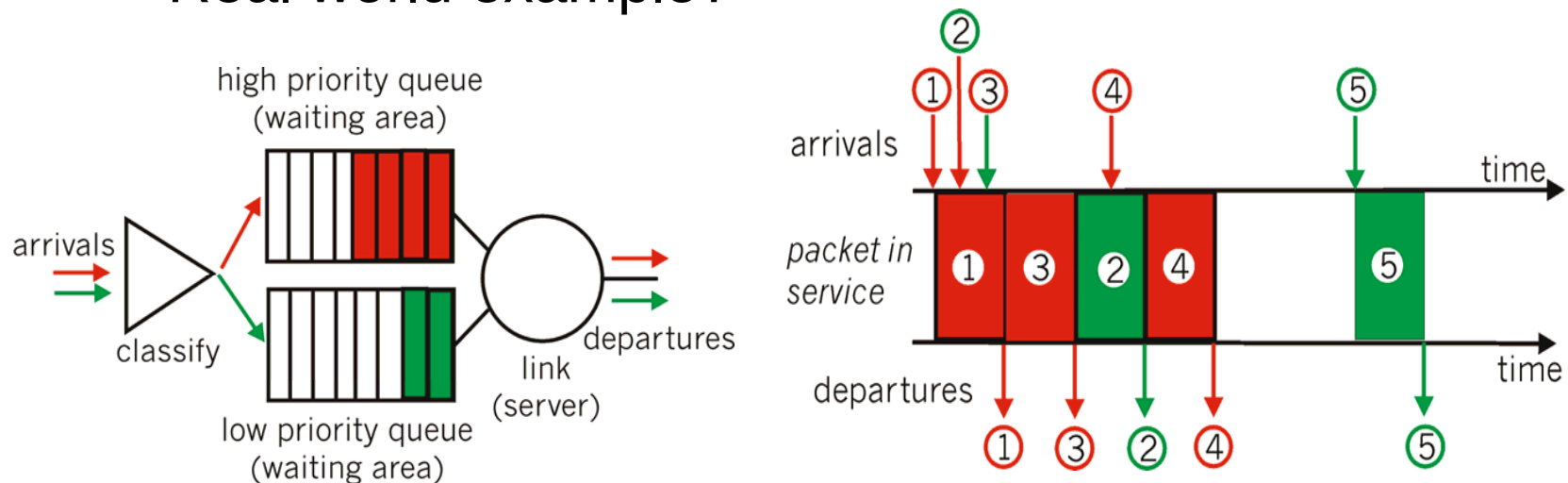
- **scheduling**: choose next packet to send on link
- **FIFO (first in first out) scheduling**: send in order of arrival to queue
 - real-world example?
 - **discard policy**: if packet arrives to full queue: who to discard?
 - Tail drop: drop arriving packet
 - priority: drop/remove on priority basis
 - random: drop/remove randomly



Scheduling Policies: more

Priority scheduling: transmit highest priority queued packet

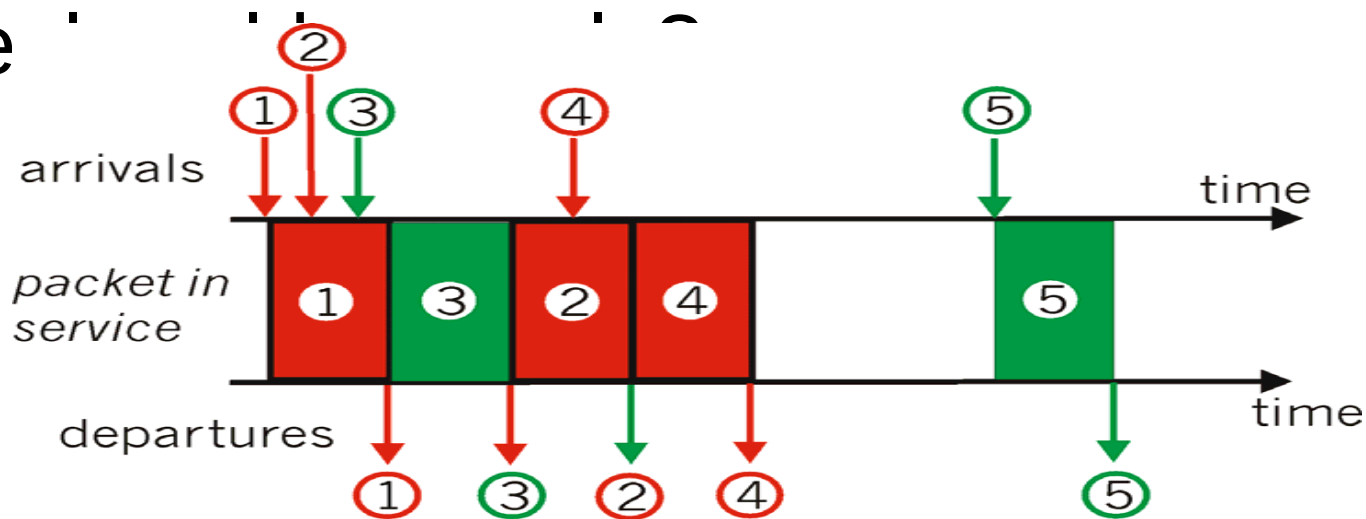
- multiple *classes*, with different priorities
 - class may depend on marking or other header info, e.g. IP source/dest, port numbers, etc..
 - Real world example?



Scheduling Policies: still more

round robin scheduling:

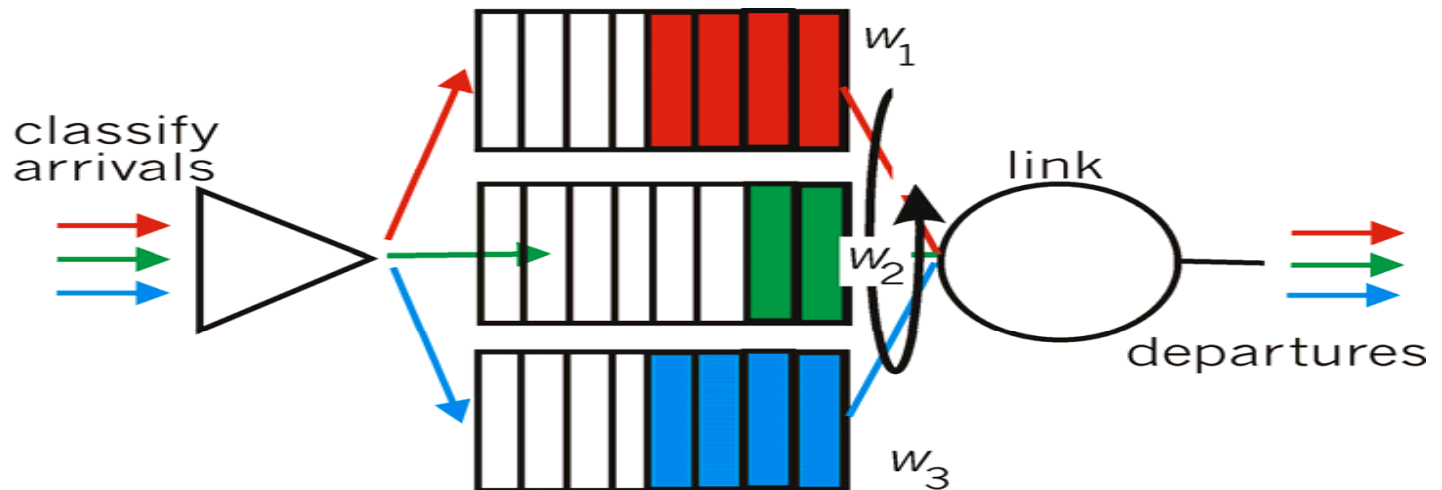
- multiple classes
- cyclically scan class queues, serving one from each class (if available)
- re



Scheduling Policies: still more

Weighted Fair Queuing:

- generalized Round Robin
- each class gets weighted amount of service in each cycle
- real-world example?



Policing Mechanisms

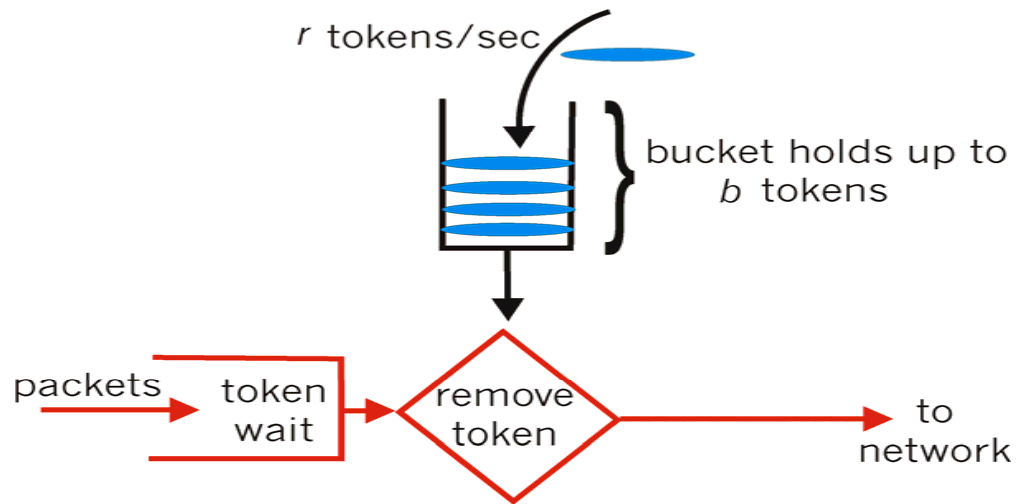
Goal: limit traffic to not exceed declared parameters

Three common-used criteria:

- *(Long term) Average Rate:* how many pkts can be sent per unit time (in the long run)
 - crucial question: what is the interval length: 100 packets per sec or 6000 packets per min have same average!
- *Peak Rate:* e.g., 6000 pkts per min. (ppm) avg.; 1500 ppm peak rate
- *(Max.) Burst Size:* max. number of pkts sent consecutively (with no intervening idle)

Policing Mechanisms

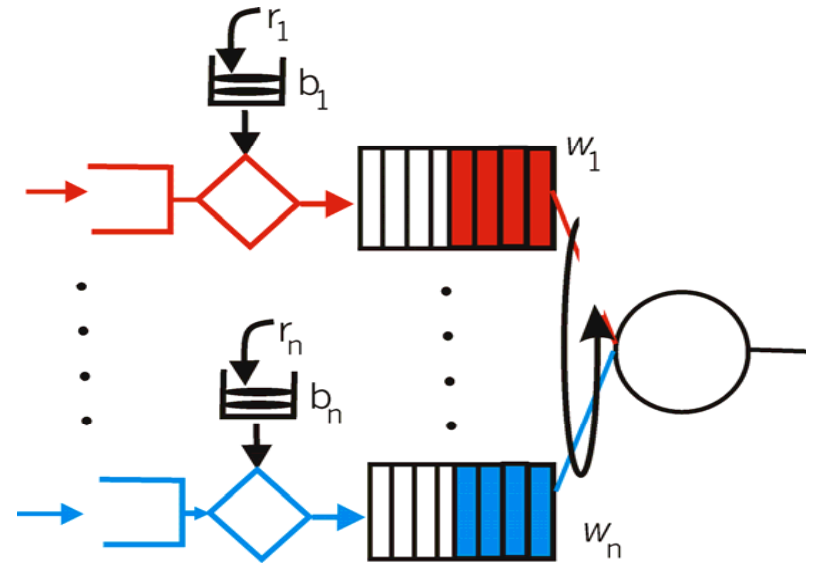
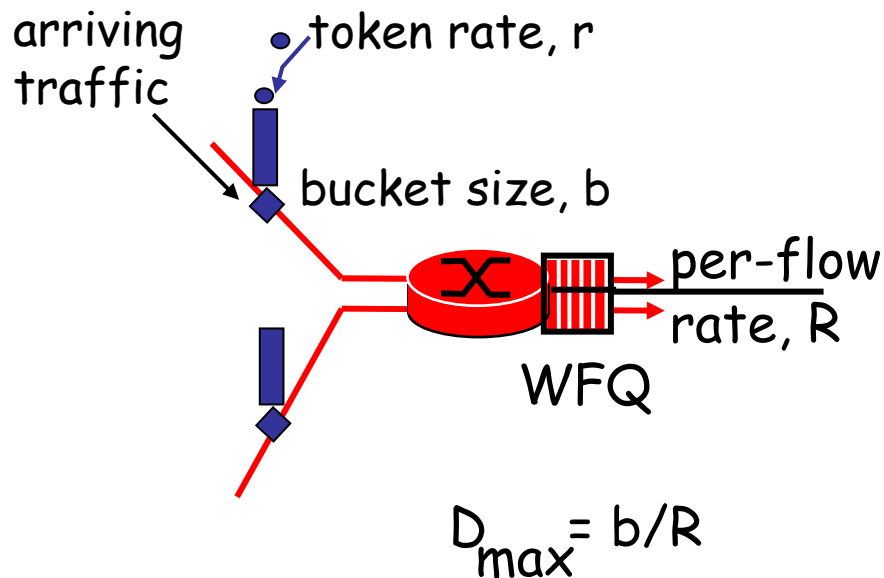
Token Bucket: limit input to specified Burst Size and Average Rate.



- bucket can hold b tokens
- tokens generated at rate r token/sec unless bucket full
- *over interval of length t : number of packets admitted less than or equal to $(r t + b)$.*

Policing Mechanisms (more)

- token bucket, WFQ combine to provide guaranteed upper bound on delay, i.e., *QoS guarantee!*



IETF Integrated Services

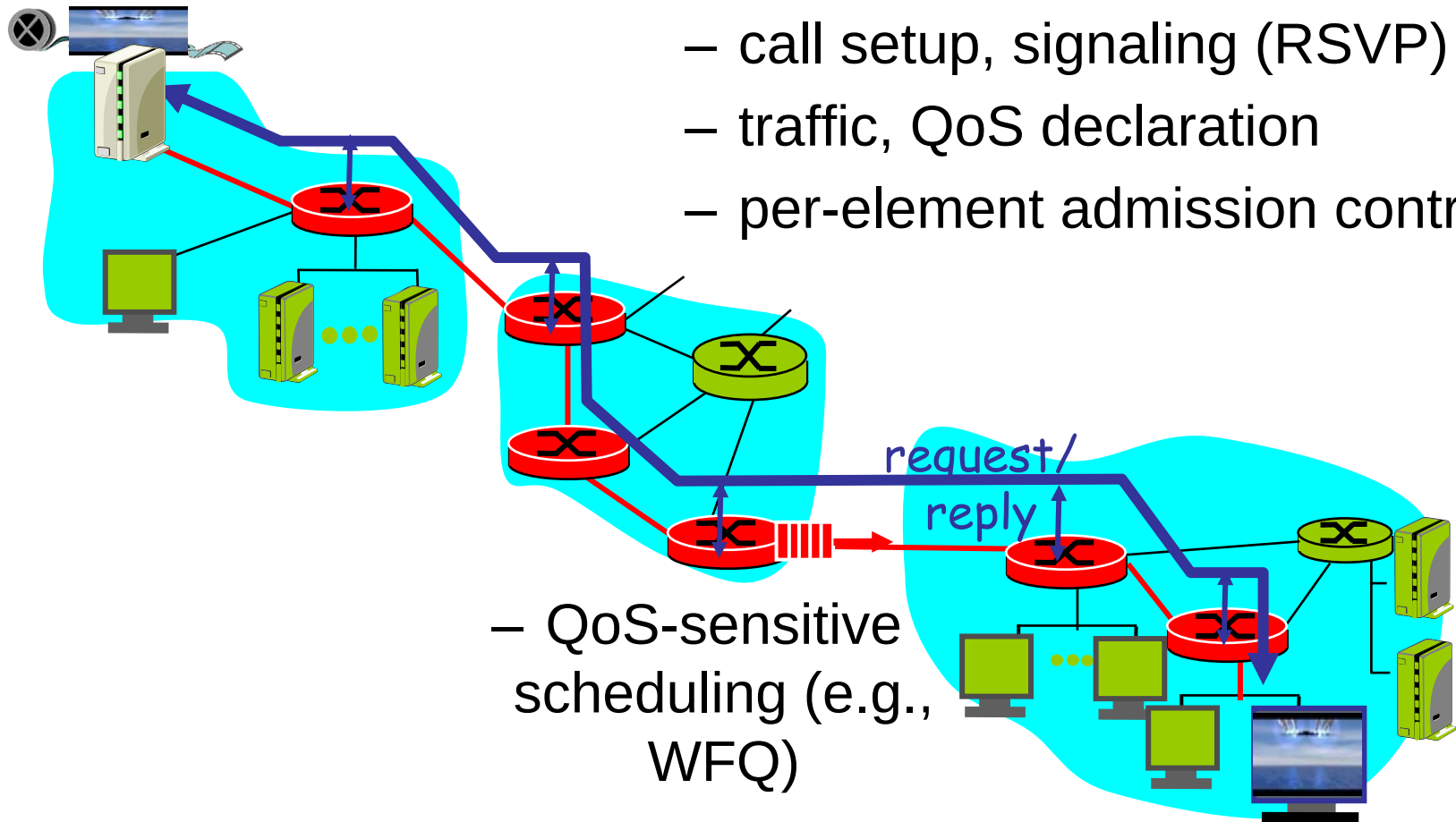
- architecture for providing QoS guarantees in IP networks for individual application sessions
- resource reservation: routers maintain state info (a la VC) of allocated resources, QoS req's
- admit/deny new call setup requests:

Question: can newly arriving flow be admitted with performance guarantees while not violates QoS guarantees made to already admitted flows?

Intserv: QoS guarantee scenario

- Resource reservation

- call setup, signaling (RSVP)
- traffic, QoS declaration
- per-element admission control



Call Admission

Arriving session must :

- declare its QOS requirement
 - **R-spec**: defines the QOS being requested
- characterize traffic it will send into network
 - **T-spec**: defines traffic characteristics
- signaling protocol: needed to carry R-spec and T-spec to routers (where reservation is required)
 - **RSVP**

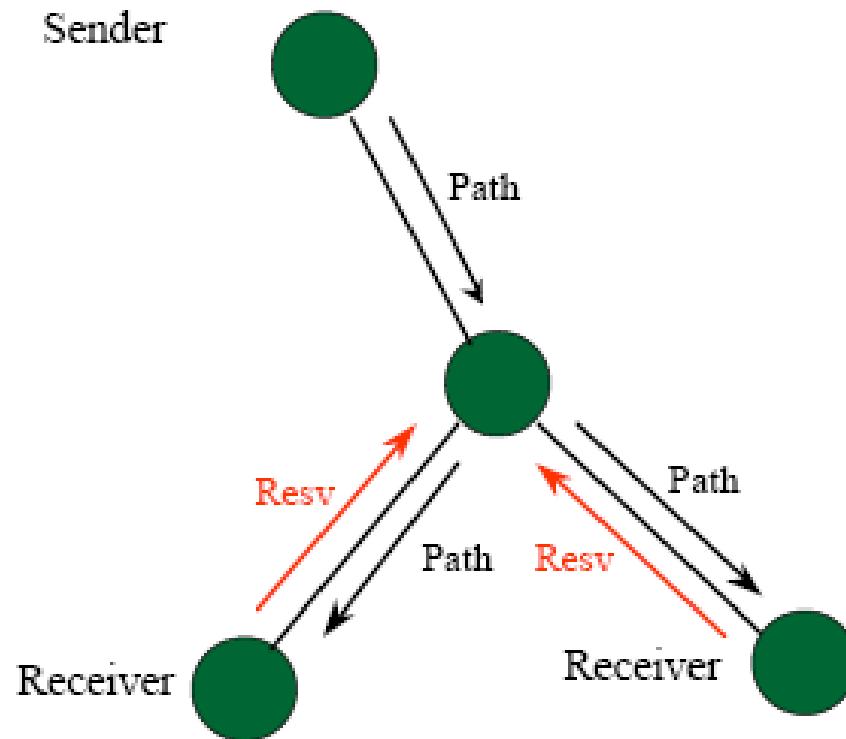
Resource Reservation Protocol

- The defacto setup protocol for IntServ is Resource Reservation Protocol (RSVP).
- The source transmits a PATH message along the routed path to the destination (unicast or multicast).
- PATH: marks the routed path across the network and collects information about the QoS viability of each router.

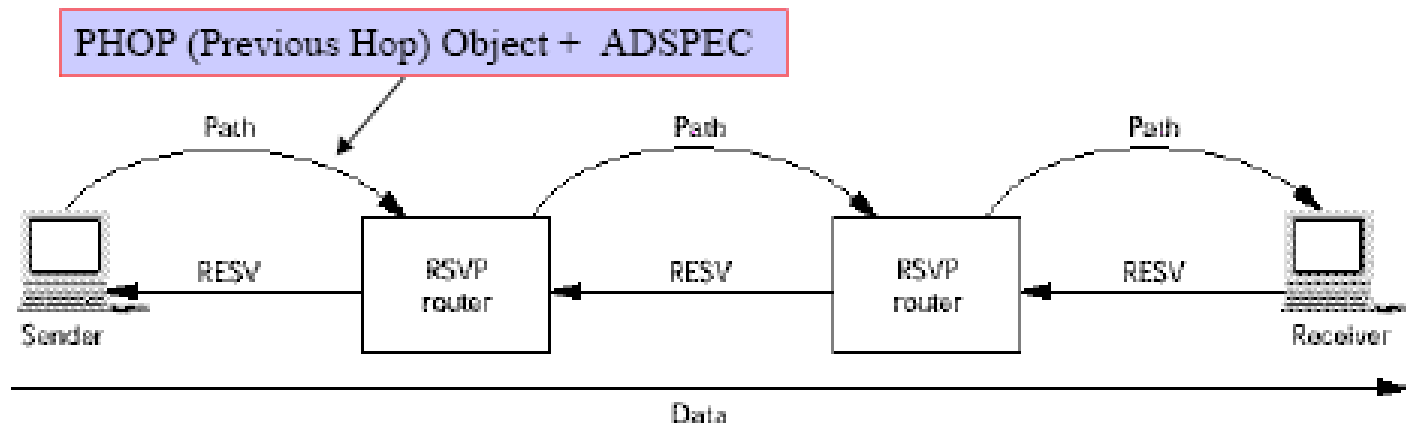
Resource Reservation Protocol

- Protocol is unidirectional. Establishes reservation state in one direction.
- Receiver initiated reservations.
- Independent from current unicast/multicast routing protocols.

RSVP basic message flow



RSVP message flow



- PHOP allows the RESV messages to find their route upstream.
- ADSPEC: summarises the path's characteristics and delivers this info to the receiver. OPWA: One Pass With Advertisement

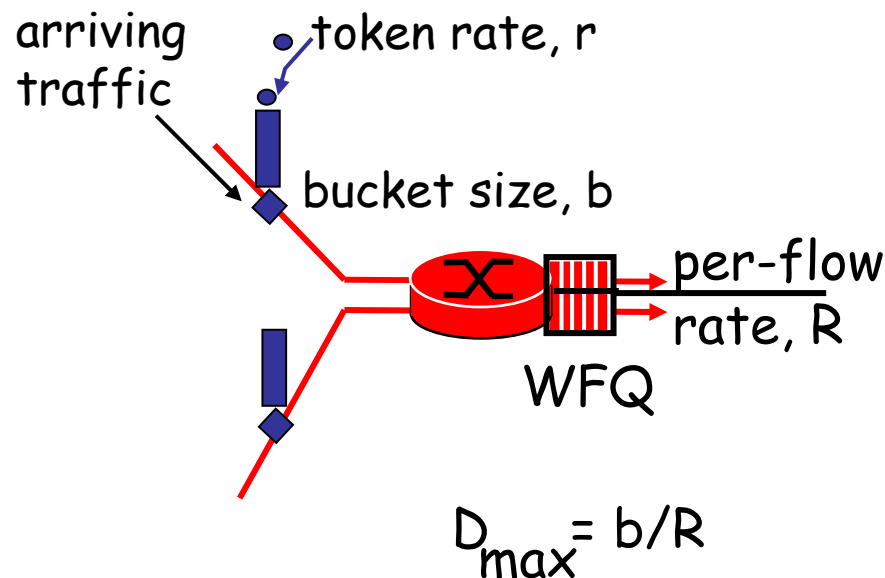
Intserv QoS: Service models [rfc2211, rfc 2212]

Guaranteed service:

- worst case traffic arrival: leaky-bucket-policed source
- simple (mathematically provable) *bound* on delay [Parekh 1992, Cruz 1988]

Controlled load service:

- "a quality of service closely approximating the QoS that same flow would receive from an unloaded network element."



IETF Differentiated Services

Concerns with Intserv:

- **Scalability:** signaling, maintaining per-flow router state difficult with large number of flows
- **Flexible Service Models:** Intserv has only two classes. Also want “qualitative” service classes
 - “behaves like a wire”
 - relative service distinction: Platinum, Gold, Silver

Diffserv approach:

- simple functions in network core, relatively complex functions at edge routers (or hosts)
- Don't define definite service classes, provide functional components to build service classes

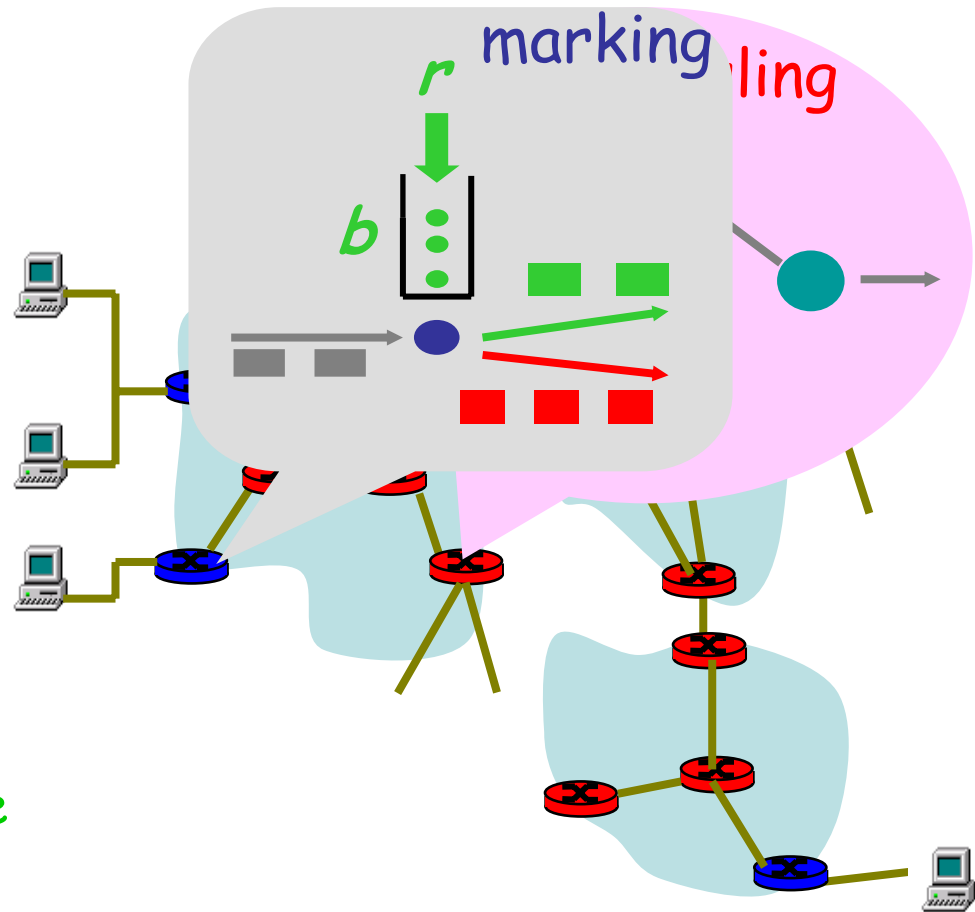
Diffserv Architecture

Edge router:

- per-flow traffic management
- marks packets as **in-profile** and **out-profile**

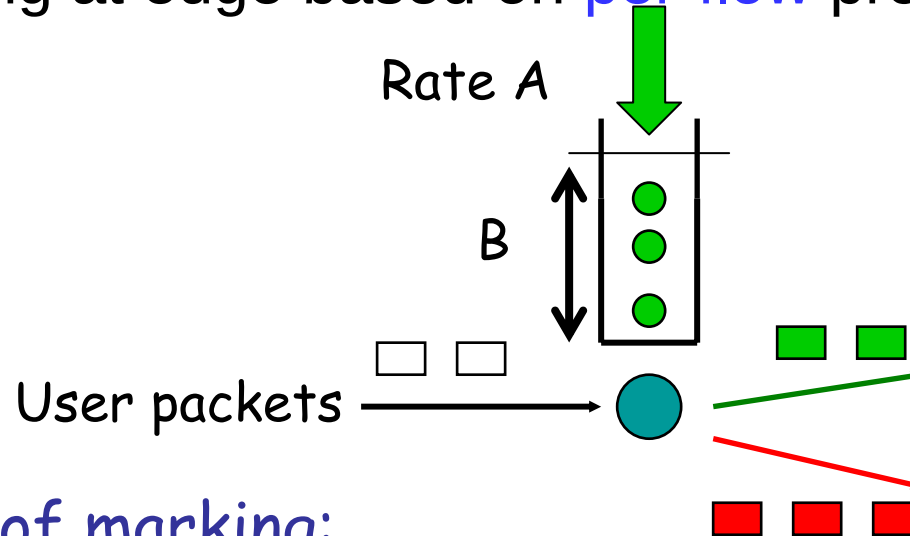
Core router:

- per class traffic management
- buffering and scheduling based on **marking** at edge
- preference given to **in-profile** packets
- Assured Forwarding



Edge-router Packet Marking

- **profile**: pre-negotiated rate A, bucket size B
- packet marking at edge based on **per-flow** profile



Possible usage of marking:

- class-based marking: packets of different classes marked differently
- intra-class marking: conforming portion of flow marked differently than non-conforming one

Classification and Conditioning

- Packet is marked in the Type of Service (TOS) in IPv4, and Traffic Class in IPv6
- 6 bits used for Differentiated Service Code Point (DSCP) and determine PHB that the packet will receive
- 2 bits are currently unused



Forwarding (PHB)

- PHB result in a different observable (measurable) forwarding performance behavior
- PHB does not specify what mechanisms to use to ensure required PHB performance behavior
- Examples:
 - Class A gets $x\%$ of outgoing link bandwidth over time intervals of a specified length
 - Class A packets leave first before packets from class B

Forwarding (PHB)

PHBs being developed:

- **Expedited Forwarding:** pkt departure rate of a class equals or exceeds specified rate
 - logical link with a minimum guaranteed rate
- **Assured Forwarding:** 4 classes of traffic
 - each guaranteed minimum amount of bandwidth
 - each with three drop preference partitions