



ΔΙΚΤΥΑ ΥΠΟΛΟΓΙΣΤΩΝ

Δρομολόγηση



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



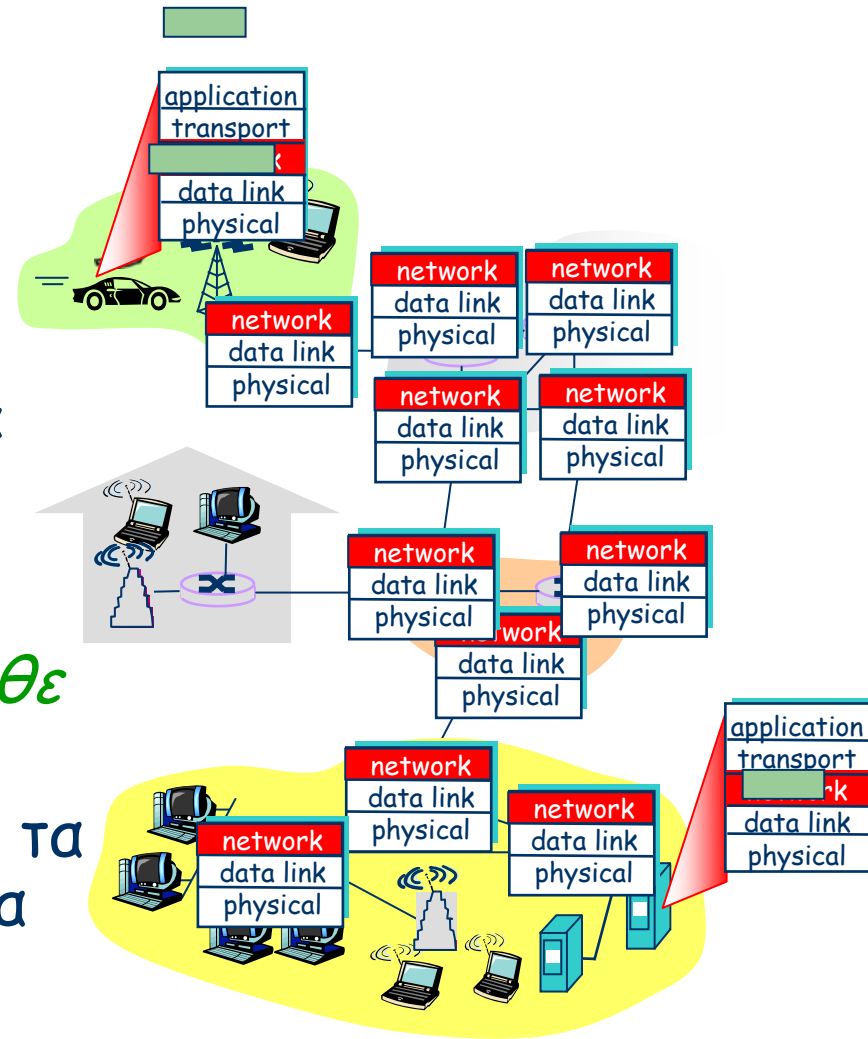
Περιεχόμενα

- Εισαγωγή και βασικές λειτουργίες του στρώματος δικτύου
 - Δίκτυα με νοητά κυκλώματα και δεδομενογράμματα
- Βασικές αρχές της δρομολόγησης
- Αλγόριθμοι δρομολόγησης
 - Διανύσματος αποστάσεων (Distance Vector)
 - Κατάστασης ζεύξεων (Link State)
- Ιεραρχική δρομολόγηση
- Πρωτόκολλα δρομολόγησης στο Internet
- Δρομολόγηση πολλαπλής διανομής

Λειτουργίες στρώματος 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

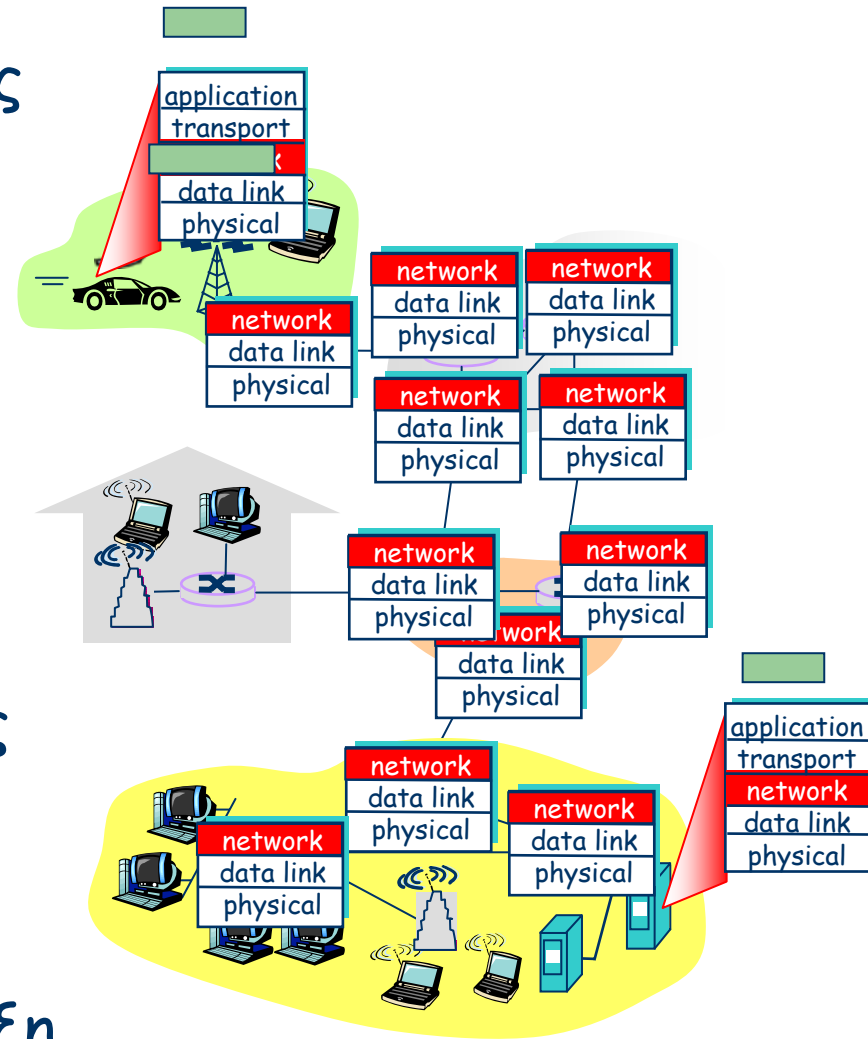


Λειτουργίες στρώματος δικτύου

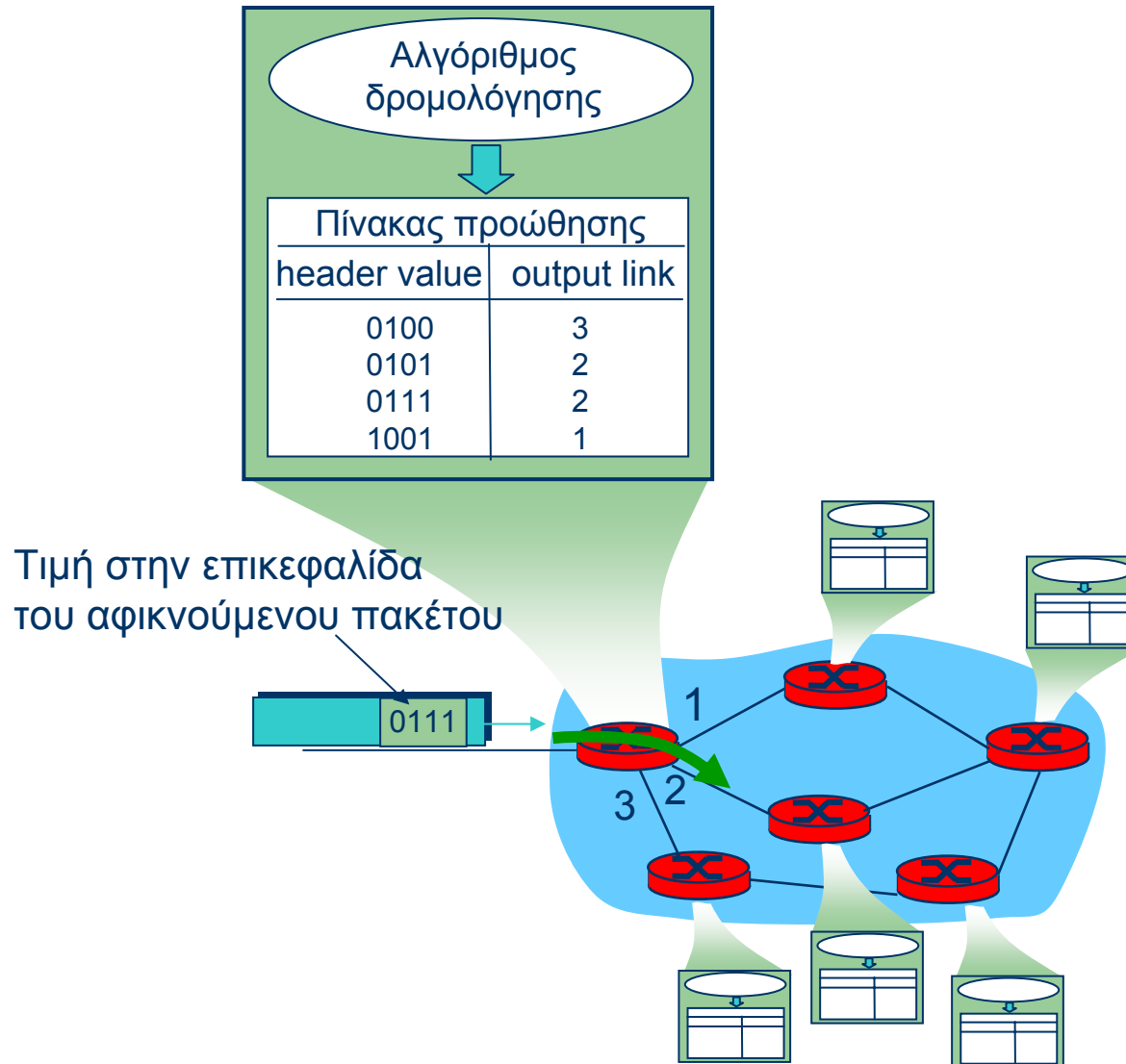


Τρεις κύριες λειτουργίες:

- **δρομολόγηση:** καθορισμός της διαδρομής των πακέτων από την πηγή στον προορισμό.
Αλγόριθμοι δρομολόγησης
- **προώθηση:** μετακίνηση πακέτων από την είσοδο του δρομολογητή στην κατάλληλη έξοδό του
- **εγκατάσταση κλήσης:** μερικές αρχιτεκτονικές δικτύου απαιτούν εγκατάσταση διαδρομής μεταξύ των δρομολογητών πριν την έναρξη αποστολής δεδομένων



Δρομολόγηση και προώθηση



Μοντέλο υπηρεσίας δικτύου



Ποιο είναι το *μοντέλο υπηρεσίας* για τον "δίαυλο" που μεταφέρει δεδομενογράμματα από τον πομπό στον δέκτη;

Παραδείγματα
υπηρεσιών με ανεξάρτητα
δεδομενογράμματα:

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

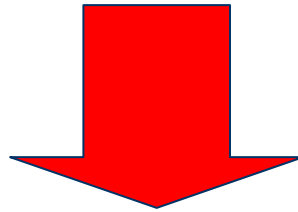
Παραδείγματα
υπηρεσιών με ροή
δεδομενογραμμάτων:

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

Υπηρεσία δικτύου με σύνδεση και χωρίς σύνδεση



Η πιο συνοπτική περιγραφή του μοντέλου υπηρεσίας που παρέχεται από το στρώμα δικτύου



νοητό κύκλωμα ή δεδομενόγραμμα;



υπηρεσία με σύνδεση



υπηρεσία χωρίς σύνδεση

Νοητά κυκλώματα



“η διαδρομή από την πηγή στον προορισμό συμπεριφέρεται ως τηλεφωνικό κύκλωμα”

- μέριμνα του δικτύου κατά μήκος της εν λόγω διαδρομής

- εγκατάσταση νοητού κυκλώματος για κάθε κλήση *πριν* την αποστολή δεδομένων, απόλυση κλήσης στο τέλος
- κάθε πακέτο μεταφέρει την ταυτότητα του νοητού κυκλώματος (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)

Υλοποίηση του νοητού κυκλώματος

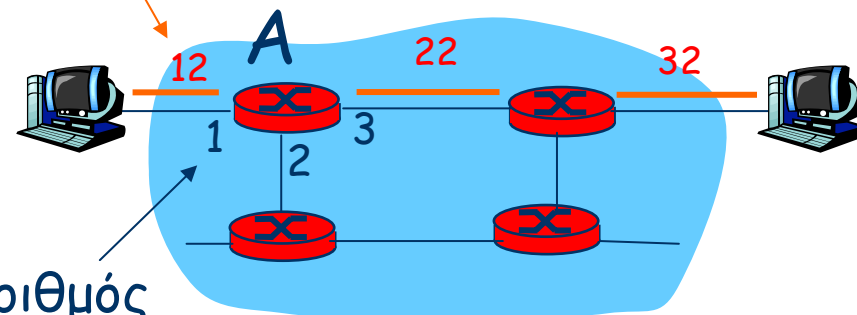


- Ένα νοητό κύκλωμα απαρτίζεται από:
 - Διαδρομή από την πηγή στον προορισμό
 - Αριθμούς VC , έναν αριθμό για κάθε ζεύξη
 - Εγγραφές στους πίνακες προώθησης στους δρομολογητές κατά μήκος της διαδρομής
- Πακέτο που ανήκει σε νοητό κύκλωμα μεταφέρει έναν αριθμό VC .
- Ο αριθμός VC πρέπει να αλλάζει σε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

Πίνακας προώθησης



Αριθμός VC



Πίνακας προώθησης του
δρομολογητή A

Αριθμός
διεπαφής

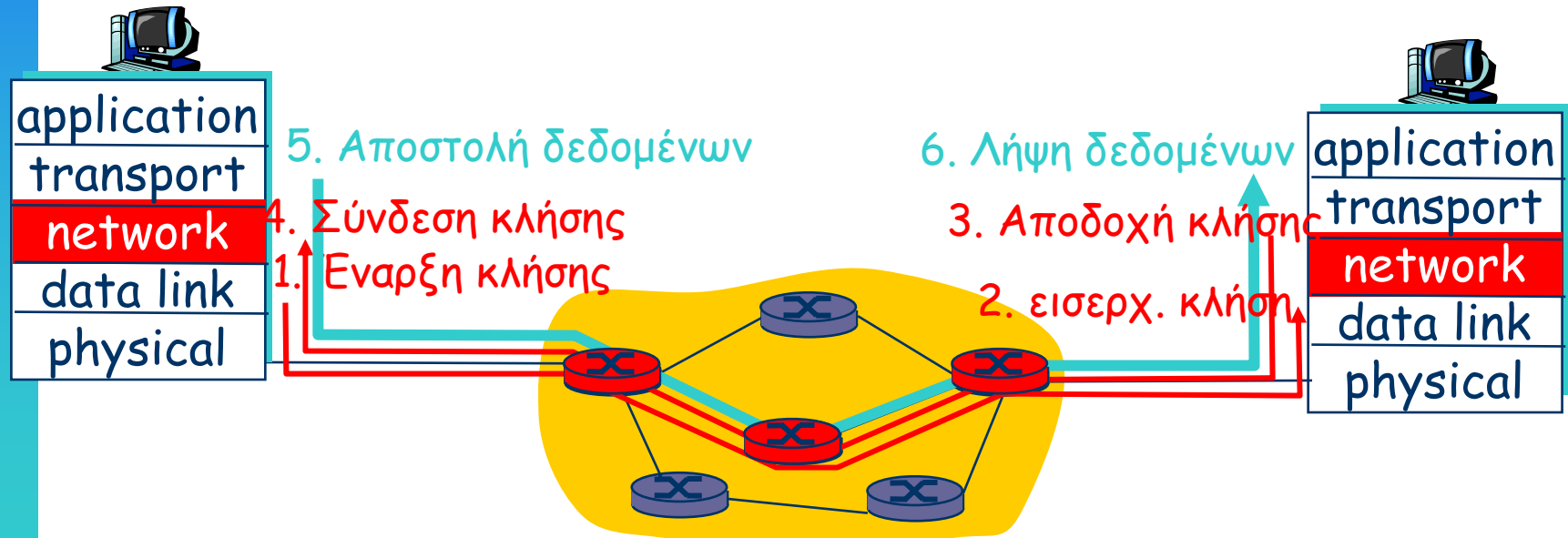
Διεπαφή εισόδου	#VC εισόδου	Διεπαφή εξόδου	#VC εξόδου
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
...	...	...	...

Οι δρομολογητές διατηρούν πληροφορία
για την κατάσταση των συνδέσεων

Νοητά κυκλώματα



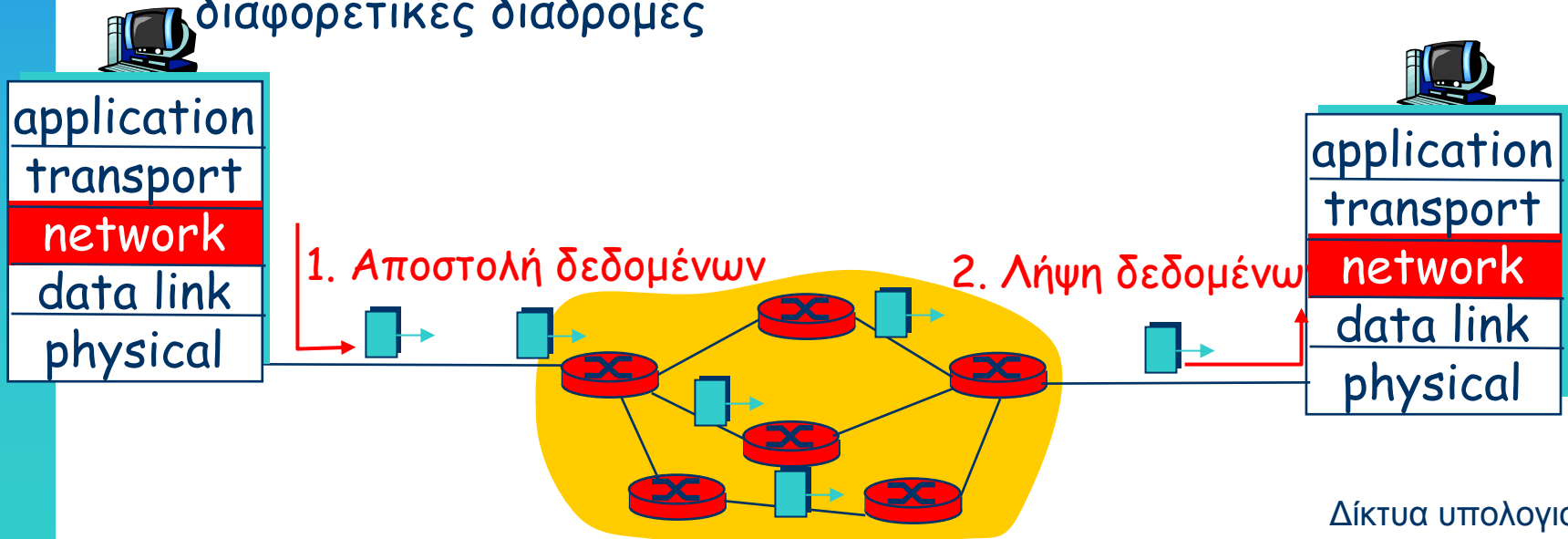
- Για την εγκατάσταση και απόλυση νοητών κυκλωμάτων χρησιμοποιούνται **πρωτόκολλα σηματοδότησης**
- χρησιμοποιούνται στο ATM, X.25,
- δεν χρησιμοποιούνται στο σημερινό Internet



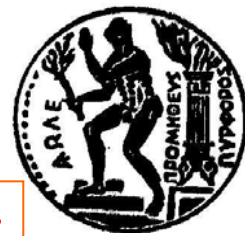
Δεδομενογράμματα: μοντέλο Internet



- 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



Πίνακας προώθησης



4 δισεκατομμύρια
δυνατές εισοδοι

Περιοχή διευθύνσεων προορισμού

Διεπαφή

11001000 00010111 00010000 00000000

έως

0

11001000 00010111 00010111 11111111

11001000 00010111 00011000 00000000

έως

1

11001000 00010111 00011000 11111111

11001000 00010111 00011001 00000000

έως

2

11001000 00010111 00011111 11111111

αλλιώς

3

Μεγαλύτερο ταίριασμα προθέματος



Πρόθεμα

11001000 00010111 00010

11001000 00010111 00011000

11001000 00010111 00011

αλλιώς

Διεπαφή

0

1

2

3

Παραδείγματα

DA: 11001000 00010111 00010110 10100001

Διεπαφή 0

DA: 11001000 00010111 00011000 10101010

Διεπαφή 1

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

Δρομολόγηση: Εισαγωγή



- Ανεξαρτήτως του 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: Εισαγωγή

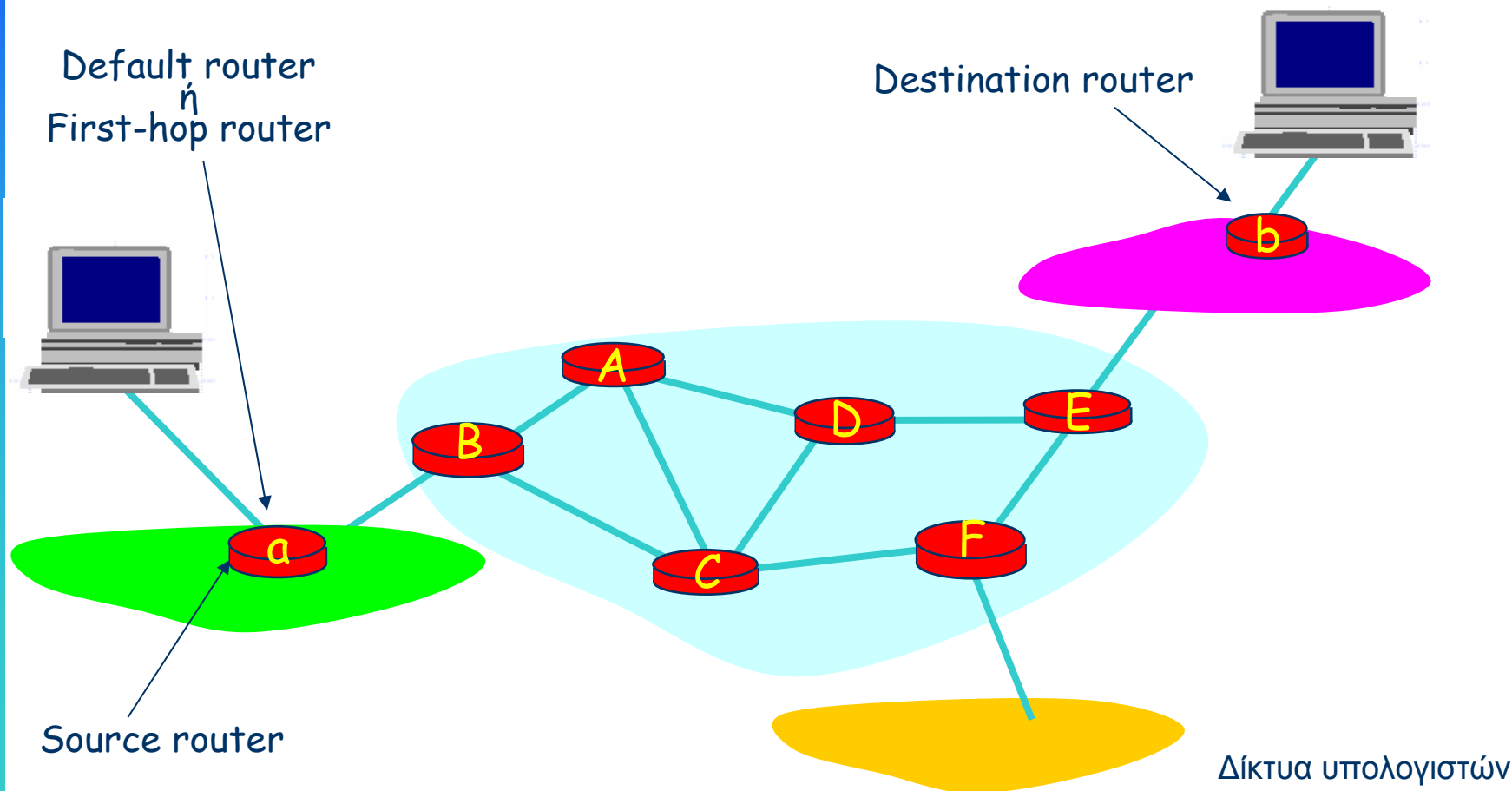


- Κύριες λειτουργίες:
 - Μετάφραση: διεύθυνση => διαδρομή
 - Συντονιζόμενες κατανεμημένες αποφάσεις: ποια είναι η απερχόμενη ζεύξη;
 - Γενικός στόχος επίδοσης: π.χ., ελαχιστοποίηση καθυστέρησης
 - Άλλα: προτεραιότητες, ασφάλεια, πολιτικές επικρατειών
- Γιατί ενδιαφέρει η δρομολόγηση;
 - Υπηρεσία του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

Δρομολόγηση: Εισαγωγή



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



Δρομολόγηση: δυσκολίες



- Επίπεδος (flat) χώρος διευθύνσεων, οπότε απαιτείται υπολογισμός των διαδρομών
 - Απαιτεί γνώση της συνολικής τοπολογίας και συντονιζόμενες κατανεμημένες 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

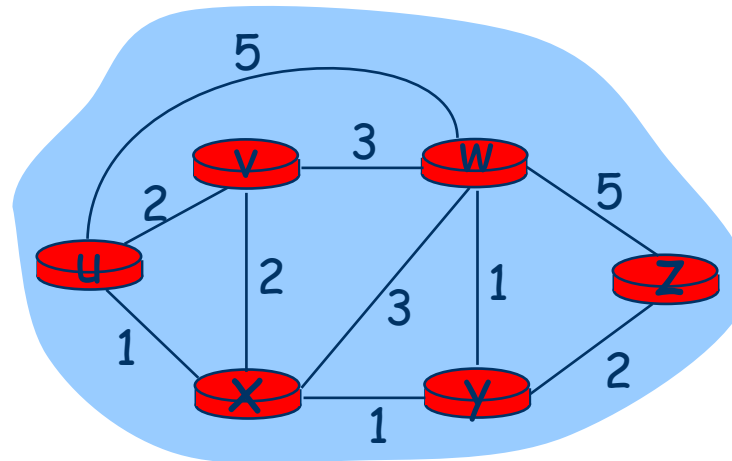


- Η καρδιά του πρωτοκόλλου δρομολόγησης είναι ο **αλγόριθμος δρομολόγησης** που καθορίζει τη διαδρομή από την πηγή στον προορισμό
- Στοιχεία ενός αλγορίθμου δρομολόγησης
 - Διαδικασία για την αποστολή και λήψη πληροφοριών πρόσβασης στους άλλους δρομολογητές του δικτύου.
 - Διαδικασία για τον υπολογισμό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

Γράφος δικτύου



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

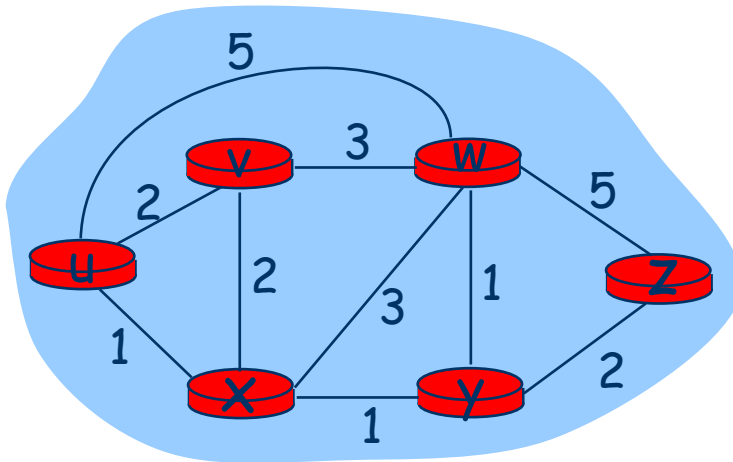


Γράφος: $G = (N, E)$

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

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

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 ελαχίστου κόστους

Ταξινόμηση αλγορίθμων δρομολόγησης



Προέλευση της πληροφορίας δρομολόγησης:

Καθολική:

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

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**

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
 - όταν αλλάζουν τα κόστη των ζεύξεων

Αλγόριθμος διανύσματος αποστάσεων



- Distance vector
- Κάθε κόμβος ξέρει την απόσταση (κόστος) προς τους άμεσα συνδεόμενους με αυτόν γείτονες
- Κάθε κόμβος στέλνει περιοδικά το δικό του διάνυσμα αποστάσεων στους γείτονές του.
- Αν όλοι οι κόμβοι ενημερώσουν τις αποστάσεις τους, οι πίνακες δρομολόγησης τελικά συγκλίνουν
- Νέοι κόμβοι αναγγέλλουν τους εαυτούς τους στους γείτονές τους

Αλγόριθμος διανύσματος αποστάσεων



Επαναληπτικός, ασύγχρονος:

κάθε τοπικός επαναυπολογισμός 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 *μόνο* τους γείτονες, 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 όπου 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:

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

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

Παράδειγμα πίνακα αποστάσεων

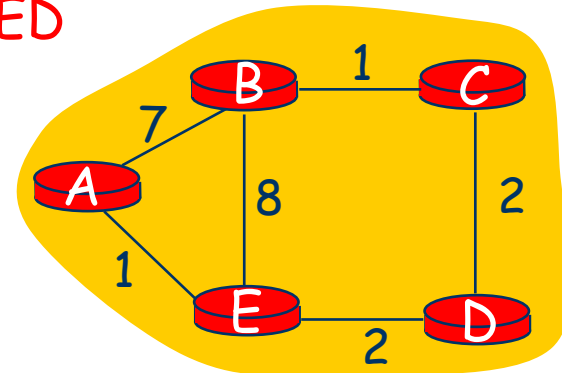


προορισμός	E	απόσταση μέσω		
		A	B	D
	A	1	14	5
	B	7	8	5
	C	6	9	4
	D	4	11	2

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

$(D,E,A)+ED$

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



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

Ο πίνακας αποστάσεων δίνει τον πίνακα δρομολόγησης



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

προορισμός

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

προορισμός

Πίνακας αποστάσεων → Πίνακας δρομολόγησης

Αλγόριθμος διανύσματος αποστάσεων



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

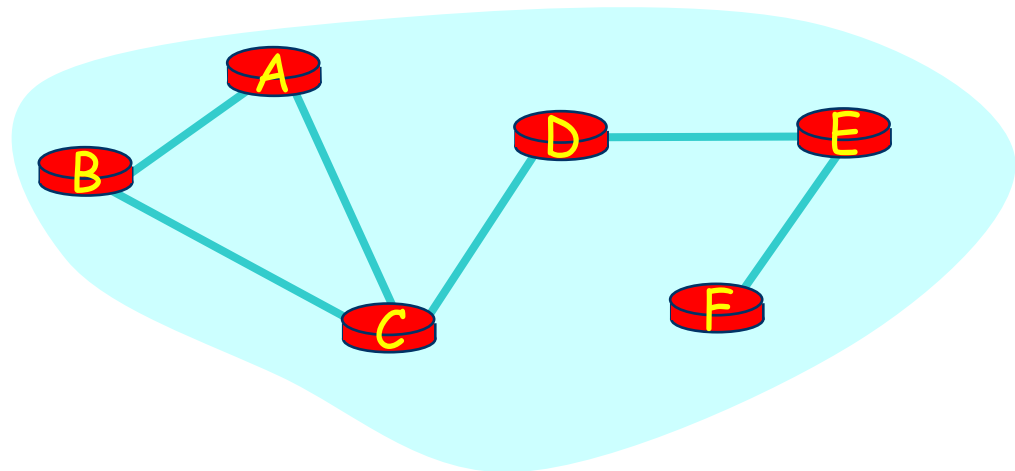
	B	
A	1	A
B		
C	1	C
D	2	C
E	3	C
F	4	C

	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		



Ενημέρωση

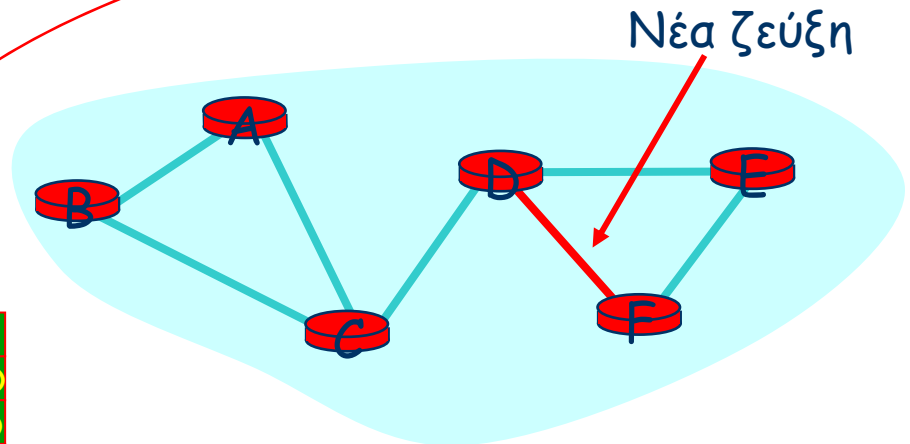


- Οι κόμβοι στέλνουν περιοδικές ενημερώσεις των διανυσμάτων αποστάσεων στους γείτονες
- Κάθε κόμβος ενημερώνει τον πίνακά του υπολογίζοντας τη συντομότερη διαδρομή

A	B	C	D	E	F
A	A1 A	A1 A	A2 C	A3 D	A4 E
B1 B	B	B1 B	B2 C	B3 D	B4 E
C1 C	C1 C	C	C1 C	C2 D	C3 E
D2 C	D2 C	D1 D	D	D1 D	D2 E
E3 C	E3 C	E2 D	E1 E	E	E1 E
F4 C	F4 C	F3 D	F2 E	F1 F	F

Ενημέρωση πίνακα
δρομολόγησης

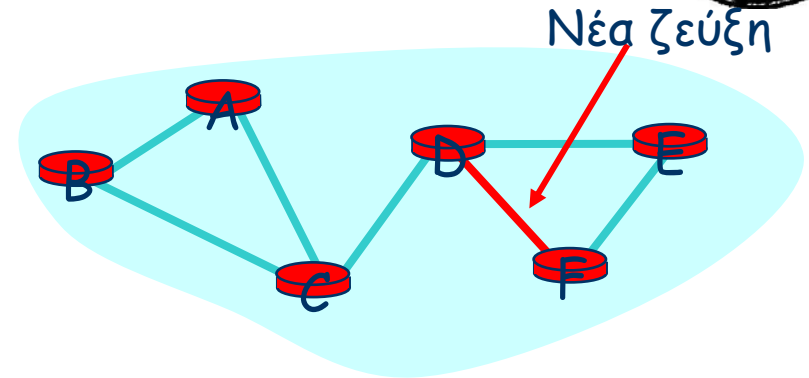
F	D	F
A4 E	A2	A3 D
B4 E	B2	B3 D
C3 E	C1	C2 D
D2 E	D	D1 D
E1 E	E1	E1 E
F	F1	F



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

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

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

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

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

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

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

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

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

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

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

Μέτρηση μέχρι το άπειρο

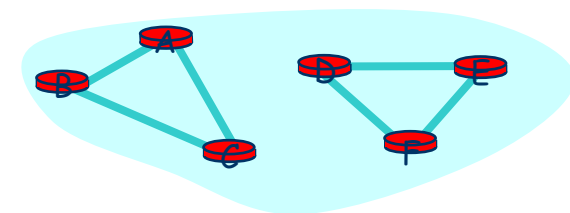


- Χάνεται η ζεύξη c-d
- Διάδοση κακών ειδήσεων
- Οι d και c το μαθαίνουν.
- Ο c λαμβάνει ενημερώσεις από τους a και b
- Ο c στέλνει ενημερώσεις στους a και b
-κ.ο.κ. μέχρι;
- Χαρακτηριστικά
 - Αργή σύγκλιση
 - Όλοι οι κόμβοι μπορεί να αλλάξουν πίνακες
 - Βρόχοι πακέτων, προσωρινές αστάθειες
- Πώς βελτιώνεται αυτό;
 - Τέχνασμα του χωρισμένου ορίζοντα

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

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

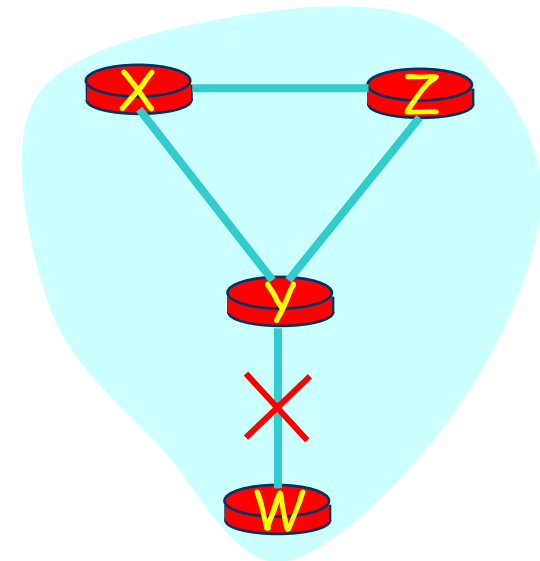
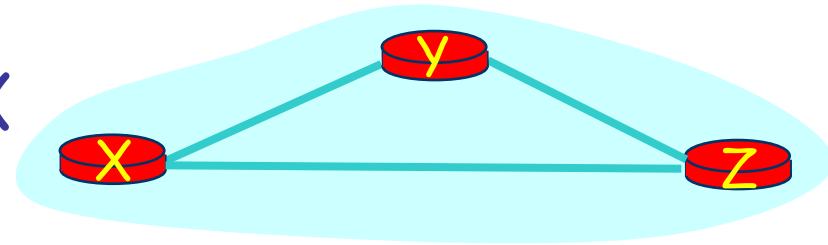
A	B	C
A	A 1 A	A 1 A
B 1 B	B	B 1 B
C 1 C	C 1 C	C
D 4 C	D 4 C	D 3 B
E 5 C	E 5 C	E 4 B
F 5 C	F 5 C	F 4 B



Τέχνασμα Poisoned reverse



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



Αλγόριθμος κατάστασης ζεύξεων



- Δύο ήταν τα βασικά προβλήματα που οδήγησαν στην αντικατάσταση του αλγόριθμου διανύσματος αποστάσεων
 - Δεν έπαιρνε υπόψη τη χωρητικότητα των γραμμών
 - Μερικές φορές ο αλγόριθμος καθυστέρουσε υπερβολικά για να συγκλίνει
- Για τους λόγους αυτούς, αντικαταστάθηκε από έναν καινούργιο αλγόριθμο, τον αλγόριθμο κατάστασης ζεύξεων (link state)

Αλγόριθμος κατάστασης ζεύξεων



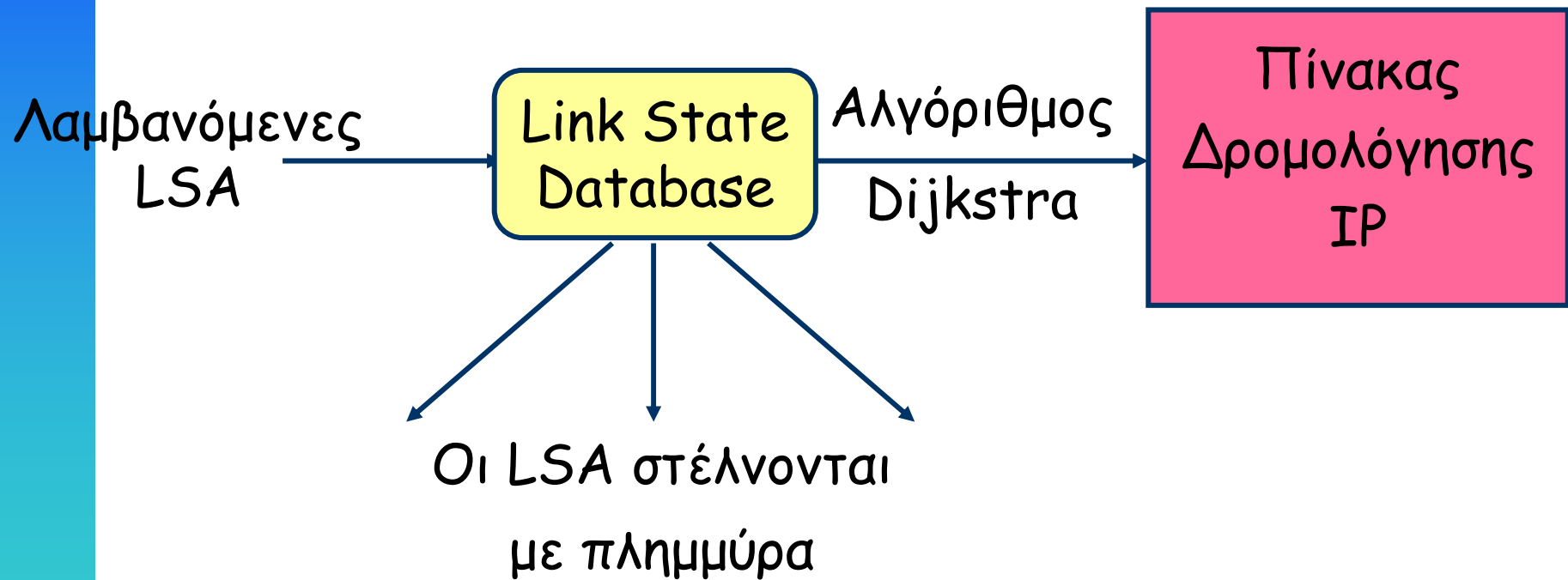
- η τοπολογία δικτύου και τα κόστη των ζεύξεων είναι γνωστά σε όλους τους κόμβους
 - επιτυγχάνεται με **εκπομπή** από κάθε δρομολογητή προς όλους τους δρομολογητές του δικτύου, του κόστους των ζεύξεων που είναι συνδεδεμένες σ' αυτόν
 - όλοι οι κόμβοι έχουν την ίδια πληροφορία
- κάθε κόμβος εκτελεί τον αλγόριθμο και υπολογίζει τις διαδρομές ελαχίστου κόστους προς όλους τους άλλους
 - δίνει τον **πίνακα 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

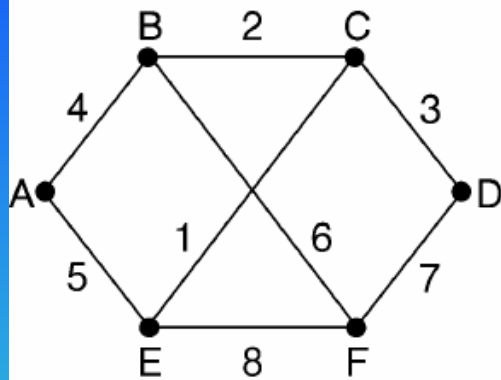


1. Κάθε δρομολογητής αναπτύσσει σχέση με τους γείτονές του
2. Υπολογίζει την καθυστέρηση ή το κόστος για καθέναν από τους γείτονές του
3. Παράγει *αναγγελίες κατάστασης ζεύξεων (LSA)*, οι οποίες διανέμονται σε όλους τους δρομολογητές
 - LSA = (link id, κατάσταση ζεύξης, κόστος, γείτονες)
4. Διατηρεί 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

Λειτουργία του πρωτοκόλλου δρομολόγησης Link State



Κατασκευή πακέτων κατάστασης ζεύξεων



(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)

Διανομή πακέτων κατάστασης ζεύξεων



- 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

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



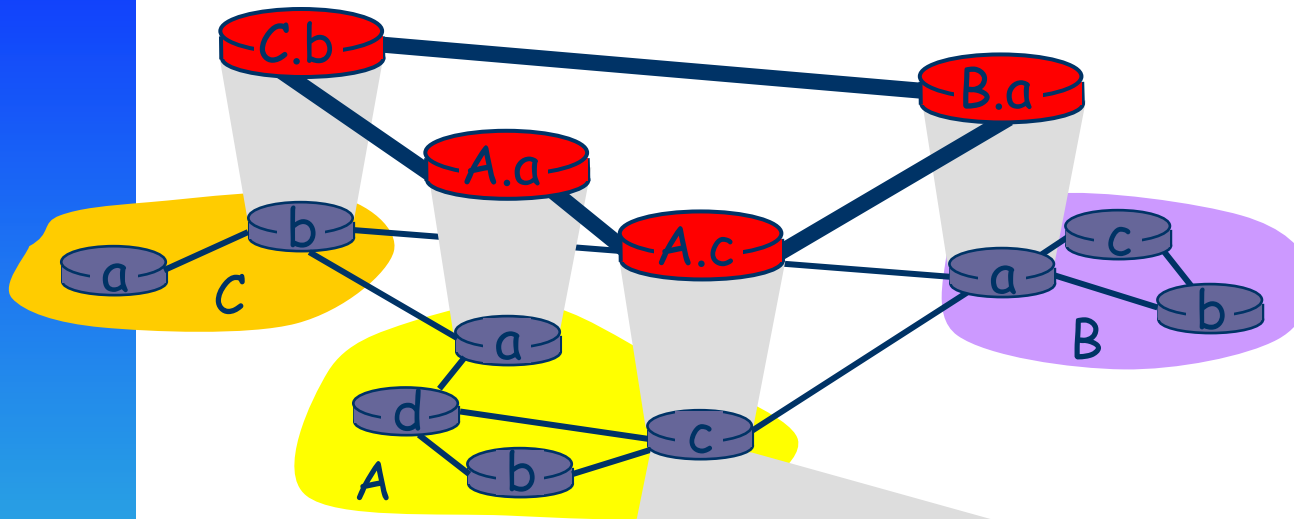
Ιεραρχική δρομολόγηση

- ομαδοποιεί τους δρομολογητές σε περιοχές, **αυτόνομα συστήματα (AS)**
- δρομολογητές στο ίδιο AS τρέχουν το ίδιο πρωτόκολλο δρομολόγησης
 - **"intra-AS"** πρωτόκολλο δρομολόγησης
- δρομολογητές σε διαφορετικά AS μπορεί να τρέχουν διαφορετικό **intra-AS** πρωτόκολλο δρομολόγησης

δρομολογητές πύλες

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

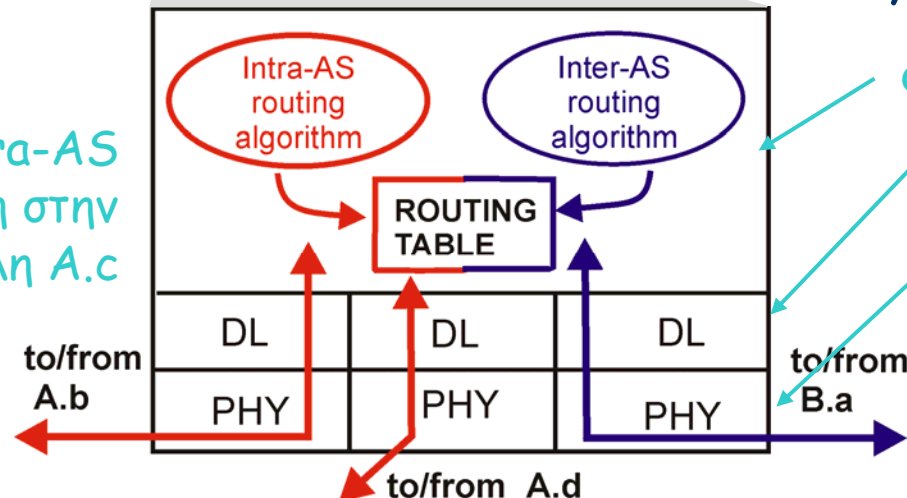
Δρομολόγηση Intra-AS και Inter-AS



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

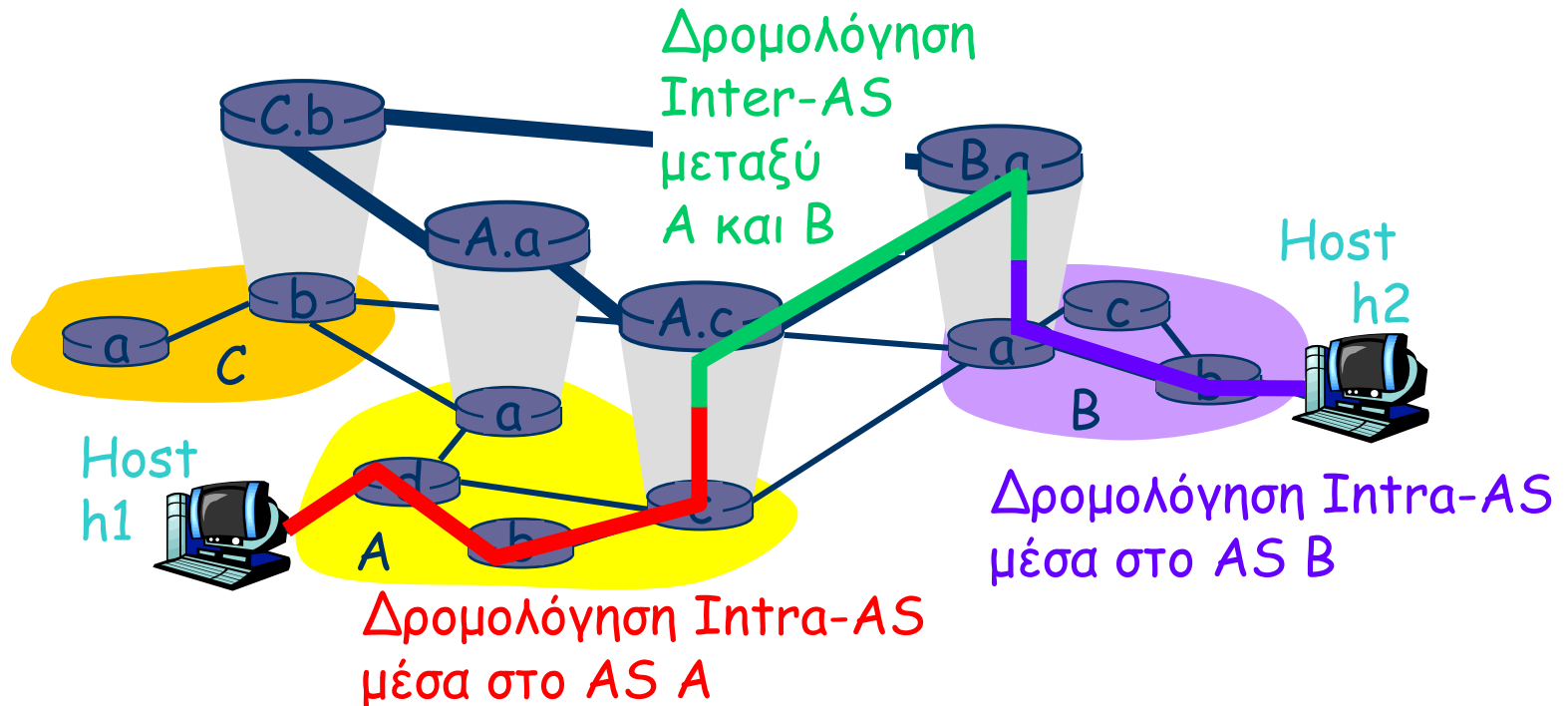


στρώμα δικτύου

στρώμα ζεύξης

φυσικό στρώμα

Δρομολόγηση Intra-AS και Inter-AS

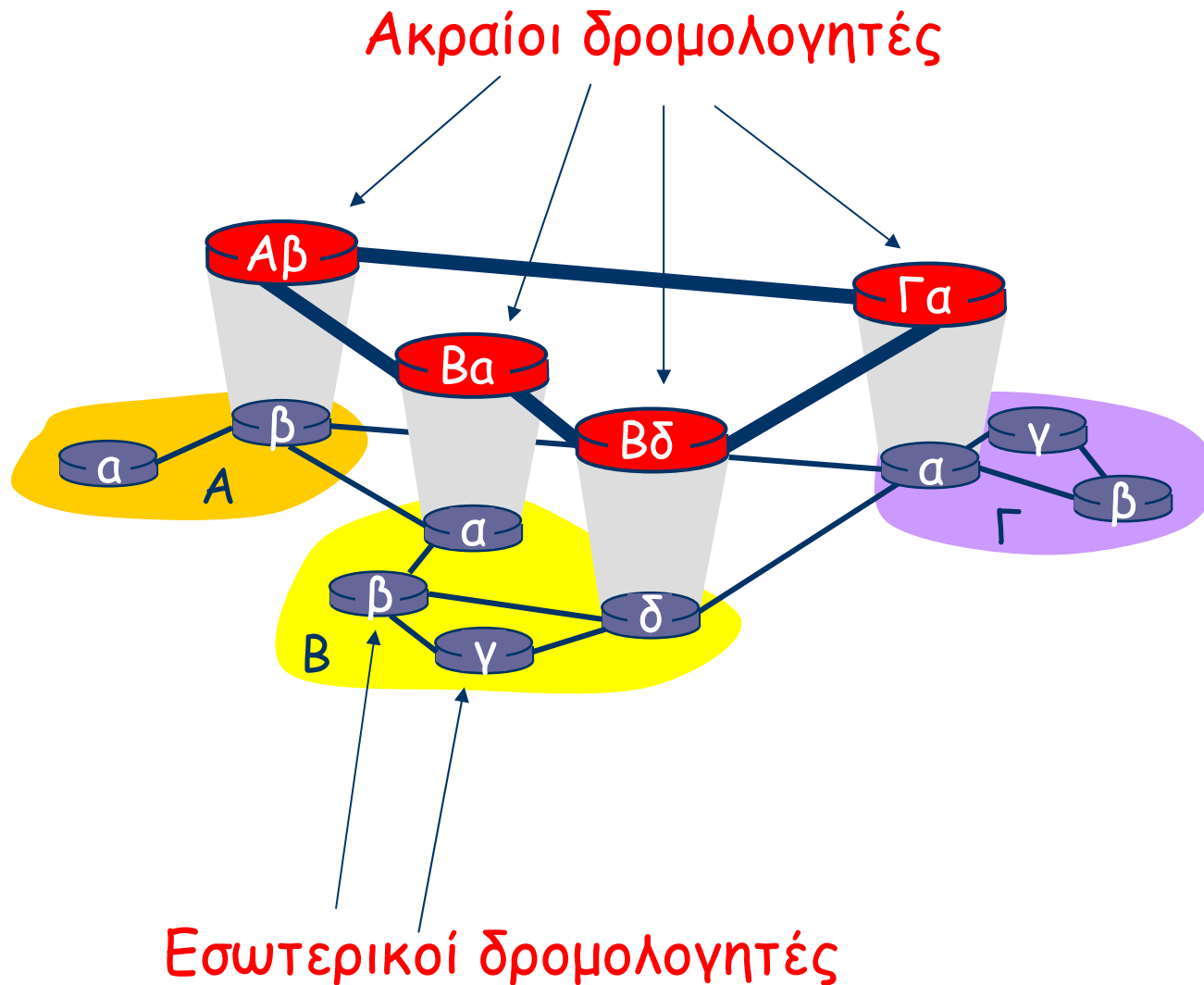




Δρομολόγηση στο Internet

- 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
- Δρομολόγηση δύο επιπέδων:
 - **Intra-AS**: ο διαχειριστής επιλέγει τον αλγόριθμο δρομολόγησης μέσα στο δίκτυο
 - **Inter-AS**: ενιαίο πρότυπο για την δρομολόγηση inter-AS, πρωτόκολλο BGP

Ιεραρχία AS στο Internet



Πρωτόκολλα Δρομολόγησης Intra-AS

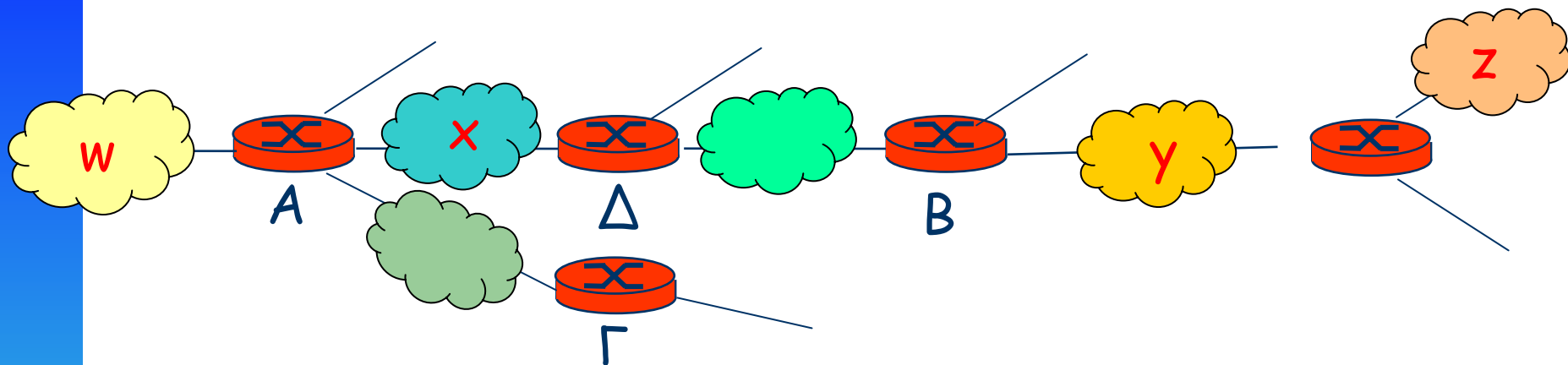


- Γνωστά και ως **Interior Gateway Protocols (IGP)**
- Τα γνωστότερα πρωτόκολλα δρομολόγησης Intra-AS:
 - **RIP**: Routing Information Protocol
 - **OSPF**: Open Shortest Path First
 - **IGRP**: Interior Gateway Routing Protocol (ιδιωτικό της Cisco)



- **RIP: 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 από 25 το πολύ δίκτυα 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 του Δ

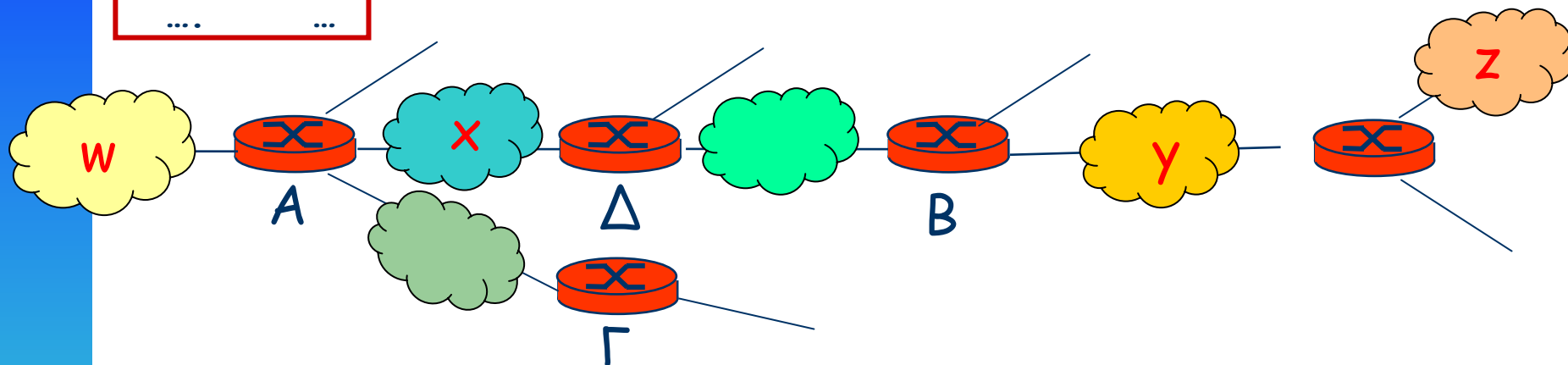
RIP: Παράδειγμα



Προορ. βήμ.

w	-
x	-
z	3
...	...

Διάνυσμα
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 του Δ

Δίκτυα υπολογιστών

RIP: Ανακάλυψη διακοπής ζεύξης

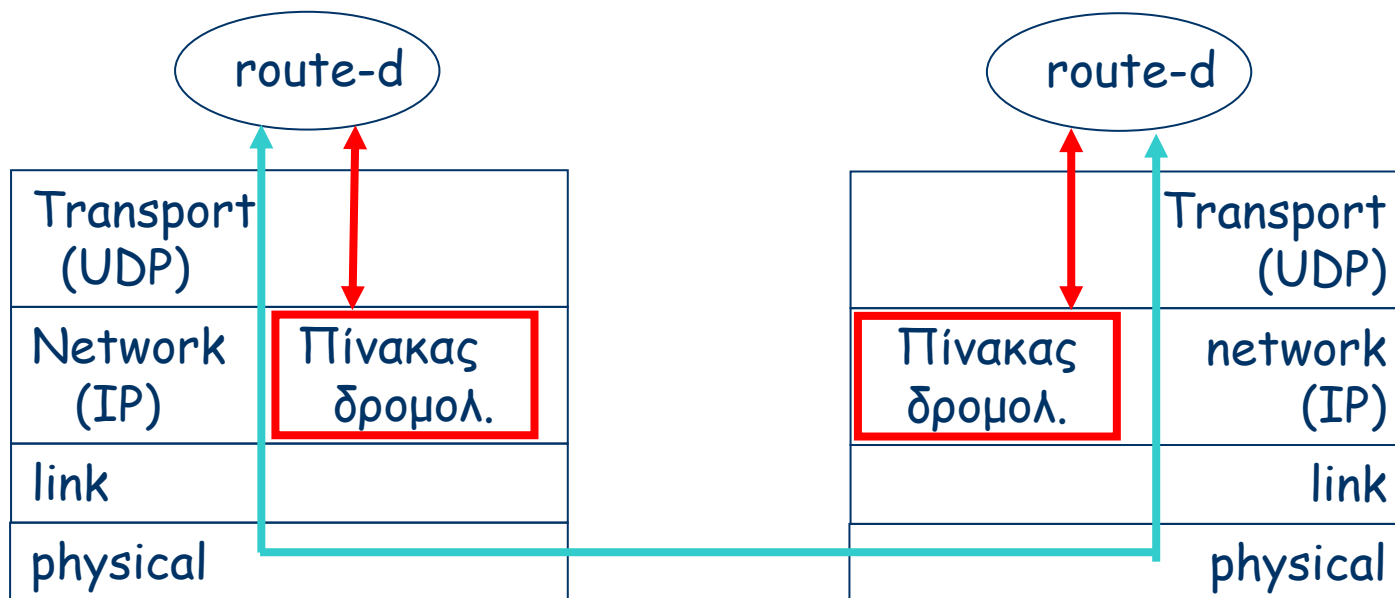


- Αν δεν ακουστεί αγγελία μετά 3 min, η ζεύξη προς τον γείτονα θεωρείται νεκρή
 - οι διαδρομές μέσω του γείτονα ακυρώνονται
 - στέλνονται νέες αγγελίες προς τους γείτονες
 - και οι γείτονες με τη σειρά τους στέλνουν νέες αγγελίες (αν αλλάξουν οι πίνακες)
 - η πληροφορία για τη διακοπή ζεύξης διαδίδεται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

Διαχείριση πινάκων στο RIP



- 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: *girofflee.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



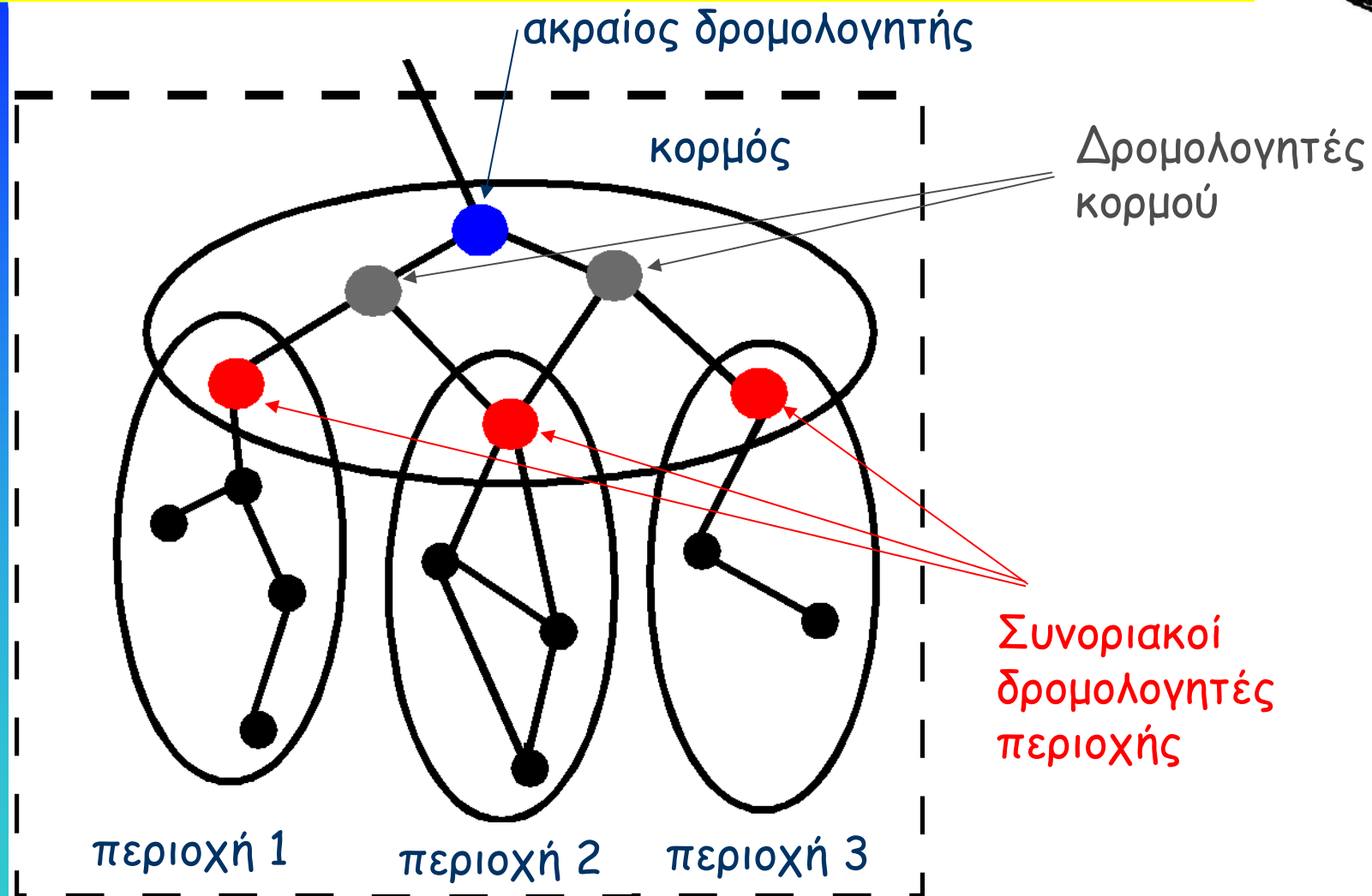
- **OSPF: Open Shortest Path First**
- "open": δημόσια διαθέσιμο
- Χρησιμοποιεί αλγόριθμο κατάστασης ζεύξεων
 - διασπορά πακέτων κατάστασης ζεύξεων
 - τοπολογία σε κάθε κόμβο
 - υπολογισμός διαδρομής με αλγόριθμο Dijkstra
- Η αναγγελία OSPF έχει μια εγγραφή για κάθε γειτονικό δρομολογητή
- Οι αγγελίες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

Εξελιγμένες ιδιότητες του OSPF



- 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 τύπους υπηρεσιών (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

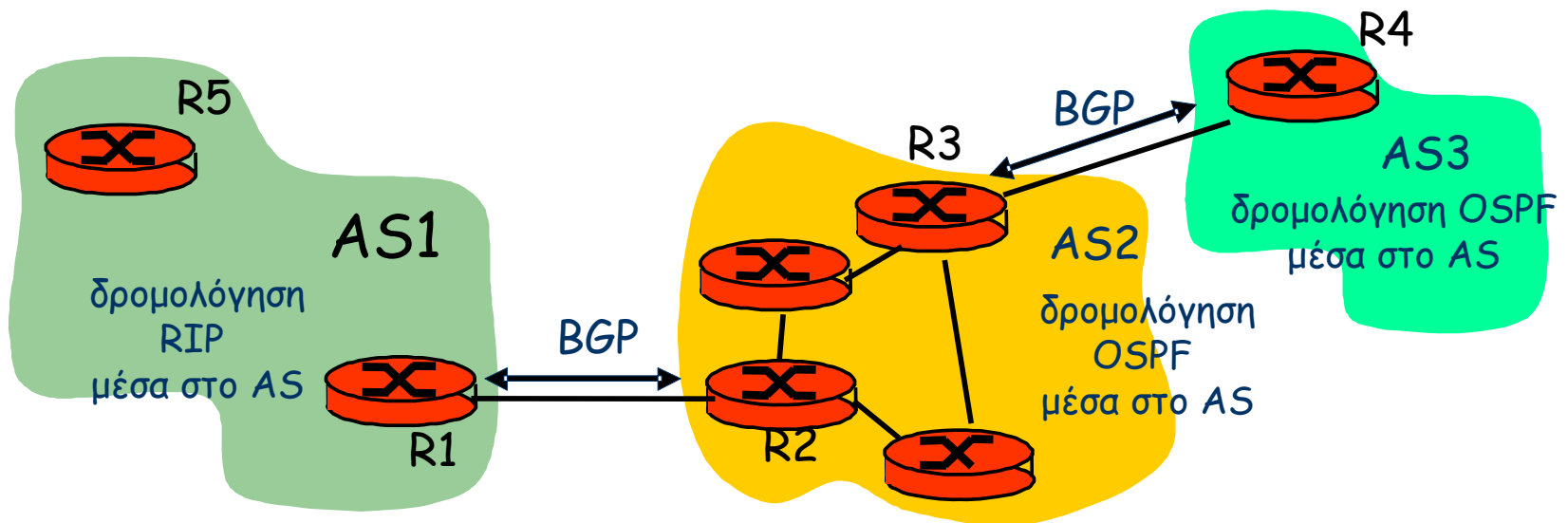


Ιεραρχικό OSPF



- **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, και τις αναγγέλλουν προς άλλους Area Border routers.
- **Δρομολογητές κορμού (Backbone routers):** τρέχουν OSPF δρομολόγηση περιοριζόμενη στο backbone.
- **Ακραίοι δρομολογητές (Boundary routers):** συνδέουν με άλλα AS.

BGP: δρομολόγηση μεταξύ AS



BGP: Δρομολόγηση μεταξύ AS

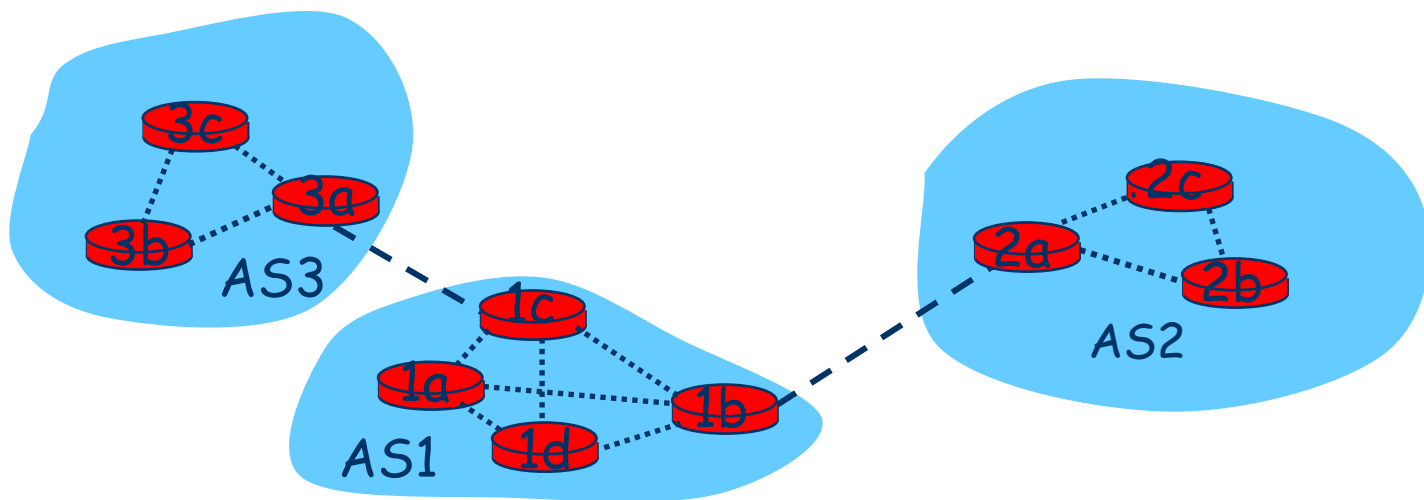


- **BGP (Border Gateway Protocol):** είναι το πρότυπο που ισχύει
- Το BGP παρέχει σε κάθε AS τα μέσα για:
 - να αποκτήσει πληροφορίες προσβασιμότητας από τα γειτονικά AS
 - να διαδίδει τη πληροφορία προσβασιμότητας σε όλους τους εσωτερικούς δρομολογητές του
 - να καθορίζει "καλές" διαδρομές προς τα άλλα υποδίκτυα βάσει της πληροφορίας προσβασιμότητας και της πολιτικής δρομολόγησης
- Επιτρέπει σε ένα υποδίκτυο 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 οποιαδήποτε πακέτα που θα απευθύνονται στο πρόθεμα αυτό προς το πρόθεμα αυτό.
 - Το AS2 μπορεί να συνοψίζει προθέματα στην αναγγελία του



----- σύνοδος eBGP

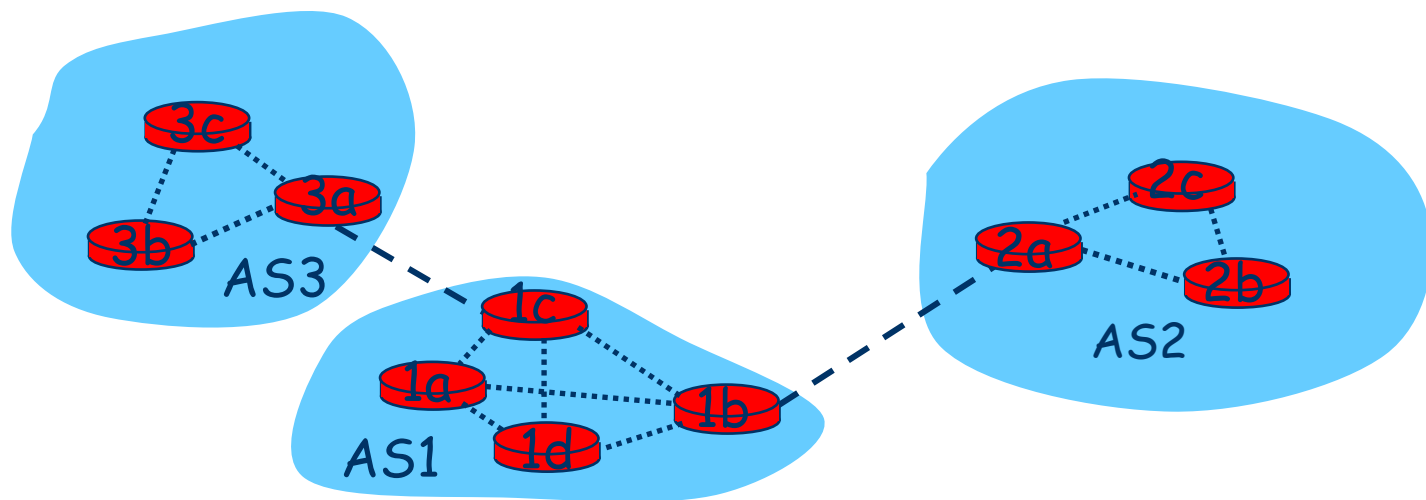
..... σύνοδος iBGP

Δίκτυα υπολογιστών

Διανομή πληροφορίας προσβασιμότητας



- Στη 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.



----- σύνοδος eBGP

..... σύνοδος iBGP

Δίκτυα υπολογιστών

Διαδρομές BGP



- Όταν αναγγέλλεται ένα πρόθεμα, η αναγγελία περιέχει και χαρακτηριστικά BGP.
 - prefix + attributes = "route"
- Δύο ενδιαφέροντα χαρακτηριστικά:
 - AS-PATH: περιέχει τα AS μέσω των οποίων πέρασε η αναγγελία για το υπόψη πρόθεμα: AS67 AS 17
 - NEXT-HOP: καθορίζει τον συγκεκριμένο εσωτερικό δρομολογητή του AS προς το next-hop AS. (Μπορεί να μεσολαβούν πολλές ζεύξεις από το τρέχον AS μέχρι το next-hop-AS.)
- Όταν ένας δρομολογητής πύλη λαμβάνει αναγγελία διαδρομής, χρησιμοποιεί πολιτική εισαγωγής για να τη δεχθεί ή να την απορρίψει.

BGP επιλογή διαδρομής



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

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

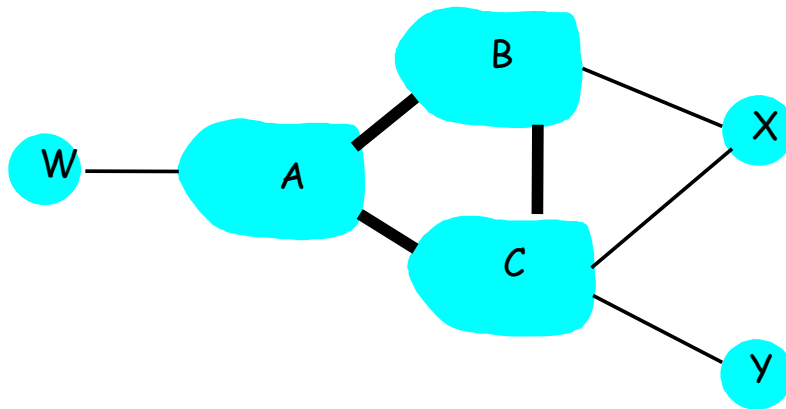


- Έστω ότι η πύλη X στέλνει προς την ομότιμη πύλη W τη διαδρομή της προς Z
- η W μπορεί και να μην επιλέξει τη διαδρομή που προσφέρεται από τη X
 - κόστος, πολιτική (ανταγωνιστικό AS), αποφυγή βρόχου.
- αν η W επιλέξει τη διαδρομή που αναγγέλλεται από την X , τότε:

$$\text{Path}(W,Z) = \text{Path}(W,X), \text{Path}(X,Z)$$

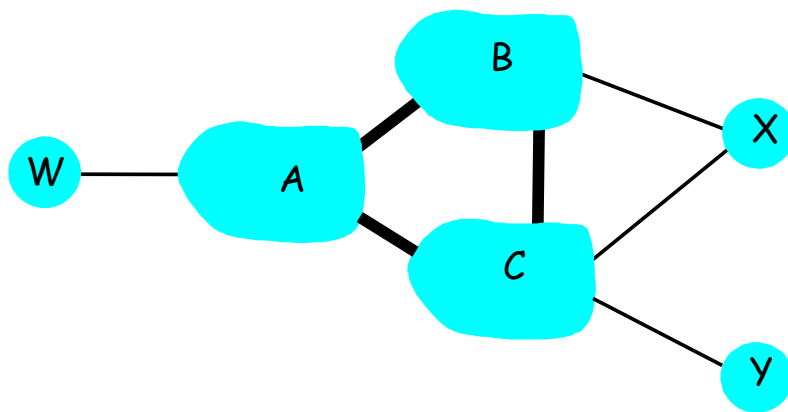
- η X μπορεί να ελέγχει την εισερχόμενη κίνηση ελέγχοντας τις αναγγελίες προς τις ομότιμες πύλες:
 - π.χ., αν δεν θέλει δρομολόγηση προς τη Z δεν θα αναγγέλλει διαδρομές προς τη Z

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
 $\Rightarrow$ 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!



Ένας δρομολογητής BGP:

- Λαμβάνει και φιλτράρει αγγελίες διαδρομών από άμεσα συνδεόμενους 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:

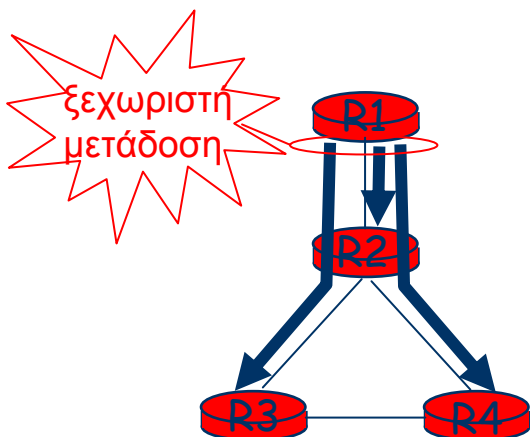
- Η ιεραρχική δρομολόγηση εξοικονομεί μέγεθος πινάκων, περιορίζει το φορτίο ενημέρωσης

Επίδοση:

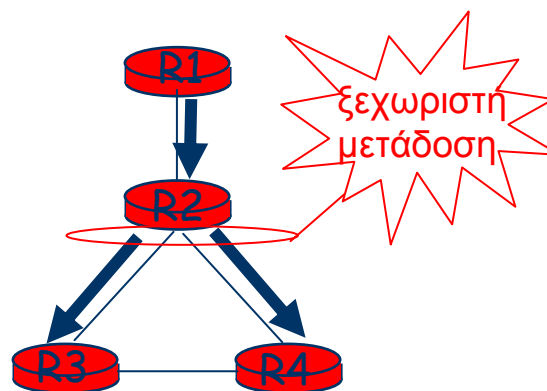
- Intra-AS: μπορεί να εστιάσει στην επίδοση
- Inter-AS: η πολιτική μπορεί να επισκιάζει την επίδοση



- Παράδοση των πακέτων από την πηγή προς όλους τους άλλους κόμβους
- Οι ξεχωριστές μεταδόσεις από την πηγή είναι αναποτελεσματικές



Ξεχωριστές μεταδόσεις
από την πηγή



Ξεχωριστές μεταδόσεις
στο δίκτυο

- Ξεχωριστές μεταδόσεις στο δίκτυο: πώς κάθε κόμβος καθορίζει τις διευθύνσεις των ληπτών;

Ξεχωριστές μεταδόσεις στο δίκτυο

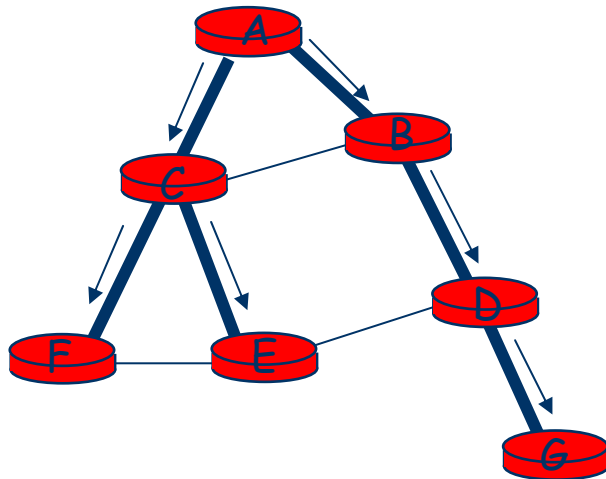


- **Πλημμύρα**: όταν ένας κόμβος λαμβάνει ένα πακέτο εκπομπής, στέλνει αντίγραφο σε όλους τους γείτονες
 - Προβλήματα: κύκλοι και πληθώρα πακέτων
- **Ελεγχόμενη πλημμύρα**: ο κόμβος εκπέμπει το πακέτο μόνο αν δεν έχει εκπέμψει το ίδιο πακέτο προηγουμένως
 - Ο κόμβος παρακολουθεί τις ταυτότητες των πακέτων που ήδη εξέπεμψε
 - Προώθηση ανάστροφης διαδρομής (reverse path forwarding, RPF): προωθεί μόνο τα πακέτα που καταφθάνουν από την συντομότερη διαδρομή μεταξύ πηγής και προορισμού
- **Επικαλύπτον δένδρο**
 - Δεν λαμβάνονται πλεονάζοντα πακέτα από οιονδήποτε κόμβο

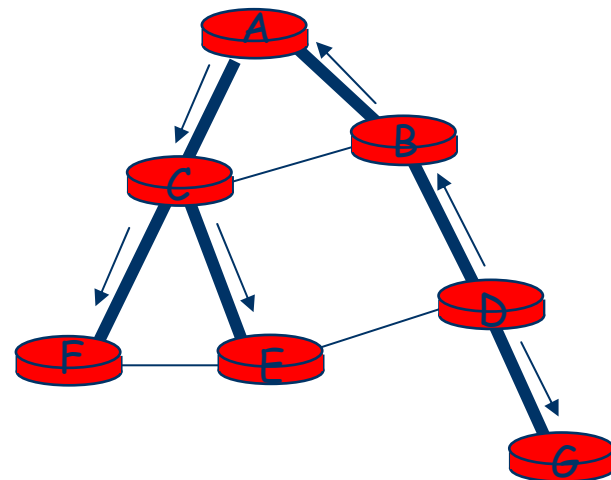
Επικαλύπτον δένδρο



- Αρχικά, κατασκευάζεται το επικαλύπτον δένδρο
- Οι κόμβοι προωθούν αντίγραφα του πακέτου μόνο κατά μήκος του επικαλύπτοντος δέντρου



(α) Η εκπομπή αρχίζει στον A

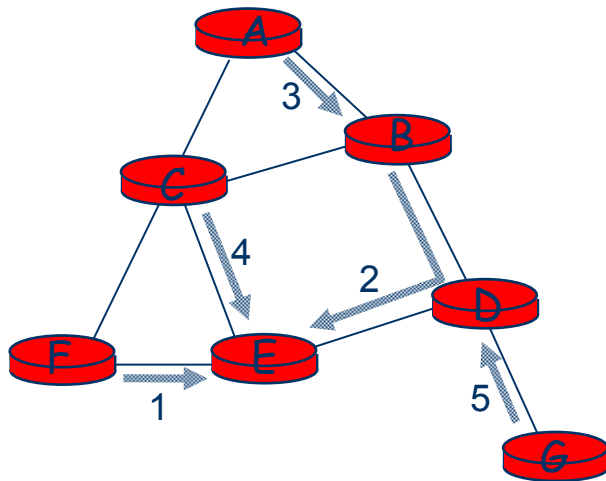


(β) Η εκπομπή αρχίζει στον D

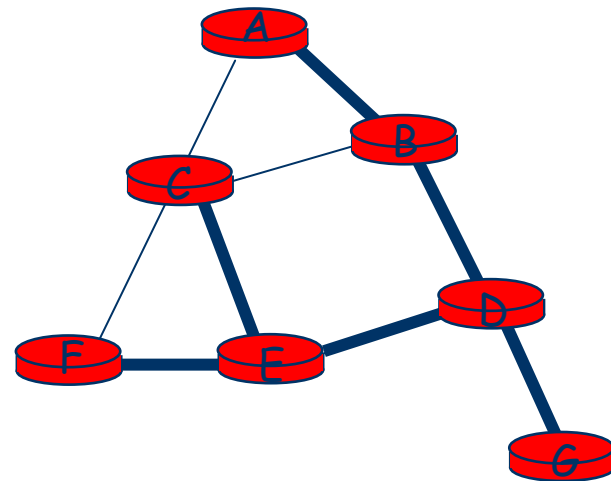
Κατασκευή επικαλύπτοντος δέντρου



- Κεντρικός κόμβος
- Κάθε κόμβος στέλνει unicast μήνυμα join στον κεντρικό κόμβο
 - Το μήνυμα προωθείται μέχρι να φθάσει σε κόμβο που ανήκει ήδη στο επικαλύπτον δέντρο



(α) Σταδιακή κατασκευή του επικαλύπτοντος δέντρου



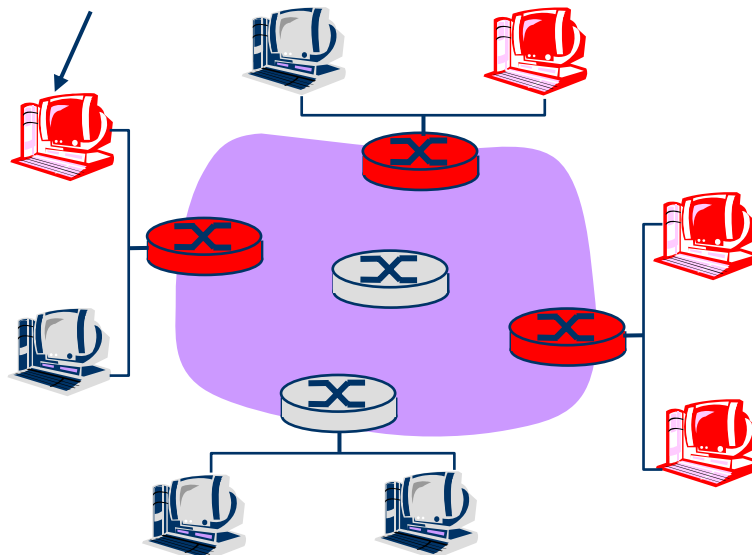
(β) Επικαλύπτον δέντρο

Πολλαπλή διανομή



- **Πολλαπλή διανομή (multicast):** αποστολή δεδομενογραμμάτων σε πολλούς δέκτες με **μία μόνο** διαδικασία μετάδοσης
 - αναλογία: ένας καθηγητής σε πολλούς σπουδαστές
- **Πώς επιτυγχάνεται;**

πομπός
πολλαπλής διανομής



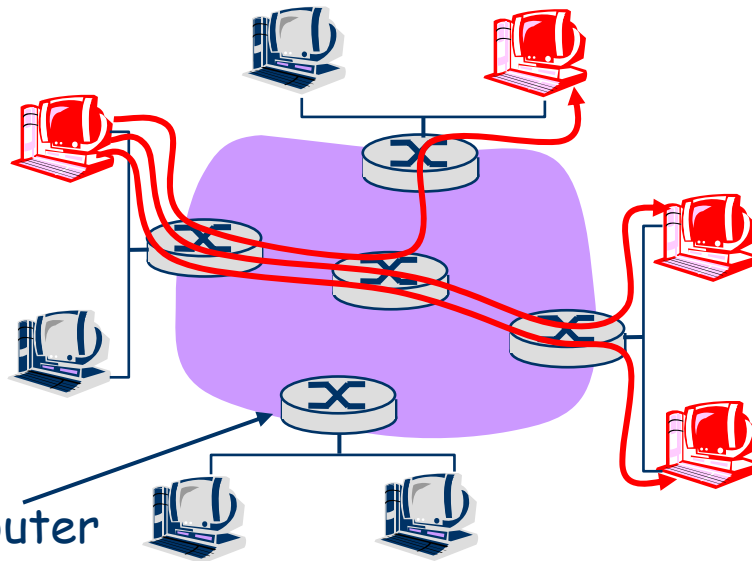
δέκτης
πολλαπλής διανομής

Πολλαπλή διανομή



Πολλαπλή διανομή μέσω απλών μεταδόσεων

- ο πομπός στέλνει N δεδομενογράμματα unicast, ένα για καθένα από τους N δέκτες



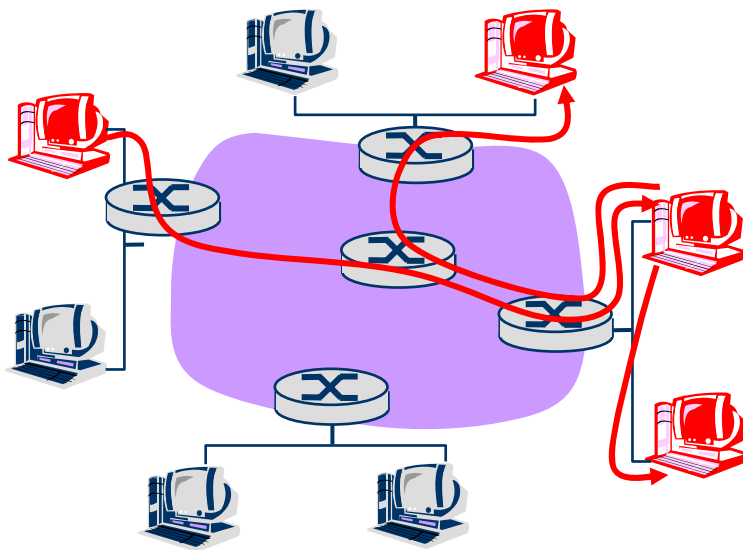
οι router
προωθούν με απλή μετάδοση
τα δεδομενογράμματα

δέκτης πολλαπλής διανομής

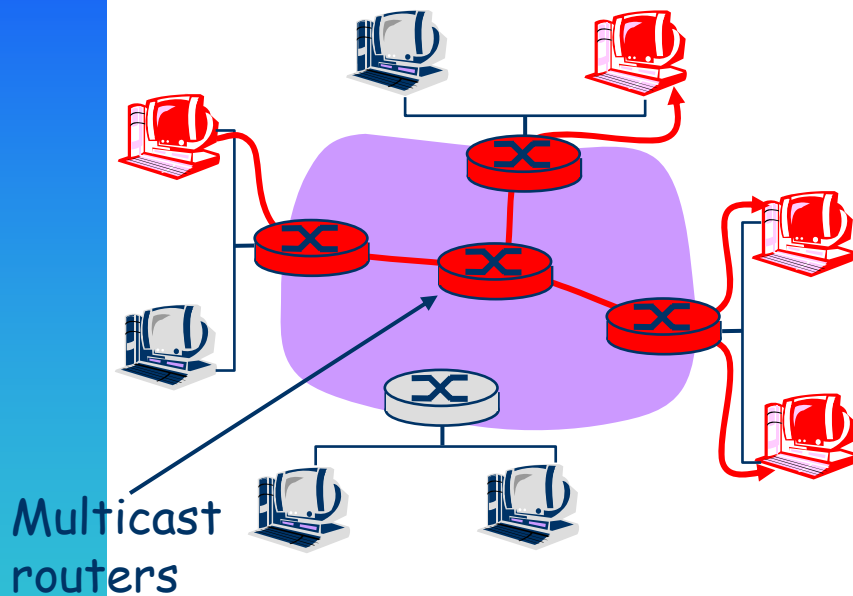


Multicast στρώματος εφαρμογής

- τα ακραία συστήματα που εμπλέκονται στην πολλαπλή διανομή αντιγράφουν και προωθούν μεταξύ τους δεδομενογράμματα unicast



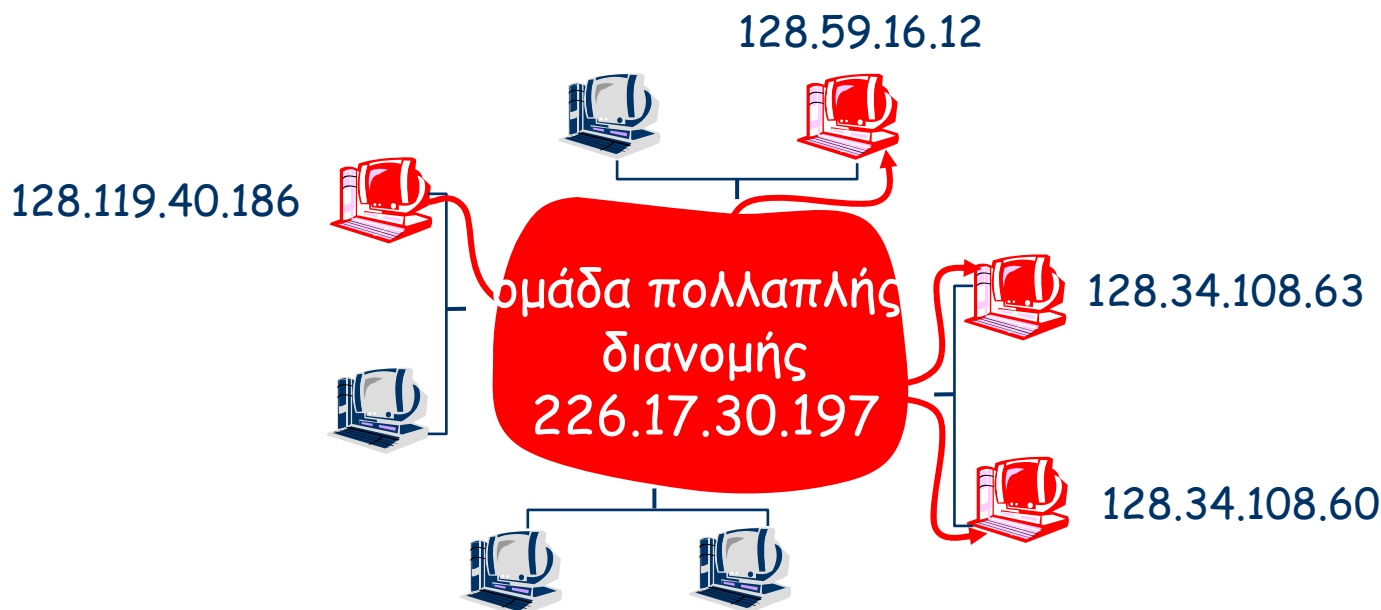
Πολλαπλή διανομή



Πολλαπλή διανομή δικτύου

- Οι δρομολογητές μετέχουν ενεργά στην πολλαπλή διανομή, κάνοντας αντίγραφα των πακέτων και προωθώντας τα στους δέκτες multicast

Πολλαπλή διανομή στο Internet



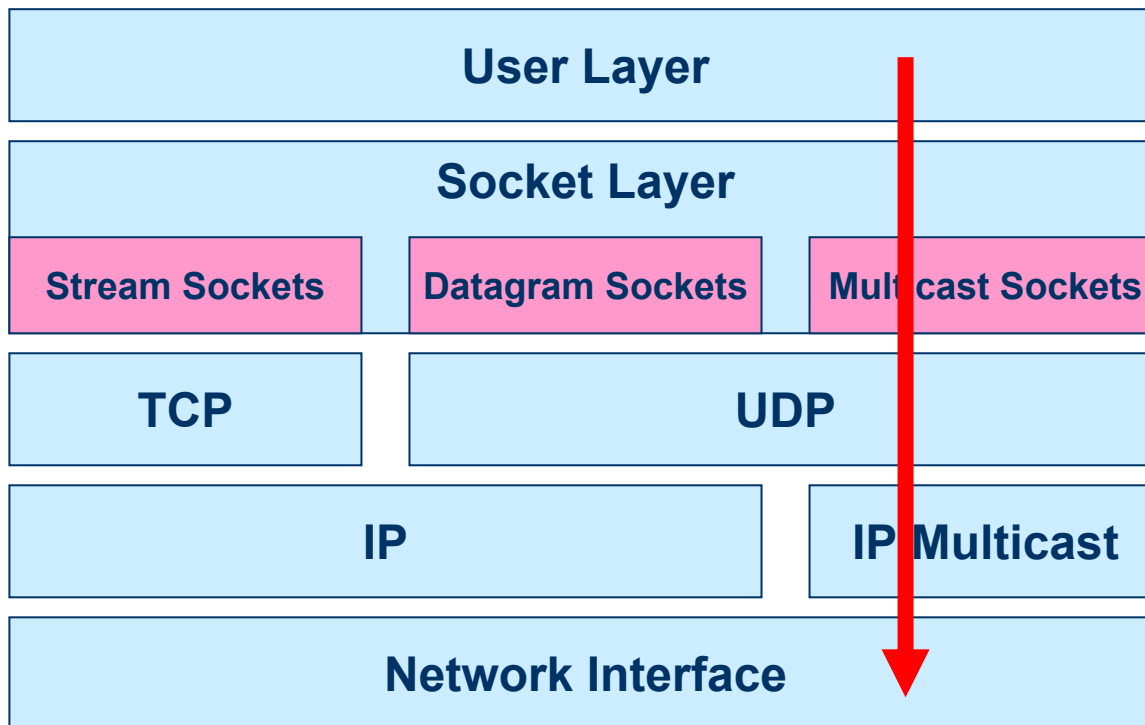
ομάδα πολλαπλής διανομής: χρήση έμμεσης λύσης

- οι host στέλνουν τα δεδομενογράμματα IP στην ομάδα πολλαπλής διανομής
- οι δρομολογητές προωθούν τα δεδομενογράμματα πολλαπλής διανομής στους host που έχουν "ενταχθεί" στην ομάδα πολλαπλής διανομής
- η πολλαπλή διανομή IP υποστηρίζει μόνο UDP

Πολλαπλή διανομή στο Internet



- Η πολλαπλή διανομή στο IP υποστηρίζει μόνο UDP
- Δεν υπάρχει πολλαπλή διανομή TCP !



Πολλαπλή διανομή στο Internet



- Τρεις είναι οι κύριες συνιστώσες της υπηρεσίας πολλαπλής διανομής στο Internet:

- Διευθυνσιοδότηση
- Διαχείριση ομάδων πολλαπλής διανομής
- Δρομολόγηση πολλαπλής διανομής

Διευθύνσεις πολλαπλής διανομής



- Διευθύνσεις Internet class D έχουν κρατηθεί για πολλαπλή διανομή:



← 28 bits →

- Για τους host:
 - οποιοσδήποτε μπορεί να "ενταχθεί" σε ομάδα πολλαπλής διανομής
 - οποιοσδήποτε μπορεί να στείλει σε ομάδα multicast
 - δεν υπάρχει αναγνωριστικό στρώματος δικτύου στους host των μελών
- απαίτηση: ύπαρξη υποδομής για την παράδοση των δεδομενογραμμάτων πολλαπλής διανομής σε όλους τους host που εντάχθηκαν στην υπόψη ομάδα πολλαπλής διανομής

Διευθύνσεις πολλαπλής διανομής



Όλες οι διευθύνσεις Class D είναι διευθύνσεις multicast:

Class D	1	1	1	0	multicast group id
					28 bit

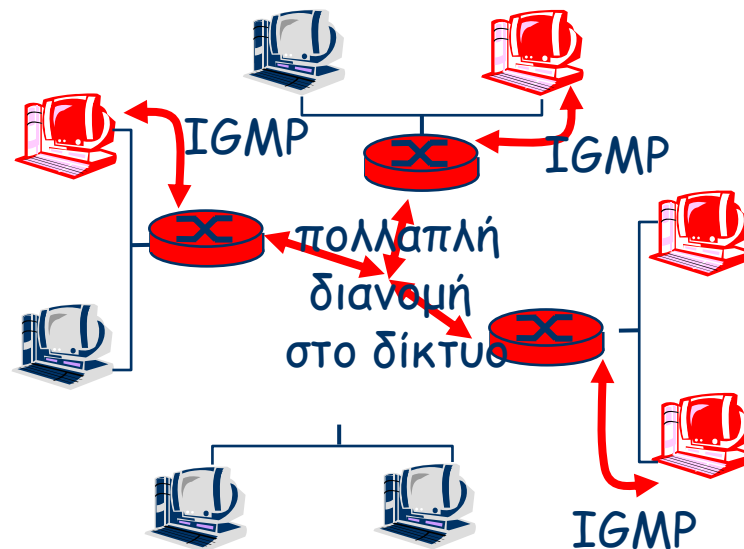
Class	Από	Έως
D	224 .0.0.0	239 .255.255.255

- Οι διευθύνσεις πολλαπλής διανομής εκχωρούνται δυναμικά
- Δεδομένογραμμα IP που στέλνεται σε διεύθυνση multicast προωθείται σε όλους που έχουν ενταχθεί στην ομάδα multicast
- Αν τελειώσει κάποια εφαρμογή, η διεύθυνση multicast (έμμεσα) ελευθερώνεται.

Ένταξη σε ομάδα πολλαπλής διανομής



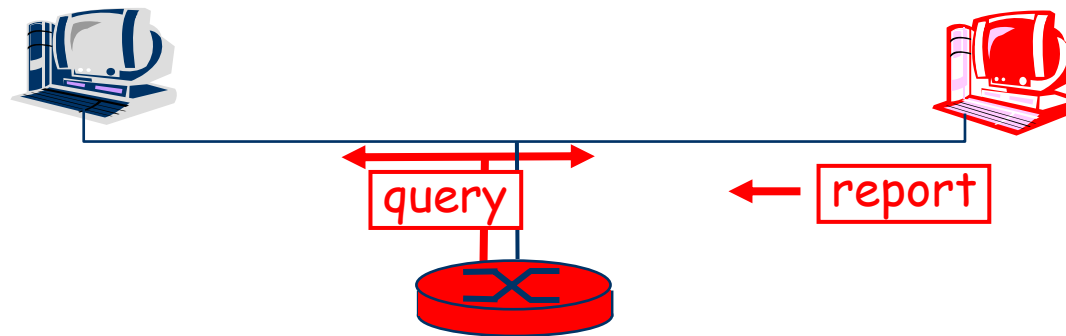
- τοπικά: ο host πληροφορεί τον τοπικό δρομολογητή πολλαπλής διανομής ότι επιθυμεί να ενταχθεί στην ομάδα. IGMP (Internet Group Management Protocol)
- στο δίκτυο: ο τοπικός δρομολογητής επικοινωνεί με άλλους δρομολογητές για να λάβει τη ροή των δεδομενογραμμάτων πολλαπλής διανομής
 - ειδικά πρωτόκολλα (π.χ. DVMRP, MOSPF, PIM)



IGMP: Internet Group Management Protocol



- host: στέλνει αναφορά IGMP, όταν κάποια εφαρμογή εντάσσεται σε ομάδα πολλαπλής διανομής
 - Membership Report
 - ο host δεν είναι ανάγκη να "διαγραφεί" άμεσα όταν εγκαταλείπει την ομάδα
- router: στέλνει IGMP query σε τακτά διαστήματα
 - ο host που ανήκει σε κάποια ομάδα multicast πρέπει να απαντήσει στην ερώτηση





IGMP version 1

- router: μήνυμα Host Membership Query εκπέμπεται στο LAN προς όλους τους host
- host: μήνυμα Host Membership Report για να δηλώσει συμμετοχή
 - τυχαία καθυστέρηση πριν την απάντηση
 - έμμεση αποχώρηση μη απαντώντας στο Query
- RFC 1112

IGMP v2: οι προσθήκες περιλαμβάνουν

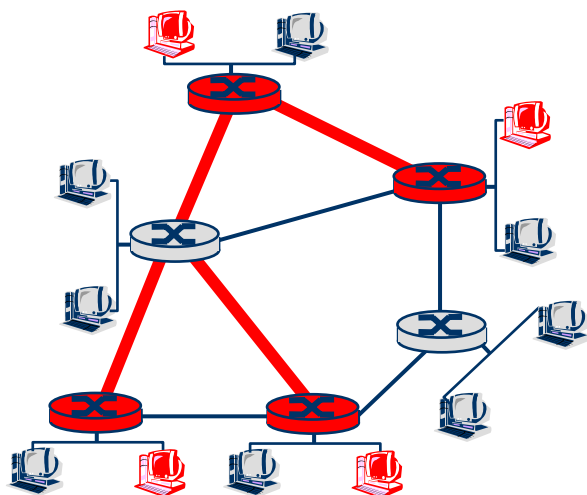
- group-specific Query
- μήνυμα Leave Group
 - ο τελευταίος host που απαντάει στο Query μπορεί να στείλει άμεσο μήνυμα Leave Group
 - ο router πραγματοποιεί group-specific query για να δει αν κάποιοι host έμειναν στην ομάδα
 - RFC 2236

IGMP v3: σε εξέλιξη ως Internet draft

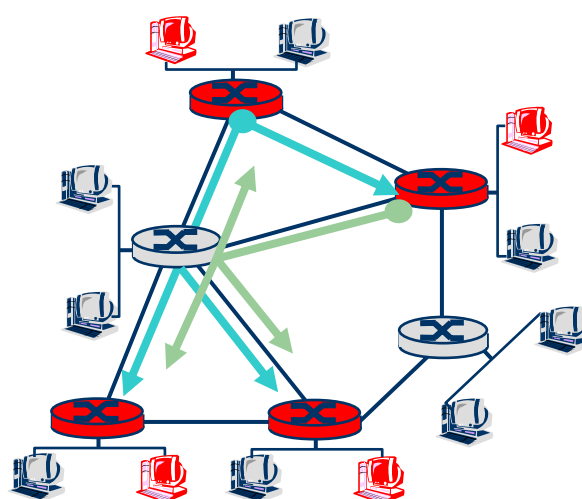
Δρομολόγηση πολλαπλής διανομής



- **Στόχος:** εύρεση δέντρου (ή δέντρων) πολλαπλής διανομής που συνδέει δρομολογητές που έχουν τοπικά μέλη πολλαπλής διανομής
 - **δέντρο:** δεν χρησιμοποιούνται όλες οι διαδρομές μεταξύ των δρομολογητών
 - **κοινό δέντρο:** ίδιο δέντρο για όλα τα μέλη της ομάδας
 - **βασισμένο στην πηγή:** διαφορετικό δέντρο για κάθε πηγή προς τους δέκτες



Κοινό δέντρο



Δέντρο βασισμένο στη πηγή

Δέντρα πολλαπλής διαδρομής



- **κοινό δέντρο:**

- ελάχιστο επικαλύπτον δέντρο

- χρειάζεται πληροφορία για όλο το δίκτυο
 - υπολογιστική πολυπλοκότητα
 - μονολιθικό: επαναυπολογισμός, όταν ένας δρομολογητής εισέρχεται ή εγκαταλείπει

- κεντρικά δέντρα

- **δέντρο βασισμένο στην πηγή:**

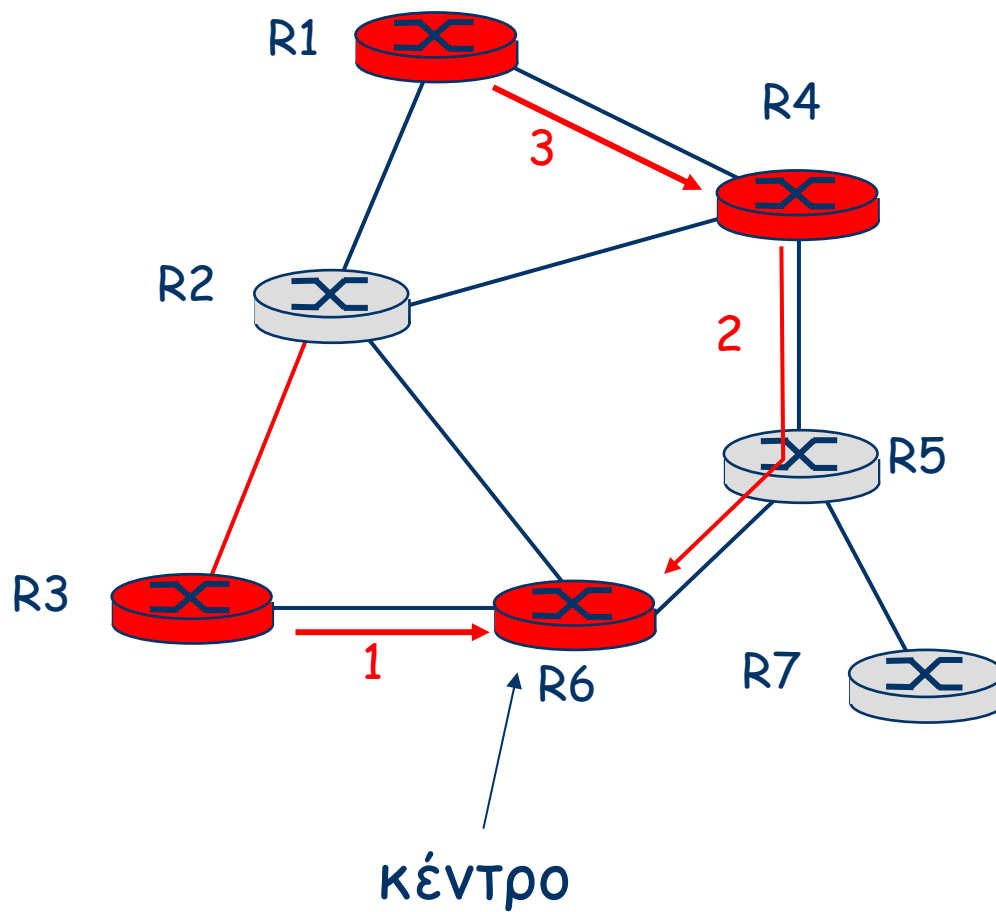
- δέντρα συντομότερης διαδρομής (αλγ. Dijkstra)
 - προώθηση ανάστροφης διαδρομής

Κεντρικά δέντρα



- κοινό δέντρο διανομής για όλους
- ένας δρομολογητής ορίζεται ως "κέντρο" του δέντρου
- για σύνδεση:
 - ο ακραίος δρομολογητής στέλνει μήνυμα *join* προς τον κεντρικό δρομολογητή
 - το μήνυμα *join* εξετάζεται από τους ενδιάμεσους δρομολογητές και προωθείται προς τον κεντρικό
 - το μήνυμα *join* είτε διέρχεται από υπάρχοντα κλάδο του κέντρου ή φθάνει στο κέντρο
 - η διαδρομή που κάνει το μήνυμα *join* αποτελεί νέο κλάδο του δέντρου για τον υπόψη δρομολογητή

Κεντρικά δέντρα

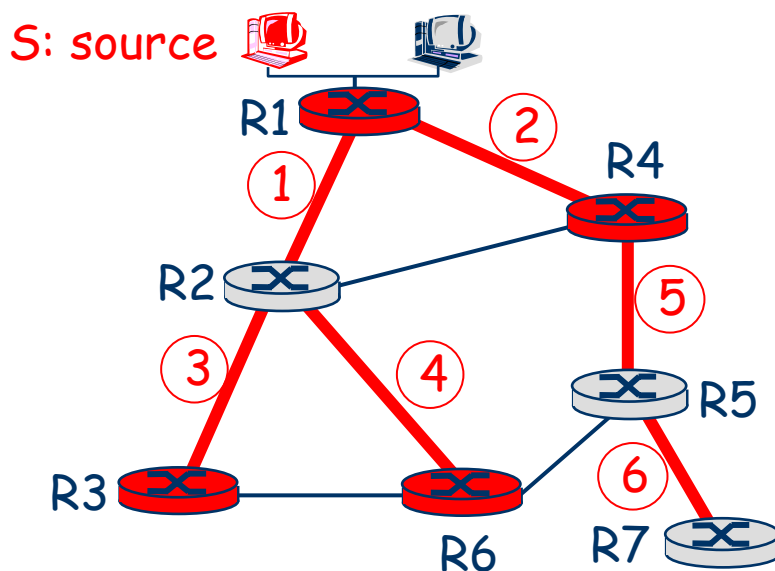


Δέντρα βασισμένα στην πηγή



Δέντρο συντομότερων διαδρομών

- δέντρο προώθησης πολλαπλής διανομής: δέντρο των συντομότερων διαδρομών από την πηγή προς όλους τους δέκτες
 - αλγόριθμος Dijkstra



Δέντρα βασισμένα στην πηγή



Πρώθηση ανάστροφης διαδρομής

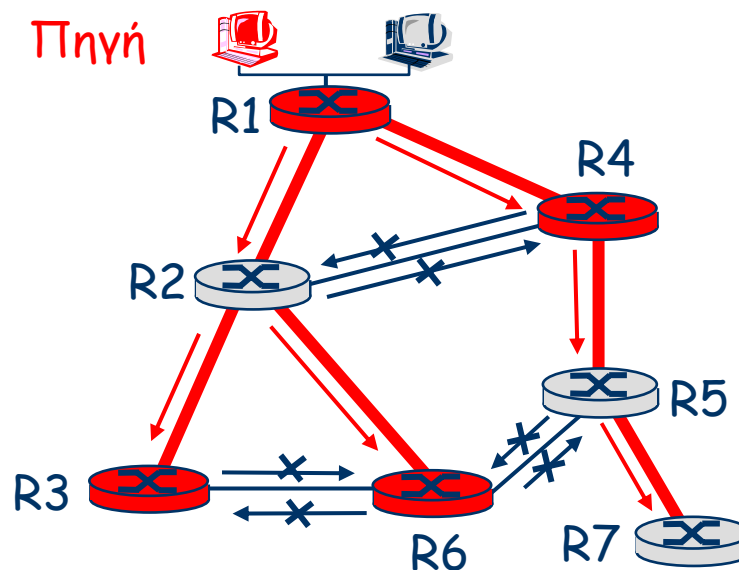
- ❑ βασίζεται στο ότι ο δρομολογητής γνωρίζει τη συντομότερη διαδρομή από τον αποστολέα
- ❑ κάθε δρομολογητής έχει απλή διαδικασία προώθησης:

if το δεδομένογραμμα πολλαπλής διανομής ελήφθη από την συντομότερη διαδρομή από τον αποστολέα)

then στείλ' το σε όλες τις απερχόμενες ζεύξεις

else αγνόησέ το

Πρώθηση ανάστροφης διαδρομής

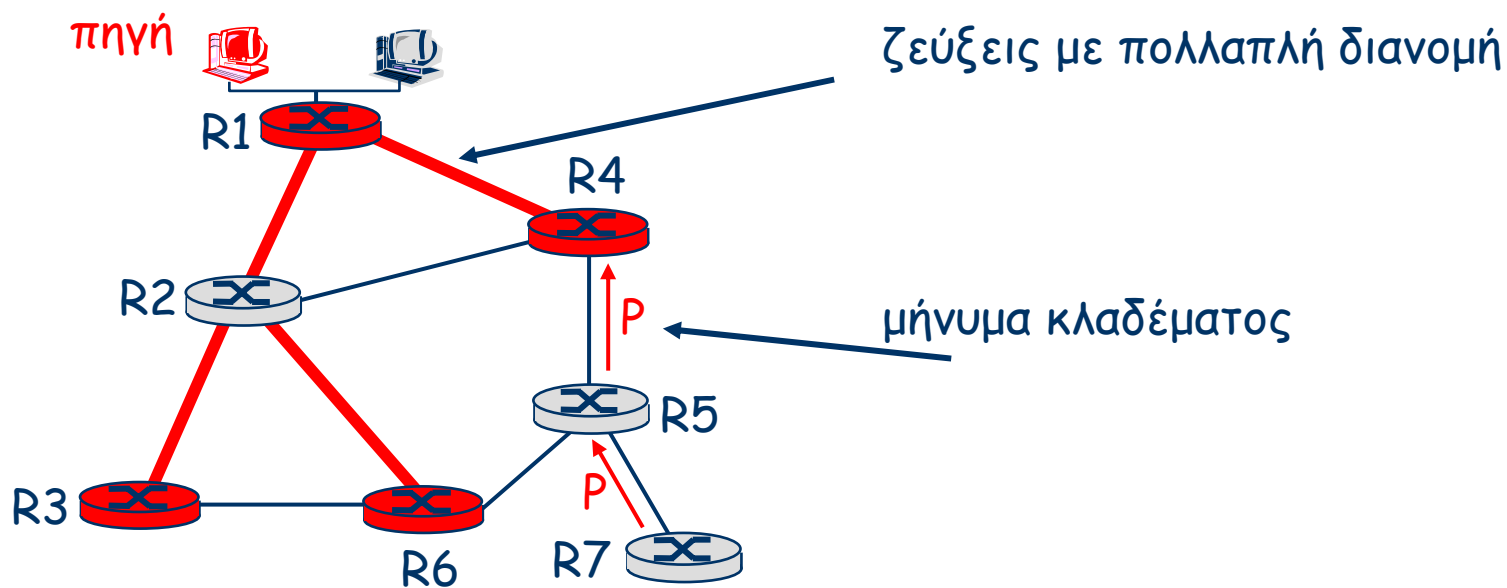


- το αποτέλεσμα είναι ένα *ανάστροφο* δέντρο ελάχιστων διαδρομών (SPT) βασισμένο στην πηγή
 - μπορεί να υπάρξει κακή επιλογή με ασύμμετρες ζεύξεις

Πρώθηση ανάστροφης διαδρομής



- Κλάδεμα (pruning)
- το δέντρο πρώθησης περιέχει υποδέντρα που δεν έχουν μέλη πολλαπλής διαδρομής
 - δεν απαιτείται πρώθηση στο υποδέντρο
 - στέλνονται μηνύματα "κλαδέματος" προς τα άνω από δρομολογητή που δεν έχει μέλη προς τα κάτω



Δρομολόγηση πολλαπλής διανομής στο Internet



- **DVMRP**: distance vector multicast routing protocol, RFC1075
- **πλημμύρα και κλάδεμα**: προώθηση ανάστροφης διαδρομής, δέντρο βασισμένο στην πηγή
 - δέντρο προώθησης ανάστροφης διαδρομής βασισμένο σε πίνακες δρομολόγησης του DVMRP, που κατασκευάζονται με επικοινωνία των δρομολογητών DVMRP
 - δεν λαμβάνονται υπόψη οι υποκείμενες unicast διαδρομές
 - το αρχικό δεδομενόγραμμα προς την ομάδα mcast πλημμυρίζεται παντού μέσω RPF
 - οι δρομολογητές που δεν επιθυμούν την ομάδα, στέλνουν μήνυμα κλαδέματος

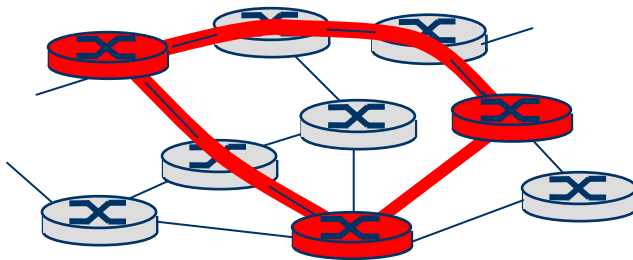
DVMRP: συνέχεια



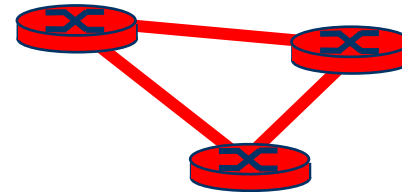
- ο δρομολογητής DVMRP περιοδικά (κάθε 1 min.) "ξεχνάει" τα κλαδεμένα τμήματα:
 - τα mcast δεδομένα οδεύουν πάλι σε όλους τους κλάδους
 - οι downstream δρομολογητές ξανακλαδεύουν ή αλλιώς συνεχίζουν να λαμβάνουν δεδομένα
- οι δρομολογητές μπορούν γρήγορα να ξανασυνδεθούν στο δέντρο
- Επίσης το DVMRP
 - εφαρμόζεται συνήθως σε εμπορικούς δρομολογητές
 - χρησιμοποιείται στην Mbone δρομολόγηση



Πώς συνδέονται "νησίδες" δρομολογητών πολλαπλής διαδρομής σε μεγάλο πλήθος δρομολογητών unicast;



φυσική τοπολογία



λογική τοπολογία

- ❑ το mcast δεδομένογραμμα ενθυλακώνεται σε κανονικό δεδομένογραμμα
- ❑ κανονικό δεδομένογραμμα IP στέλνεται μέσω σήραγγας με κανονικό IP unicast στον δρομολογητή mcast προορισμού
- ❑ ο δρομολογητής mcast προορισμού κάνει αποθυλάκωση για να λάβει το δεδομένογραμμα mcast

PIM: Protocol Independent Multicast



- δεν εξαρτάται από κανένα αλγόριθμο δρομολόγησης unicast που τρέχει από κάτω (λειτουργεί με όλους)
- δύο διαφορετικά σενάρια πολλαπλής διανομής:

Πυκνή διάταξη:

- ❑ τα μέλη της ομάδας εμφανίζουν μεγάλη πυκνότητα και βρίσκονται κοντά μεταξύ τους
- ❑ το εύρος ζώνης είναι πιο άφθονο

Αραιή διάταξη:

- ❑ πολλά δίκτυα με λίγα μέλη ομάδας σε σχέση με τον αριθμό των συνδεδεμένων δικτύων
- ❑ μέλη ομάδας σε μεγάλη διασπορά
- ❑ το εύρος ζώνης δεν είναι άφθονο

Επακόλουθα από την πυκνή - αραιή διάταξη



Πυκνή διάταξη

- Θεωρείται ότι οι δρομολογητές ανήκουν στην ομάδα μέχρι αυτοί να διαγραφούν άμεσα
- η κατασκευή του δέντρου mcast εξαρτάται από τα δεδομένα (π.χ., RPF)
- σπατάλη εύρους ζώνης και επεξεργασίας στους δρομολογητές μη-μέλη

Αραιή διάταξη

- δεν θεωρείται κανείς δρομολογητής ως μέλος μέχρι να ενταχθεί άμεσα
- η κατασκευή του δέντρου mcast εξαρτάται από τον δέκτη (π.χ., center-based)
- συντηρητική χρησιμοποίηση του εύρους ζώνης και της επεξεργασίας σε δρομολογητές μη-μέλη

RIM- Πυκνή διάταξη



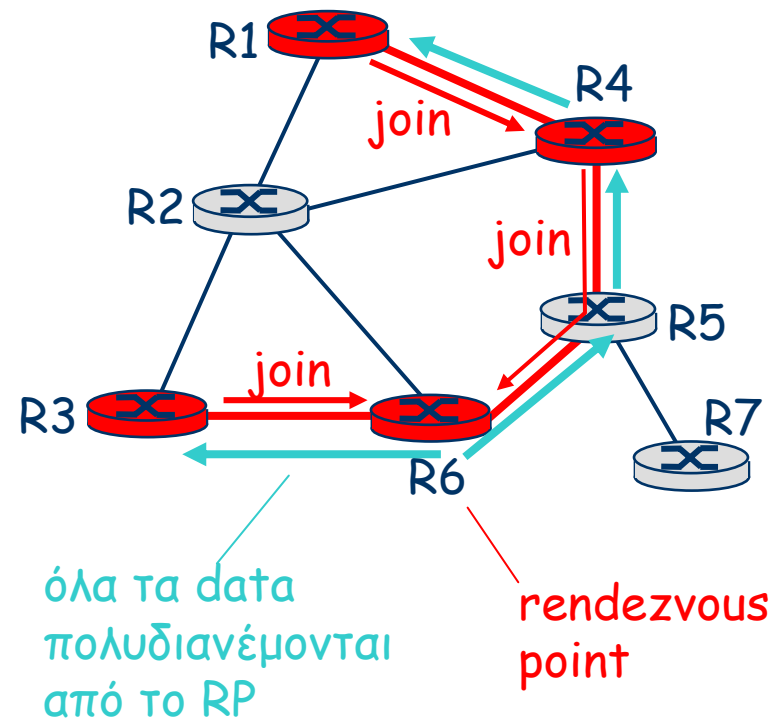
RPF με πλημμύρα και κλάδεμα, παρόμοια με το DVMRP αλλά

- ❑ το πρωτόκολλο unicast που τρέχει από κάτω παρέχει πληροφορία RPF για το εισερχόμενο δεδομένογραμμα
- ❑ λιγότερο πολύπλοκη (λιγότερο αποτελεσματική) ροή προς τα κάτω από εκείνη του DVMRP
- ❑ έχει μηχανισμούς πρωτοκόλλου για τον δρομολογητή για να ανιχνεύει αν είναι ακραίο μέλος ομάδας



PIM - Αραιή διάταξη

- βασίζεται στο κέντρο
- ο δρομολογητής στέλνει μήνυμα *join* στο rendezvous point (RP)
 - οι ενδιάμεσοι δρομολογητές ενημερώνονται και προωθούν το *join*
- μετά την ένταξη μέσω του RP, ο δρομολογητής μπορεί να μπει στο δέντρο πηγής
 - αυξημένη επίδοση: μικρότερη συγκέντρωση, μικρότερες διαδρομές



όλα τα data
πολυδιανέμονται
από το RP

rendezvous
point