



ΔΙΚΤΥΑ ΥΠΟΛΟΓΙΣΤΩΝ

Δρομολόγηση

Περίληψη



- Σύντομη αναφορά στις βασικές λειτουργίες του στρώματος δικτύου
- Βασικές αρχές της δρομολόγησης
- Δρομολογητές
- Μελέτη των δύο κυριότερων αλγορίθμων δρομολόγησης
- Εξειδίκευση και εφαρμογή των αλγορίθμων δρομολόγησης στο Internet

Περιεχόμενα

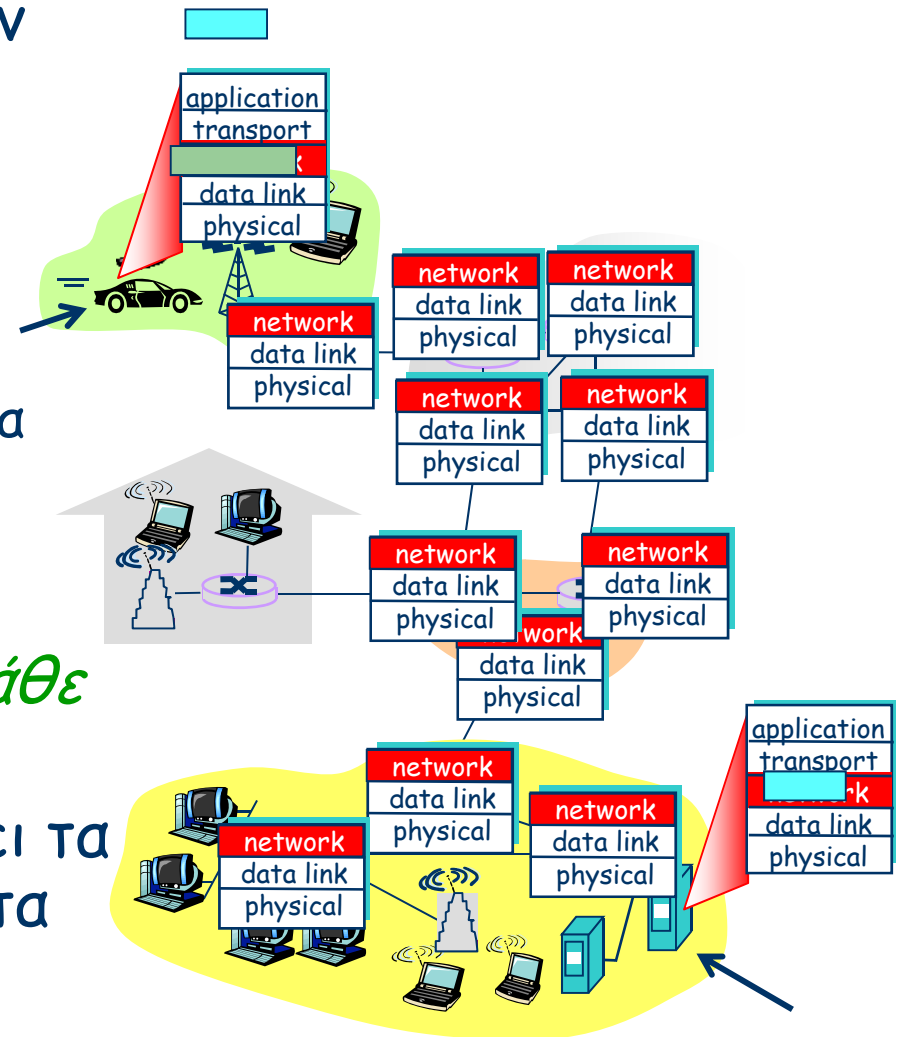


- Εισαγωγή και βασικές λειτουργίες του στρώματος δικτύου
 - Δίκτυα με νοητές συνδέσεις και με δεδομενογράμματα
- Βασικές αρχές της δρομολόγησης
- Δρομολογητές
 - Αρχιτεκτονική
 - Λειτουργίες
- Αλγόριθμοι δρομολόγησης
 - Διανύσματος αποστάσεων (Distance Vector)
 - Κατάστασης ζεύξεων (Link State)
- Ιεραρχική δρομολόγηση
- Πρωτόκολλα δρομολόγησης στο Internet

Στρώμα δικτύου



- Μεταφορά τεμαχίων από τον host αποστολής στον host λήψης
 - στην πλευρά αποστολής, τα τεμάχια ενθυλακώνονται σε δεδομενογράμματα
 - στην πλευρά λήψης, τα τεμάχια παραδίδονται στο στρώμα μεταφοράς
- Πρωτόκολλα στρώματος δικτύου σε *κάθε host* και *κάθε δρομολογητή*
- Κάθε δρομολογητής εξετάζει τα πεδία επικεφαλίδας σε όλα τα δεδομενογράμματα IP που διέρχονται από αυτόν

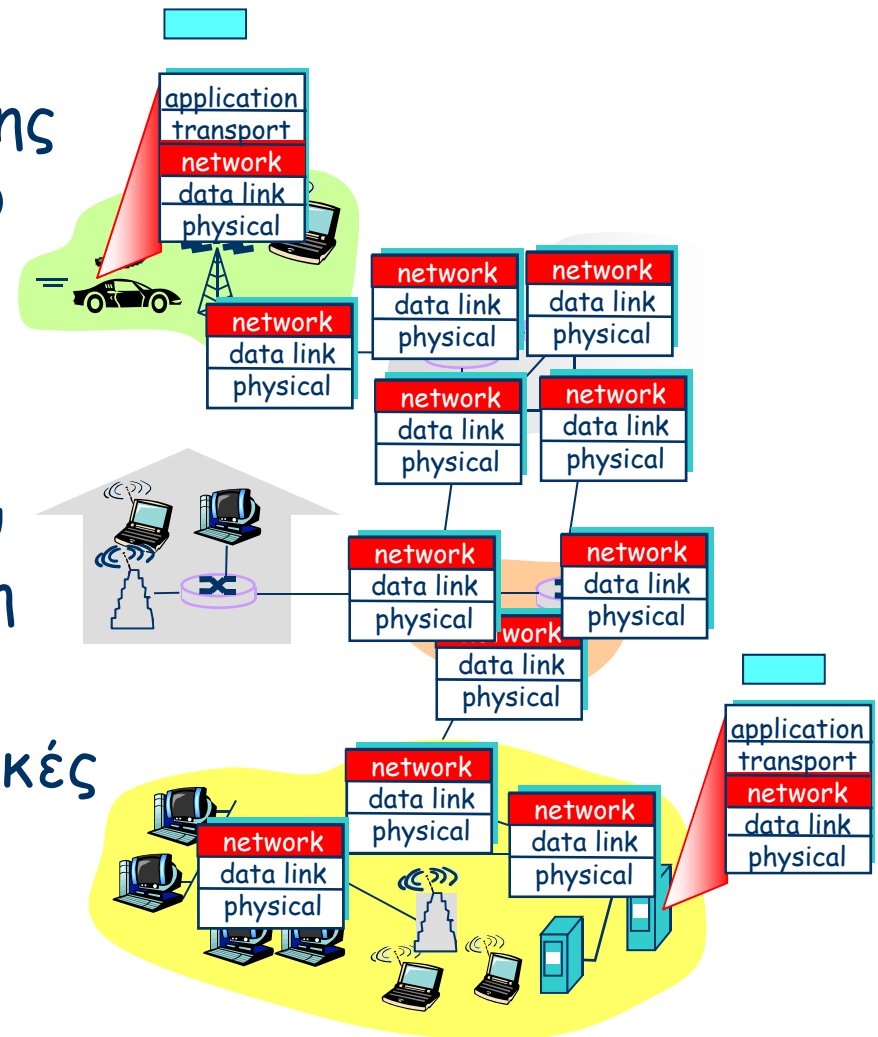


Λειτουργίες στρώματος δικτύου



Τρεις κύριες λειτουργίες:

- **Δρομολόγηση:** καθορισμός της διαδρομής των πακέτων από την πηγή στον προορισμό.
Αλγόριθμοι δρομολόγησης
- **Πρωώθηση:** μετακίνηση πακέτων από την είσοδο του δρομολογητή στην κατάλληλη έξοδο του
- **Εγκατάσταση σύνδεσης:** μερικές αρχιτεκτονικές δικτύου απαιτούν εγκατάσταση διαδρομής μεταξύ των δρομολογητών πριν την έναρξη αποστολής δεδομένων



Σχέση δρομολόγησης με προώθηση



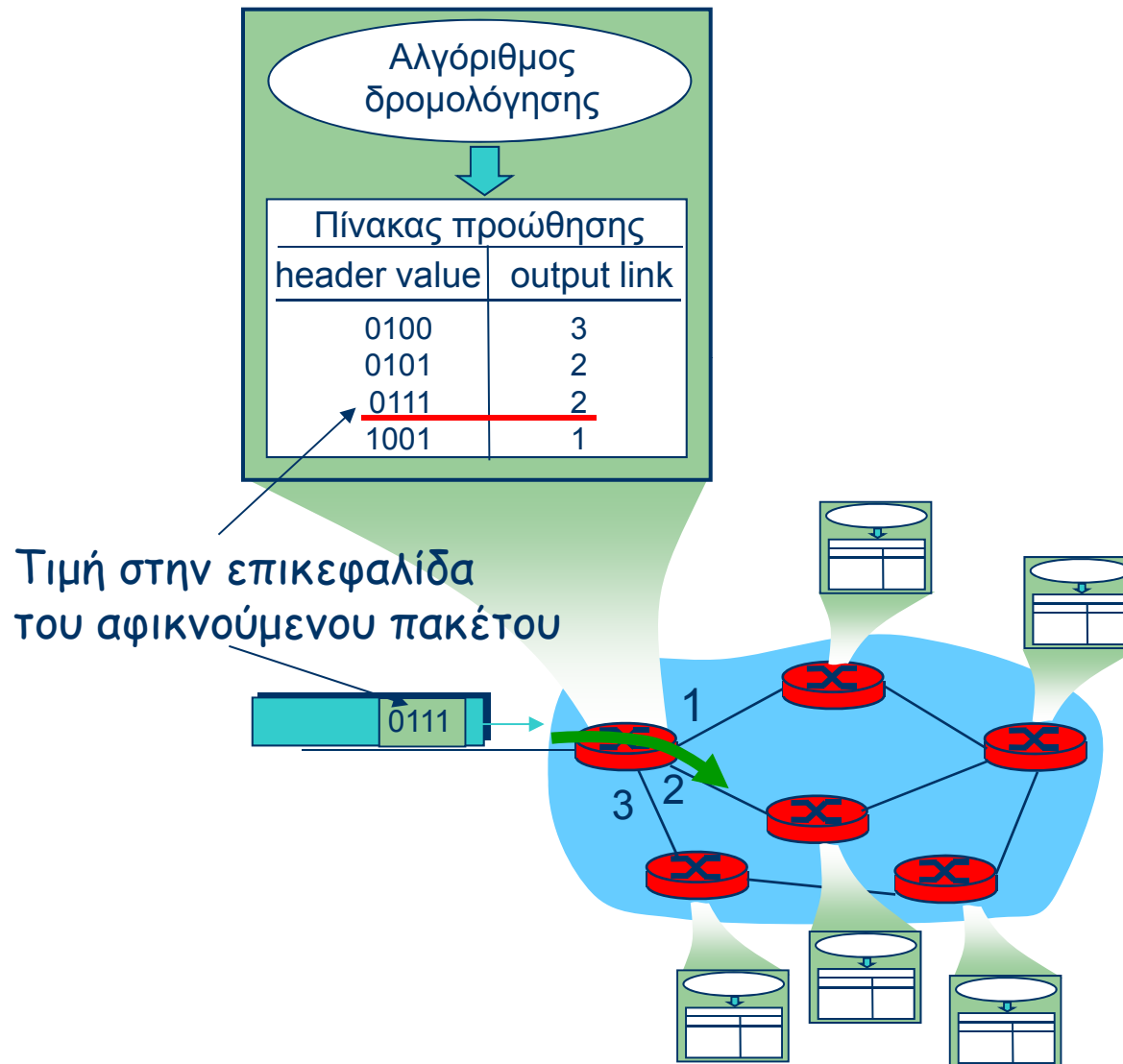
Οι λειτουργίες δρομολόγησης περιλαμβάνουν:

- Τον προσδιορισμό των διαδρομών
- Την ενημέρωση των πινάκων δρομολόγησης
- Την εφαρμογή των πρωτοκόλλων δρομολόγησης
- Στους εμπορικούς δρομολογητές εκτελούνται από τον κεντρικό επεξεργαστή του δρομολογητή (*route processor*)

Η προώθηση αποτελεί επεξεργασία σε επίπεδο πακέτων IP

- Στους εμπορικούς δρομολογητές γίνεται κατανεμημένα
- Το μεγαλύτερο μέρος στο υλικό των διεπαφών

Δρομολόγηση και προώθηση



Εγκατάσταση σύνδεσης



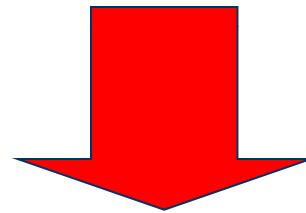
- Η 3^η κύρια λειτουργία σε *κάποιες* αρχιτεκτονικές δικτύου:
 - ATM, frame relay, X.25
- Πριν τη ροή των δεδομενογραμμάτων, οι δύο ακραίοι host και οι μεσολαβούντες δρομολογητές εγκαθιστούν νοητή σύνδεση (virtual connection, VC)
 - εμπλέκονται και δρομολογητές
- Υπηρεσία με σύνδεση στα στρώματα δικτύου και μεταφοράς:
 - **στρώμα δικτύου:** μεταξύ δύο host (μπορεί επίσης να εμπλέκονται δρομολογητές στην περίπτωση που έχουνε νοητές συνδέσεις)
 - **στρώμα μεταφοράς:** μεταξύ δύο διαδικασιών

Μοντέλο υπηρεσίας δικτύου



Ποιο είναι το *μοντέλο υπηρεσίας* που παρέχει το στρώμα δικτύου;

Συνοπτική περιγραφή



νοητές συνδέσεις ή δεδομενογράμματα;

υπηρεσία με σύνδεση

υπηρεσία χωρίς σύνδεση

Μοντέλο υπηρεσίας δικτύου



Παραδείγματα

Υπηρεσίες με ανεξάρτητα δεδομενογράμματα:

- Εξασφαλισμένη παράδοση
- Εξασφαλισμένη παράδοση με καθυστέρηση μικρότερη από x msec

Υπηρεσίες με ροή δεδομενογραμμάτων:

- Παράδοση δεδομενογραμμάτων με τη σειρά
- Εξασφαλισμένο ελάχιστο εύρος ζώνης για τη ροή
- Περιορισμοί στις μεταβολές της απόστασης μεταξύ πακέτων

Νοητές συνδέσεις



Η διαδρομή από την πηγή στον προορισμό συμπεριφέρεται ως τηλεφωνικό κύκλωμα.

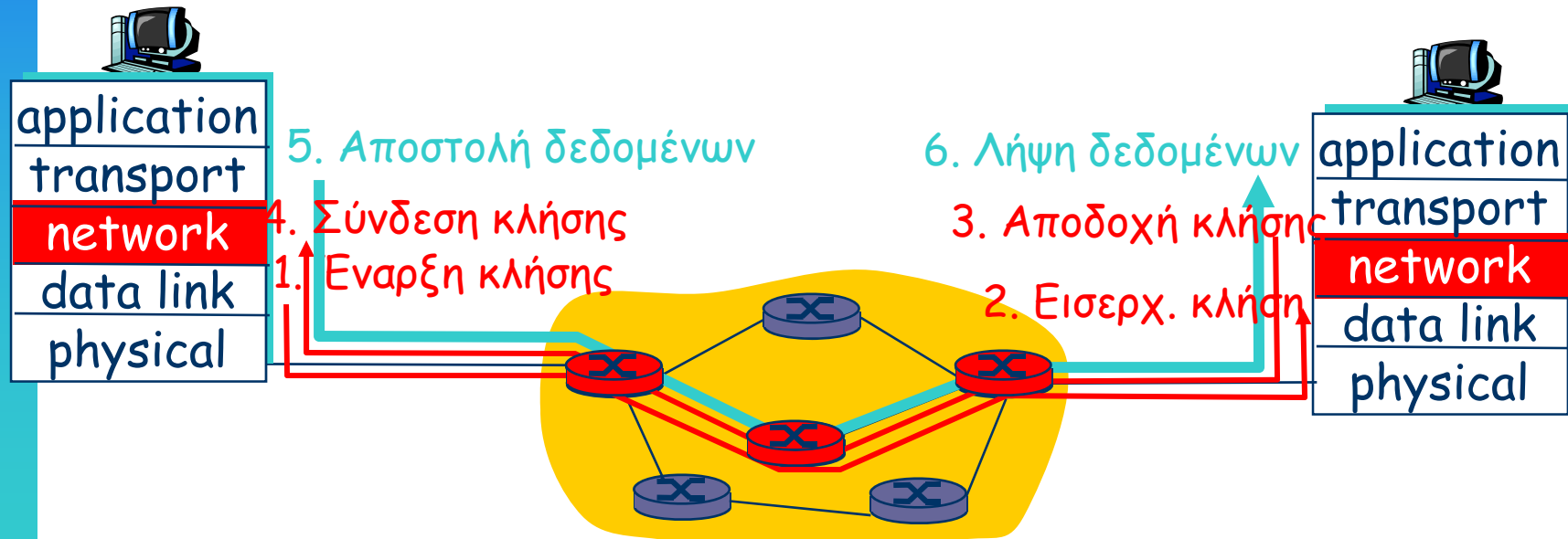
- μέριμνα του δικτύου κατά μήκος της εν λόγω διαδρομής

- Εγκατάσταση νοητής σύνδεσης για κάθε κλήση *πριν* την αποστολή δεδομένων, απόλυση σύνδεσης στο τέλος.
- Κάθε πακέτο μεταφέρει την ταυτότητα της νοητής σύνδεσης (όχι την ταυτότητα του host προορισμού).
- Κάθε δρομολογητής στη διαδρομή πομπός-δέκτης κρατάει πληροφορίες για κάθε διερχόμενη σύνδεση.
- Μπορεί να *εκχωρηθούν* στο νοητό κύκλωμα πόροι της ζεύξης και του δρομολογητή (εύρος ζώνης, buffers).

Νοητές συνδέσεις



- Για την εγκατάσταση και απόλυση νοητών συνδέσεων χρησιμοποιούνται **πρωτόκολλα σηματοδοσίας**
 - χρησιμοποιούνται στο ATM, X.25,
 - δεν χρησιμοποιούνται στο σημερινό Internet

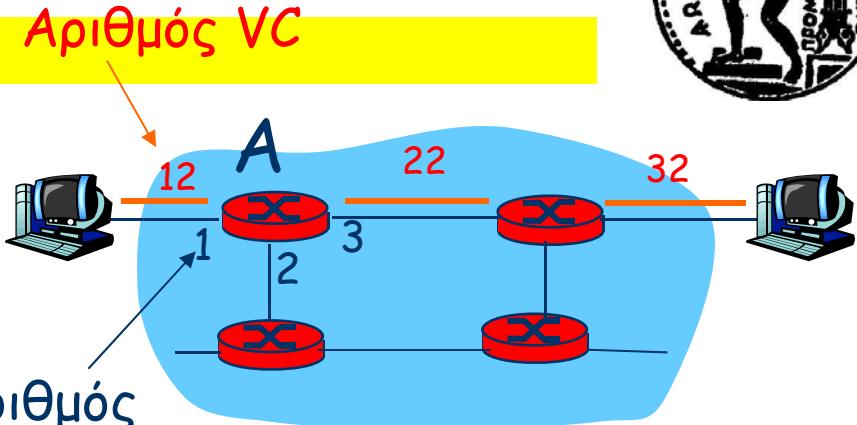


Νοητές συνδέσεις: Υλοποίηση



- Μια νοητή σύνδεση απαρτίζεται από:
 - Διαδρομή από την πηγή στον προορισμό
 - Αριθμούς VC, έναν αριθμό για κάθε ζεύξη
 - Εγγραφές στους πίνακες προώθησης των δρομολογητών κατά μήκος της διαδρομής
- Πακέτο που ανήκει σε νοητή σύνδεση μεταφέρει έναν αριθμό VC.
- Ο αριθμός VC πρέπει να αλλάζει σε κάθε ζεύξη.
 - Νέος αριθμός VC προκύπτει από τον πίνακα προώθησης

Νοητές συνδέσεις: Πίνακας προώθησης



Αριθμός διεπαφής

Πίνακας προώθησης του δρομολογητή A

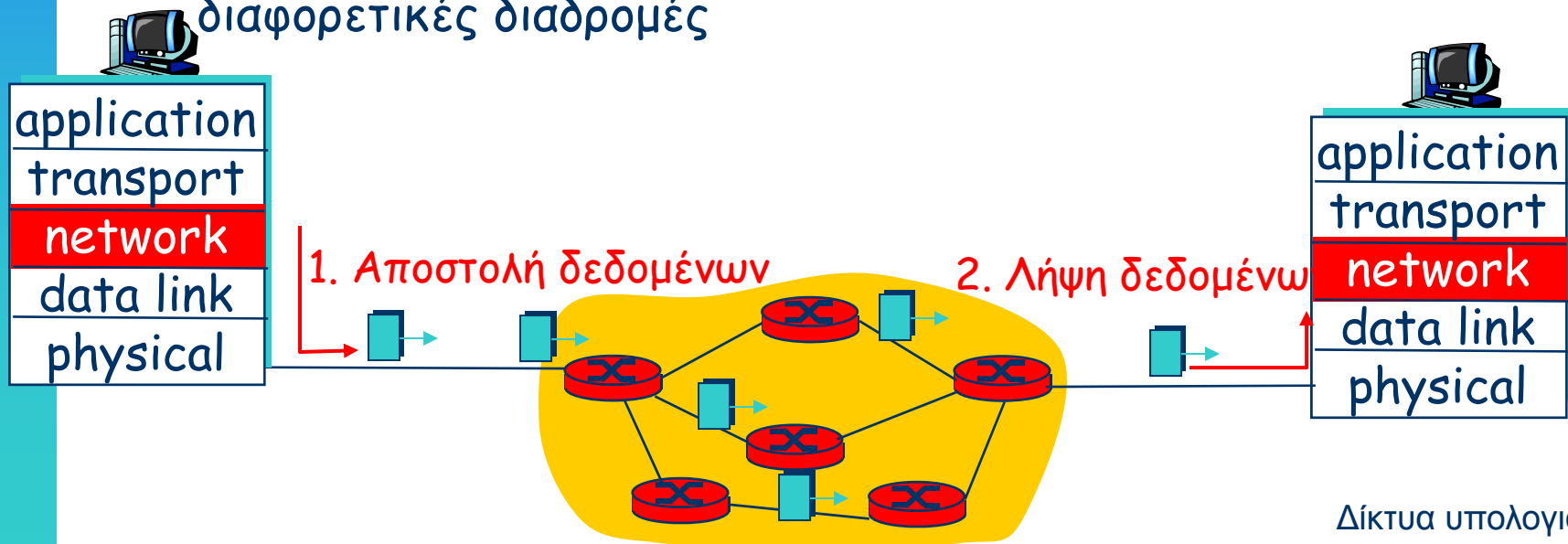
Διεπαφή εισόδου	#VC εισόδου	Διεπαφή εξόδου	#VC εξόδου
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
...

Οι δρομολογητές διατηρούν πληροφορία για την κατάσταση των συνδέσεων

Δεδομενογράμματα: μοντέλο Internet



- Δεν εγκαθίσταται κλήση στο στρώμα δικτύου
- Οι δρομολογητές δεν έχουν στοιχεία για τις συνδέσεις απ' άκρη σ' άκρη
 - δεν υπάρχει η έννοια της "σύνδεσης" στο στρώμα δικτύου
- Τα πακέτα προωθούνται χρησιμοποιώντας τη διεύθυνση του host προορισμού
 - πακέτα μεταξύ των ίδιων ακραίων ζευγών μπορεί να ακολουθούν διαφορετικές διαδρομές



Δεδομενογραμματα: Πίνακας προώθησης

4 δισεκατομμύρια
δυνατές εισοδοι



Περιοχή διευθύνσεων προορισμού	Διεπαφή
11000000 00000000 00000000 00000000 έως 11001111 11111111 11111111 11111111	0
00000100 01010011 10000000 00000000 έως 00000100 01010011 11111111 11111111	1
11001001 00001010 00000000 00000000 έως 11001001 00001010 00000111 11111111	3
11001001 00001010 00000110 00000000 έως 11001001 00001010 00000111 11111111	2

Γιατί δεδομενογράμματα στο Internet;



- Ανταλλαγή δεδομένων μεταξύ υπολογιστών
 - "ελαστική" υπηρεσία, όχι αυστηρές χρονικές απαιτήσεις
 - "έξυπνες" τερματικές διατάξεις (υπολογιστές)
 - μπορούν να προσαρμόζονται, να πραγματοποιούν ελέγχους, διόρθωση λαθών
- Απλό δίκτυο, πολυπλοκότητα στα άκρα
- Πολλοί τύποι ζεύξεων
 - διαφορετικά χαρακτηριστικά
 - δύσκολη η ομοιόμορφη υπηρεσία

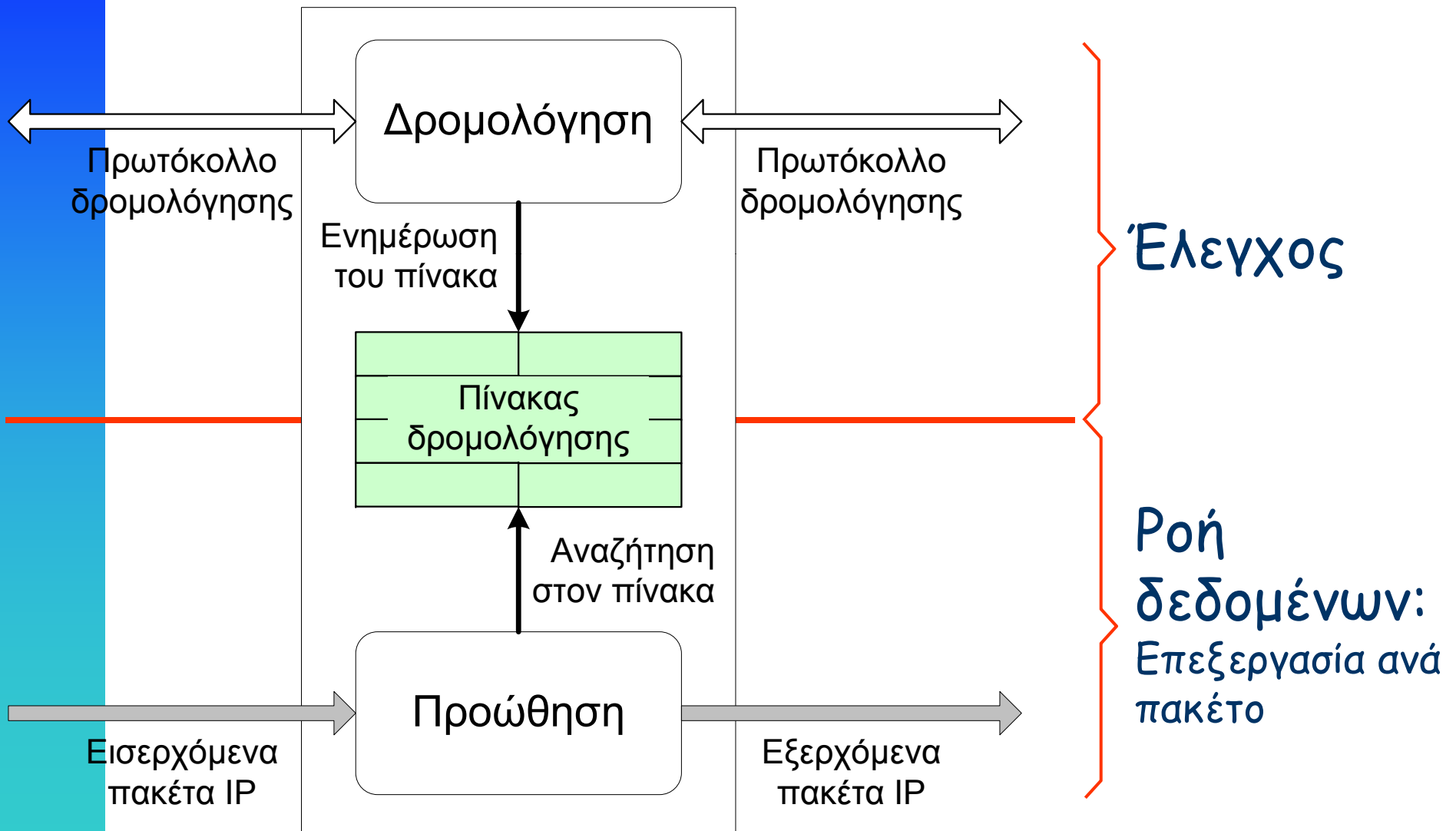
Δρομολογητές: Λειτουργίες



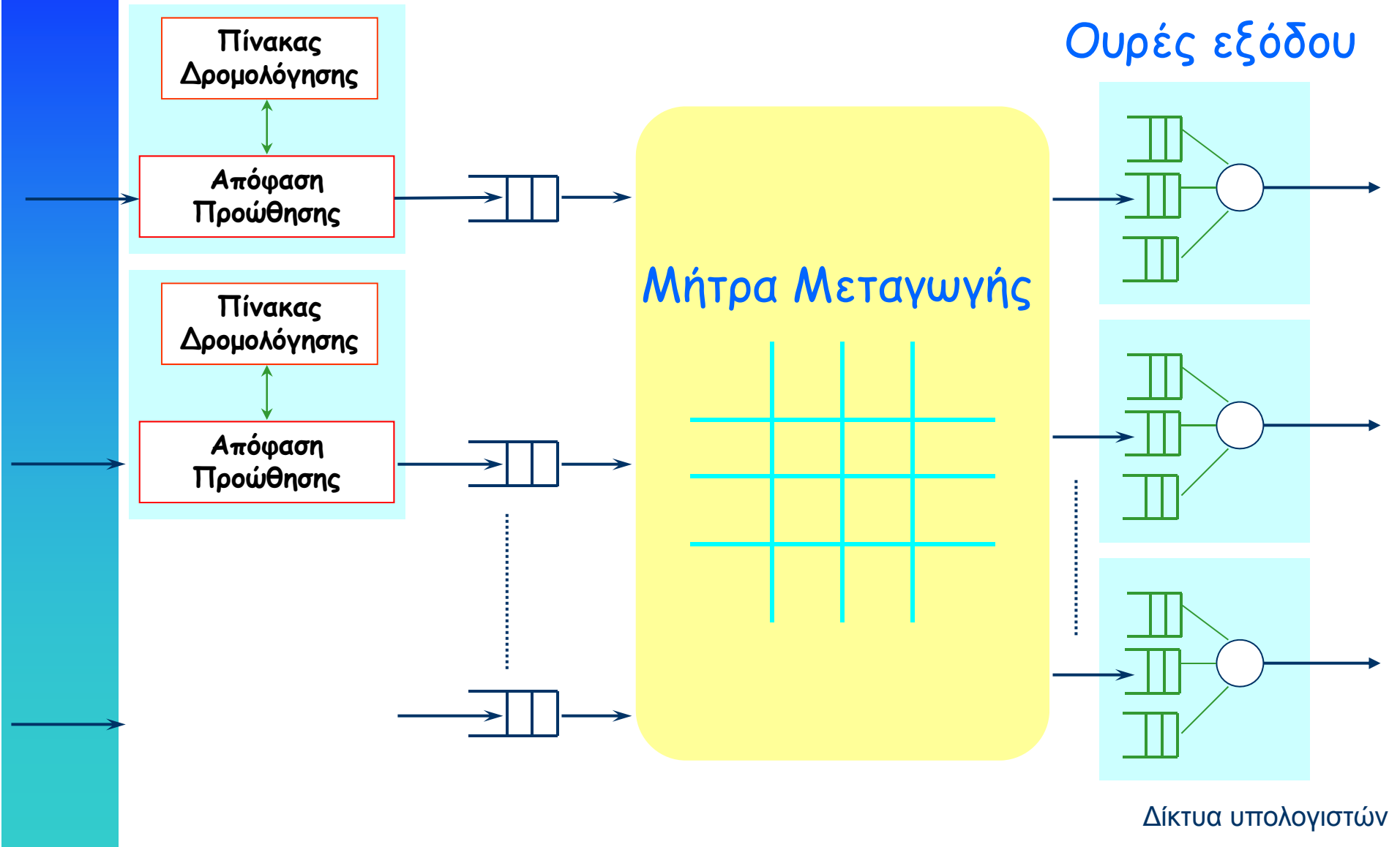
Δύο βασικές λειτουργίες

- **Λειτουργίες ελέγχου:** Εκτέλεση των αλγορίθμων δρομολόγησης (RIP, OSPF, BGP) και ενημέρωση του πίνακα δρομολόγησης
- **Λειτουργίες ροής δεδομένων:** Προώθηση των πακέτων IP από την γραμμή εισόδου στη γραμμή εξόδου. Επεξεργασία ανά πακέτο IP.

Δρομολογητές: Λειτουργίες



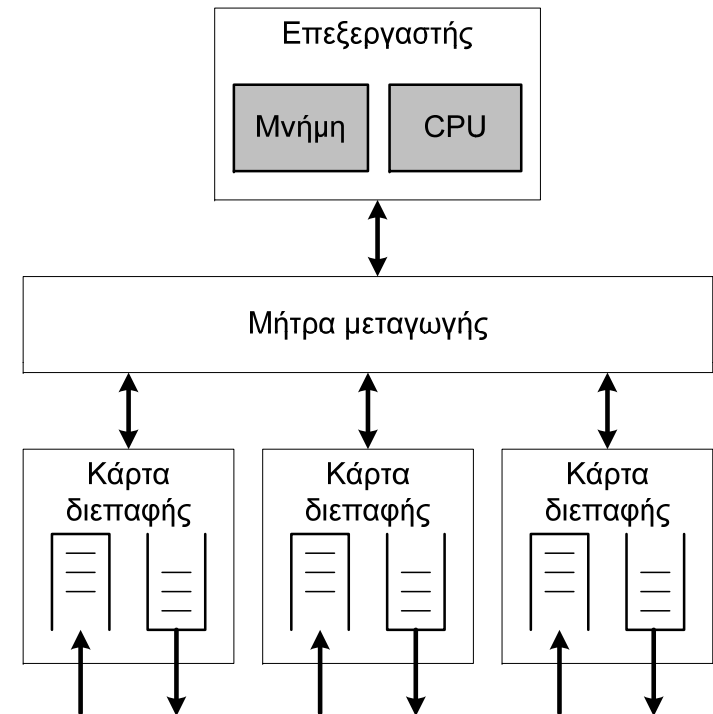
Δρομολογητές: Αρχιτεκτονική



Δρομολογητές: Υλικό



- Το **υλικό** περιλαμβάνει:
 - τις δικτυακές διεπαφές
 - τη μήτρα μεταγωγής και
 - τον επεξεργαστή
- **Δρομολογητές PC:**
 - Η μήτρα μεταγωγής είναι ο δίαυλος PCI
 - Οι δικτυακές διεπαφές είναι οι κάρτες Ethernet
 - Η προώθηση και η δρομολόγηση γίνονται από τον κεντρικό επεξεργαστή
- **Εμπορικοί δρομολογητές:**
 - Η μήτρα μεταγωγής είναι ένα πολύπλοκο δίκτυο μεταγωγής
 - Οι δικτυακές διεπαφές είναι προηγμένες κατασκευές υλικού
 - Ο επεξεργαστής είναι υπεύθυνος για τις λειτουργίες ελέγχου
 - Σχεδόν όλη η λειτουργία προώθησης γίνεται στις διεπαφές



Δρομολογητές: Επεξεργασία πακέτου IP



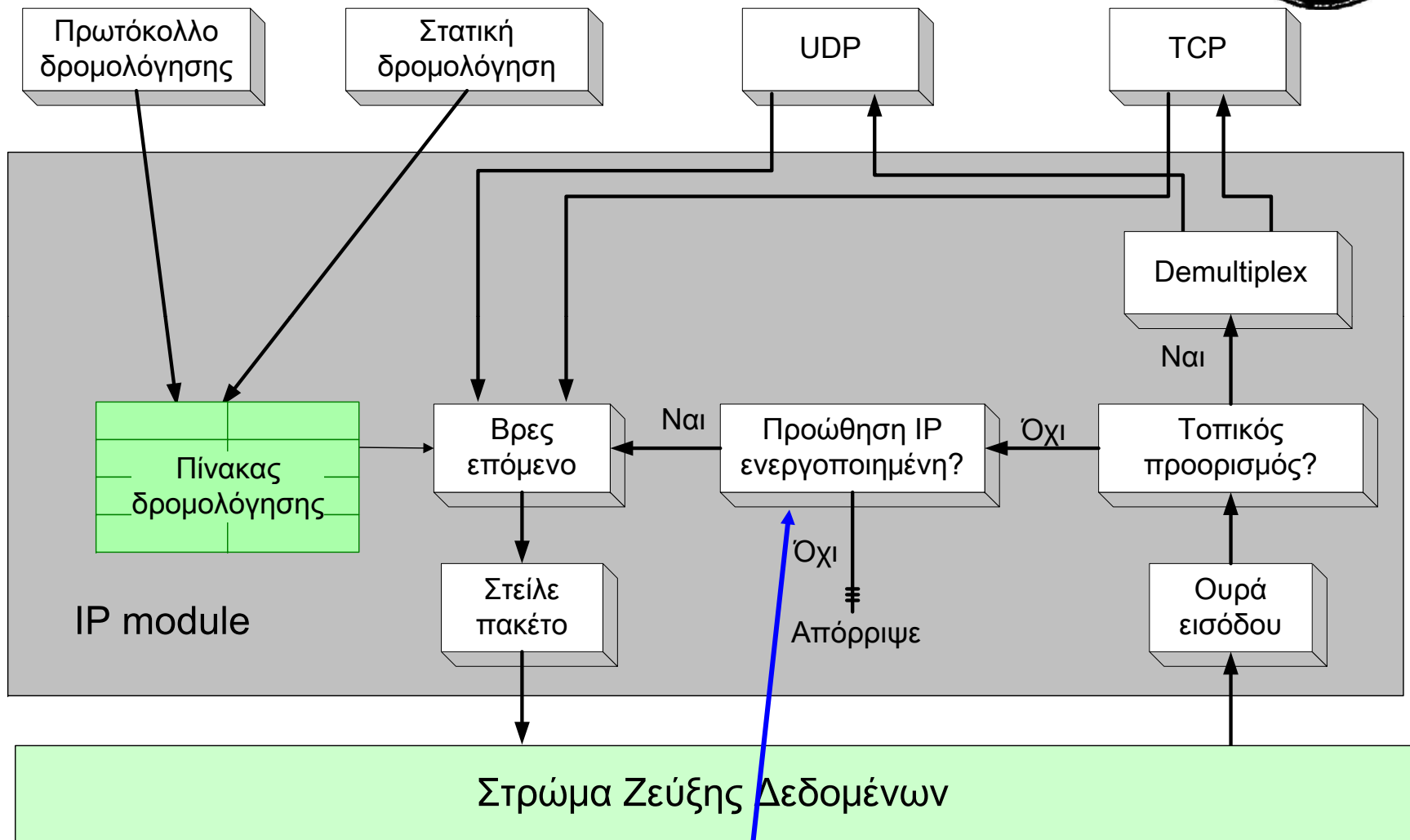
- Ο δρομολογητής \longrightarrow λαμβάνει ένα πακέτο IP
 1. Εγκυρότητα επικεφαλίδας IP
 2. Επεξεργασία επιλογών στην επικεφαλίδα IP
 3. Διεύθυνση IP προορισμού
 4. Αναζήτηση στον πίνακα δρομολόγησης
 5. Μείωση TTL
 6. Θρυμματισμός (εάν απαιτείται)
 7. Υπολογισμός του πεδίου ελέγχου
 8. Προώθηση στο επόμενο βήμα
 9. Αποστολή μηνύματος ICMP (εάν απαιτείται)

Πρώθηση πακέτου IP



- Η επεξεργασία των πακέτων IP στους δρομολογητές και υπολογιστές είναι παρόμοια
- **Βασική διαφορά:** Η πρώθηση IP είναι ενεργοποιημένη στους δρομολογητές, ενώ είναι συνήθως απενεργοποιημένη στους υπολογιστές
- **Ενεργοποιημένη πρώθηση IP**
→ εάν το πακέτο δεν προορίζεται για το τοπικό σύστημα, θα αποσταλεί σε κάποιο άλλο
- **Απενεργοποιημένη πρώθηση IP**
→ εάν το πακέτο δεν προορίζεται για το τοπικό σύστημα, θα απορριφθεί

Πρώθηση πακέτου IP



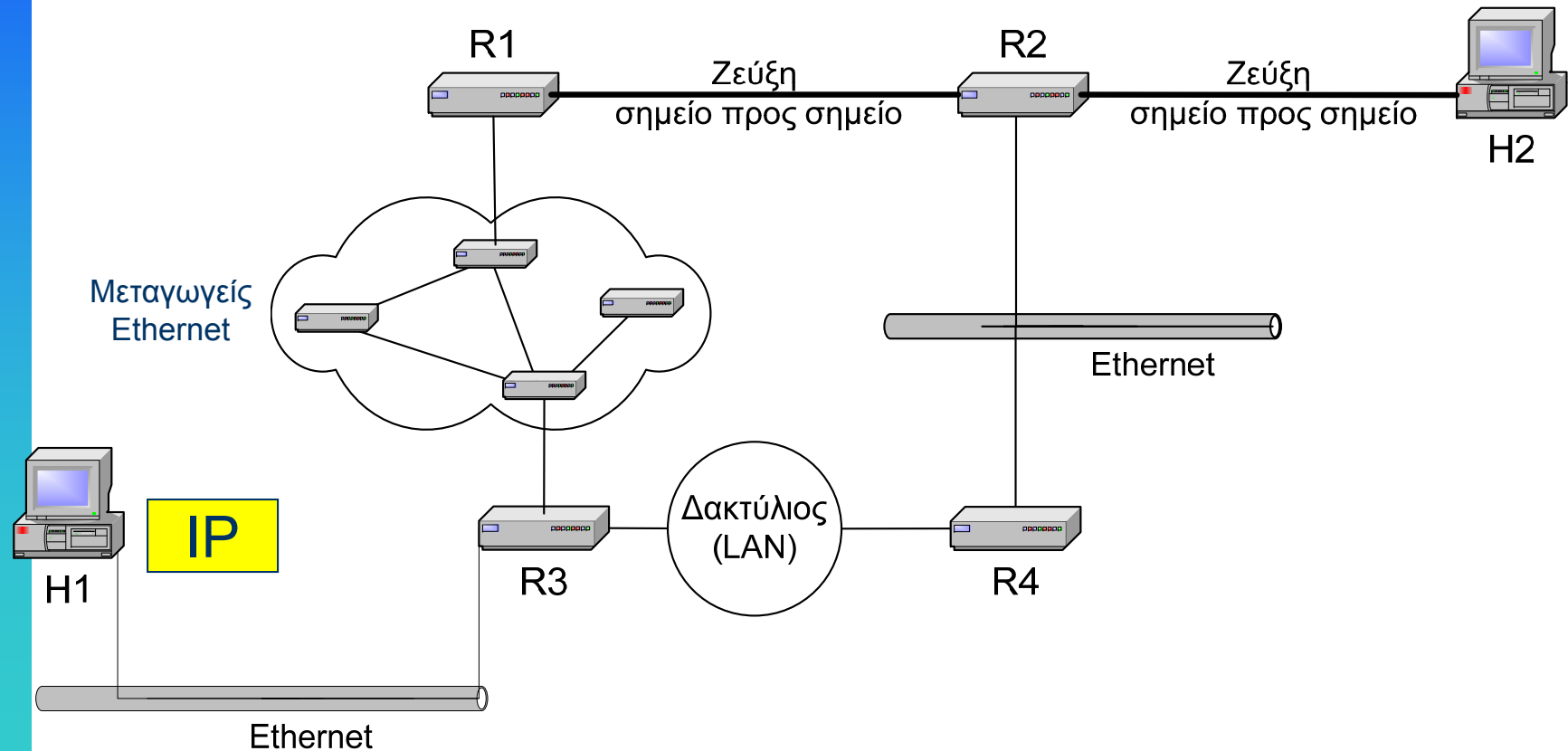
Δρομολογητής: ενεργοποιημένη
Υπολογιστής: απενεργοποιημένη

Πρώθηση πακέτου IP



- Όψη στο στρώμα ζεύξης δεδομένων:

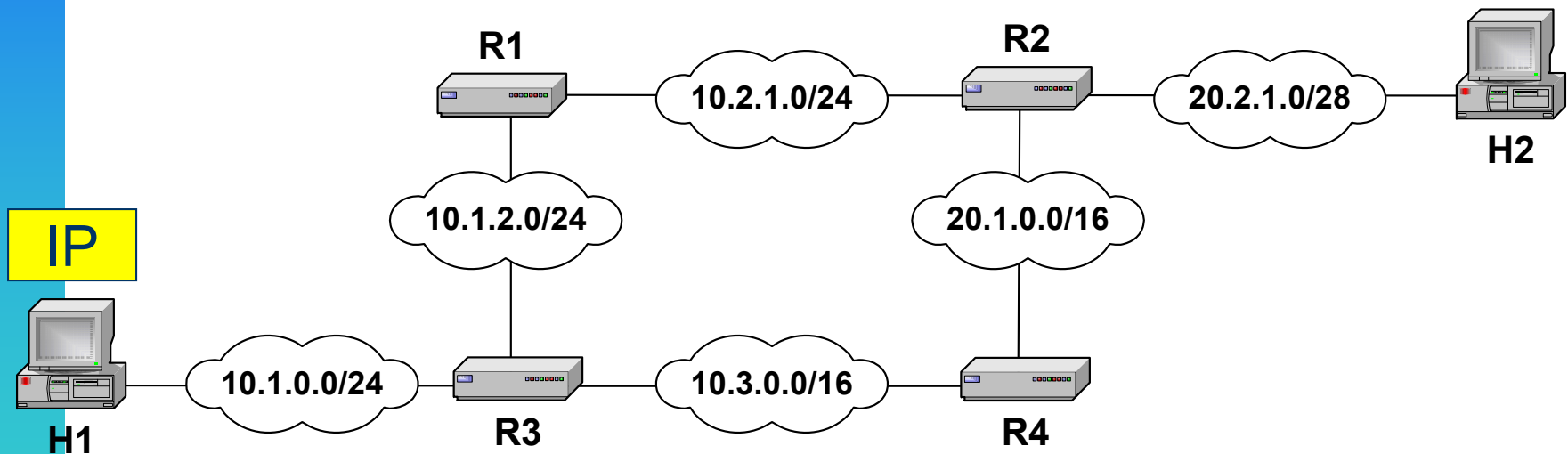
- Το δίκτυο είναι μια συλλογή τοπικών δικτύων (LAN) ή ζεύξεων σημείο προς σημείο ή δικτύων μεταγωγής που διασυνδέονται μέσω δρομολογητών



Πρώθηση πακέτου IP



- Όψη στο στρώμα IP :
 - Το δίκτυο IP είναι μια λογική οντότητα με ένα αριθμό δικτύου
 - Αναπαριστάμε το δίκτυο IP με "σύννεφο"
 - Η προώθηση των πακέτων βλέπει τα σύννεφα και αγνοεί τις ζεύξεις



Παραδοχές για την από άκρο σε άκρο παράδοση πακέτων



Για να παραδοθεί επιτυχώς ένα πακέτο IP πρέπει να ισχύουν οι ακόλουθες συνθήκες:

- Η διεύθυνση δικτύου (το μέρος δικτύου της διεύθυνσης IP) πρέπει να αντιστοιχεί σε ένα μοναδικό δίκτυο στρώματος ζεύξης δεδομένων
- Οι δρομολογητές και υπολογιστές που έχουν την ίδια διεύθυνση δικτύου πρέπει να μπορούν να επικοινωνήσουν απευθείας μέσω του στρώματος ζεύξης δεδομένων
- Κάθε δίκτυο στρώματος ζεύξης δεδομένων συνδέεται με τουλάχιστον ένα άλλο δίκτυο στρώματος ζεύξης δεδομένων μέσω δρομολογητή

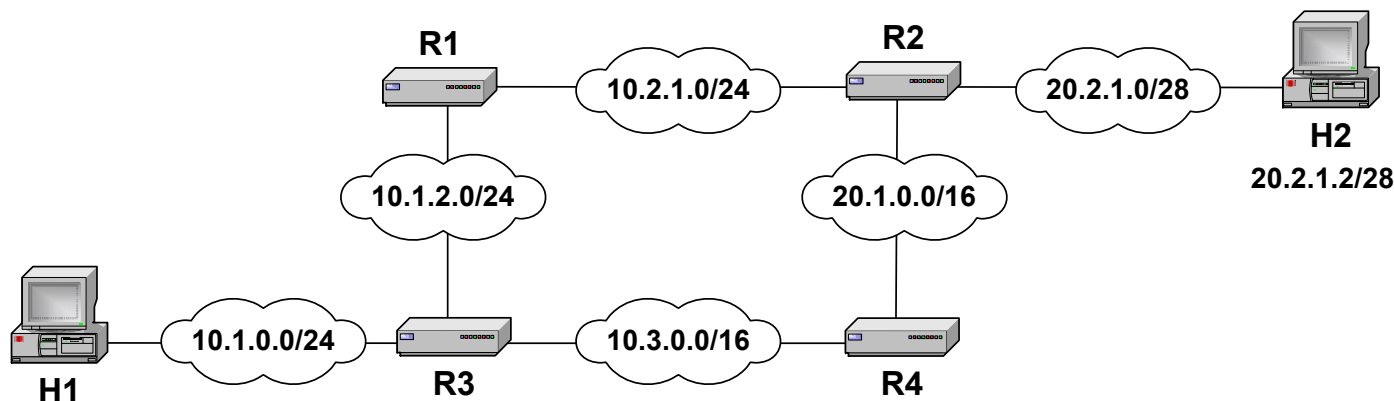
Πρώθηση με πίνακες δρομολόγησης



Προορισμός	Επ. Βήμα
10.1.0.0/24	R3
10.1.2.0/24	απευθείας
10.2.1.0/24	απευθείας
10.3.1.0/24	R3
20.2.0.0/16	R2
30.1.1.0/28	R2

Προορισμός	Επ. Βήμα
10.1.0.0/24	R1
10.1.2.0/24	R1
10.2.1.0/24	απευθείας
10.3.1.0/24	R4
20.1.0.0/16	απευθείας
20.2.1.0/28	απευθείας

Προορισμός	Επ. Βήμα
10.1.0.0/24	R2
10.1.2.0/24	R2
10.2.1.0/24	R2
10.3.1.0/24	R2
20.1.0.0/16	R2
20.2.1.0/28	απευθείας



προς:
20.2.1.2

Προορισμός	Επ. Βήμα
10.1.0.0/24	απευθείας
10.1.2.0/24	R3
10.2.1.0/24	R3
10.3.1.0/24	R3
20.1.0.0/16	R3
20.2.1.0/28	R3



Προορισμός	Επ. Βήμα
10.1.0.0/24	απευθείας
10.1.2.0/24	απευθείας
10.2.1.0/24	R4
10.3.1.0/24	απευθείας
20.1.0.0/16	R4
20.2.1.0/28	R4



Προορισμός	Επ. Βήμα
10.1.0.0/24	R3
10.1.2.0/24	R3
10.2.1.0/24	R2
10.3.1.0/24	απευθείας
20.1.0.0/16	απευθείας
20.2.1.0/28	R2



Πίνακας δρομολόγησης



- Κάθε δρομολογητής διατηρεί έναν πίνακα δρομολόγησης που δείχνει πώς προωθούνται τα απερχόμενα πακέτα
- Εγγραφές στον πίνακα δρομολόγησης
 1. **Διεύθυνση προορισμού:** πού πηγαίνει το πακέτο IP;
 2. **Επόμενο βήμα:** πώς θα αποσταλεί το πακέτο IP;
 3. **Διεπαφή:** ποια είναι η θύρα εξόδου;
- Το επόμενο βήμα και η διεπαφή συνήθως συμπτύσσονται σε μια στήλη
- Οι πίνακες δρομολόγησης ορίζονται έτσι ώστε τα πακέτα να πλησιάζουν τον προορισμό τους

Πίνακας δρομολόγησης



Προορισμός	Επόμενο βήμα	Διεπαφή
10.2.1.0/24	απευθείας	eth0
10.1.2.0/24	R1	eth1
20.2.1.0/28	απευθείας	serial0
10.3.1.0/24	R4	eth1
20.1.0.0/16	απευθείας	eth1

Τα πακέτα IP μπορούν να παραδοθούν "απευθείας", δηλαδή, ο προορισμός βρίσκεται σε απευθείας συνδεδεμένο υποδίκτυο ή να προωθηθούν σε άλλον **δρομολογητή** για προορισμό σε όχι απευθείας συνδεδεμένο υποδίκτυο

Πίνακας δρομολόγησης: Αναζήτηση



- Η αναζήτηση εκτελείται όταν απαιτείται η αποστολή ενός πακέτου IP
- Χρησιμοποιείται η διεύθυνση IP του προορισμού ως κλειδί για την αναζήτηση στον πίνακα δρομολόγησης
- Το αποτέλεσμα της αναζήτησης είναι η διεύθυνση IP του επόμενου βήματος και/ή το όνομα της διεπαφής

Διεύθυνση προορισμού	Επόμενο βήμα/διεπαφή
διεύθυνση δικτύου ή διεύθυνση υπολογιστή ή διεύθυνση βρόχου επιστροφής ή προκαθορισμένη πύλη	διεύθυνση IP του δρομολογητή στο επόμενο βήμα ή όνομα δικτυακής διεπαφής

Πίνακας δρομολόγησης: Τύποι εγγραφών



- **Διαδρομή προς δίκτυο**
 - Η διεύθυνση προορισμού είναι μια διεύθυνση δικτύου (π.χ., 10.0.2.0/24)
 - Οι περισσότερες εγγραφές είναι διαδρομές προς δίκτυα
- **Διαδρομή προς υπολογιστή**
 - Η διεύθυνση προορισμού είναι η διεύθυνση μιας διεπαφής (π.χ., 10.0.1.2/32)
 - Για να ορισθούν διαδρομές προς συγκεκριμένους υπολογιστές
- **Προκαθορισμένη (default) διαδρομή**
 - Όταν δεν υπάρχει ταίριασμα με άλλη εγγραφή
 - Ο δρομολογητής της προκαθορισμένης διαδρομής είναι η προκαθορισμένη πύλη
- **Διεύθυνση βρόχου επιστροφής**
 - Εγγραφή για τη διεύθυνση 127.0.0.1
 - Δηλώνει ως επόμενο βήμα τη διεπαφή του βρόχου επιστροφής (IoO)

Αναζήτηση στον πίνακα δρομολόγησης: ταίριασμα μεγαλύτερου προθέματος



- **Ταίριασμα μεγαλύτερου προθέματος:**

αναζήτηση στον πίνακα δρομολόγησης της εγγραφής που ταιριάζει και έχει διεύθυνση IP με το μεγαλύτερο πρόθεμα

1. Αναζήτηση ταιριάσματος όλων των 32 bit
 2. Αναζήτηση ταιριάσματος των 31 bit
 - ...
 32. Αναζήτηση ταιριάσματος με 0 bit
- **Διαδρομή προς υπολογιστή, Διεύθυνση βρόχου επιστροφής**
→ ταίριασμα των 32 bit
 - **Προκαθορισμένη διαδρομή (0.0.0.0/0)**
→ ταίριασμα των 0 bit

- Πακέτο IP προς 147.102.38.50



Διεύθυνση προορισμού	Επόμενο βήμα
10.0.0.0/8	R1
147.102.0.0/16	R2
147.102.64.0/20	R3
147.102.192.0/20	R3
147.102.38.0/24	R4
147.102.38.56/32	R3
0.0.0.0/0	R5

- Το ταίριασμα με μεγαλύτερο πρόθεμα της διεύθυνσης 147.102.38.50 είναι με 24 bit στην εγγραφή 147.102.38.0/24



- Θα αποσταλεί στον δρομολογητή R4

Ταίριασμα μεγαλύτερου προθέματος



Πίνακας προώθησης

	πρόθεμα	διεπαφή
προορισμός	192.0.0.0/4	0
201.10.7.17	4.83.128.0/17	1
201: 11001001	201.10.0.0/21	3
10: 00001010	201.10.6.0/23	2
7: 00000111		
17: 00010001		

Ταίριασμα μεγαλύτερου προθέματος



Πίνακας προώθησης

προορισμός	πρόθεμα	διεπαφή
201.10.7.17 →	192.0.0.0/4	0
201: 11001001	4.83.128.0/17	1
10: 00001010	201.10.0.0/21	3
7: 00000111	201.10.6.0/23	2
17: 00010001		

192: 11000000

Ταίριασμα μεγαλύτερου προθέματος



Πίνακας προώθησης

	πρόθεμα	διεπαφή
προορισμός 201.10.7.17 →	192.0.0.0/4	0
	4.83.128.0/17	1
201: 11001001 10: 00001010 7: 00000111 17: 00010001	201.10.0.0/21	3
	201.10.6.0/23	2

4: 00000100
83: 01010011
128: 10000000

Ταίριασμα μεγαλύτερου προθέματος



Πίνακας προώθησης

πρόθεμα διεπαφή

προορισμός

201.10.7.17 →

201: **11001001**
10: **00001010**
7: **00000111**
17: 00010001

192.0.0.0/4	0
4.83.128.0/17	1
201.10.0.0/21	3
201.10.6.0/23	2

201: **11001001**
10: **00001010**
0: **00000000**

Ταίριασμα μεγαλύτερου προθέματος



Πίνακας προώθησης

πρόθεμα διεπαφή

192.0.0.0/4	0
4.83.128.0/17	1
201.10.0.0/21	3
201.10.6.0/23	2

προορισμός

201.10.7.17



201: **11001001**
10: **00001010**
7: **00000111**
17: 00010001

201: **11001001**
10: **00001010**
6: **00000110**

Δρομολόγηση πακέτων IP

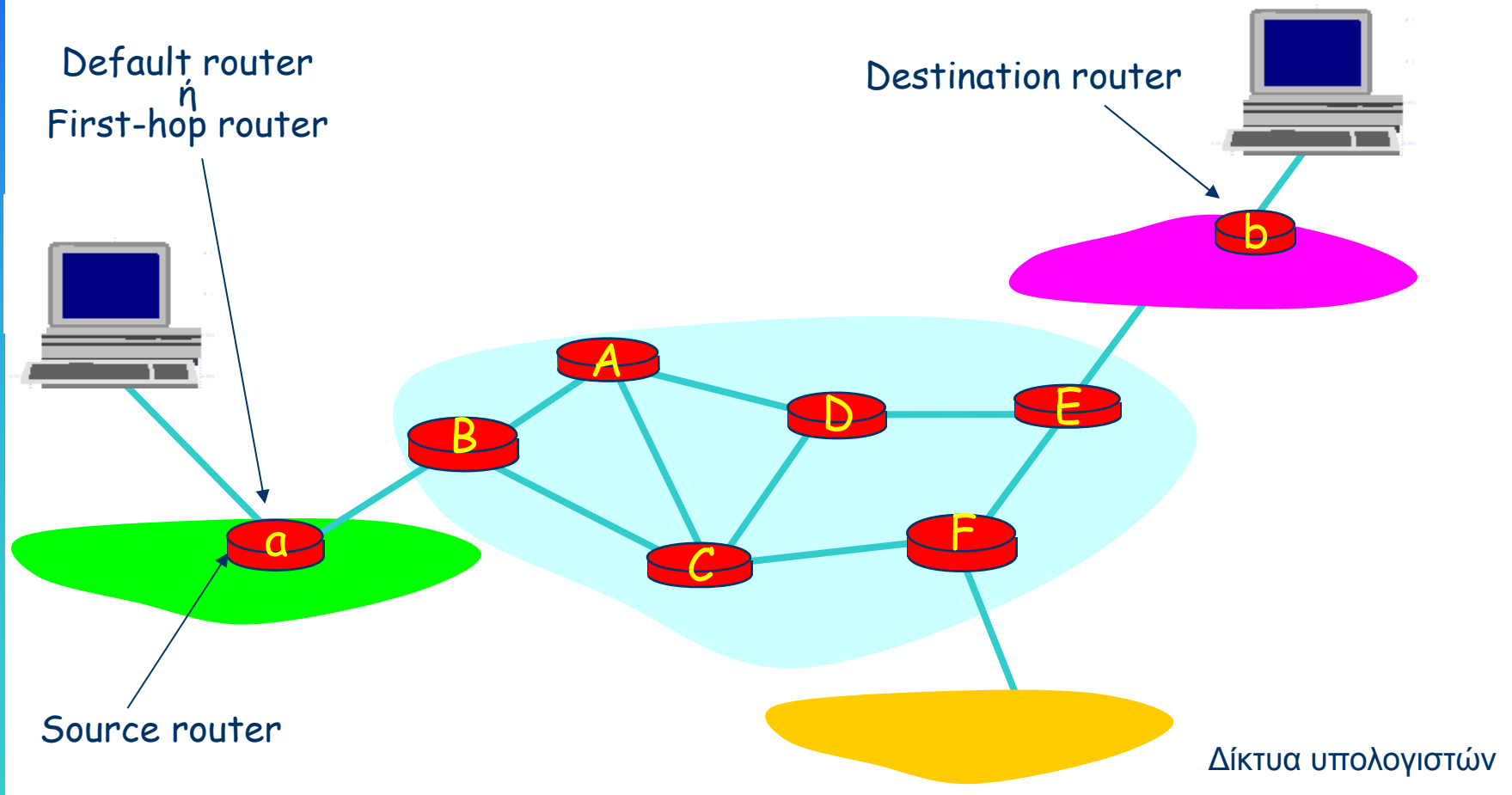


- Ανεξάρτητα από το αν παρέχει υπηρεσία με ή χωρίς σύνδεση, το στρώμα δικτύου πρέπει να *καθορίζει τη διαδρομή των πακέτων*
- Τούτο είναι δουλειά του **πρωτοκόλλου δρομολόγησης** του στρώματος δικτύου
- Στόχος του πρωτοκόλλου δρομολόγησης είναι η παράδοση των πακέτων από τον δρομολογητή πηγής στον δρομολογητή προορισμού με *αποτελεσματικό τρόπο*

Δρομολόγηση πακέτων IP



Στόχος του πρωτοκόλλου δρομολόγησης: καθορισμός μιας "καλής" διαδρομής (ακολουθία δρομολογητών) μέσω του δικτύου από τον δρομολογητή πηγής προς τον δρομολογητή προορισμού



Δρομολόγηση πακέτων IP



- Κύριες λειτουργίες:
 - Μετάφραση: διεύθυνση => διαδρομή
 - Συντονιζόμενες κατανεμημένες αποφάσεις: ποια είναι η απερχόμενη ζεύξη σε κάθε κόμβο;
 - Γενικός στόχος επίδοσης: π.χ., ελαχιστοποίηση καθυστέρησης
 - Άλλα: προτεραιότητες, ασφάλεια, πολιτικές επικρατειών
- Γιατί ενδιαφέρει η δρομολόγηση;
 - Υπηρεσία του δικτύου κορμού
 - Αποτελεσματική χρήση των πόρων του δικτύου

Δρομολόγηση πακέτων IP



Δυσκολίες

- Επίπεδος (flat) χώρος διευθύνσεων, οπότε απαιτείται υπολογισμός των διαδρομών
 - Απαιτεί γνώση της συνολικής τοπολογίας και συντονιζόμενες καταναεμημένες αποφάσεις
- Αντιμετώπιση των αλλαγών στην τοπολογία
 - Πώς να κοινοποιηθούν οι ενημερώσεις και να ελαχιστοποιηθούν οι χρονικές ασυμφωνίες
- Αποτελεσματική χρησιμοποίηση των πόρων του δικτύου
 - π.χ., ελαχιστοποίηση της συμφόρησης, καθυστέρηση
- Επεκτασιμότητα για μεγάλα δίκτυα με πολλές επικράτειες
- Υποστήριξη πολιτικών ευέλικτης δρομολόγησης
- Υποστήριξη εφαρμογών με απαιτήσεις QoS

Αλγόριθμος δρομολόγησης



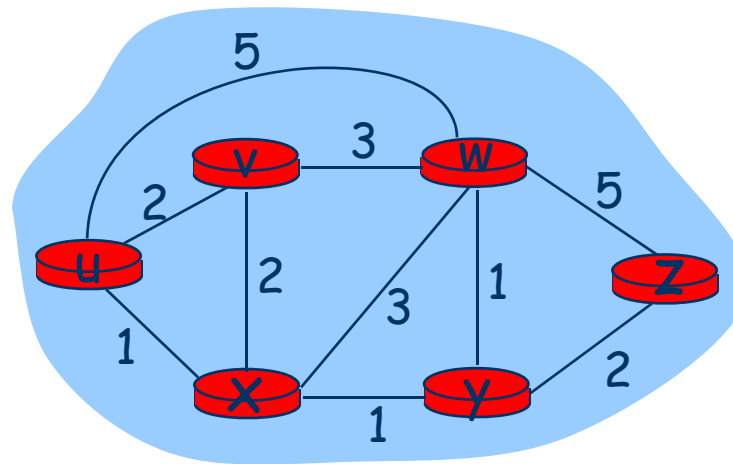
- Η καρδιά του πρωτοκόλλου δρομολόγησης είναι ο **αλγόριθμος δρομολόγησης** που καθορίζει τη διαδρομή από την πηγή στον προορισμό
- Στοιχεία ενός αλγορίθμου δρομολόγησης
 - Διαδικασία για την αποστολή και λήψη πληροφοριών πρόσβασης στους άλλους δρομολογητές του δικτύου.
 - Διαδικασία για τον υπολογισμό των βέλτιστων διαδρομών
 - Οι διαδρομές υπολογίζονται χρησιμοποιώντας κάποιον αλγόριθμο ελάχιστης διαδρομής
 - Διαδικασίες για αντίδραση σε αλλαγές της τοπολογίας και κοινοποίησης των αλλαγών αυτών

Αλγόριθμος δρομολόγησης



Γράφος δικτύου

Για τη διατύπωση του αλγορίθμου δρομολόγησης χρησιμοποιείται ο γράφος του δικτύου.



Γράφος: $G = (N,E)$

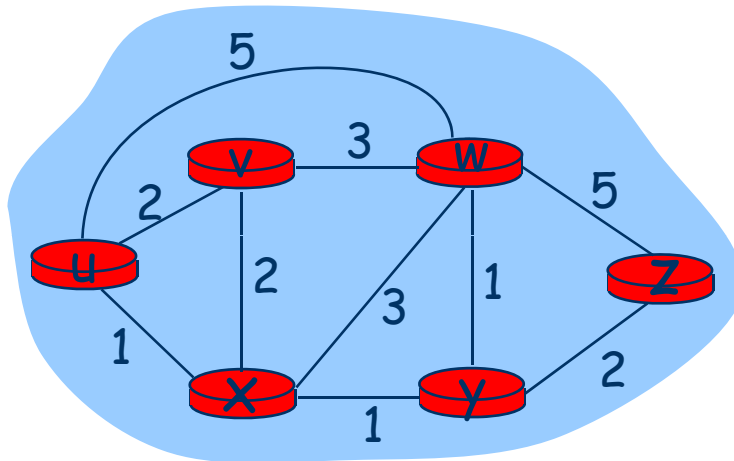
Σύνολο δρομολογητών $N = \{ u, v, w, x, y, z \}$

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

Αλγόριθμος δρομολόγησης



Γράφος δικτύου: κόστη



➤ $c(x,x') =$ κόστος ζεύξης (x,x')

- π.χ., $c(w,z) = 5$

➤ το κόστος μπορεί να είναι πάντα 1, ή αντιστρόφως ανάλογο του εύρους ζώνης, ή ανάλογο της συμφόρησης

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

Ποια είναι η διαδρομή ελάχιστου κόστους μεταξύ u και z ;

Ο αλγόριθμος δρομολόγησης βρίσκει τη διαδρομή ελάχιστου κόστους

Ταξινόμηση αλγορίθμων δρομολόγησης



Πρόελευση της πληροφορίας δρομολόγησης:

Τοπική:

- οι διαδρομές υπολογίζονται μόνο με βάση την τοπική τοπολογία και κίνηση
- ο δρομολογητής γνωρίζει τους φυσικά συνδεδεμένους γείτονες και το κόστος των ζεύξεων προς τους γείτονες
- επαναληπτική διαδικασία υπολογισμού, ανταλλαγή πληροφοριών με τους γείτονες
- **Αλγόριθμοι διανύσματος αποστάσεων**

Καθολική:

- οι διαδρομές υπολογίζονται με βάση την συνολική τοπολογία
- όλοι οι δρομολογητές ξέρουν την πλήρη τοπολογία και το κόστος των ζεύξεων
- **Αλγόριθμοι κατάστασης ζεύξεων**

Ταξινόμηση αλγορίθμων δρομολόγησης



Στατική ή δυναμική δρομολόγηση;

Στατική:

- εκ των προτέρων καθορισμένες διαδρομές
- οι διαδρομές αλλάζουν αργά με το χρόνο

Δυναμική:

- οι διαδρομές αλλάζουν πιο συχνά
- επανακαθορισμός και αναδιάταξη των διαδρομών
 - περιοδική ενημέρωση
 - ενημέρωση, όταν αλλάζουν τα κόστη των ζεύξεων

Αλγόριθμος διανύσματος αποστάσεων



- Διάνυσμα αποστάσεων: *Distance vector*
- Κάθε κόμβος ξέρει την απόσταση (κόστος) προς τους άμεσα συνδεδεμένους με αυτόν γείτονες
- Κάθε κόμβος στέλνει περιοδικά το δικό του διάνυσμα αποστάσεων στους γείτονές του.
- Όταν ένας κόμβος X λάβει νέο διάνυσμα αποστάσεων από γείτονα, ενημερώνει το δικό του διάνυσμα σύμφωνα με τη σχέση

$$d_X(Y) = \min_V \{c(X, V) + d_V(Y)\}, \quad \forall Y \in N$$

- Μετά από μερικές αλλαγές οι αποστάσεις συγκλίνουν στις ελαχίστου κόστους

Αλγόριθμος διανύσματος αποστάσεων



Επαναληπτικός, ασύγχρονος:

- κάθε τοπικός επανυπολογισμός προκαλείται από:
 - τοπική αλλαγή κόστους ζεύξης
 - νέο διάνυσμα αποστάσεων από τον γείτονα: αλλαγή διαδρομής ελαχίστου κόστους από τον γείτονα
- *τερματίζει μόνος του*, δεν απαιτείται εντολή για να σταματήσει

Κατανεμημένος:

- κάθε κόμβος ειδοποιεί τους γείτονες, *μόνον* όταν η ελαχίστου κόστους διαδρομή προς *οποιαδήποτε* κατεύθυνση αλλάζει
 - οι γείτονες τότε ειδοποιούν τους γείτονές τους αν είναι αναγκαίο

Κάθε κόμβος:

περιμένει για (αλλαγή κόστους τοπικής ζεύξης ή μήνυμα από γείτονα)

ξαναυπολογίζει τον πίνακα αποστάσεων

αν η διαδρομή ελαχίστου κόστους προς οιονδήποτε προορισμό έχει αλλάξει, *ειδοποιεί* τους γείτονες

Αλγόριθμος διανύσματος αποστάσεων



Σε κάθε κόμβο διατηρείται πίνακας αποστάσεων στον οποίο υπάρχει:

- μία σειρά για κάθε δυνατό προορισμό
- μία στήλη για κάθε άμεσα συνδεδεμένο γειτονικό κόμβο
- παράδειγμα: στον κόμβο X , η καταχώρηση για τον προορισμό Y , υπολογίζεται ως εξής:

$$d_X(Y) = c(X, V) + \min d_V(Y)$$

- όπου το \min λαμβάνεται για όλους τους γείτονες V του X
Ο γειτονικός κόμβος για τον οποίο ισχύει το $\min d_X(Y)$ είναι ο next hop στον πίνακα προώθησης του X για τον προορισμό Y

Αλγόριθμος διανύσματος αποστάσεων

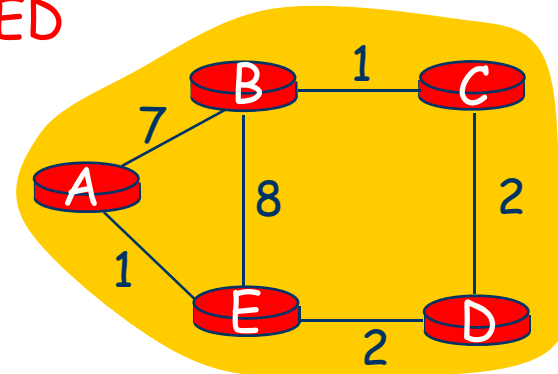


Πίνακας αποστάσεων: παράδειγμα

Ε	απόσταση μέσω		
	A	B	D
A	1	14	5
B	7	8	5
C	6	9	4
D	4	11	2

προορισμός

$(D,E,A)+ED$



$(B,C,D,E,A)+EB$

$(A,E,D,C,B)+EA$

- Κάθε κόμβος έχει τη δική του
- σειρά, για κάθε δυνατό προορισμό
 - στήλη, για κάθε άμεσα συνδεδεμένο γειτονικό κόμβο

Αλγόριθμος διανύσματος αποστάσεων



Από τον πίνακα αποστάσεων προκύπτει ο πίνακας προώθησης/δρομολόγησης

Ε	απόσταση μέσω		
	A	B	D
A	1	14	5
B	7	8	5
C	6	9	4
D	4	11	2

προορισμός

Ε	Εξερχόμενη ζεύξη προς χρήση
A	A,1
B	D,5
C	D,4
D	D,2

προορισμός

Πίνακας αποστάσεων → Πίνακας προώθησης/δρομολόγησης

Αλγόριθμος διανύσματος αποστάσεων



	A	
A		
B	1	B
C	1	C
D	2	C
E	3	C
F	4	C

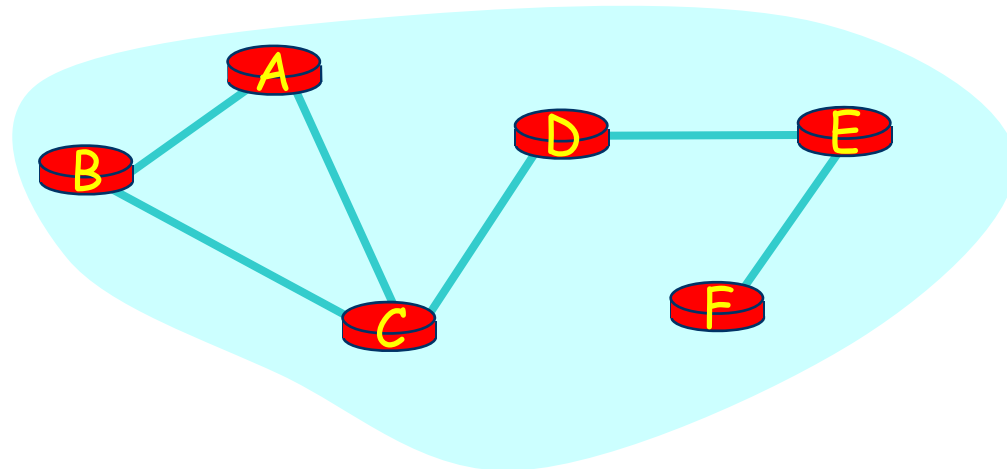
	B	
A	1	A
B		
C	1	C
D	2	C
E	3	C
F	4	C

	C	
A	1	A
B	1	B
C		
D	1	D
E	2	D
F	3	D

	D	
A	2	C
B	2	C
C	1	C
D		
E	1	E
F	2	E

	E	
A	3	D
B	3	D
C	2	D
D	1	D
E		
F	1	F

	F	
A	4	E
B	4	E
C	3	E
D	2	E
E	1	E
F		



Αλγόριθμος διανύσματος αποστάσεων

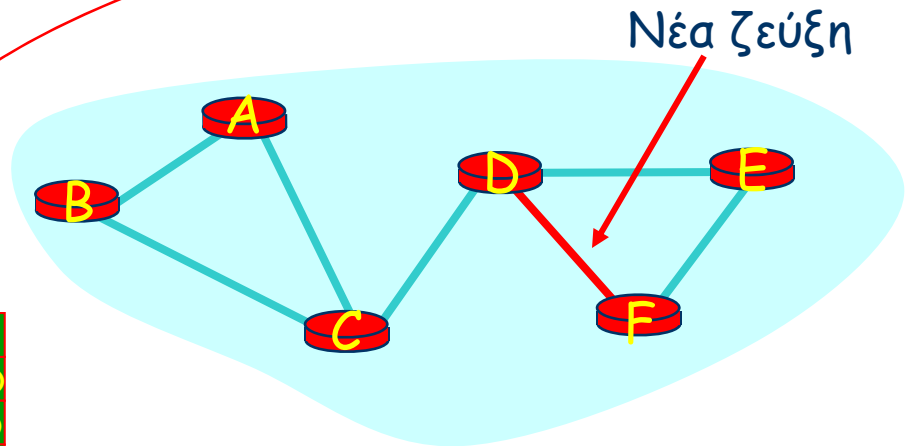
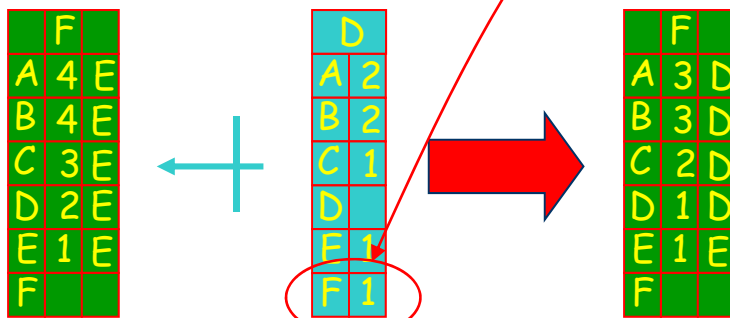


Ενημέρωση πινάκων

- Οι κόμβοι στέλνουν ενημερώσεις των διανυσμάτων αποστάσεων στους γείτονες
- Κάθε κόμβος ενημερώνει τον πίνακά του υπολογίζοντας τη συντομότερη διαδρομή

	A	B	C	D	E	F
A						
B	1 B					
C	1 C					
D	2 C					
E	3 C					
F	4 C					
B						
C						
D						
E						
F						

Ενημέρωση πίνακα δρομολόγησης

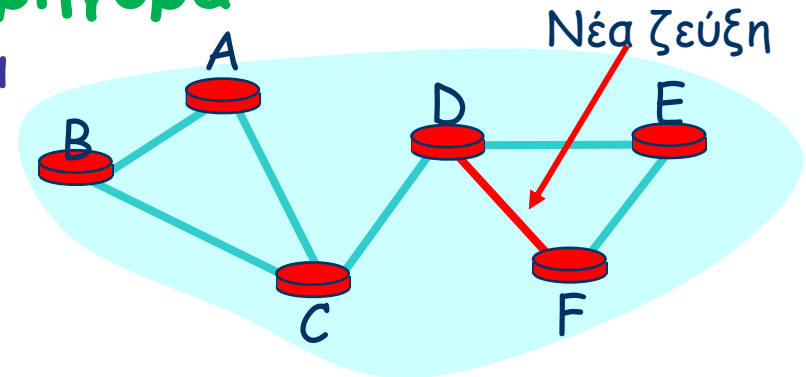


Αλγόριθμος διανύσματος αποστάσεων



Τα καλά νέα διαδίδονται γρήγορα

- Νέα ζεύξη, η D-F εγκαθίσταται
- Διάδοση καλών ειδήσεων
- Βήμα 1: οι D και F μαθαίνουν για τη νέα ζεύξη
- Βήμα 2: ο D στέλνει πίνακα στους C, E, F
- Βήμα 3: ο C στέλνει πίνακα στους A, B, D
- Χαρακτηριστικά
 - Ταχεία σύγκλιση
 - Μόνο οι κόμβοι που ωφελούνται αλλάζουν πίνακες
 - Τα πακέτα δεν κάνουν βρόχους



A	B	C	D	E	F
A		A1 A	A2 C	A3 D	A3 D
B1 B	B	B1 B	B2 C	B3 D	B3 D
C1 C	C1 C	C	C1 C	C2 D	C2 D
D2 C	D2 C	D1 D	D	D1 D	D1 D
E3 C	E3 C	E2 D	E1 E	E	E1 E
F4 C	F4 C	F2 D	F1 E	F1 F	F

A	B	C	D	E	F
A		A1 A	A2 C	A3 D	A3 D
B1 B	B	B1 B	B2 C	B3 D	B3 D
C1 C	C1 C	C	C1 C	C2 D	C2 D
D2 C	D2 C	D1 D	D	D1 D	D1 D
E3 C	E3 C	E2 D	E1 E	E	E1 E
F3 C	F3 C	F2 D	F1 E	F1 F	F

Αλγόριθμος διανύσματος αποστάσεων



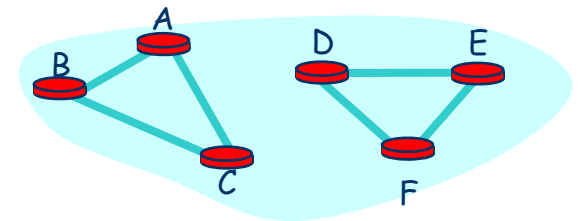
Μέτρημα μέχρι το άπειρο

- Χάνεται η ζεύξη C-D
- Διάδοση κακών ειδήσεων
- Οι D και C το μαθαίνουν.
- Ο C λαμβάνει ενημερώσεις από τους A και B
- Ο C στέλνει ενημερώσεις στους A και B
-κ.ο.κ. μέχρι;
- Χαρακτηριστικά
 - Αργή σύγκλιση
 - Όλοι οι κόμβοι μπορεί να αλλάξουν πίνακες
 - Βρόχοι πακέτων, προσωρινές αστάθειες
- Πώς βελτιώνεται αυτό;
 - Τέχνασμα poisoned reverse

	A	B	C	D	E	F
A						
B	1 B					
C	1 C					
D	2 C					
E	3 C					
F	3 C					

	A	B	C
A			
B	1 B		
C	1 C		
D	2 C		
E	3 C		
F	3 C		

	A	B	C
A			
B	1 B		
C	1 C		
D	4 C		
E	5 C		
F	5 C		

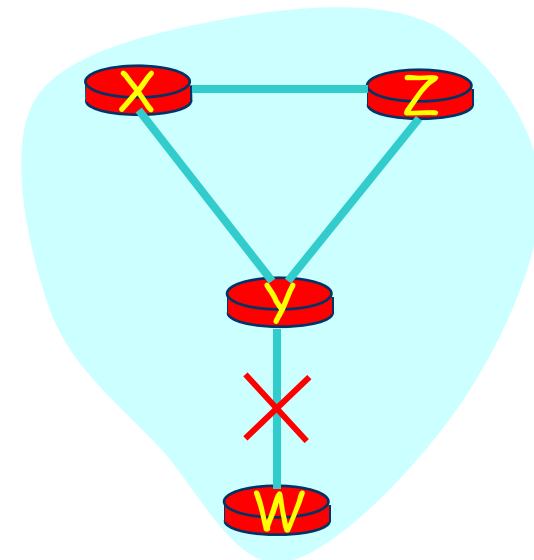
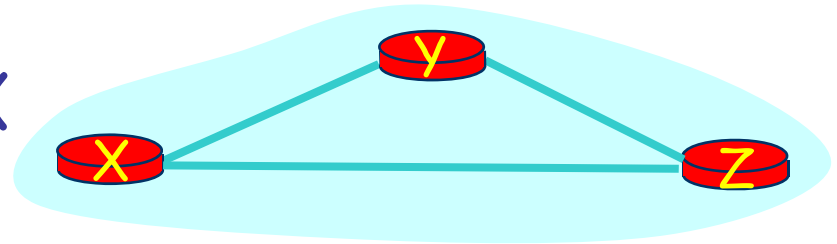


Αλγόριθμος διανύσματος αποστάσεων



Τέχνασμα Poisoned reverse

- Αν ο Z δρομολογεί προς X μέσω του Y, τότε ο Z θα αναγγέλλει προς Y ότι η απόστασή του από τον X είναι ∞
- Το poisoned reverse μερικές φορές αποτυγχάνει



Αλγόριθμος κατάστασης ζεύξεων



- Δύο είναι τα βασικά μειονεκτήματα του αλγόριθμου διανύσματος αποστάσεων
 - Δεν λαμβάνει υπόψη τη χωρητικότητα των γραμμών
 - Μερικές φορές ο αλγόριθμος καθυστερεί υπερβολικά για να συγκλίνει
- Για τους λόγους αυτούς, επινοήθηκε ένας καινούργιος αλγόριθμος, ο αλγόριθμος **κατάστασης ζεύξεων (link state)**

Αλγόριθμος κατάστασης ζεύξεων



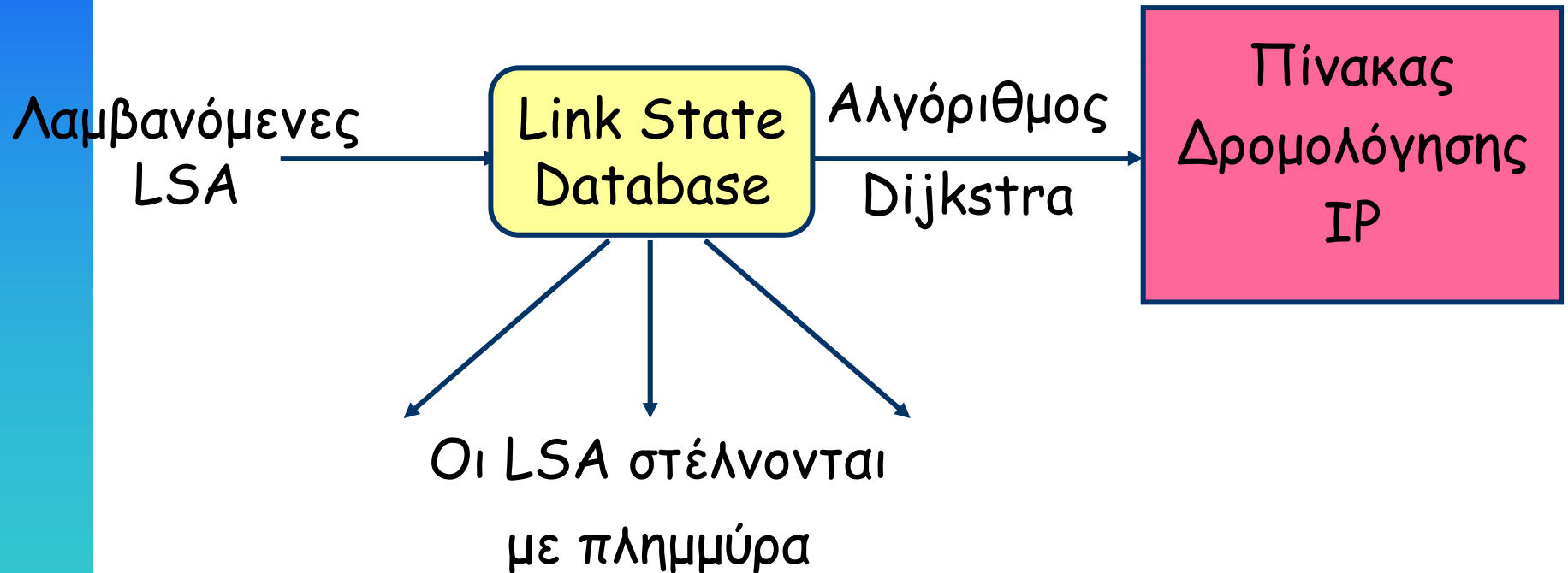
- Η τοπολογία δικτύου και τα κόστη των ζεύξεων είναι γνωστά σε όλους τους κόμβους
 - επιτυγχάνεται με **εκπομπή**, από κάθε δρομολογητή προς όλους τους δρομολογητές του δικτύου, του κόστους των ζεύξεων που είναι συνδεδεμένες σ' αυτόν
 - όλοι οι κόμβοι έχουν την ίδια πληροφορία
- Κάθε κόμβος εκτελεί τον αλγόριθμο και υπολογίζει τις διαδρομές ελαχίστου κόστους προς όλους τους άλλους
 - Αποτέλεσμα: ο **πίνακας δρομολόγησης** για τον κόμβο
- **Ο αλγόριθμος Dijkstra**: επαναληπτικός, μετά k επαναλήψεις είναι γνωστή η διαδρομή ελαχίστου κόστους για k προορισμούς.

Link State Routing: Βασικές αρχές

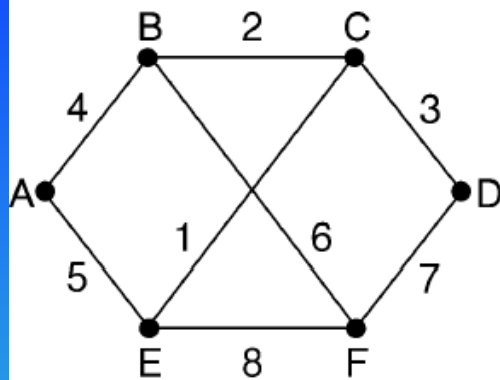


1. Κάθε δρομολογητής αναπτύσσει σχέση με τους γείτονές του
2. Υπολογίζει την καθυστέρηση ή το κόστος για καθέναν από τους γείτονές του
3. Παράγει *αναγγελίες κατάστασης ζεύξεων (LSA)*, οι οποίες διανέμονται σε όλους τους δρομολογητές
 - LSA = (link id, κατάσταση ζεύξης, κόστος, γείτονες)
4. Διατηρεί *database* με όλες τις λαμβανόμενες LSA (*database τοπολογίας* ή *link state database*), που περιγράφει το δίκτυο ως γράφο με τα βάρη των κλάδων
5. Κάθε δρομολογητής χρησιμοποιεί την link state database του για να τρέξει έναν αλγόριθμο ελάχιστης διαδρομής (αλγόριθμος Dijkstra) και να βρει την ελάχιστη διαδρομή προς κάθε προορισμό

Link State Routing: Λειτουργία



Link State Routing: Αναγγελίες



(α)

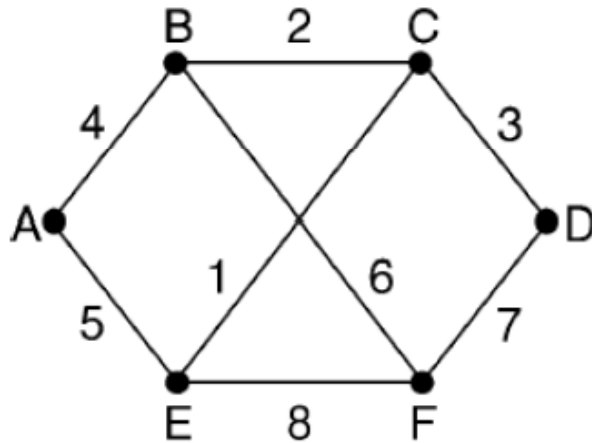
Πακέτα Κατάστασης Ζεύξεων

A	B	C	D	E	F
Seq.	Seq.	Seq.	Seq.	Seq.	Seq.
Age	Age	Age	Age	Age	Age
B 4	A 4	B 2	C 3	A 5	B 6
E 5	C 2	D 3	F 7	C 1	D 7
	F 6	E 1		F 8	E 8

(β)

- Αύξοντες αριθμοί 32 bit για να μην γίνεται αναδίπλωση γρήγορα
- Ηλικία κάθε πακέτου. Το πεδίο ηλικίας μειώνεται σε κάθε δρομ/τή
- Προσωρινή αναμονή των πακέτων σε κάθε δρομολογητή
- Επαληθεύσεις, για ασφάλεια έναντι λαθών

Link State Routing: Διανομή αναγγελιών



Πηγή	α/α	Ηλικία	Σημαίες αποστολής			Σημαίες ACK			Δεδομένα
			A	C	F	A	C	F	
A	21	60	0	1	1	1	0	0	
F	21	60	1	1	0	0	0	1	
E	21	59	0	1	0	1	0	1	
C	20	60	1	0	1	0	1	0	
D	21	59	1	0	0	0	1	1	

Κόμβος B

Ιεραρχική δρομολόγηση



Έως τώρα η μελέτη για τη δρομολόγηση ήταν ιδανική

- όλοι οι δρομολογητές ίδιοι
- το δίκτυο "επίπεδο"

... δεν συμβαίνει το ίδιο στην πράξη

Κλίμακα

με 200×10^6 προορισμούς:

- αδυναμία αποθήκευσης όλων των προορισμών στους πίνακες δρομολόγησης!
- η ανταλλαγή των πινάκων δρομολόγησης θα πλημμύριζε τις ζεύξεις!

Διοικητική αυτονομία

- διαδίκτυο = δίκτυο από δίκτυα
- κάθε διαχειριστής δικτύου μπορεί να θέλει να ελέγχει τη δρομολόγηση στο δίκτυό του

Ιεραρχική δρομολόγηση



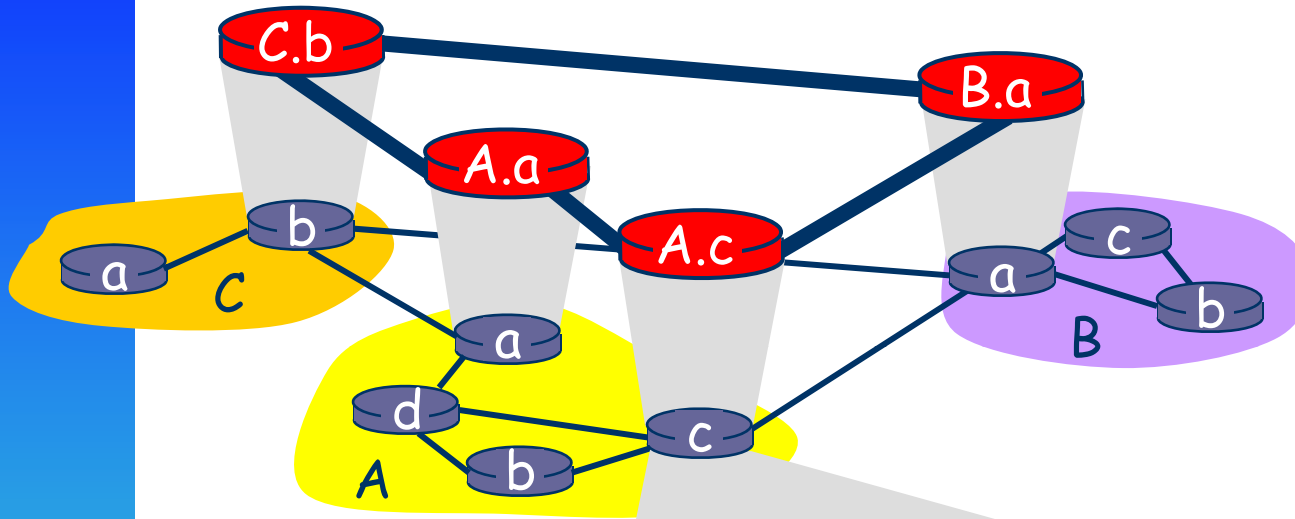
- Το Internet αποτελείται από **Αυτόνομα Συστήματα (AS)** διασυνδεδεμένα μεταξύ τους:
 - **Stub AS**: μικρή επιχείρηση, μία σύνδεση με άλλο AS
 - **Multihomed AS**: μεγάλη επιχείρηση (όχι transit), πολλαπλές συνδέσεις με άλλα AS
 - **Transit AS**: πάροχος, διασυνδέει πολλά AS
- δρομολογητές στο ίδιο AS τρέχουν το ίδιο "**intra-AS**" πρωτόκολλο δρομολόγησης
- δρομολογητές σε διαφορετικά AS μπορεί να τρέχουν διαφορετικό **intra-AS** πρωτόκολλο δρομολόγησης

δρομολογητές πύλες

- ειδικοί δρομολογητές στο AS
- **υπεύθυνοι** για δρομολόγηση προς προορισμούς εκτός AS
 - τρέχουν πρωτόκολλο δρομολόγησης **inter-AS** με τους άλλους δρομολογητές πύλες



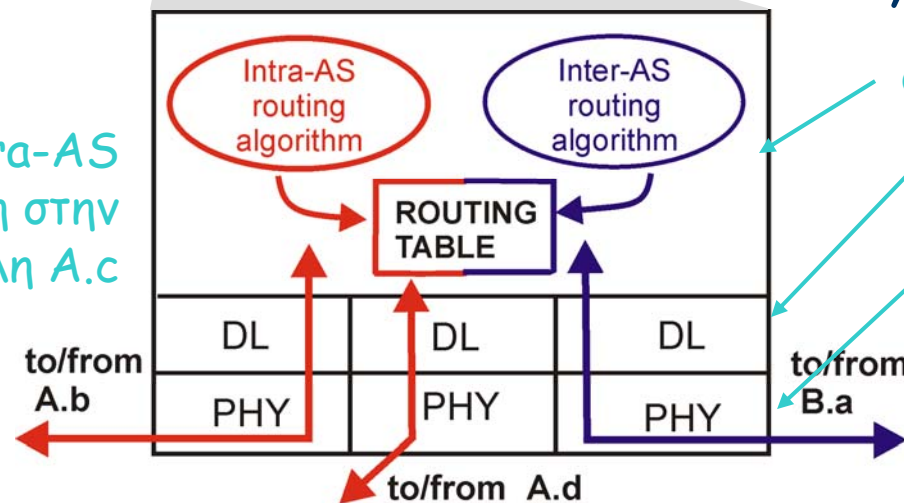
Δρομολόγηση Intra-AS και Inter-AS



Πύλες:

- πραγματοποιούν δρομολόγηση **inter-AS** μεταξύ τους
- πραγματοποιούν δρομολόγηση **intra-AS** με τους άλλους δρομολογητές στο AS τους

inter-AS, intra-AS
δρομολόγηση στην
πύλη A.c

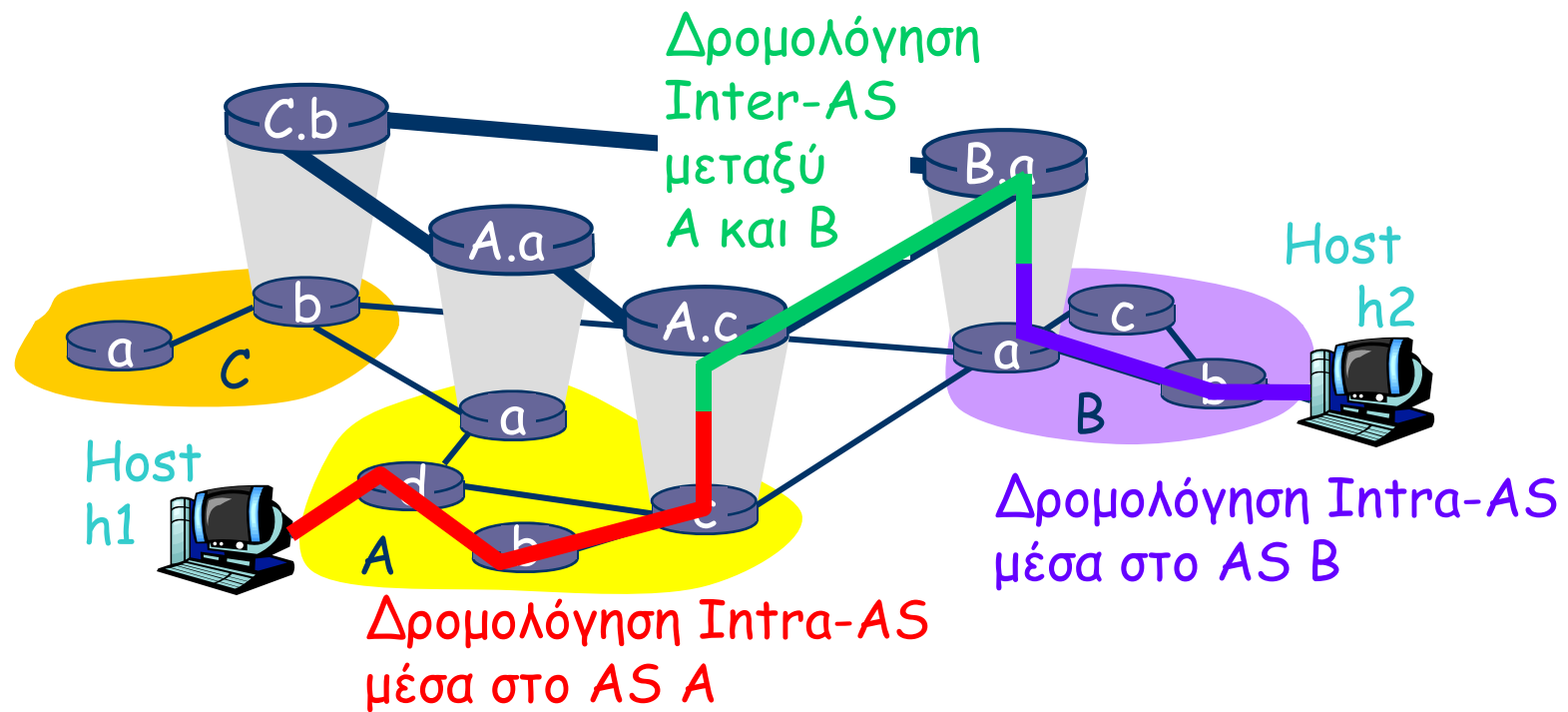


στρώμα δικτύου

στρώμα ζεύξης

φυσικό στρώμα

Δρομολόγηση Intra-AS και Inter-AS





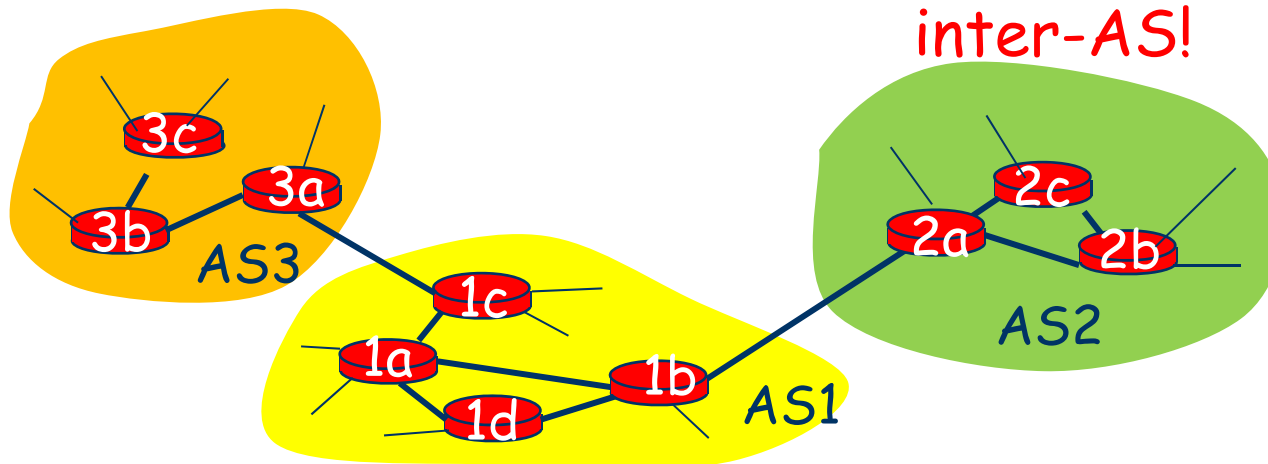
Δρομολόγηση Inter-AS

- έστω ότι κάποιος router στο AS1 λαμβάνει δεδομένογραμμα με προορισμό εκτός του AS1:
 - ο router θα πρέπει να προωθήσει το πακέτο στον gateway router, αλλά ποιον;

Το AS1 πρέπει:

1. να μάθει ποιοι προορισμοί είναι προσβάσιμοι μέσω του AS2 και ποιοι μέσω του AS3
2. να διαδώσει αυτή την πληροφορία πρόσβασης σε όλους τους routers στο AS1

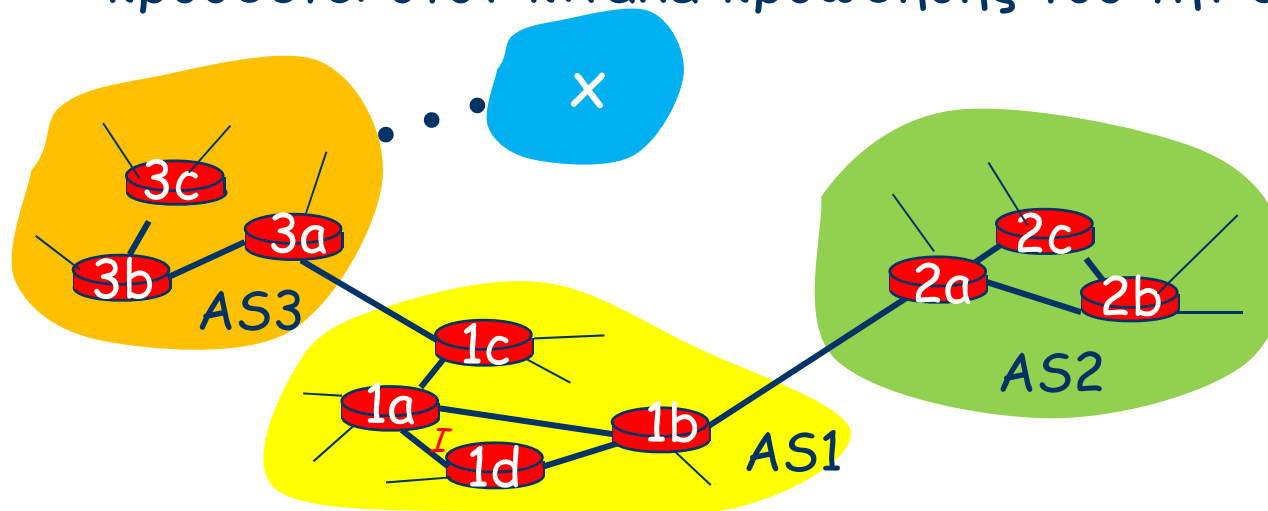
Έργο της δρομολόγησης inter-AS!



Παράδειγμα διάρθρωσης πίνακα προώθησης



- Έστω ότι το AS1 μαθαίνει (μέσω του πρωτοκόλλου inter-AS) ότι υποδίκτυο X είναι προσβάσιμο μέσω του AS3 (gateway 1c) αλλά όχι μέσω του AS2 (gateway 1b).
- Το πρωτόκολλο inter-AS διαδίδει την πληροφορία πρόσβασης σε όλους τους εσωτερικούς δρομολογητές.
- Ο δρομολογητής 1d γνωρίζει από τη δρομολόγηση intra-AS ότι η I διεπαφή του είναι στην διαδρομή ελαχίστου κόστους προς τον 1c.
 - προσθέτει στον πίνακα προώθησής του την εγγραφή (X,I)



Πρωτόκολλα Δρομολόγησης Intra-AS



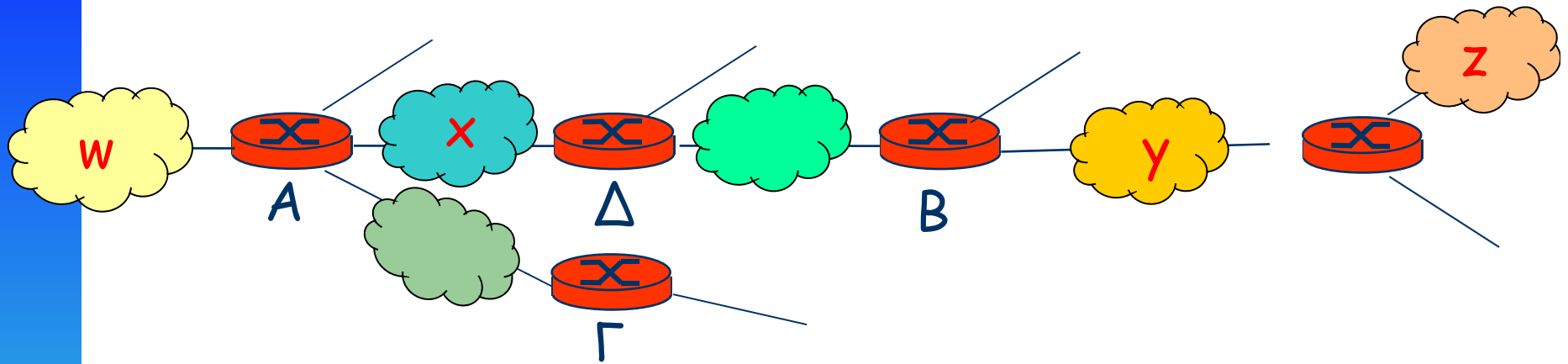
- Γνωστά και ως **Interior Gateway Protocols (IGP)**
- Τα γνωστότερα πρωτόκολλα δρομολόγησης Intra-AS:
 - **RIP**: Routing Information Protocol
 - **OSPF**: Open Shortest Path First
 - **IGRP**: Interior Gateway Routing Protocol (ιδιωτικό της Cisco)

RIP



- **RIP: Routing Information Protocol**
- Χρησιμοποιεί αλγόριθμο διανύσματος αποστάσεων
- Ως μέτρο των αποστάσεων χρησιμοποιείται ο αριθμός των βημάτων [max = 15 βήματα (hops)]
- Τα διανύσματα αποστάσεων ανταλλάσσονται μεταξύ γειτόνων κάθε 30 sec μέσω απαντητικού μηνύματος, που ονομάζεται **αναγγελία RIP (RIP advertisement)**
- Κάθε αγγελία είναι μια λίστα που περιέχει μέχρι 25 υποδίκτυα προορισμού μέσα στο AS

RIP: Παράδειγμα



Δίκτυο προορισμού	Επόμενος δρομολογ.	Αριθμός βημάτων
W	A	2
Y	B	2
Z	B	7
X	--	1
....

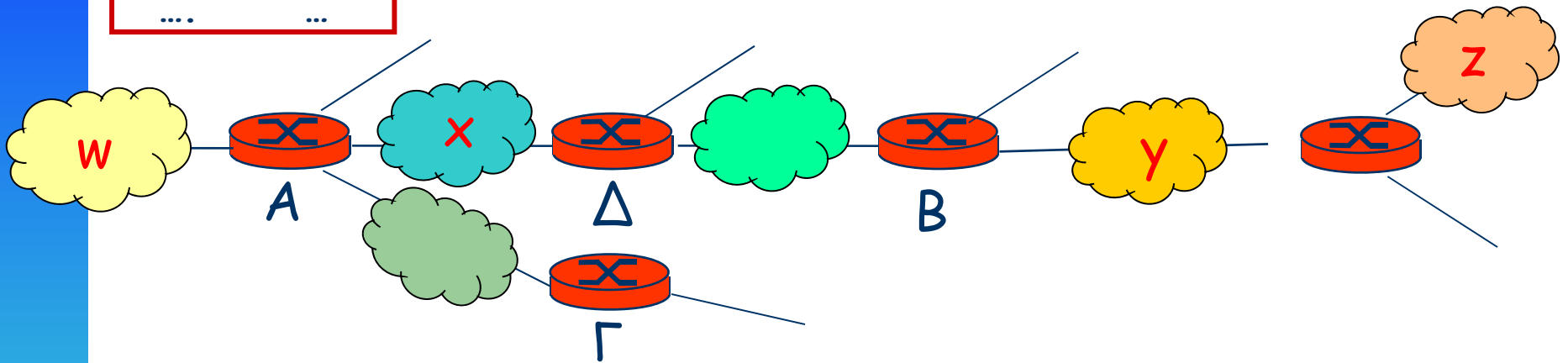
Πίνακας δρομολόγησης/προώθησης του Δ

RIP: Παράδειγμα



Προορ.	βήμ.
w	-
x	-
z	3
...	...

Διάνυση
από A προς Δ



Δίκτυο προορισμού	Επόμενος δρομολογ.	Αριθμός βημάτων
w	A	2
y	B	2
z	B A	7 4
x	--	1
...

Πίνακας δρομολόγησης/προώθησης του Δ

Δίκτυα υπολογιστών

RIP: Ανακάλυψη διακοπής ζεύξης

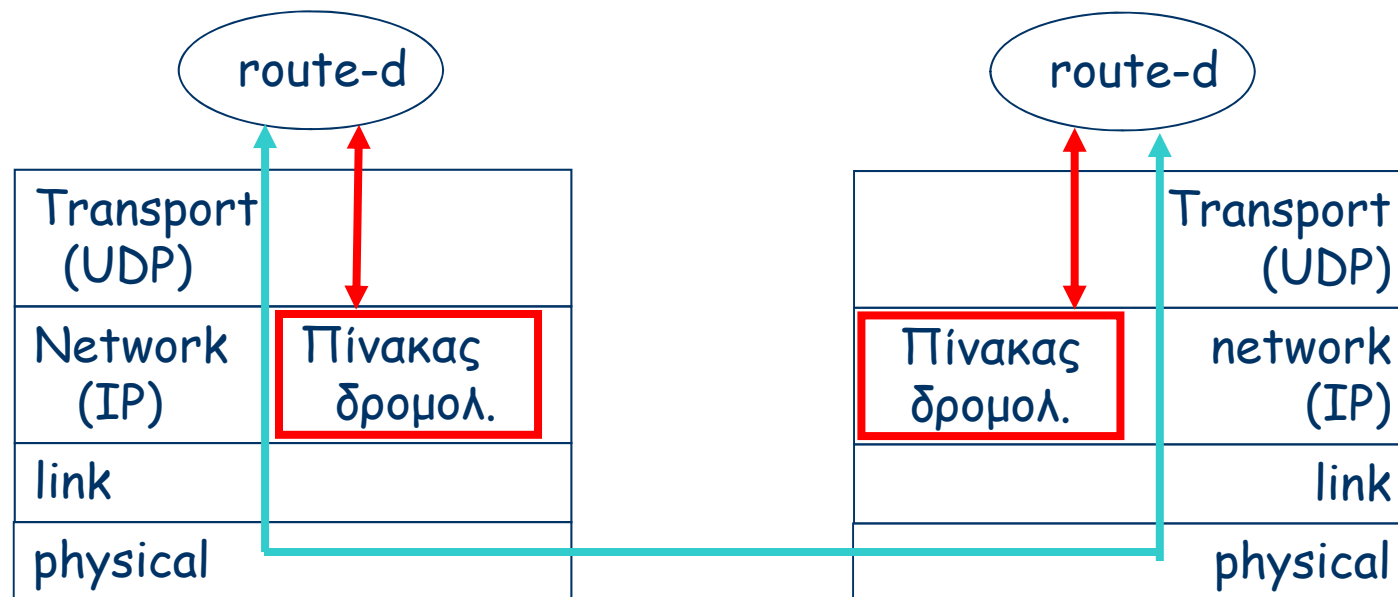


- Αν δεν ακουστεί αγγελία μετά 3 min, η ζεύξη προς τον γείτονα θεωρείται νεκρή
 - οι διαδρομές μέσω του γείτονα ακυρώνονται
 - στέλνονται νέες αγγελίες προς τους γείτονες
 - και οι γείτονες με τη σειρά τους στέλνουν νέες αγγελίες (αν αλλάξουν οι πίνακες)
 - η πληροφορία για τη διακοπή ζεύξης διαδίδεται σ' όλο το δίκτυο
 - χρησιμοποιείται poisoned reverse για να παρεμποδιστούν οι βρόχοι (άπειρη απόσταση = 16 βήματα)

RIP: Διαχείριση πινάκων



- Η διαχείριση των πινάκων δρομολόγησης RIP γίνεται από διαδικασία του στρώματος εφαρμογών που ονομάζεται route-d (daemon)
- Οι αγγελίες στέλνονται με πακέτα UDP, περιοδικά επαναλαμβανόμενα



RIP: Παράδειγμα πίνακα



Router: *giroflée.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	

- Συνδέονται τρία δίκτυα κατηγορίας C (LAN)
- Ο δρομολογητής γνωρίζει μόνο τις διαδρομές για τα συνδεδεμένα LAN
- Χρησιμοποιείται Default δρομολογητής για προώθηση
- Διεύθυνση πολλαπλής διανομής: 224.0.0.0
- Διεπαφή βρόχου επιστροφής (loopback) για debugging

OSPF



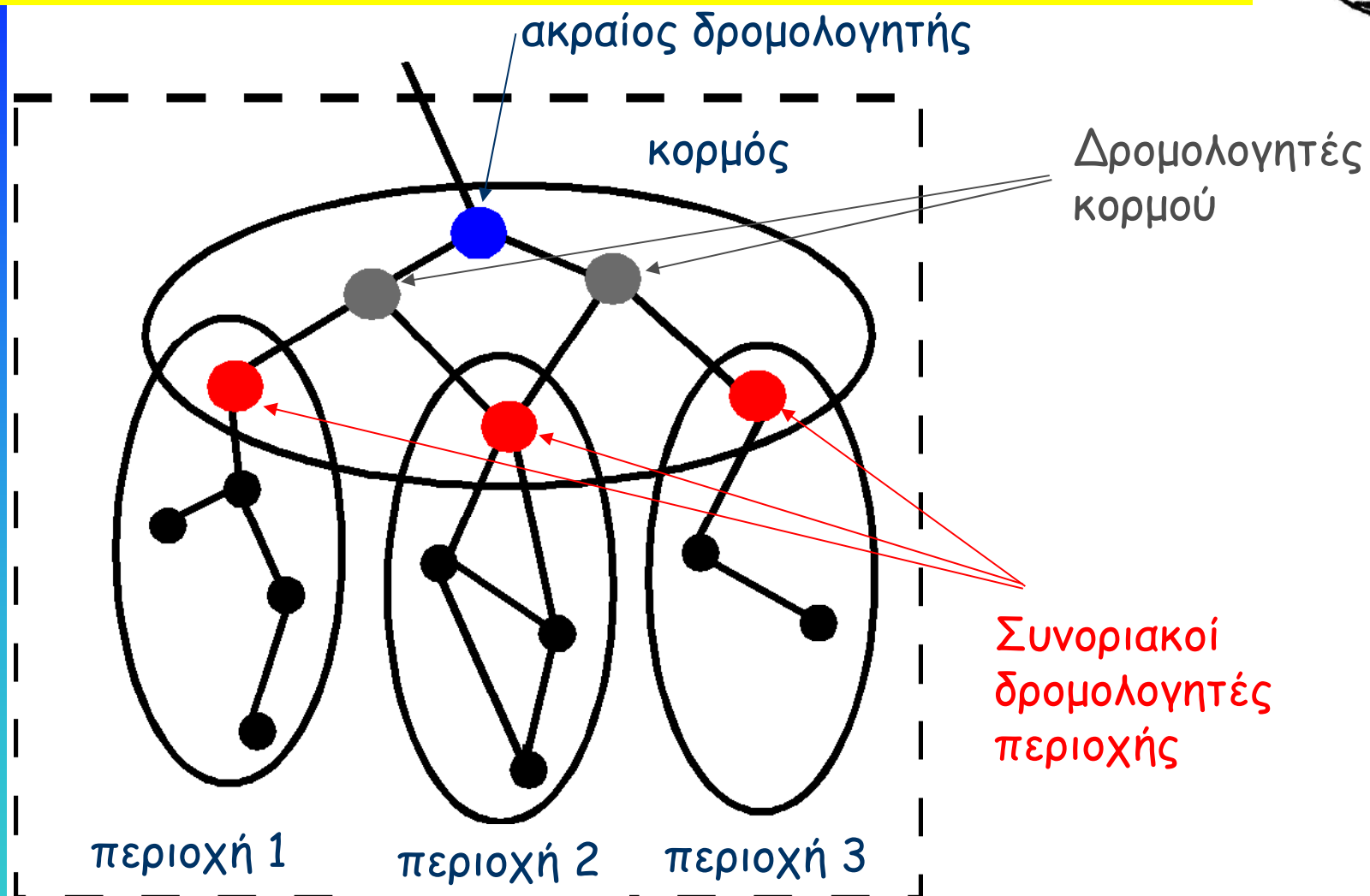
- **OSPF: Open Shortest Path First**
- "open": δημόσια διαθέσιμο
- Χρησιμοποιεί αλγόριθμο κατάστασης ζεύξεων
 - διασπορά πακέτων κατάστασης ζεύξεων
 - τοπολογία σε κάθε κόμβο
 - υπολογισμός διαδρομής με αλγόριθμο Dijkstra
- Η αναγγελία OSPF έχει μια εγγραφή για κάθε γειτονικό δρομολογητή
- Οι αγγελίες διασπείρονται σε **όλο** το AS (με πλημμύρα)
 - Μεταφέρονται από μηνύματα OSPF απευθείας πάνω στο IP (upper-layer protocol: 89)

OSPF: Εξελιγμένες ιδιότητες



- Ασφάλεια: πιστοποιείται η αυθεντικότητα όλων των μηνυμάτων OSPF
- Επιτρέπονται περισσότερες τις μιας διαδρομές με το ίδιο κόστος (μόνο μία για το RIP)
- Για κάθε ζεύξη υπάρχουν πολλαπλές μετρικές για διαφορετικούς TOS (π.χ., το κόστος της δορυφορικής ζεύξης τίθεται "χαμηλό" για best effort και "υψηλό" για πραγματικό χρόνο)
- Ολοκληρωμένη υποστήριξη απλής μετάδοσης και πολλαπλής διανομής:
 - Το Multicast OSPF (MOSPF) χρησιμοποιεί την ίδια βάση δεδομένων για την τοπολογία με το OSPF
- Ιεραρχικό OSPF σε μεγάλες επικράτειες

Ιεραρχικό OSPF

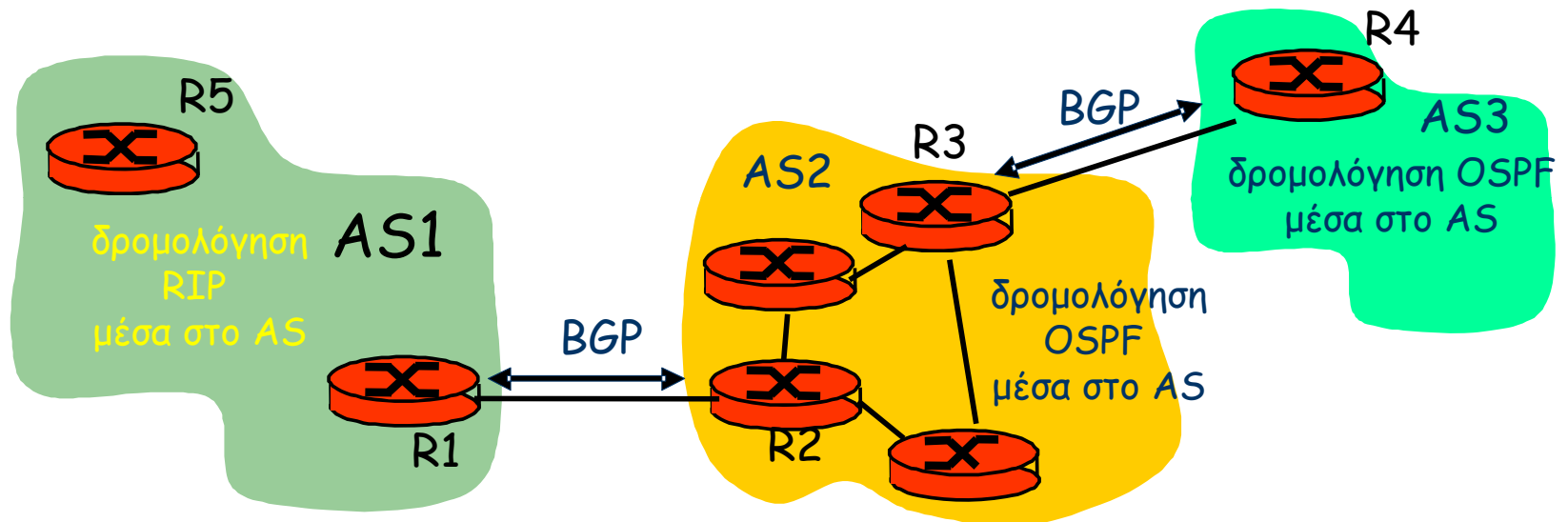


Ιεραρχικό OSPF



- **Ιεραρχία δύο επιπέδων:** τοπική περιοχή, δίκτυο κορμού.
 - αγγελίες κατάστασης ζεύξεων μόνο στην περιοχή
 - κάθε κόμβος γνωρίζει λεπτομερώς την τοπολογία της περιοχής
 - για τα δίκτυα άλλων περιοχών γνωρίζει μόνο την κατεύθυνση (συντομότερη διαδρομή).
- **Συνοριακοί δρομολογητές περιοχής (Area border routers):** "συνοψίζουν" αποστάσεις για τα δίκτυα της περιοχής τους και τις αναγγέλλουν προς άλλους Area Border routers.
- **Δρομολογητές κορμού (Backbone routers):** τρέχουν OSPF δρομολόγηση περιορισόμενη στο backbone.
- **Ακραίοι δρομολογητές (Boundary routers):** συνδέουν με άλλα AS.

BGP: δρομολόγηση μεταξύ AS



BGP: Δρομολόγηση μεταξύ AS

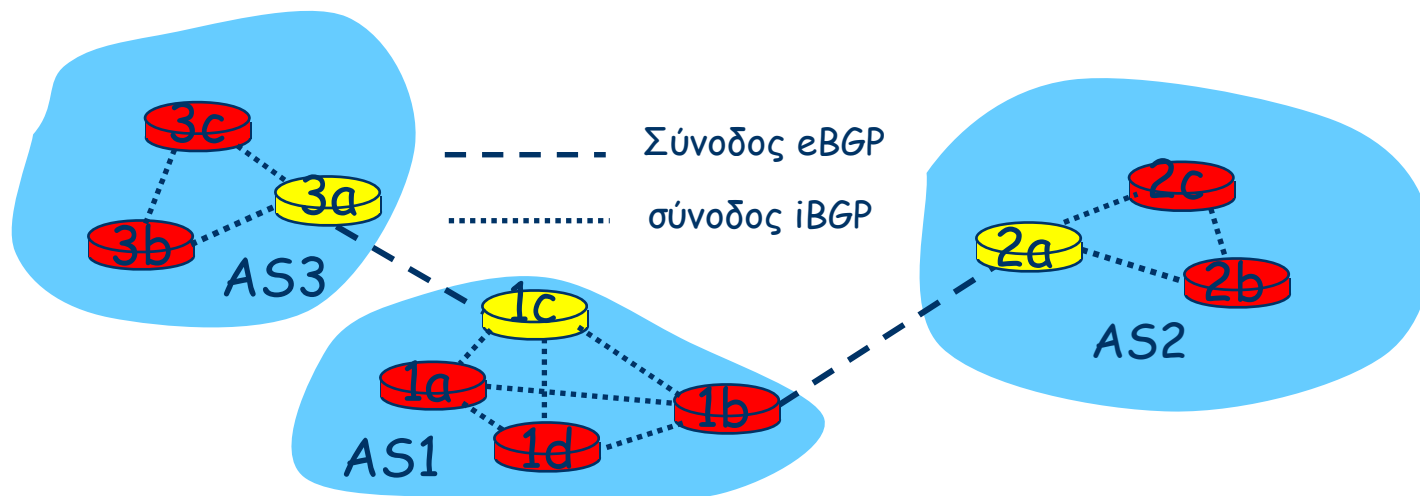


- **BGP (Border Gateway Protocol):** είναι το πρότυπο που ισχύει
- Το BGP παρέχει σε κάθε AS τα μέσα για:
 - να αποκτήσει πληροφορίες προσβασιμότητας από τα γειτονικά AS
 - να διαδίδει τη πληροφορία προσβασιμότητας σε όλους τους εσωτερικούς δρομολογητές του
 - να καθορίζει "καλές" διαδρομές προς τα άλλα AS βάσει της πληροφορίας προσβασιμότητας και της πολιτικής δρομολόγησης
- Επιτρέπει σε ένα AS να αναγγείλει την ύπαρξή του στο υπόλοιπο Internet.

BGP: Δρομολόγηση μεταξύ AS



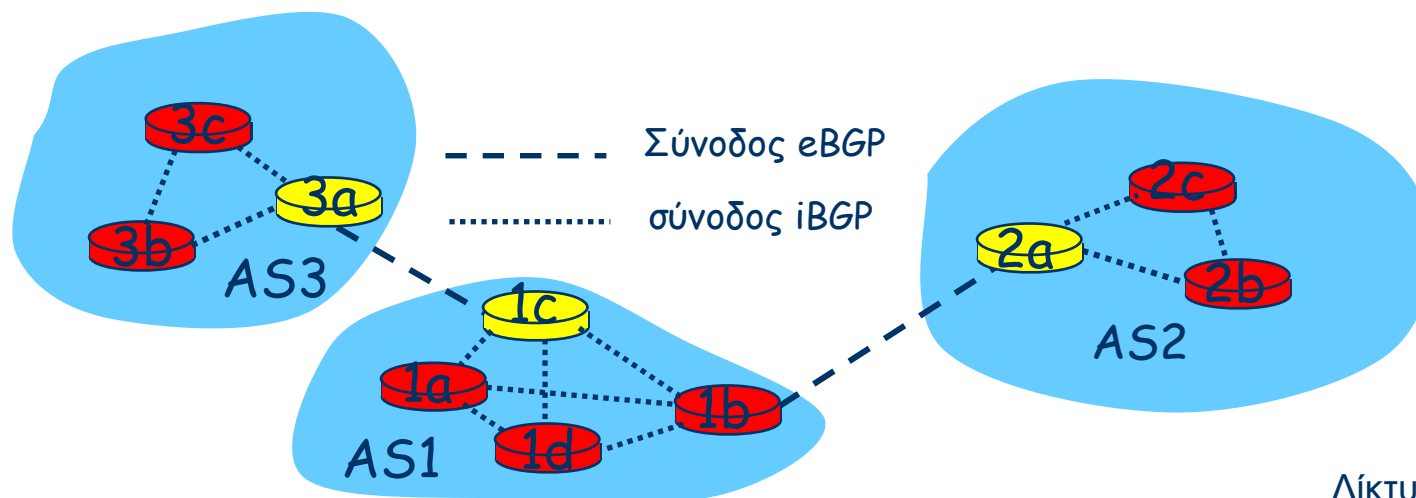
- Ζεύγη δρομολογητών (BGP peers) ανταλλάσσουν πληροφορίες δρομολόγησης πάνω από ημι-μόνιμες συνδέσεις TCP: **σύνοδοι BGP**
- Οι σύνοδοι BGP δεν αντιστοιχούν σε φυσικές ζεύξεις.
- Όταν το AS2 αναγγέλλει ένα πρόθεμα στο AS1, το AS2 **υπόσχεται** ότι θα προωθεί πακέτα προς το υπόψη πρόθεμα.
 - Το AS2 μπορεί να συνοψίζει προθέματα στην αναγγελία του



BGP: Διανομή πληροφορίας



- Χρησιμοποιώντας τη σύνοδο eBGP μεταξύ 3a και 1c, το AS3 στέλνει πληροφορία προσβασιμότητας στο AS1.
 - Ο 1c μπορεί στη συνέχεια, χρησιμοποιώντας iBGP να διανείμει αυτή τη νέα πληροφορία σε όλους τους δρομολογητές του AS1
 - Ο 1b μπορεί μετά να επαναγγείλει τη νέα πληροφορία προσβασιμότητας στο AS2 μέσω της συνόδου eBGP μεταξύ 1b και 2a
- Όταν ένας δρομολογητής μαθαίνει για ένα νέο πρόθεμα, δημιουργεί μια εγγραφή για το πρόθεμα αυτό στον πίνακα προώθησής του.

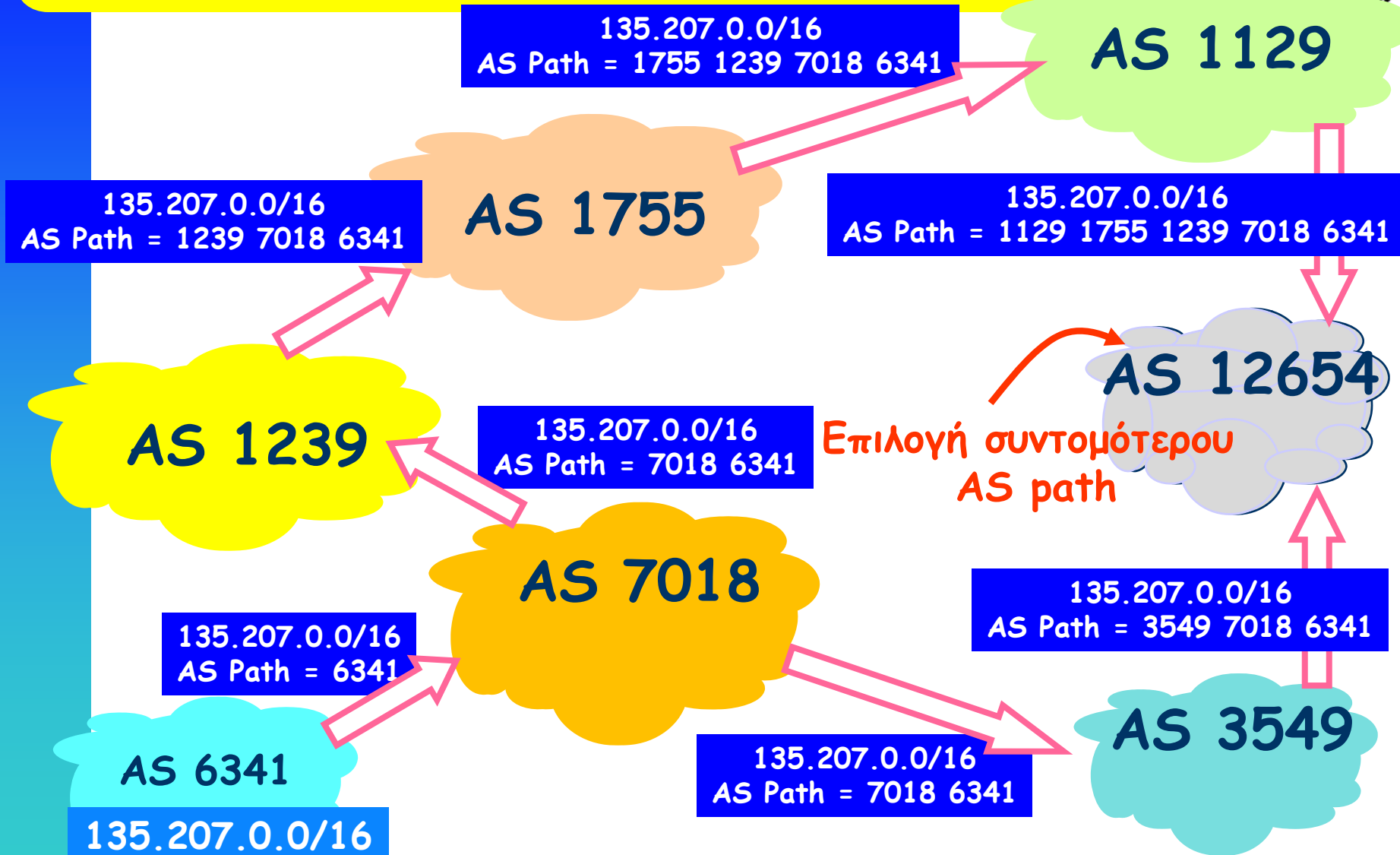


BGP: Διαδρομές BGP



- Όταν αναγγέλλεται ένα πρόθεμα, η αναγγελία περιέχει και χαρακτηριστικά BGP.
 - prefix + attributes = "route"
- Δύο ενδιαφέροντα attributes:
 - **AS-PATH**: περιέχει τα AS μέσω των οποίων πέρασε η αναγγελία για το υπόψη πρόθεμα: AS 67 AS 17
 - **NEXT-HOP**: πληροφορεί το AS που λαμβάνει την αναγγελία για το ποιος δρομολογητής από το αποστέλλον AS είναι ο next-hop. (Μπορεί να υπάρχουν πολλές ζεύξεις από το τρέχον AS προς το next-hop-AS.)
- Όταν ένας δρομολογητής πύλη λαμβάνει αναγγελία διαδρομής, χρησιμοποιεί **πολιτική εισαγωγής** για να τη δεχθεί ή να την απορρίψει.

BGP: AS-PATH attribute



Αρχή αναγγελίας

Δίκτυα υπολογιστών

BGP: Επιλογή διαδρομής



- Ένας δρομολογητής μπορεί να πληροφορηθεί για περισσότερες από μία διαδρομές προς κάποιο πρόθεμα. Πρέπει τότε να επιλέξει διαδρομή.
- Κανόνες επιλογής διαδρομής:
 - Πολιτική δρομολόγησης
 - Συντομότερη AS-PATH
 - Πλησιέστερος NEXT-HOP router
 - Επιπρόσθετα κριτήρια
- Ένας δρομολογητής μπορεί να μη θέλει να αναγγέλλει κάποιες διαδρομές

Μηνύματα BGP



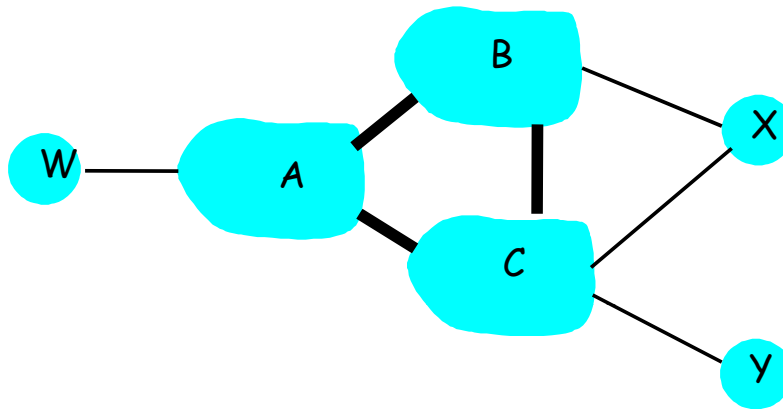
- Τα μηνύματα BGP ανταλλάσσονται με χρήση TCP.
- Μηνύματα BGP:
 - **OPEN**: ανοίγει σύνδεση TCP προς ομότιμη οντότητα και πιστοποιεί την αυθεντικότητα του αποστολέα
 - **UPDATE**: αναγγέλλει νέα διαδρομή (ή αποσύρει παλαιά)
 - **KEEPALIVE**: διατηρεί ανοικτή τη σύνδεση όταν δεν υπάρχουν μηνύματα UPDATE. Επίσης, επαληθεύει αίτηση OPEN
 - **NOTIFICATION**: αναφέρει σφάλματα προηγούμενου μηνύματος. Επίσης, χρησιμοποιείται για κλείσιμο σύνδεσης

BGP: Δρομολόγηση μεταξύ AS



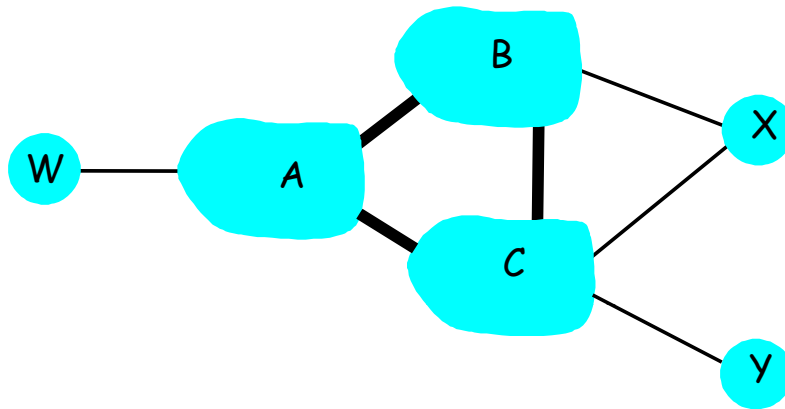
- Έστω ότι η πύλη X στέλνει προς την ομότιμη πύλη W τη διαδρομή της προς Z
- η W μπορεί και να μην επιλέξει τη διαδρομή που προσφέρεται από τη X
 - κόστος, πολιτική (ανταγωνιστικό AS), αποφυγή βρόχου.
- αν η W επιλέξει τη διαδρομή που αναγγέλλεται από την X , τότε:
$$\text{Path}(W,Z) = \text{Path}(W,X), \text{Path}(X,Z)$$
- η X μπορεί να ελέγχει την εισερχόμενη κίνηση ελέγχοντας τις αναγγελίες προς τις ομότιμες πύλες:
 - π.χ., αν δεν θέλει δρομολόγηση προς τη Z δεν θα αναγγέλλει διαδρομές προς τη Z

BGP: ένα απλό σενάριο



- A,B,C είναι **δίκτυα παρόχων**
- X,W,Y είναι πελάτες
- το X είναι συνδεδεμένο σε δύο δίκτυα παρόχων
 - το X δεν θέλει διαδρομή από το B προς το C μέσω X
⇒ το X δεν αναγγέλλει προς το B διαδρομή προς το C

BGP: ένα απλό σενάριο



- το A αναγγέλλει στο B τη διαδρομή AW
- το B αναγγέλλει στο X τη διαδρομή BAW
- πρέπει το B να αναγγείλει στο C τη διαδρομή BAW;
 - Όχι! Το B δεν κερδίζει τίποτε από τη διαδρομή CBAW, καθώς ούτε το W ούτε το C είναι πελάτες του B
 - το B θέλει να ωθήσει το C να δρομολογεί στο w μέσω του A
 - το B θέλει να δρομολογεί **μόνο** προς/από τους πελάτες του!

BGP: Λειτουργία δρομολογητή BGP



Ένας δρομολογητής BGP:

- Λαμβάνει και φιλτράρει αγγελίες διαδρομών από άμεσα συνδεδεμένους γείτονες
- Επιλέγει διαδρομές
 - Για τον προορισμό X, ποια διαδρομή (από αρκετές αναγγελλόμενες) θα επιλεγεί;
- Στέλνει αγγελίες διαδρομών προς τους γείτονες

Γιατί δρομολόγηση Intra- και Inter-AS:



Πολιτική:

- Inter-AS: ο διαχειριστής θέλει να ελέγχει πώς δρομολογείται η κίνηση του δικτύου του και ποιος δρομολογεί μέσω του δικτύου του.
- Intra-AS: μία επικράτεια διαχείρισης, οπότε δεν χρειάζεται πολιτική δρομολόγησης

Κλίμακα:

- Η ιεραρχική δρομολόγηση εξοικονομεί μέγεθος πινάκων, περιορίζει το φορτίο ενημέρωσης

Επίδοση:

- Intra-AS: μπορεί να εστιάσει στην επίδοση
- Inter-AS: η πολιτική μπορεί να επισκιάζει την επίδοση