ΔΙΚΤΥΑ ΥΠΟΛΟΓΙΣΤΩΝ

Στρώμα μεταφοράς στο Διαδίκτυο

Περίληψη

- Αρχές λειτουργίας του στρώματος μεταφοράς και βασικές υπηρεσίες του.
- Πρωτόκολλα στρώματος μεταφοράς στο Internet:
  - UDP: μεταφορά χωρίς σύνδεση
  - TCP: μεταφορά με σύνδεση

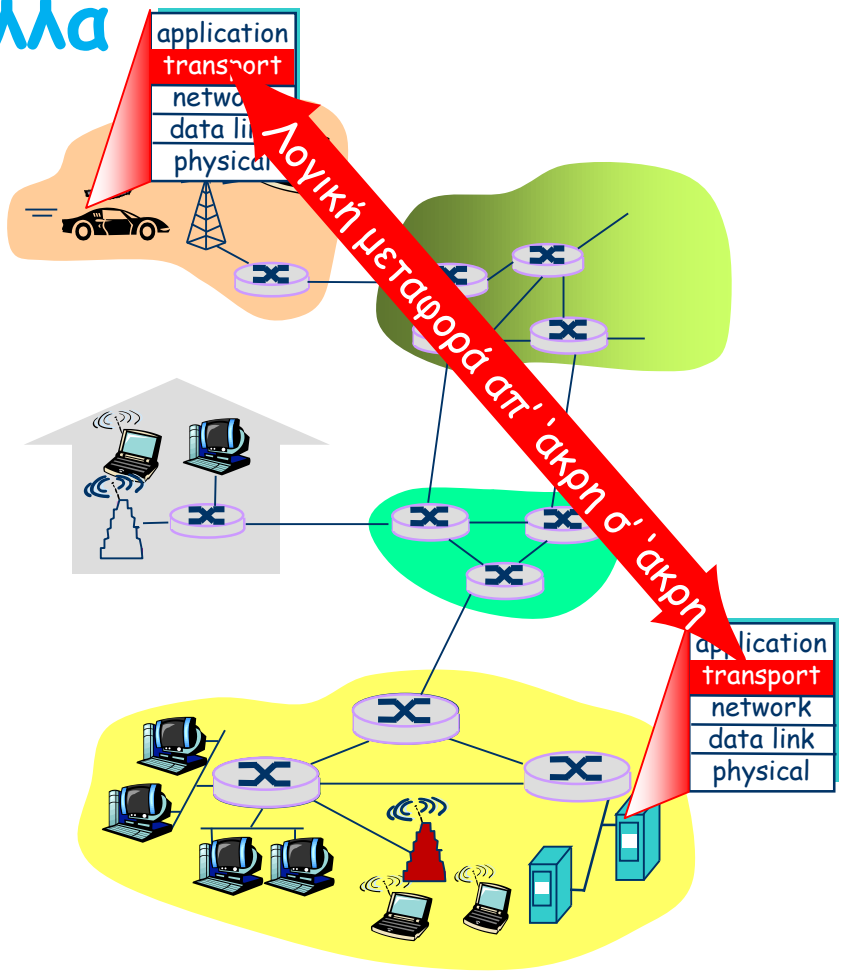
Περιεχόμενα

- Υπηρεσίες και πρωτόκολλα του στρώματος μεταφοράς
- Πολυπλεξία/αποπολυπλεξία
- Μεταφορά χωρίς σύνδεση: UDP
- Αξιόπιστη μεταφορά δεδομένων
- Μεταφορά με σύνδεση: TCP
  - υπηρεσία συρμού byte
  - δομή τεμαχίου
  - αξιόπιστη μετάδοση στο TCP
  - διαχείριση συνδέσεων
  - μεταφορά δεδομένων
  - έλεγχος ροής

Στρώμα μεταφοράς

Υπηρεσίες και πρωτόκολλα

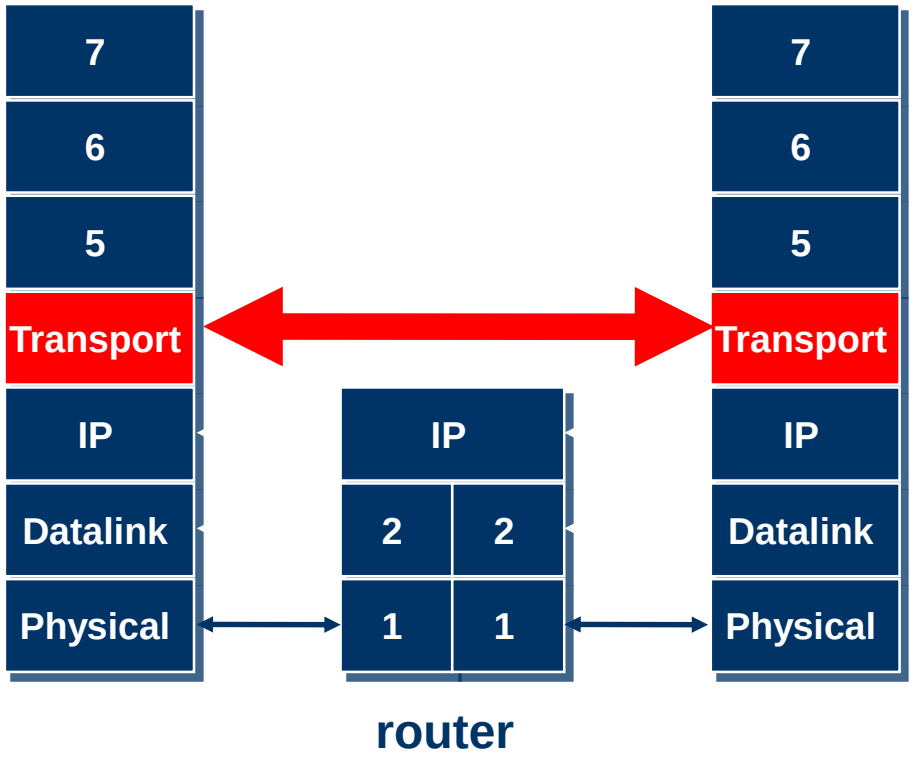
- Παρέχει *λογική επικοινωνία* μεταξύ διαδικασιών εφαρμογής που τρέχουν σε διαφορετικούς host.
- Τα πρωτόκολλα μεταφοράς τρέχουν στους host:
  - Εκπομπή: χωρίζει τα μηνύματα εφαρμογής σε τεμάχια και τα διοχετεύει στο στρώμα δικτύου.
  - Λήψη: επανασυναρμολογεί τα τεμάχια σε μηνύματα και τα διοχετεύει στο στρώμα εφαρμογής.
- Υπάρχουν περισσότερα από ένα πρωτόκολλα μεταφοράς διαθέσιμα στις εφαρμογές:
  - Internet: TCP and UDP



Στρώμα μεταφοράς

Πρωτόκολλα μεταφοράς

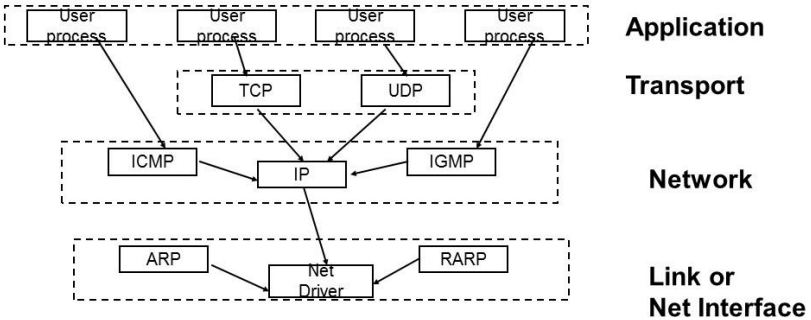
- Είναι τα πρώτα πρωτόκολλα, αρχίζοντας από το φυσικό στρώμα, που λειτουργουν από άκρη σε άκρη.
- Η επικεφαλίδα του πρωτοκόλλου μεταφοράς που δημιουργείται στην πηγή εξετάζεται μόνο από τον προορισμό.
- Οι δρομολογητές θεωρούν την επικεφαλίδα του πρωτοκόλλου μεταφοράς ως μέρος του ωφέλιμου φορτίου του πακέτου IP.



Στρώμα μεταφοράς

Στρώμα μεταφοράς και στρώμα δικτύου

- *Στρώμα δικτύου:* λογική επικοινωνία μεταξύ host.
- *Στρώμα μεταφοράς:* λογική επικοινωνία μεταξύ διαδικασιών εφαρμογής
  - βασίζεται στις υπηρεσίες δικτύου και τις βελτιώνει.

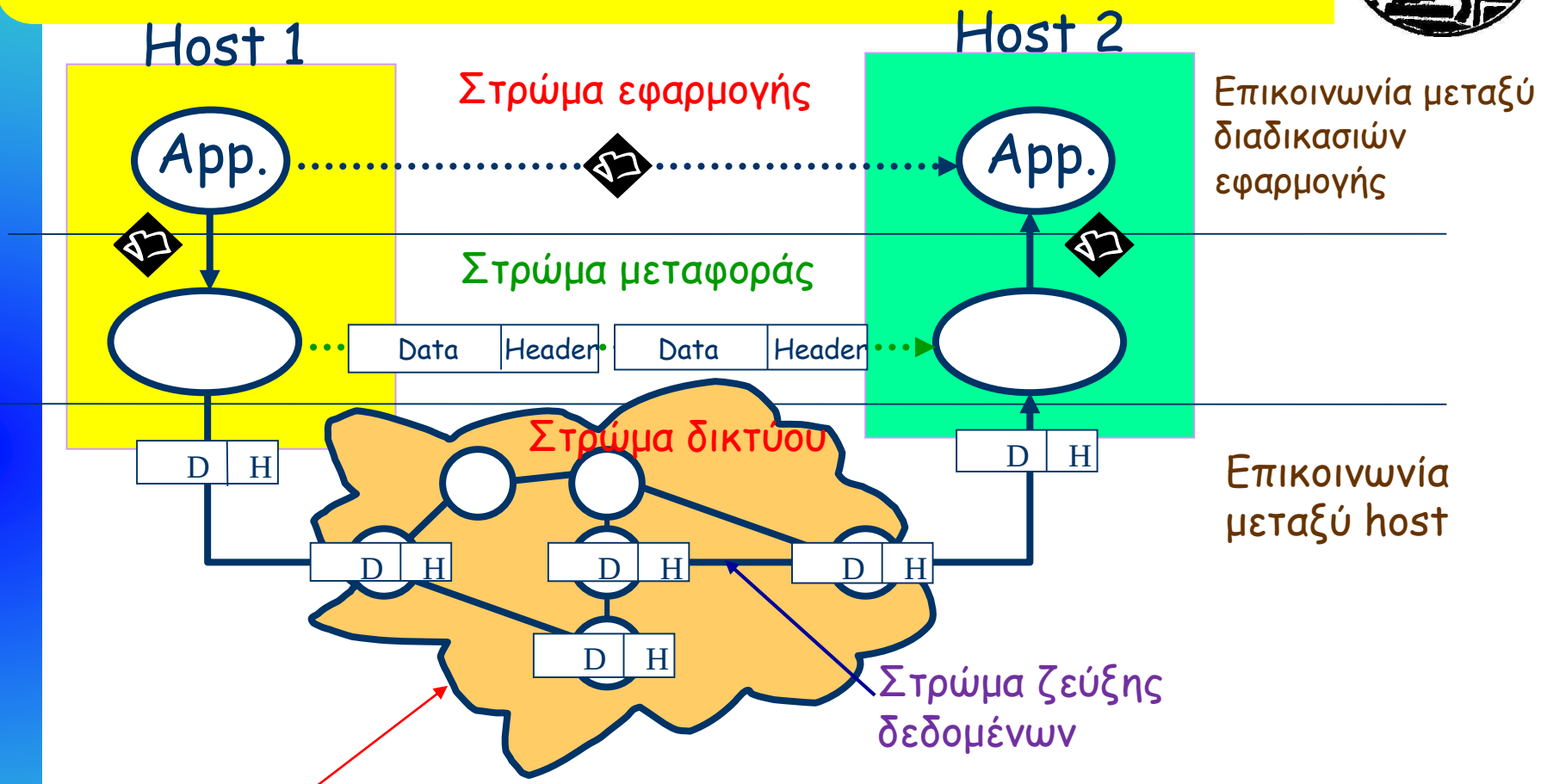


ICMP: Internet Control Message Protocol
IGMP: Internet Group Management Protocol
IP : Internet Protocol
ARP : Address Resolution Protocol
RARP: Reverse Address Resolution Protocol
TCP : Transmission Control Protocol
UDP : User Datagram Protocol

Στρώμα μεταφοράς

- Παρέχει *αξιόπιστη* ή *μη αξιόπιστη* μεταφορά δεδομένων μεταξύ διαδικασιών εφαρμογής.
- Υποστηρίζει μηνύματα *αυθαίρετου μήκους.*
- Παρέχει τρόπο απόφασης για το ποια τεμάχια πηγαίνουν σε ποιες εφαρμογές (πολυπλεξία/αποπολυπλεξία).
- Ρυθμίζει πότε οι host πρέπει να στέλνουν.

Στρώμα μεταφοράς

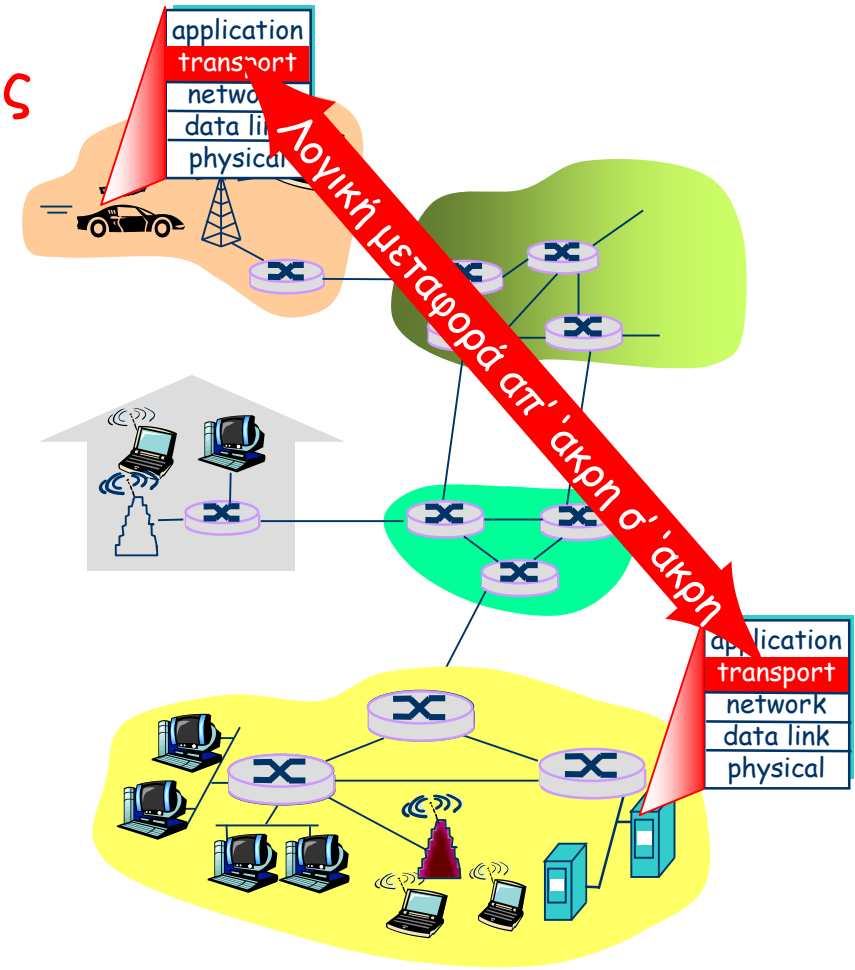


- Υπηρεσία best-effort
- Πακέτα περιορισμένου μήκους
- Πακέτα καθυστερούν, χάνονται,...
- Επικοινωνία μεταξύ host

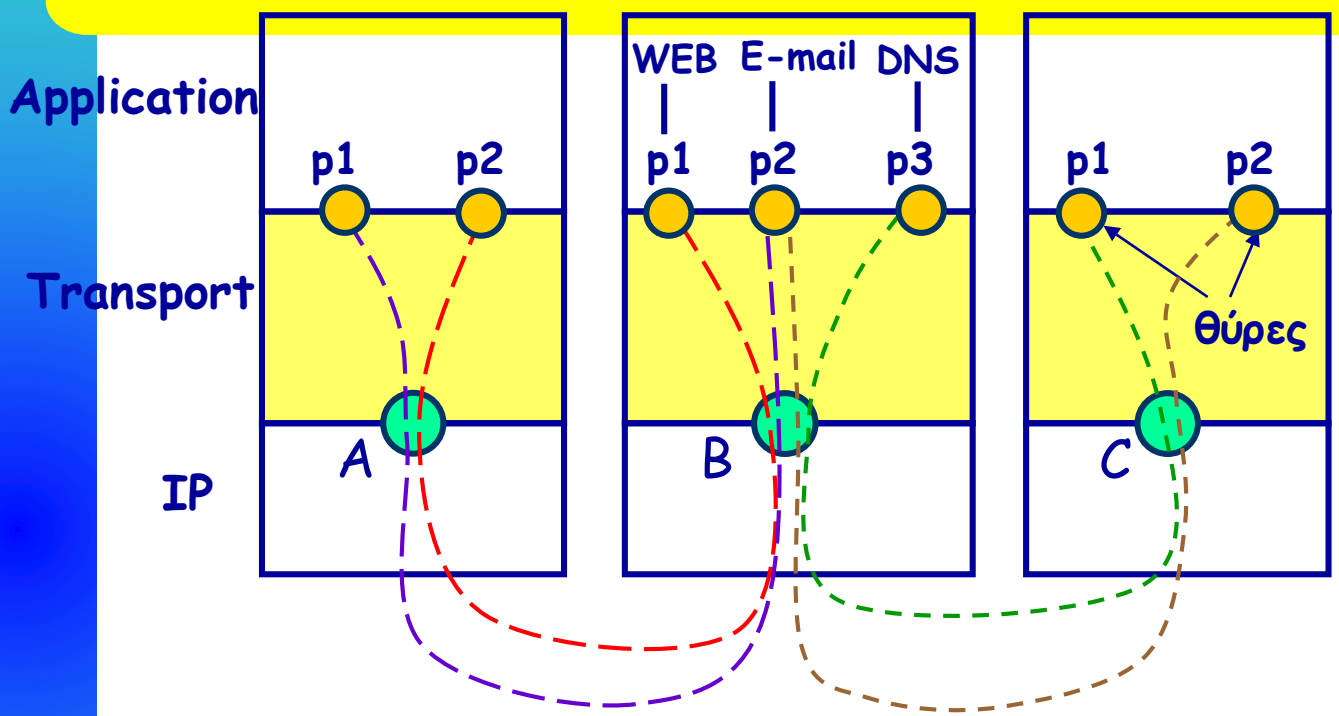
- Ποια εφαρμογή λαμβάνει ποια πακέτα;
- Με ποιο ρυθμό πρέπει να στέλνουν οι host στο δίκτυο;

Στρώμα μεταφοράς στο Internet

- Μη αξιόπιστη παράδοση χωρίς διατήρηση της σειράς: UDP
  - Μινιμαλιστική επέκταση της υπηρεσίας "best-effort" του IP
- Αξιόπιστη παράδοση με διατήρηση της σειράς: TCP
  - Εγκατάσταση σύνδεσης
  - Έλεγχος ροής
  - Έλεγχος συμφόρησης
- Υπηρεσίες που δεν προσφέρονται:
  - Εξασφάλιση καθυστέρησης
  - Εξασφάλιση εύρους ζώνης



Στρώμα μεταφοράς στο Internet



- UDP: αναξιόπιστη μεταφορά
- TCP: αξιόπιστη, με τη σειρά

- Και το TCP και το UDP επεκτείνουν την επικοινωνία IP μεταξύ host σε επικοινωνία μεταξύ διαδικασιών εφαρμογής: πολυπλεξία/αποπολυπλεξία του στρώματος μεταφοράς.
- Ποια πακέτα λαμβάνει η κάθε εφαρμογή;
  - Λύση: αντιστοίχηση κάθε υποδοχής σε μια θύρα.
  - Κάθε τεμάχιο έχει ειδικά πεδία για τους αριθμούς θυρών.
  - Οι υποδοχές προσδιορίζονται μονοσήμαντα.

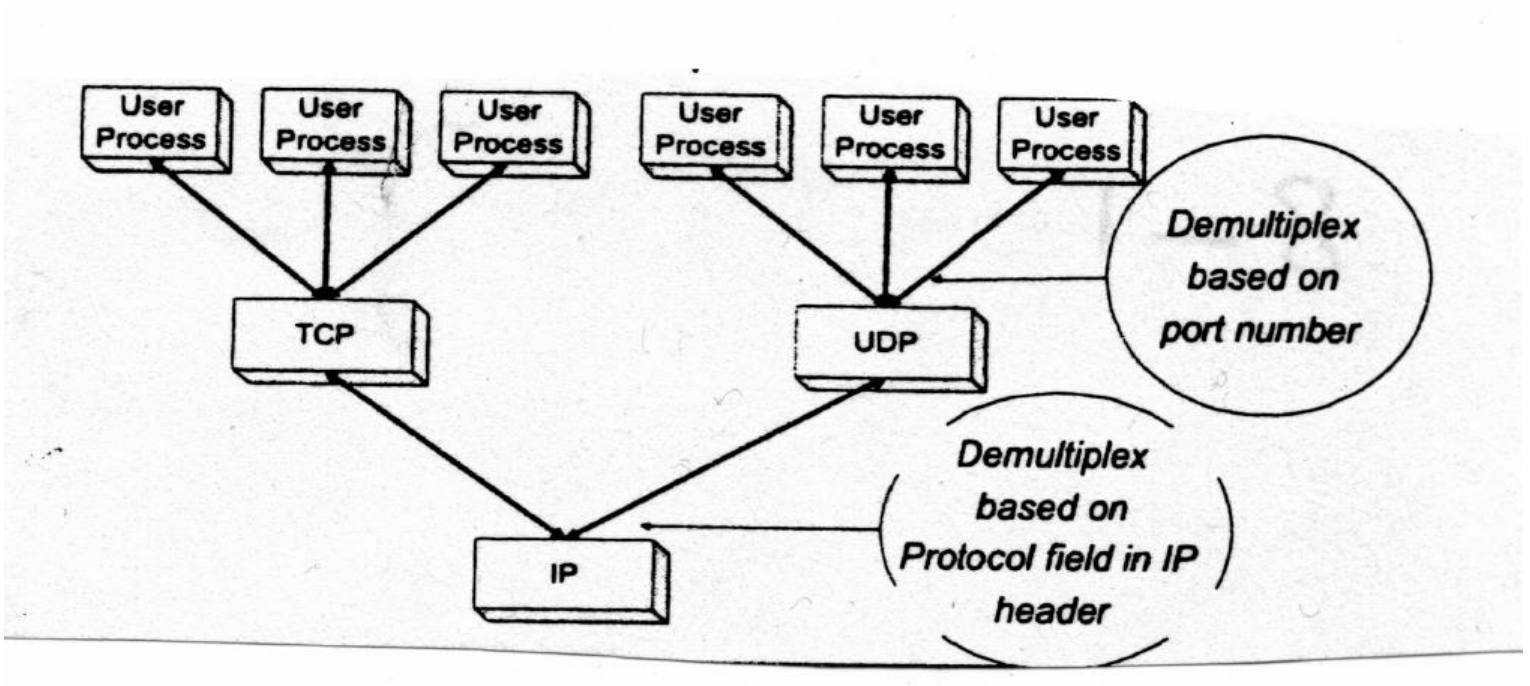
Στρώμα μεταφοράς στο Internet

Θύρες

- Χώρος διευθύνσεων θυρών των 16-bit για το UDP και το TCP.
- Ο client πρέπει να ξέρει τη θύρα του server.
- Πασίγνωστες θύρες (0-1023): όλοι συμφωνούν ποιες υπηρεσίες τρέχουν σ' αυτές τις θύρες.
  - π.χ., telnet:23, SMTP:25, DNS:53, http:80.
- Εφήμερες θύρες (1024-65535): δίδονται στους client.

Πολυπλεξία/Αποπολυπλεξία

- UDP and TCP use port numbers to identify applications
- At the transport layer a globally unique address is identified by: <IP address, port number>



Στρώμα μεταφοράς στο Internet

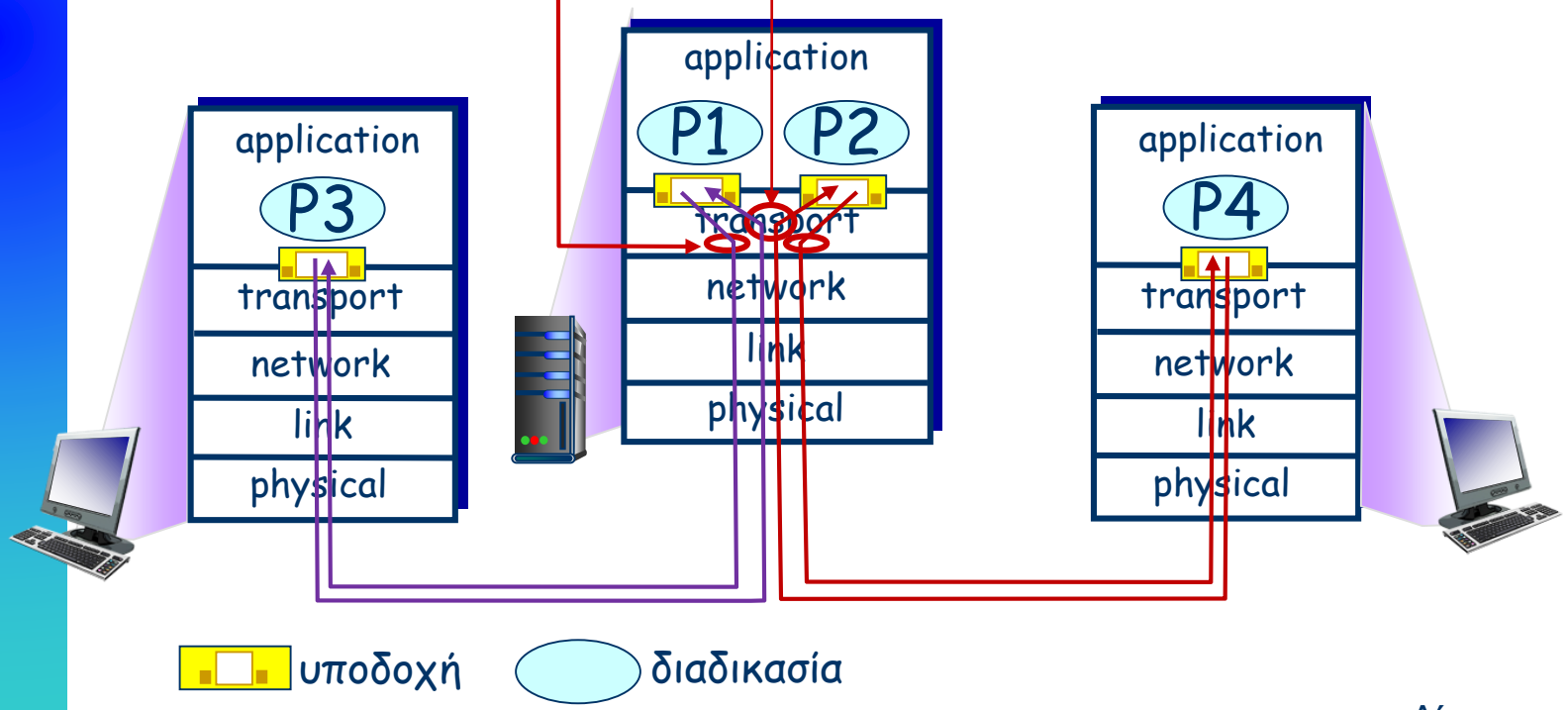
Πολυπλεξία/αποπολυπλεξία

πολυπλεξία στον πομπό:

χειρίζεται δεδομένα από πολλές υποδοχές, προσθέτει επικεφαλίδες μεταφοράς (χρησιμοποιούνται μετά για αποπολυπλεξία)

αποπολυπλεξία στον δέκτη:

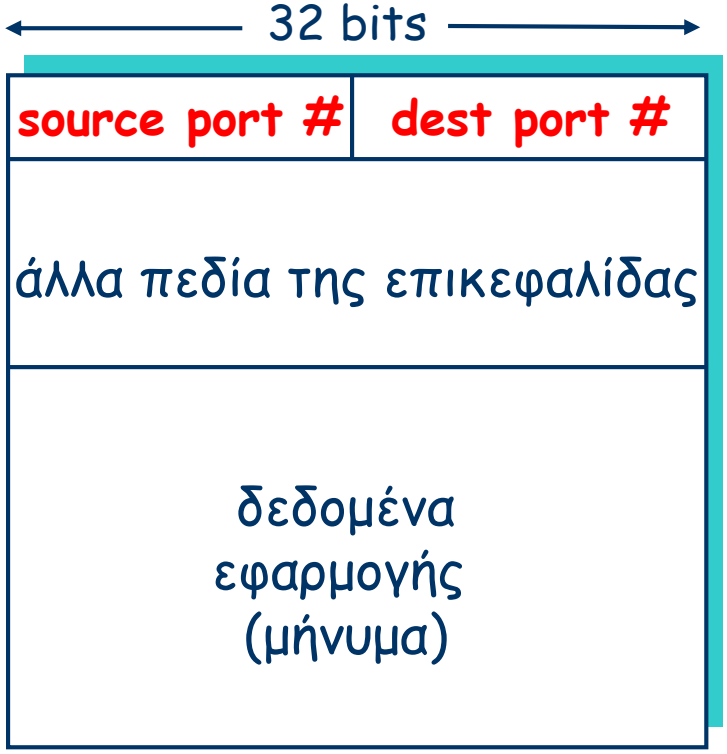
χρησιμοποιεί την πληροφορία της επικεφαλίδας για να παραδώσει τα λαμβανόμενα τεμάχια στη σωστή υποδοχή



Στρώμα μεταφοράς στο Internet

Πώς λειτουργεί η αποπολυπεξία

- Ο host λαμβάνει πακέτα IP:
  - Κάθε πακέτο IP έχει διευθύνσεις source IP και destination IP.
  - Κάθε πακέτο IP μεταφέρει ένα τεμάχιο στρώματος μεταφοράς.
  - Κάθε τεμάχιο έχει αριθμούς source port και destination port.
- Ο host χρησιμοποιεί τις IP διευθύνσεις και τους αριθμούς θυρών για να κατευθύνει το τεμάχιο στην κατάλληλη υποδοχή.



μορφή τεμαχίου στρώματος μεταφοράς

Στρώμα μεταφοράς στο Internet

Αποπολυπλεξία χωρίς σύνδεση

- Δημιουργία υποδοχών με αριθμούς θυρών:

```
DatagramSocket mySocket1 = new DatagramSocket(9157);
DatagramSocket mySocket2 = new DatagramSocket(9222);
```

- Μια υποδοχή UDP προσδιορίζεται από το ζεύγος:

  (dest IP address, dest port number)

- Όταν ο host λαμβάνει τεμάχιο UDP:
  - Ελέγχει τον αριθμό destination port στο τεμάχιο,
  - Κατευθύνει το τεμάχιο UDP στην υποδοχή με αυτόν τον αριθμό.
- Πακέτα IP με διαφορετικές διευθύνσεις source IP και/ή διαφορετικούς αριθμούς source port, αλλά με ίδια dest IP και ίδιο dest port, οδηγούνται στην ίδια υποδοχή.

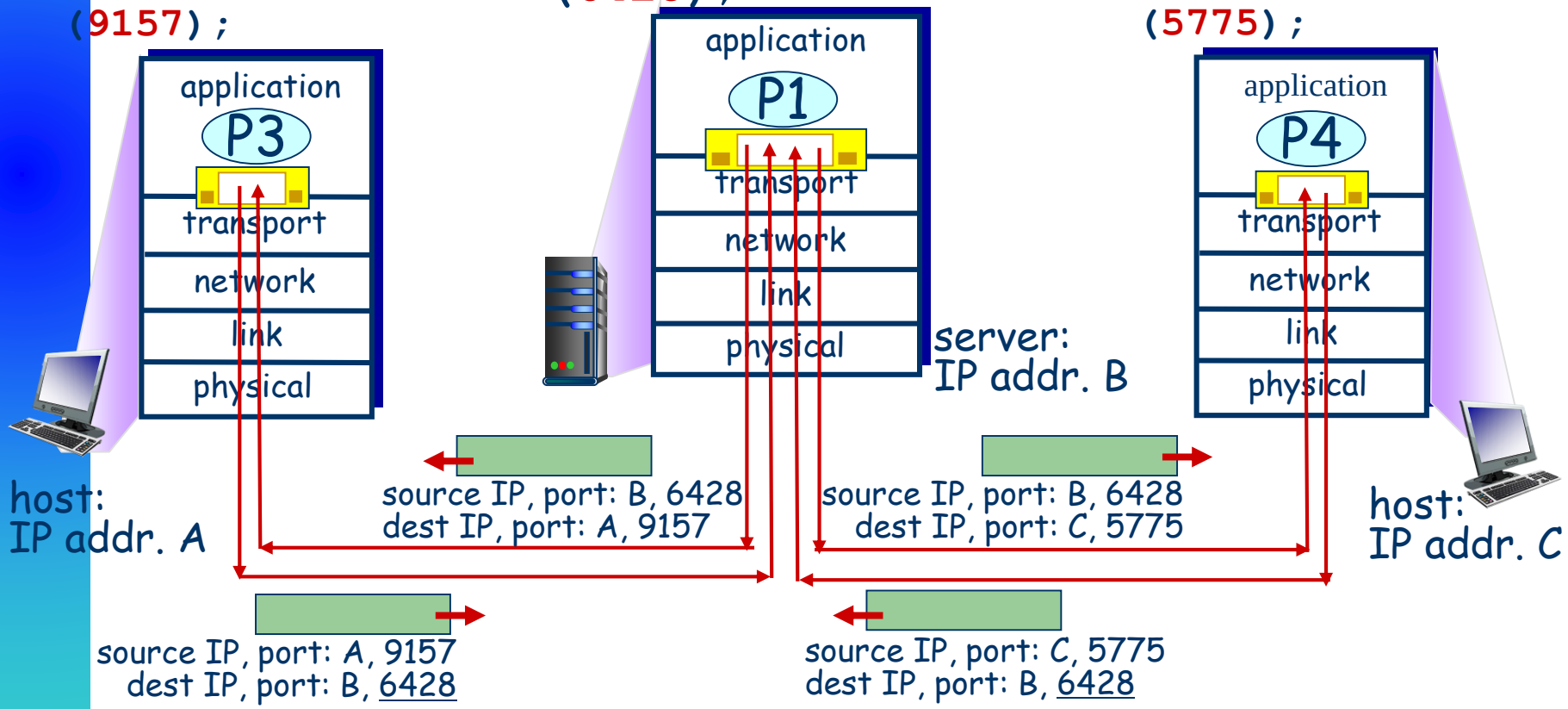
Στρώμα μεταφοράς στο Internet

Αποπολυπλεξία χωρίς σύνδεση

```
DatagramSocket
 mySocket2 = new
 DatagramSocket
 (9157);
```

```
DatagramSocket
 serverSocket = new
 DatagramSocket
 (6428);
```

```
DatagramSocket
 mySocket1 = new
 DatagramSocket
 (5775);
```



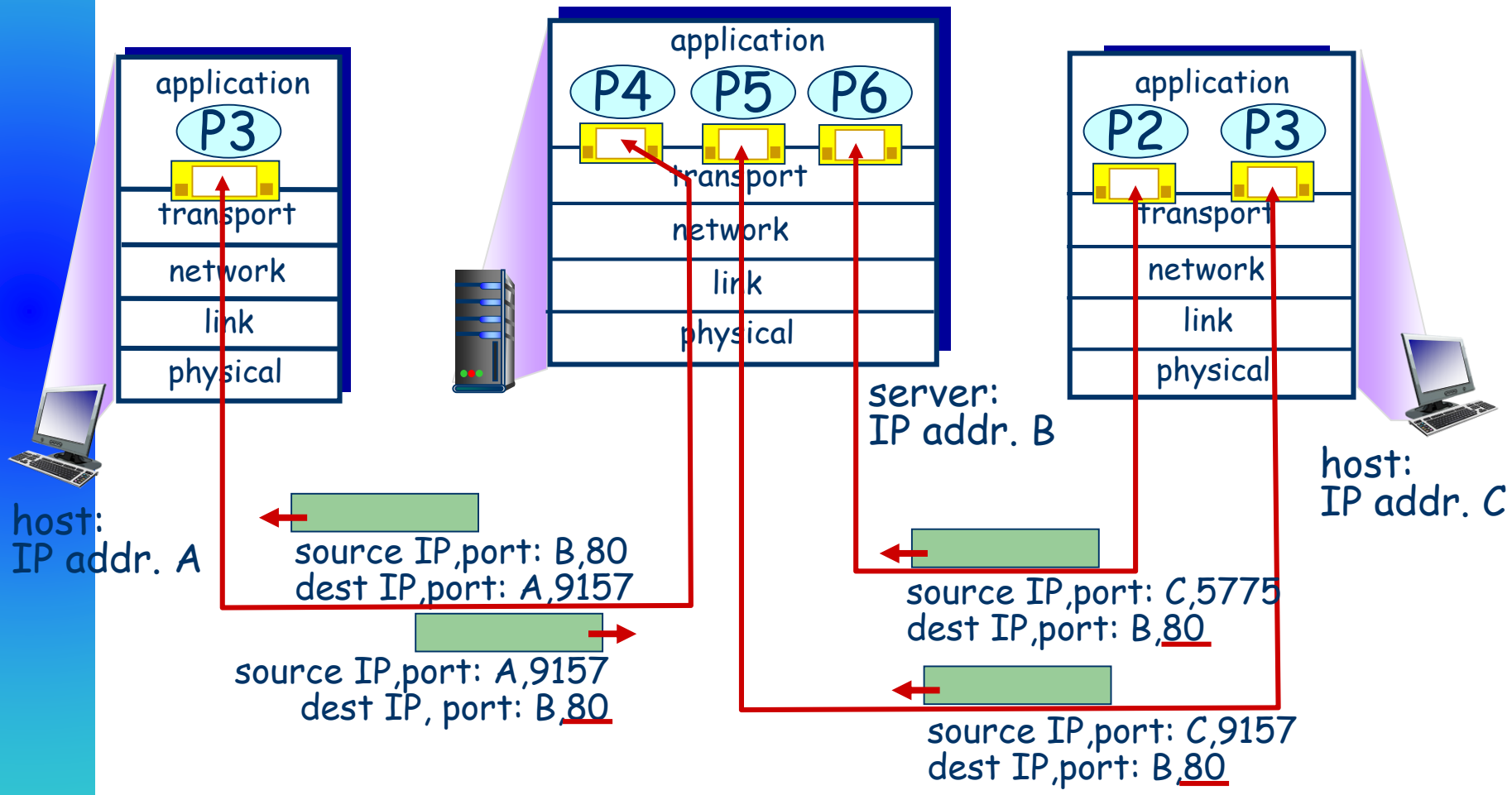
Στρώμα μεταφοράς στο Internet

Αποπολυπλεξία με σύνδεση

- Η υποδοχή TCP καθορίζεται από την τετράδα:
  - source IP address, source port number, dest IP address, dest port number
- Ο host λήψης χρησιμοποιεί και τις 4 τιμές για να κατευθύνει το τεμάχιο στην κατάλληλη υποδοχή.
- Ένας Server host μπορεί να υποστηρίζει πολλές υποδοχές TCP ταυτόχρονα:
  - Κάθε υποδοχή καθορίζεται από τη δική της τετράδα.
- Οι Web servers έχουν διαφορετικές υποδοχές για κάθε συνδεόμενο client.
  - Στο μη-επίμονο HTTP έχουμε διαφορετική υποδοχή για κάθε αίτηση/απάντηση.

Στρώμα μεταφοράς στο Internet

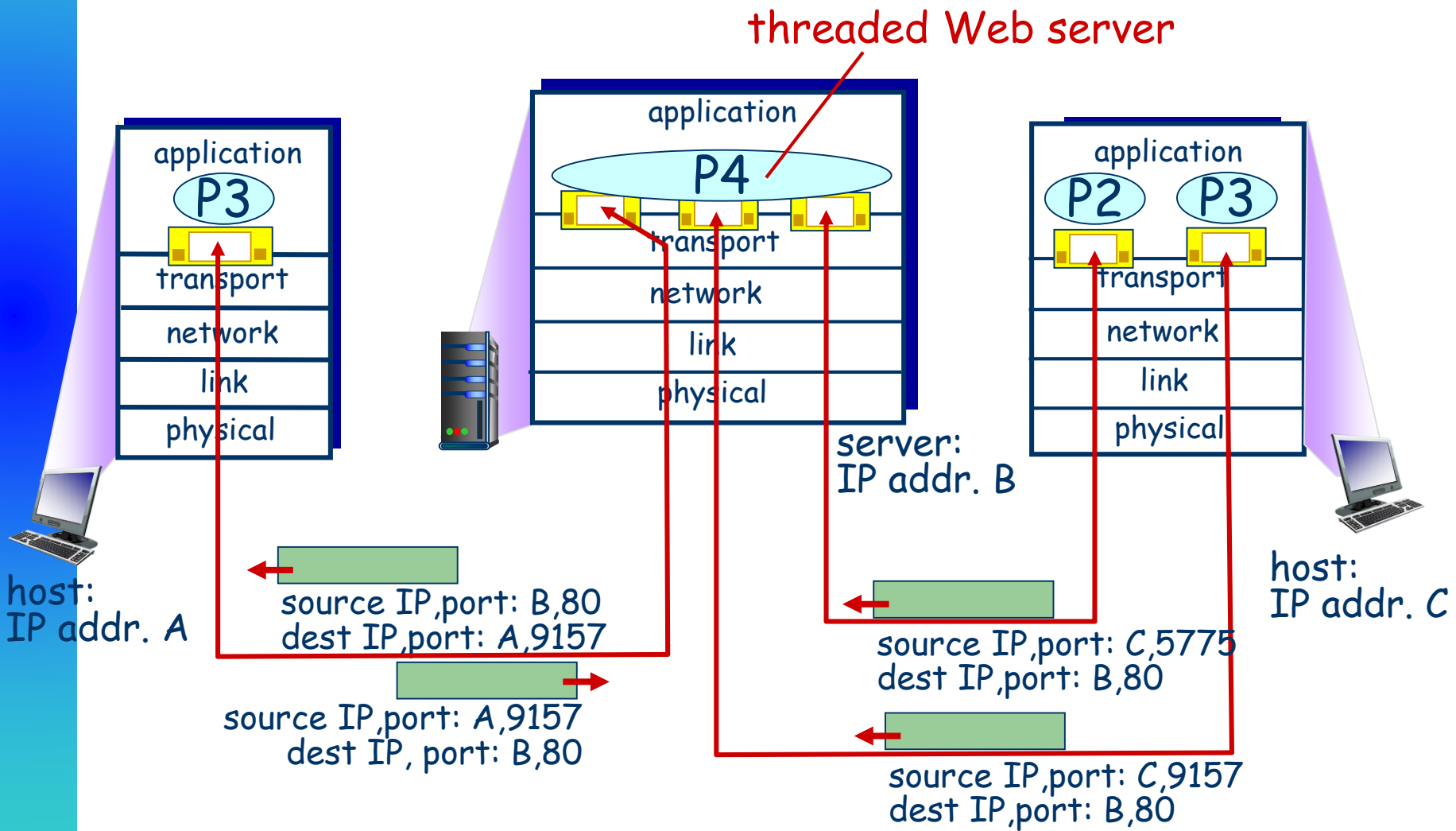
Αποπολυπλεξία με σύνδεση



τρία τεμάχια, όλα προοριζόμενα για τη διεύθυνση IP: B, dest port: 80 αποπολυπλέκονται σε *διαφορετικές* υποδοχές

Στρώμα μεταφοράς στο Internet

Αποπολυπλεξία με σύνδεση



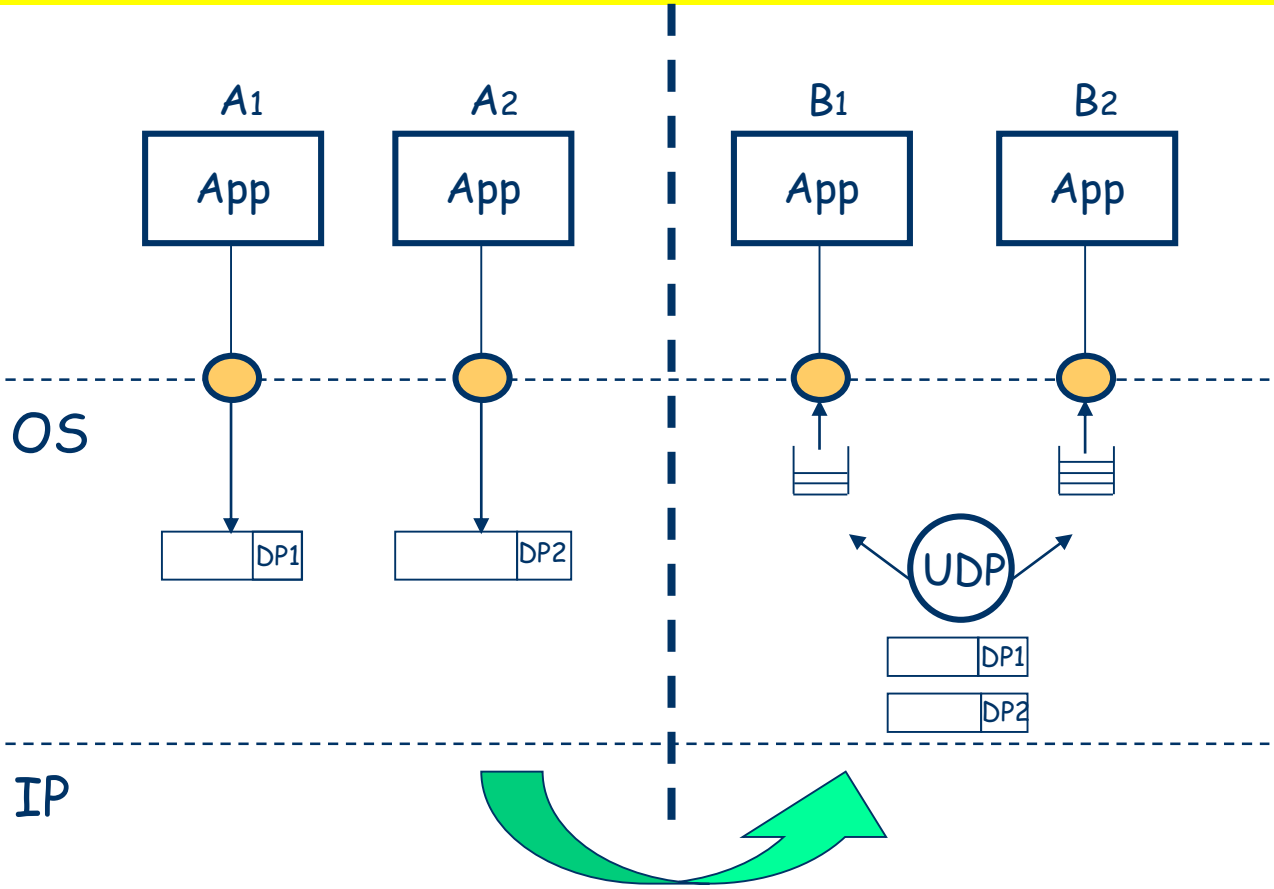
UDP: User Datagram Protocol [RFC 768]

- Λιτό πρωτόκολλο μεταφοράς στο Internet.
- Υπηρεσία "καλύτερης προσπάθειας". Τα τεμάχια UDP μπορεί:
  - Να χαθούν,
  - Να παραδοθούν εκτός σειράς στο ανώτερο στρώμα.
- *Χωρίς σύνδεση:*
  - Όχι χειραψία μεταξύ του πομπού και του δέκτη UDP.
  - Κάθε τεμάχιο UDP αντιμετωπίζεται ανεξάρτητα από τα άλλα.

Γιατί υπάρχει το UDP;

- Δεν εγκαθίσταται σύνδεση (που μπορεί να εισάγει καθυστέρηση).
- Απλό: δεν διατηρείται κατάσταση σύνδεσης στον πομπό, δέκτη.
- Μικρή επικεφαλίδα στο τεμάχιο.
- Όχι έλεγχος συμφόρησης: το UDP μπορεί να στέλνει όσο γρήγορα μπορεί

UDP



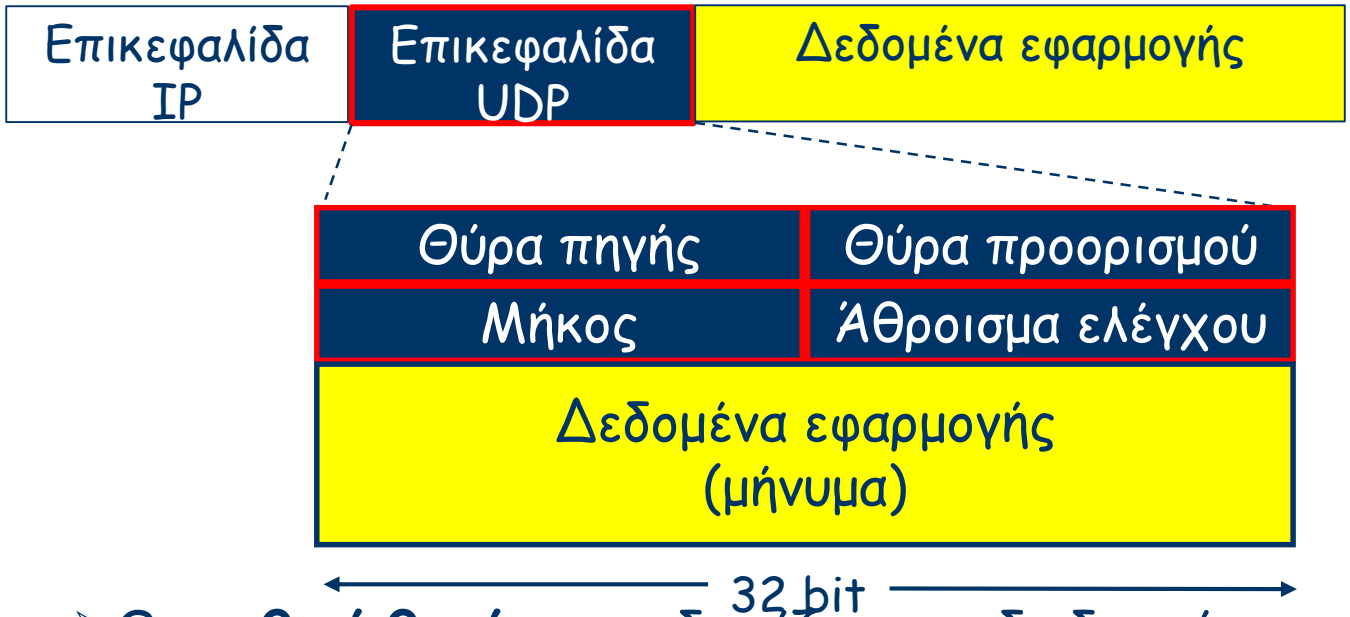
Το UDP, όπως και το TCP, χρησιμοποιεί αριθμούς θυρών για να αποπολυπλέξει τα τεμάχια.

UDP

- Μηνύματα μέχρι και 64KB.
- Παρέχει πολυπλεξία/αποπολυπλεξία στο IP.
- Πλεονεκτεί έναντι του TCP στο ότι δεν αυξάνει την καθυστέρηση απ' άκρη σ' άκρη πάνω από το IP.
- Χρησιμοποιείται σε εφαρμογές streaming multimedia:
  - Ανοχή σε απώλειες,
  - Ευαισθησία στον ρυθμό μετάδοσης.
- Άλλες χρήσεις του UDP:
  - DNS
  - SNMP
  - RIP
- Για αξιόπιστη μεταφορά με UDP, χρειάζεται προσθήκη αξιοπιστίας στο στρώμα εφαρμογής.
  - Αποκατάσταση λαθών ειδική για την εφαρμογή.

UDP

Μορφή τεμαχίου UDP



- Οι **αριθμοί θυρών** προσδιορίζουν τις διαδικασίες αποστολής και λήψης. Η μέγιστη τιμή αριθμού θύρας είναι $2^{16}-1= 65,535$.
- Το **μήκος τεμαχίου** UDP (συμπεριλαμβάνει την επικεφαλίδα UDP) είναι τουλάχιστον 8 byte (π.χ., άδειο πεδίο Data) και το πολύ 65,535 byte.
- Το **άθροισμα ελέγχου** περιλαμβάνει την επικεφαλίδα του UDP και μερικά πεδία της επικεφαλίδας IP.

UDP

Άθροισμα ελέγχου UDP

Στόχος: ανίχνευση σφαλμάτων στο μεταδιδόμενο τεμάχιο

- Συμπεριλαμβάνει την ψευδοεπικεφαλίδα.

Πομπός:

- Μεταχειρίζεται τα περιεχόμενα του τεμαχίου ως ακολουθία ακεραίων των 16-bit.
- checksum: πρόσθεση των περιεχομένων του τεμαχίου και λήψη του συμπληρώματος ως προς 1 του αθροίσματος.
- Ο πομπός τοποθετεί την τιμή του checksum στο πεδίο checksum του τεμαχίου UDP.

Δέκτης:

- Υπολογίζει το checksum του λαμβανόμενου τεμαχίου συμπεριλαμβανομένου και του πεδίου checksum.
- Ελέγχει αν το υπολογισθέν checksum ισούται με την τιμή 1111111111111111:
  - ΟΧΙ - ανιχνεύθηκε σφάλμα.
  - ΝΑΙ - δεν ανιχνεύθηκε σφάλμα. Αλλά μπορεί να υπάρχουν σφάλματα….

- Γιατί το UDP προβλέπει checksum;

UDP

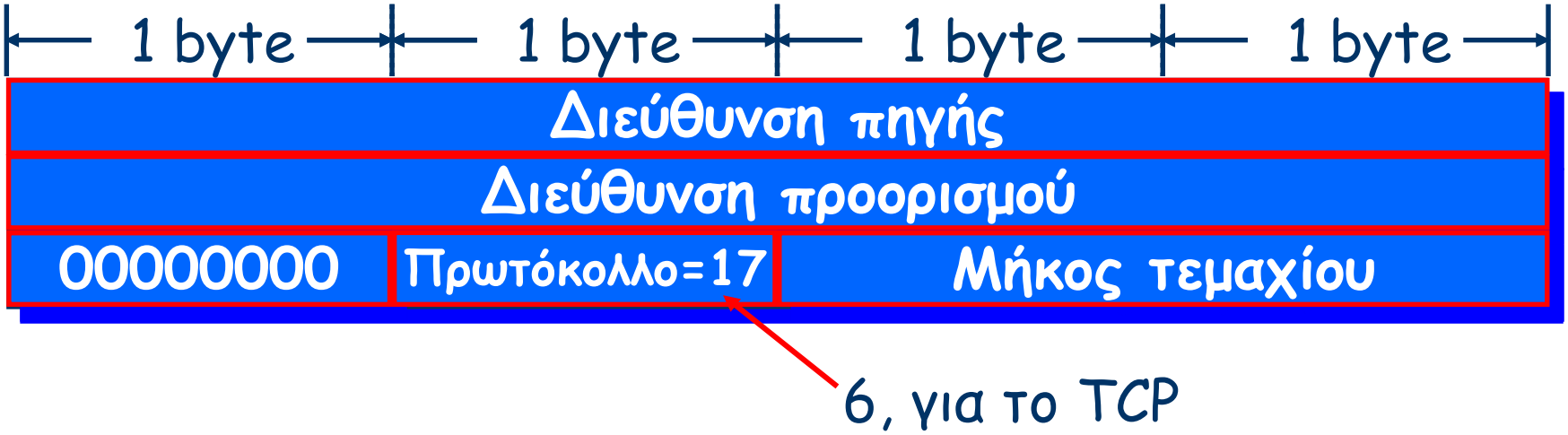
Άθροισμα ελέγχου

- Σημείωση: Κατά την πρόσθεση αριθμών, το κρατούμενο από το πιο σημαντικό bit πρέπει να προστίθεται στο αποτέλεσμα.
- Παράδειγμα: πρόσθεση δύο ακεραίων 16-bit

	1	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0
	1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
wraparound	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
sum		1	0	1	1	1	0	1	1	1	0	1	1	1	1	0	0
checksum		0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	1

UDP

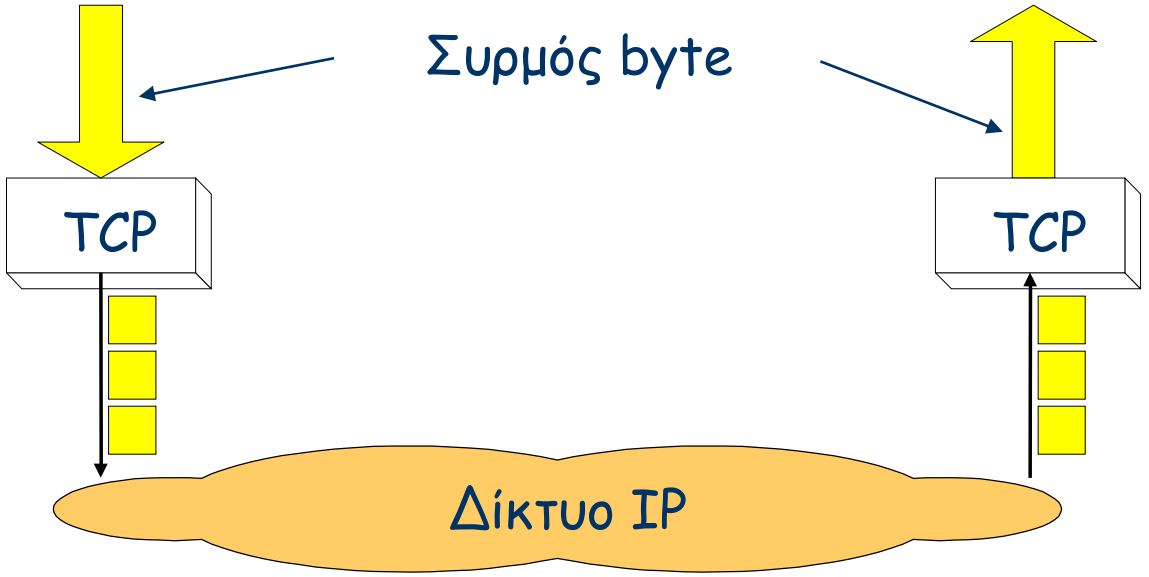
Ψευδοεπικεφαλίδα



- Η ψευδοεπικεφαλίδα εξασφαλίζει ότι το πακέτο IP έχει πραγματικά οδηγηθεί στο σωστό προορισμό, host και θύρα.
- Χρησιμοποιείται μόνο για τον υπολογισμό του Checksum και δεν μεταδίδεται.

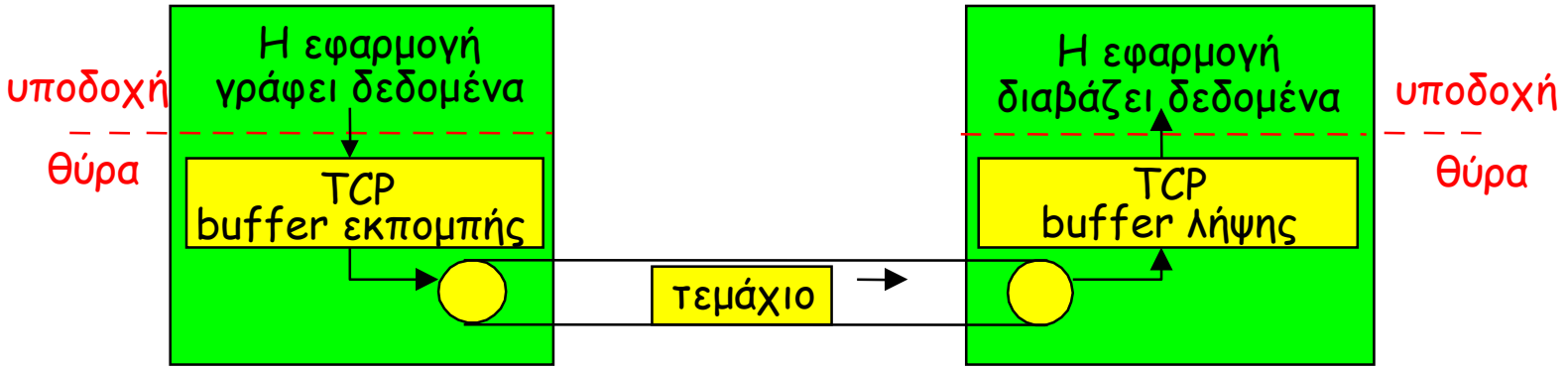
TCP: Transmission Control Protocol

- Είναι πρωτόκολλο με σύνδεση.
- Παρέχει αξιόπιστη μετάδοση συρμού byte απ' άκρη σ' άκρη πάνω από μη αξιόπιστο δίκτυο.



TCP

- Μετάδοση σημείου προς σημείο:
  - Ένας πομπός ένας δέκτης.
- Αξιόπιστη μετάδοση (παράδοση με τη σειρά) συρμού byte:
  - Το μήνυμα μπορεί να έχει αυθαίρετο μήκος.
- Συνεχής παροχή:
  - Ο έλεγχος συμφόρησης και ροής στο TCP καθορίζουν το μέγεθος του παραθύρου.
- Buffers εκπομπής και λήψης.



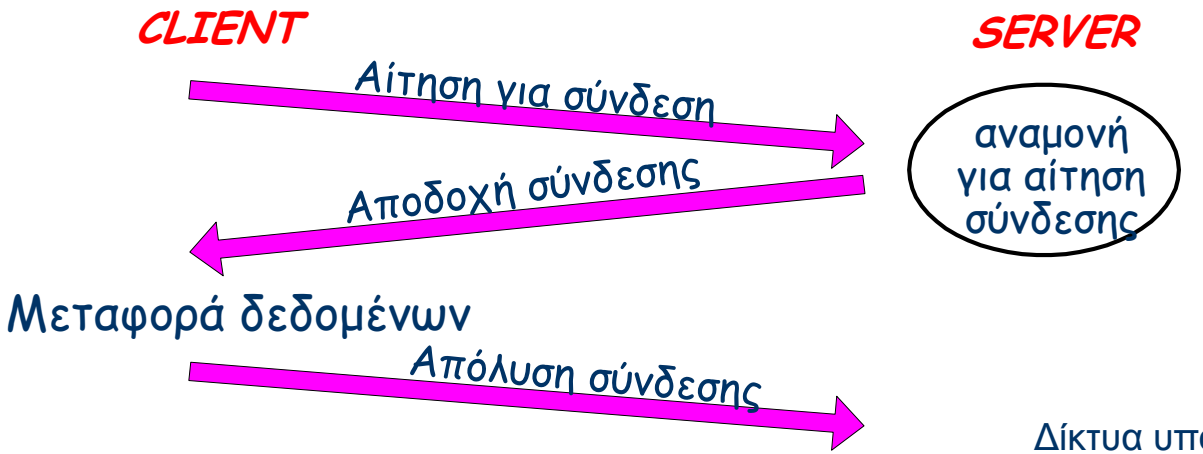
TCP

- Αμφίδρομη μετάδοση
  - Αμφίδρομη ροή δεδομένων στην ίδια σύνδεση.
  - MSS: maximum segment size
- Υπηρεσία με σύνδεση
  - Τριμερής χειραψία αρχικοποιεί την κατάσταση πομπού και δέκτη πριν την ανταλλαγή δεδομένων.
- Έλεγχος ροής
  - Ο πομπός δεν υπερχειλίζει τον δέκτη.
- Παρέχει πολυπλεξία/αποπολυπλεξία πάνω από το IP.
- Έλεγχος και αποφυγή συμφόρησης.
- Παραδείγματα χρησιμοποίησης: file transfer, chat, web, SMTP (e-mail).

TCP

Πρωτόκολλο με σύνδεση

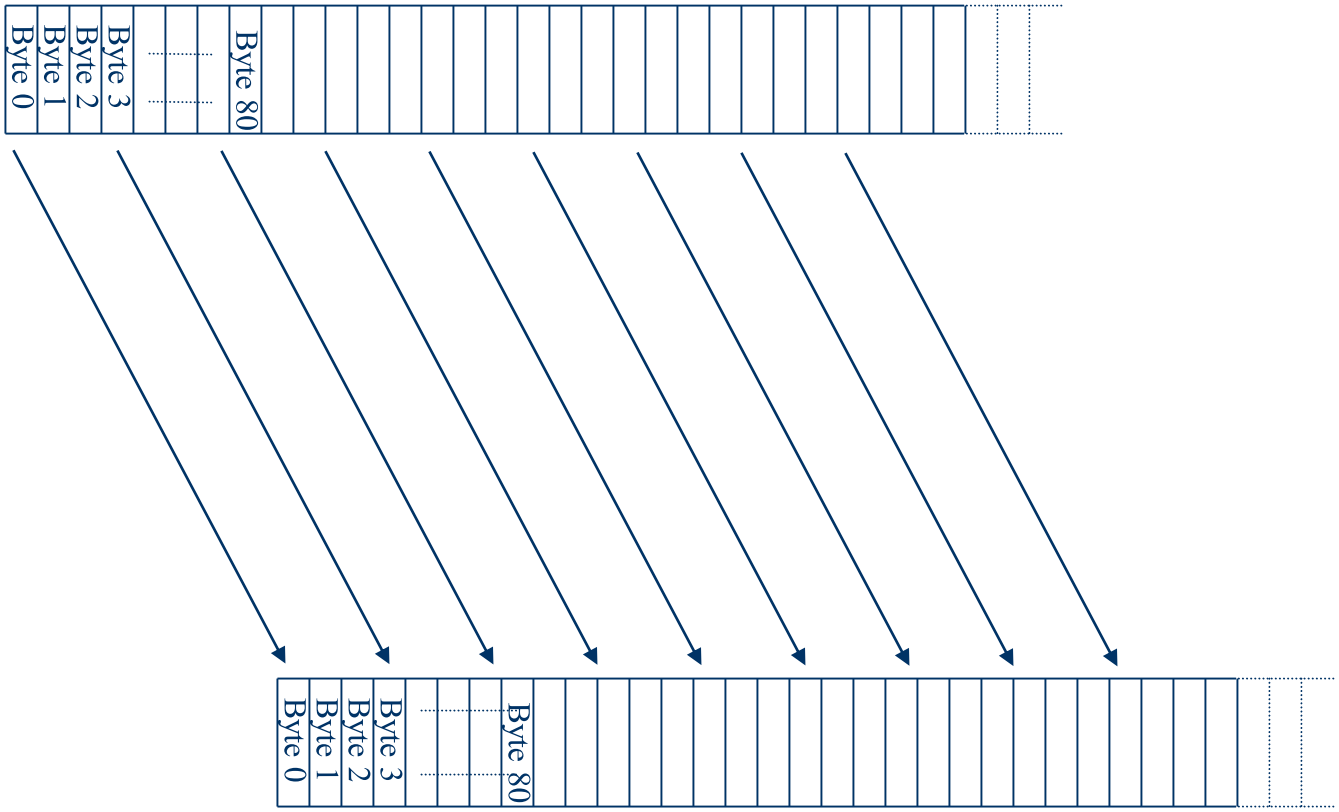
- Πριν από οποιανδήποτε μεταφορά δεδομένων, το TCP εγκαθιστά μια **σύνδεση**:
  - Η μια οντότητα TCP αναμένει για σύνδεση ("**server**")
  - Η άλλη οντότητα TCP ("**client**") συνδέεται με τον server
- Η διαδικασία εγκατάστασης σύνδεσης είναι στην πραγματικότητα πιο πολύπλοκη.
- Κάθε σύνδεση είναι αμφίδρομη, σημείου προς σημείο.



TCP

Υπηρεσία συρμού byte

Host A

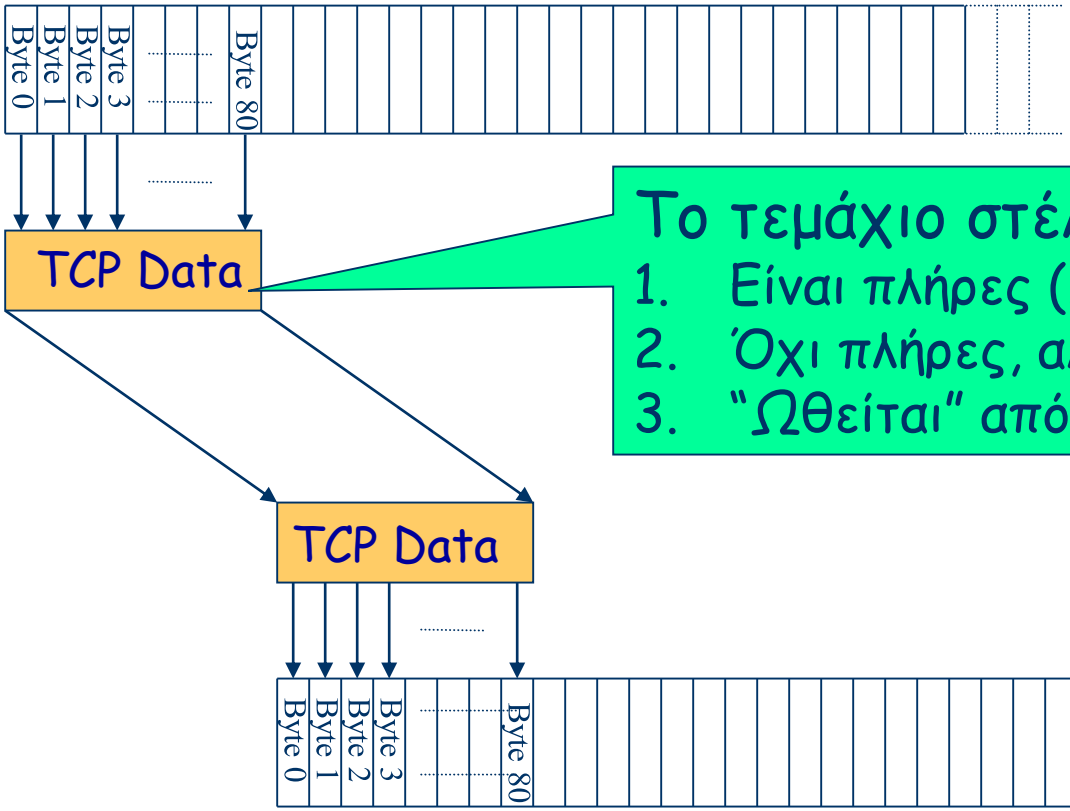


Host B

TCP

Υπηρεσία συρμού byte

Host A



Το τεμάχιο στέλνεται όταν:

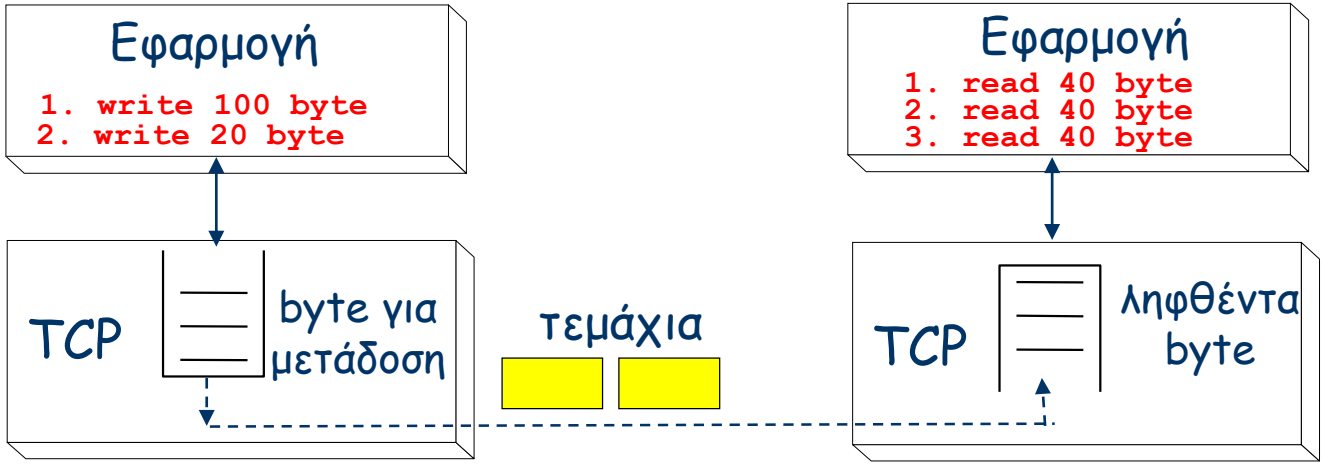
1. Είναι πλήρες (MSS byte),
2. Όχι πλήρες, αλλά λήγει ο χρόνος
3. "Ωθείται" από την εφαρμογή.

Host B

TCP

Υπηρεσία συρμού byte

- Στα κατώτερα στρώματα, το TCP παραδίδει δεδομένα σε τμήματα, τα τεμάχια.
- Στα ανώτερα στρώματα, το TCP παραδίδει δεδομένα ως ακολουθία από byte και δεν καθορίζει όρια μεταξύ των byte.
- Συνεπώς, τα ανώτερα στρώματα δεν γνωρίζουν την αρχή και το τέλος των τεμαχίων!

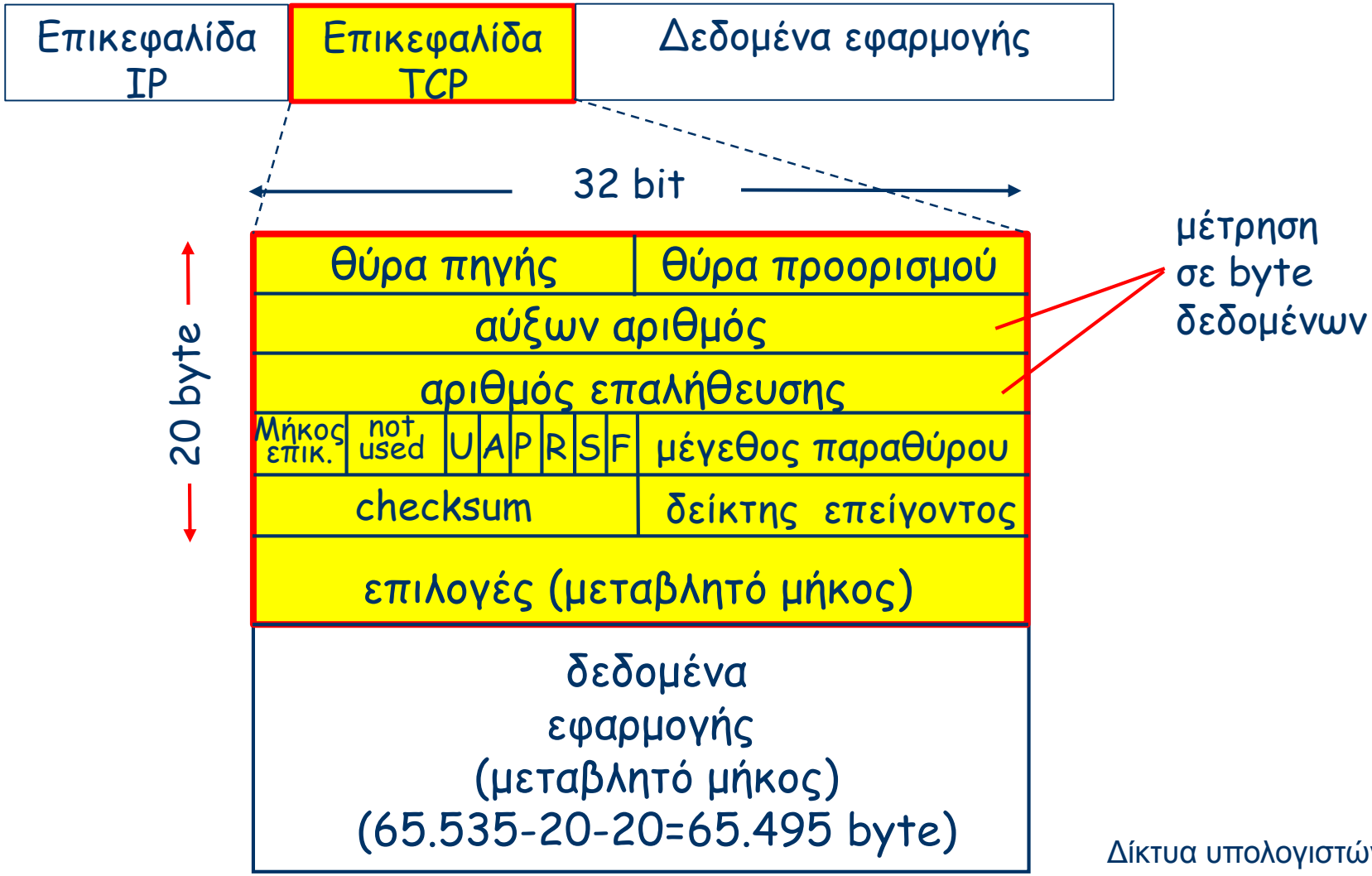


MSS: Maximum segment size

- MSS είναι ο μέγιστος αριθμός δεδομένων του στρώματος εφαρμογής που περιέχονται στο τεμάχιο και όχι το μέγιστο μέγεθος του τεμαχίου.
- Εξαρτάται από την υλοποίηση του TCP (καθοριζόμενη από το OS) και μπορεί να επιλεγεί.
- Συνήθεις τιμές: 1500, 536, 512 byte.
- Επιλέγεται το μέγιστο μέγεθος κατά τρόπο που να αποφεύγεται ο θρυμματισμός στο IP.

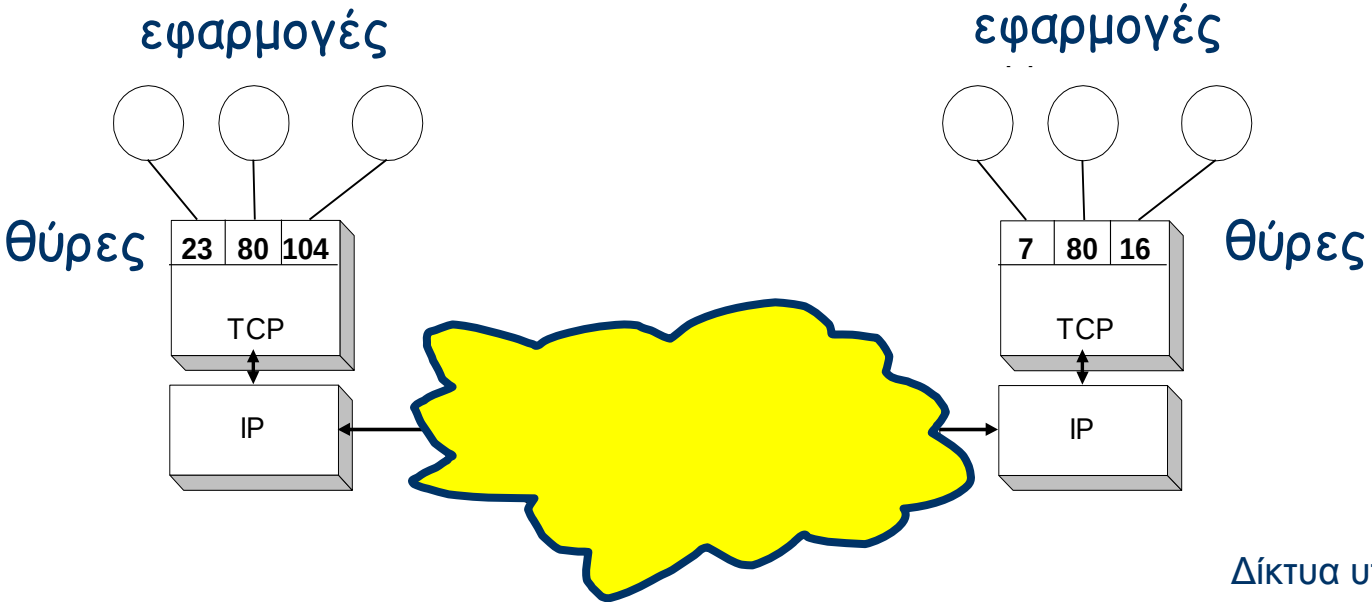
TCP

Δομή τεμαχίου



Αριθμοί Θυρών

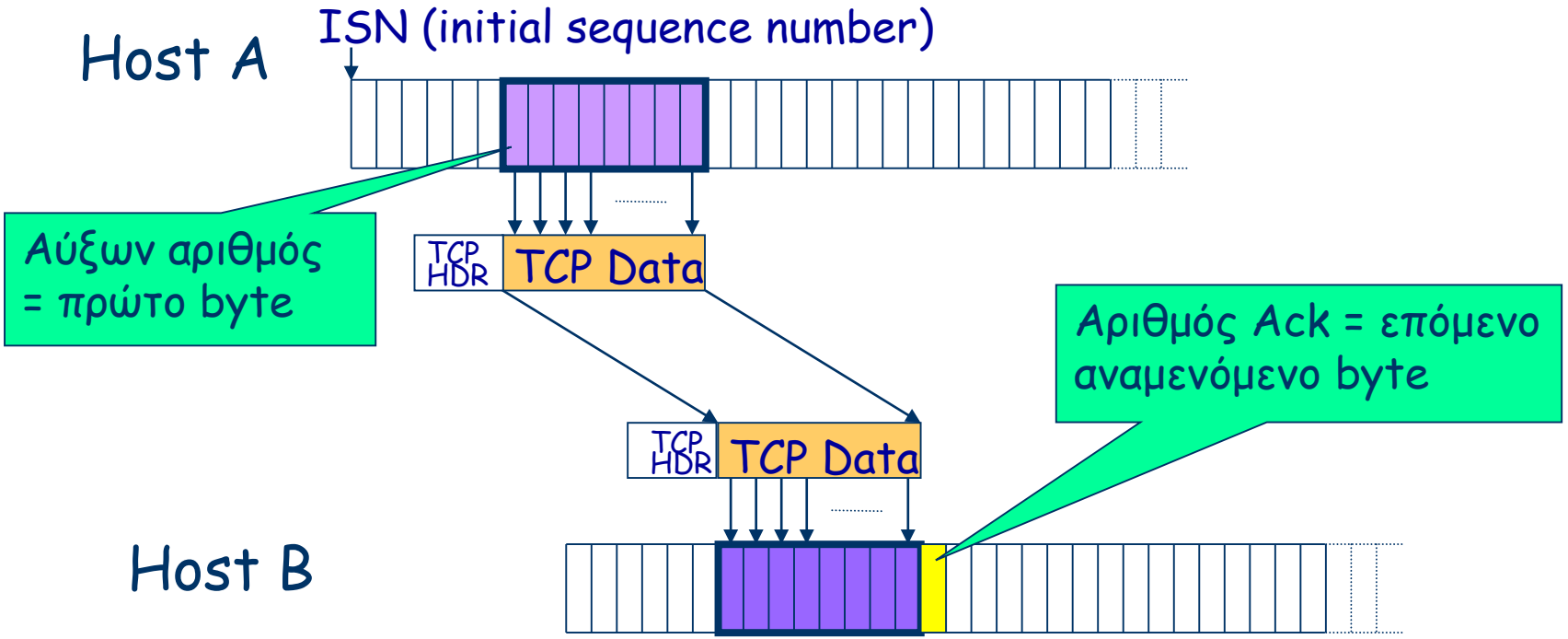
- Ο αριθμός θύρας προσδιορίζει την υποδοχή μιας εφαρμογής.
- Ένα ζεύγος (IP address, port number) προσδιορίζει το ένα άκρο μια σύνδεσης.
- Δύο ζεύγη (client IP address, client port number) και (server IP address, server port number) προσδιορίζουν μια σύνδεση TCP.



TCP

Αύξοντες αριθμοί και επαληθεύσεις

- Οι αύξοντες αριθμοί και οι επαληθεύσεις στο TCP έχουν μήκος 32 bit.
- Η περιοχή τιμών είναι 0 ≤ Sequence number ≤ 2^{32} -1 ≈ 4.3 Gbyte
- Ο client και ο server επιλέγουν ο καθένας τους τον ISN τυχαία κατά την εγκατάσταση της σύνδεσης.



TCP

Αριθμός επαλήθευσης

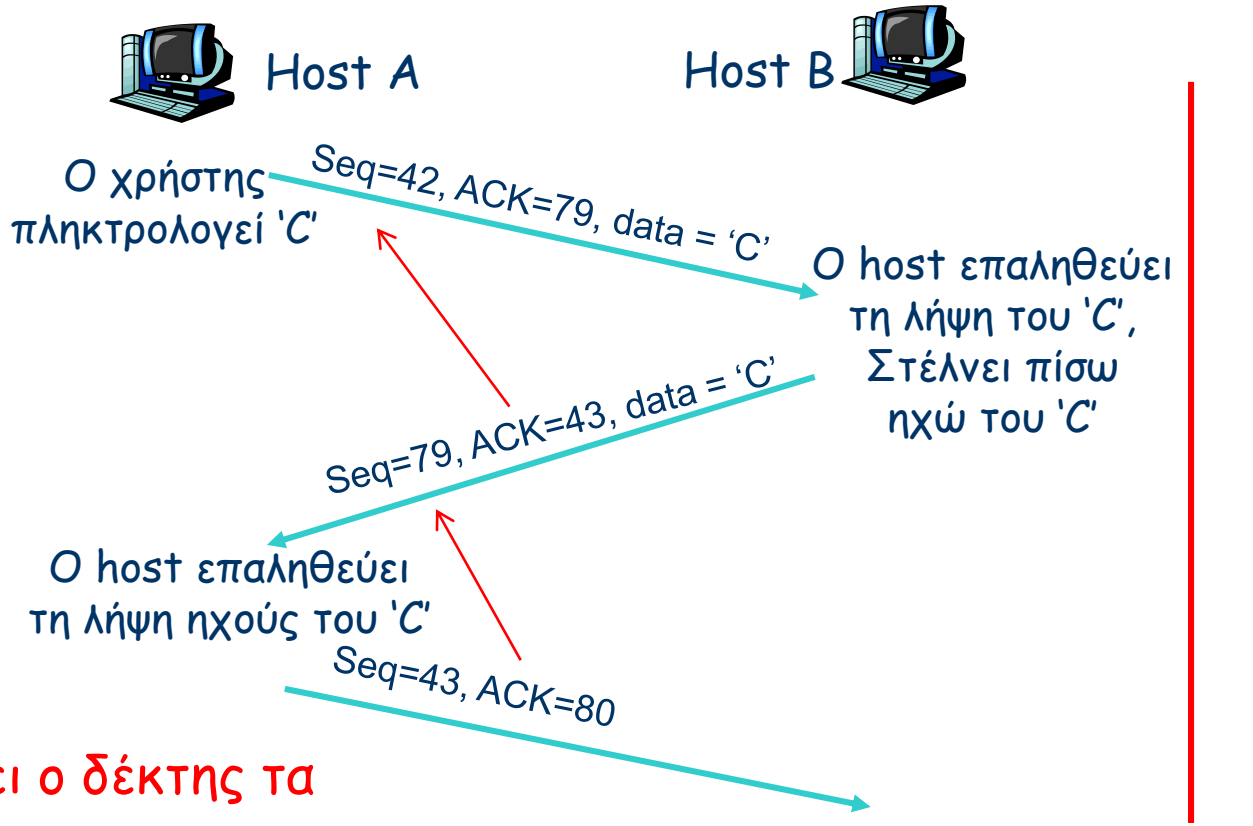
- Το TCP χρησιμοποιεί παραλλαγή του πρωτοκόλλου ολισθαίνοντος παραθύρου για τον έλεγχο ροής μεταξύ πομπού και δέκτη:
  - Όχι NACK (**N**egative **ACK**nowledgement).
  - Μόνο συσσωρευτικές ACK.
- Παράδειγμα:

  Ο πομπός στείλει δύο τεμάχια με "1...1500" και "1501...3000", αλλά ο δέκτης λαμβάνει μόνο το δεύτερο.

  Ο δέκτης δεν μπορεί να επαληθεύσει το δεύτερο τεμάχιο. Μπορεί να στείλει μόνο AckNo = 1.

TCP

Αύξοντες αριθμοί: παράδειγμα Telnet

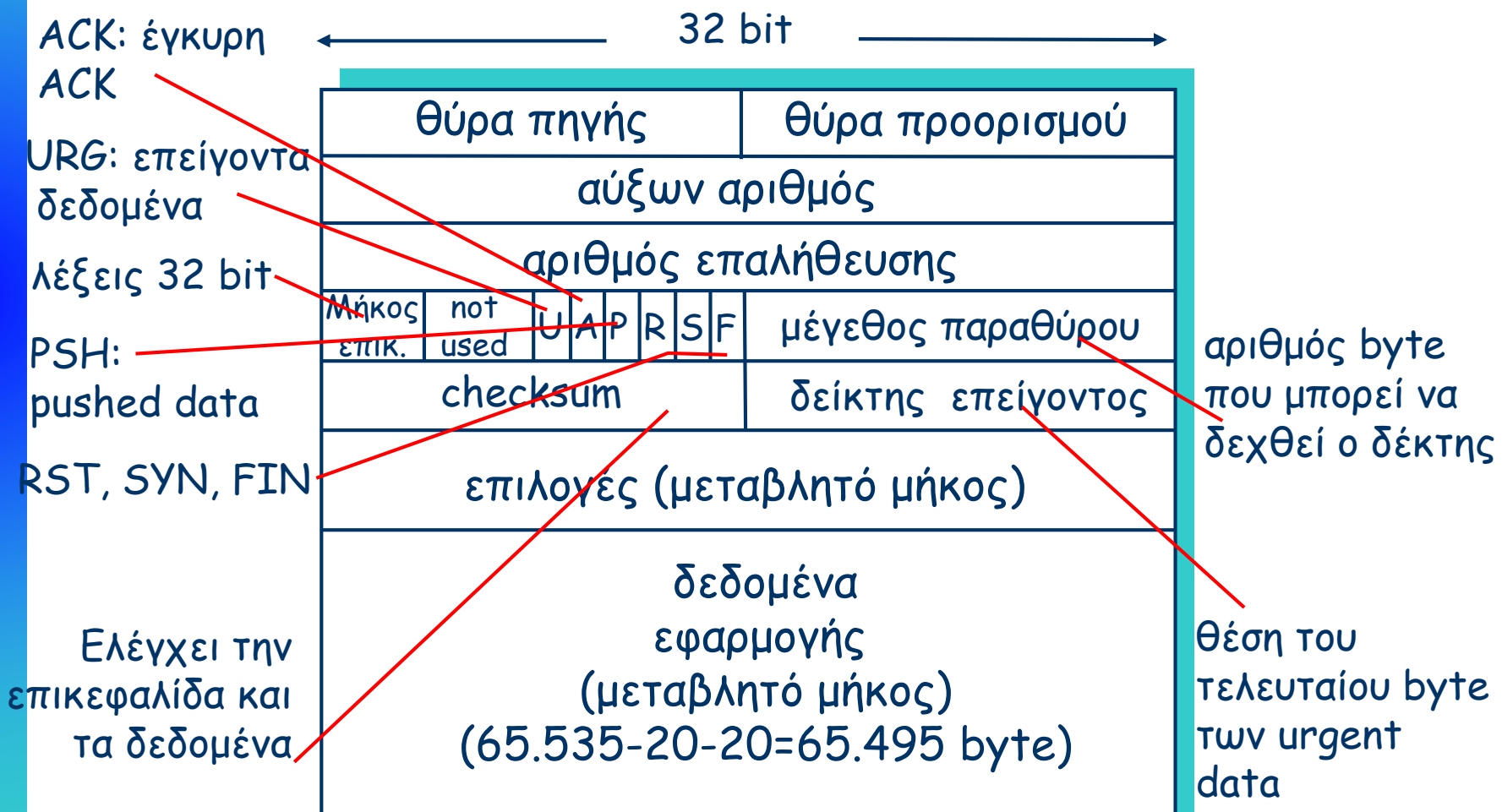


απλό σενάριο telnet

- Πώς αντιμετωπίζει ο δέκτης τα εκτός σειράς τεμάχια;
  - Οι προδιαγραφές του TCP δεν αναφέρουν τίποτε. Εξαρτάται από τον κατασκευαστή.

TCP

Δομή τεμαχίου



Προαιρετικές επιλογές (options)

- Είναι ένας τρόπος να προστεθούν επιπλέον δυνατότητες που δεν καλύπτονται από την κανονική επικεφαλίδα.

End of Options	kind=0			
	1 byte			
NOP (no operation)	kind=1			
	1 byte			
Maximum Segment Size	kind=2	len=4	maximum segment size	
	1 byte	1 byte	2 bytes	
Window Scale Factor	kind=3	len=3	shift count	
	1 byte	1 byte	1 byte	
Timestamp	kind=8	len=10	timestamp value	timestamp echo reply
	1 byte	1 byte	4 bytes	4 bytes

Χρησιμοποιείται για παραγέμισμα της επικεφαλίδας TCP ώστε να είναι πολλαπλάσιο των 4 byte.

TCP

Αξιόπιστη μετάδοση

- Το TCP δημιουργεί υπηρεσία αξιόπιστης μετάδοσης πάνω από την αναξιόπιστη υπηρεσία του IP.
- Στέλνει τεμάχια με συνεχή παροχή.
- Δύο τύποι σφαλμάτων:
  - Απωλεσθέντα τεμάχια
  - Κατεστραμμένα τεμάχια
- Το TCP έχει αθροίσματα ελέγχου για την επικεφαλίδα και τα δεδομένα. Τεμάχια με μη έγκυρα αθροίσματα ελέγχου απορρίπτονται.
- Ο δέκτης στέλνει επαληθεύσεις (ACK) για τα σωστά τεμάχια. Οι ACK μπορεί να είναι συσσωρευτικές.

TCP

Αξιόπιστη μετάδοση

- Το TCP χρησιμοποιεί μοναδικό χρονόμετρο επαναμετάδοσης.
- Οι επαναμεταδόσεις των τεμαχίων προκαλούνται από λήξεις χρόνου ή από διπλές ACK.
- Ο αριθμός της ACK είναι ο επόμενος αναμενόμενος αύξων αριθμός.
- Καθυστερημένη ACK: ο δέκτης TCP συνήθως καθυστερεί τη μετάδοση μιας ACK (για περίπου 200ms).
- Οι ACK δεν καθυστερούνται όταν τα πακέτα λαμβάνονται εκτός σειράς.

TCP: Αξιόπιστη μετάδοση

Απλοποιημένος πομπός TCP

Άφιξη δεδομένων από το στρώμα εφαρμογής:

- Δημιουργία τεμαχίου με seq #
- Ο seq # αντιστοιχεί στο πρώτο byte του συρμού δεδομένων.
- Εκκίνηση χρονομέτρου, εάν δεν ξεκίνησε ήδη (χρονόμετρο για προηγούμενο ανεπιβεβαίωτο τεμάχιο).
- Χρόνος εκπνοής: `TimeOutInterval`

Εκπνοή χρόνου:

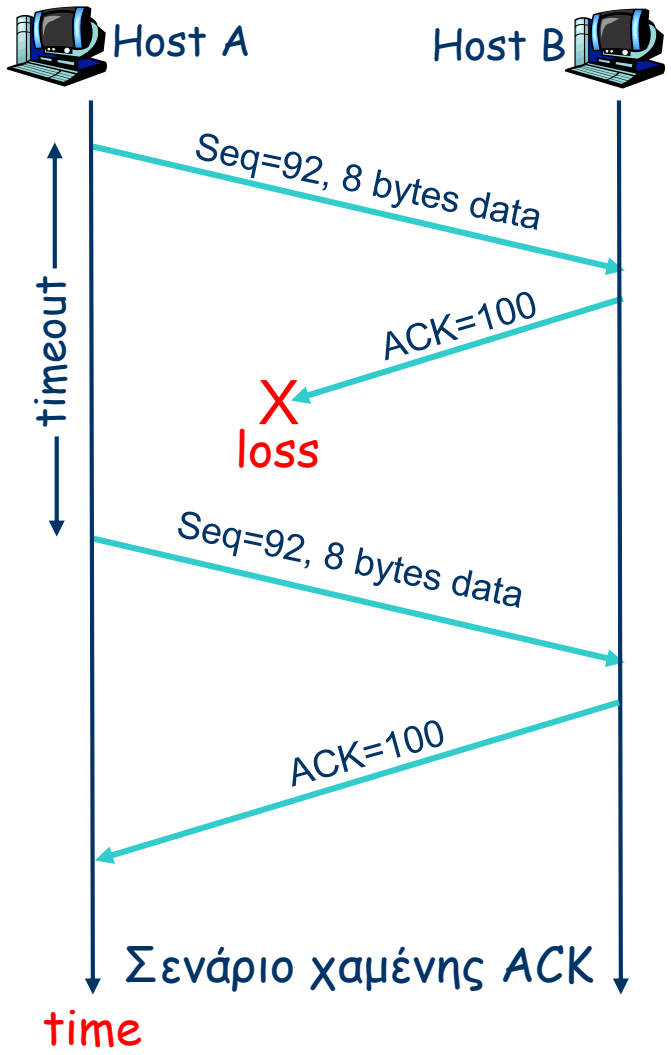
- Επανεκπομπή του τεμαχίου που προκάλεσε timeout.
- Επανεκκίνηση του timer.

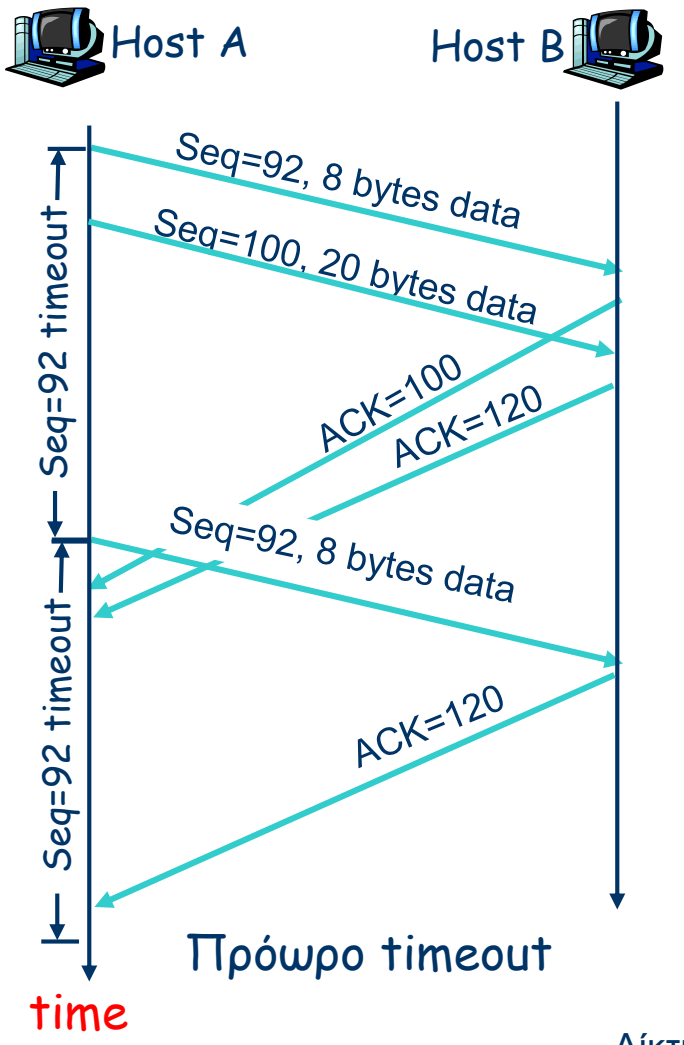
Λήψη Ack:

- Αν επαληθεύει προγούμενα ανεπαλήθευτα τεμάχια:
  - Ενημέρωση των ήδη επαληθευθέντων,
  - Εκκίνηση του timer αν υπάρχουν εκκρεμούντα τεμάχια.

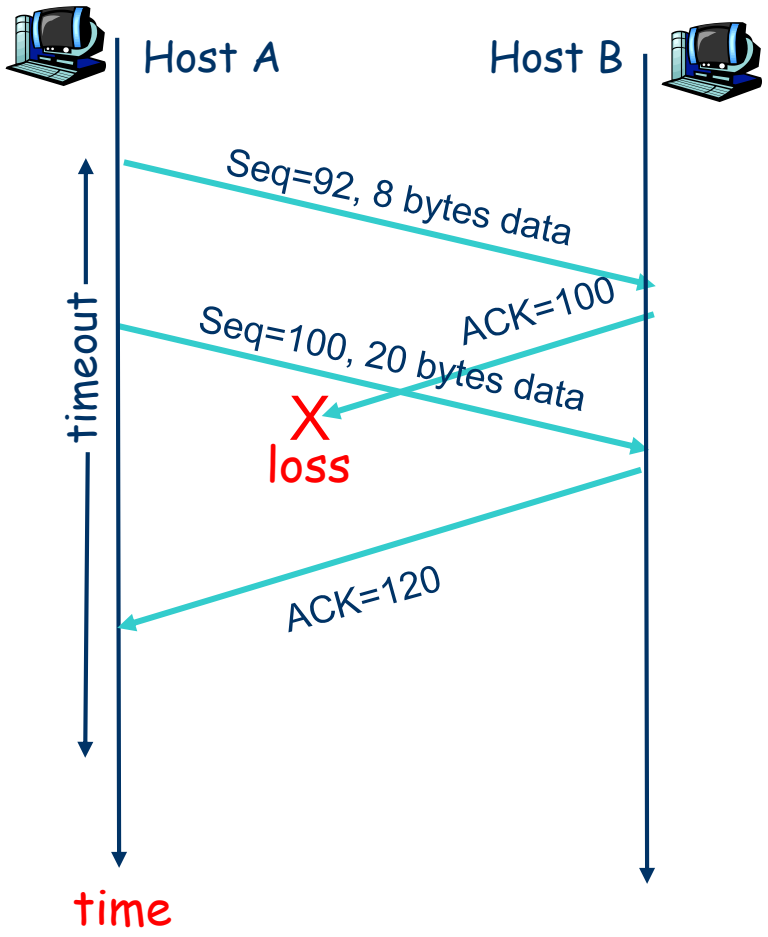
TCP: Αξιόπιστη μετάδοση

Σενάρια επαναμετάδοσης





TCP: Αξιόπιστη μετάδοση



Σενάριο συγκεντρωτικής ACK

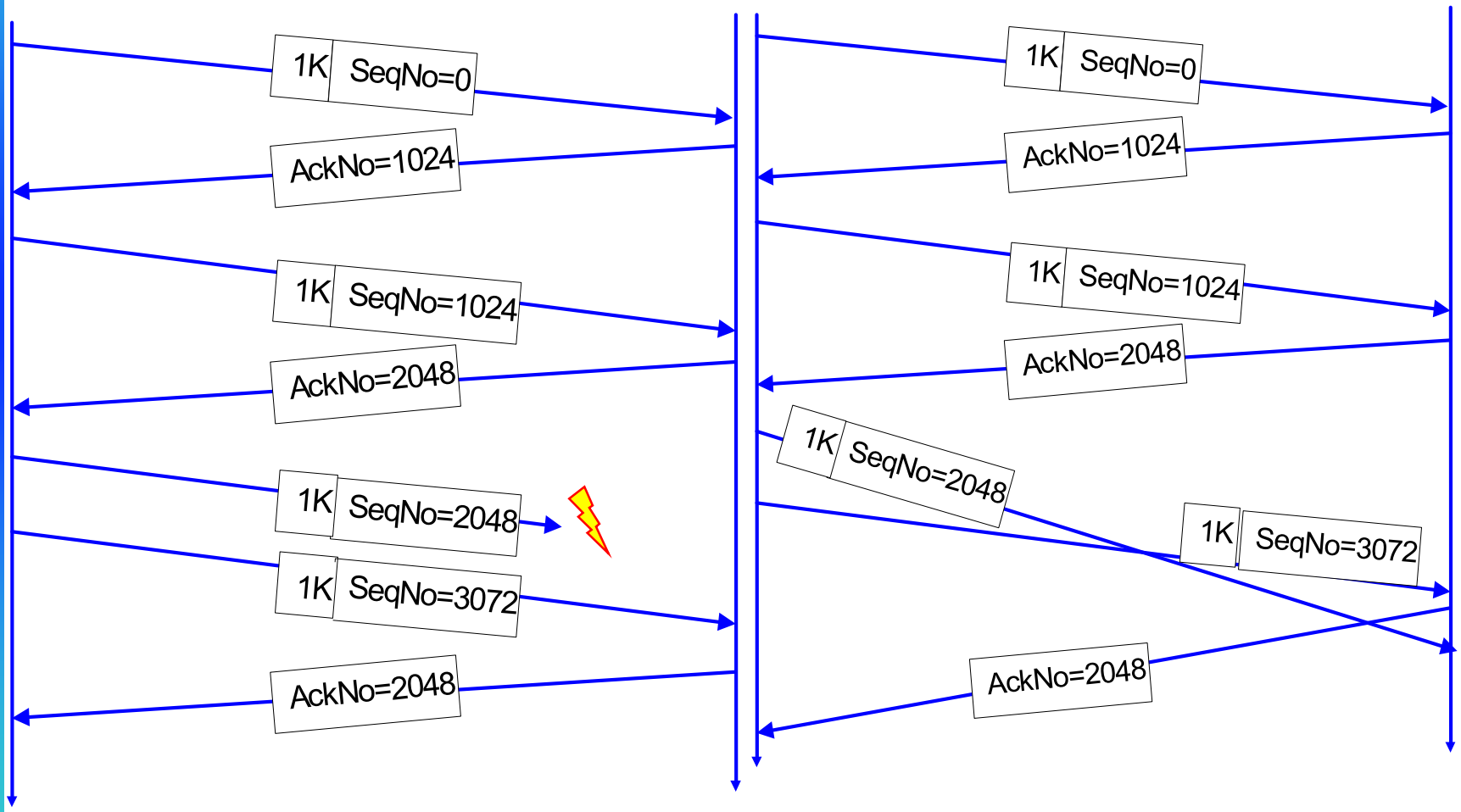
TCP: Αξιόπιστη μετάδοση

Δημιουργία επαληθεύσεων

Γεγονός	Ενέργεια του δέκτη TCP
Άφιξη τεμαχίου στην κανονική σειρά, όχι κενά, οτιδήποτε άλλο έχει ήδη επαληθευτεί.	Καθυστερημένη ACK. Αναμονή 200ms για το επόμενο τεμάχιο. Αν δεν υπάρχει επόμενο τεμάχιο, στέλνει ACK.
Άφιξη τεμαχίου στην κανονική σειρά, όχι κενά, εκκρεμεί μία καθυστερημένη ACK.	Άμεση αποστολή μιας συσσωρευτικής ACK και για τα δύο τεμάχια που αφίχθηκαν με κανονική σειρά.
Άφιξη τεμαχίου εκτός σειράς με μεγαλύτερο αύξοντα αριθμό από τον αναμενόμενο.	Αποστολή επαναληπτικής ACK, που να δείχνει τον αύξοντα αριθμό του επόμενου αναμενόμενου byte.
Άφιξη τεμαχίου που εν μέρει ή πλήρως συμπληρώνει κενό.	Άμεση αποστολή ACK, αν το τεμάχιο αρχίζει στο κατώτερο άκρο του κενού.

TCP: Αξιόπιστη μετάδοση

Επαληθεύσεις στο TCP



Απώλεια τεμαχίου

Άφιξη εκτός σειράς

TCP: Αξιόπιστη μετάδοση

Επαναμεταδόσεις στο TCP

Ένας πομπός TCP επαναμεταδίδει ένα τεμάχιο, όταν θεωρήσει ότι το υπόψη τεμάχιο έχει χαθεί:

- Δεν έχει ληφθεί ACK και έχει λήξει το χρονόμετρο.
- Έχουν ληφθεί πολλαπλές ACK για το ίδιο τεμάχιο.

TCP: Αξιόπιστη μετάδοση

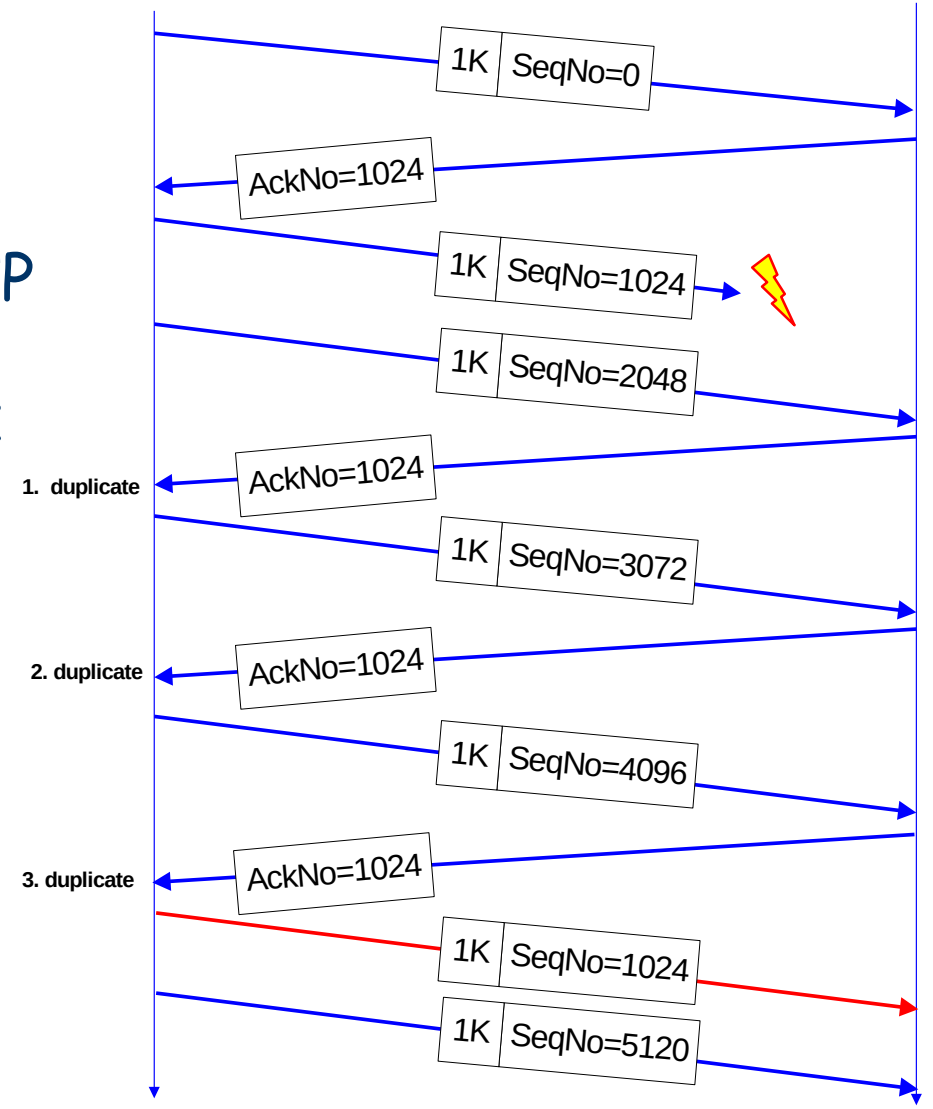
Ταχεία επαναμετάδοση

- Η περίοδος timeout είναι συχνά σχετικά μεγάλη:
  - Μεγάλη καθυστέρηση πριν την αποστολή του χαμένου πακέτου.
- Ανίχνευση των χαμένων τεμαχίων μέσω διπλών επαληθεύσεων.
  - Ο πομπός στέλνει συχνά πολλά τεμάχια το ένα πίσω απ' το άλλο.
  - Αν χαθεί τεμάχιο, θα υπάρχουν ενδεχομένως πολλές ίδιες επαληθεύσεις.
- Ταχεία επαναμετάδοση: επαναποστολή του τεμαχίου πριν τη λήξη της χρονομέτρησης.

TCP: Αξιόπιστη μετάδοση

Λήψη διπλών ACK

- Αν ληφθούν τρεις ACK στη σειρά για τα ίδια δεδομένα, ο πομπός TCP θεωρεί ότι το τεμάχιο μετά τα επαληθευόμενα δεδομένα χάθηκε.
- Τότε ο πομπός TCP επαναμεταδίδει το τεμάχιο που θεωρεί ότι χάθηκε, χωρίς να περιμένει τη λήξη χρόνου.
- Τούτο διορθώνει μεμονωμένες απώλειες τεμαχίων.



TCP: Αξιόπιστη μετάδοση

Αλγόριθμος ταχείας επαναμετάδοσης

γεγονός: λήψη ACK, με τιμή της ACK ίση με y

```
if (y > SendBase) {
    SendBase = y
    if (υπάρχουν τρέχοντα μη επαληθευθέντα ήδη τεμάχια)
        εκκίνηση χρονομετρητή
  }
else {
    αύξηση του μετρητή των διπλών ACKs που ελήφθησαν
            για το y
    if (μετρητής ληφθεισών διπλών ACKs για το y = 3)
        επανεκπομπή του τεμαχίου με αύξοντα αριθμό y
   }
```

διπλή ACK για ήδη επαληθευθέν τεμάχιο

Ταχεία επαναμετάδοση

SendBase-1: τελευταίο επαληθευθέν byte

TCP: Αξιόπιστη μετάδοση

Διαχείριση χρονομετρητών

- Το TCP χρησιμοποιεί πολλούς χρονομετρητές. Ο σπουδαιότερος είναι ο χρονομετρητής επαναμετάδοσης (retransmission timer).
- Όταν στέλνεται ένα τεμάχιο, ξεκινά ένας χρονομετρητής αναμετάδοσης.
- Αν η λήψη του τεμαχίου επαληθευτεί πριν εκπνεύσει ο χρόνος, τότε ο χρονομετρητής σταματά.
- Αν εκπνεύσει ο χρόνος πριν φθάσει η επαλήθευση, το τεμάχιο μεταδίδεται ξανά.
- Πόσο μεγάλο πρέπει να είναι το χρονικό διάστημα πριν λήξει η χρονομέτρηση;

TCP: Αξιόπιστη μετάδοση

Χρόνοι Round Trip και Timeout

Πώς τίθεται η τιμή του timeout επαναμετάδοσης (Retransmission Timeout, RTO) στο TCP;

- μεγαλύτερη από RTT;
  - αλλά το RTT μεταβάλλεται
- πολύ μικρή; ⇒ πρόωρο timeout
  - άσκοπες επαναμεταδόσεις
- πολύ μεγάλη; ⇒ αργή αντίδραση, όταν χάνεται τεμάχιο

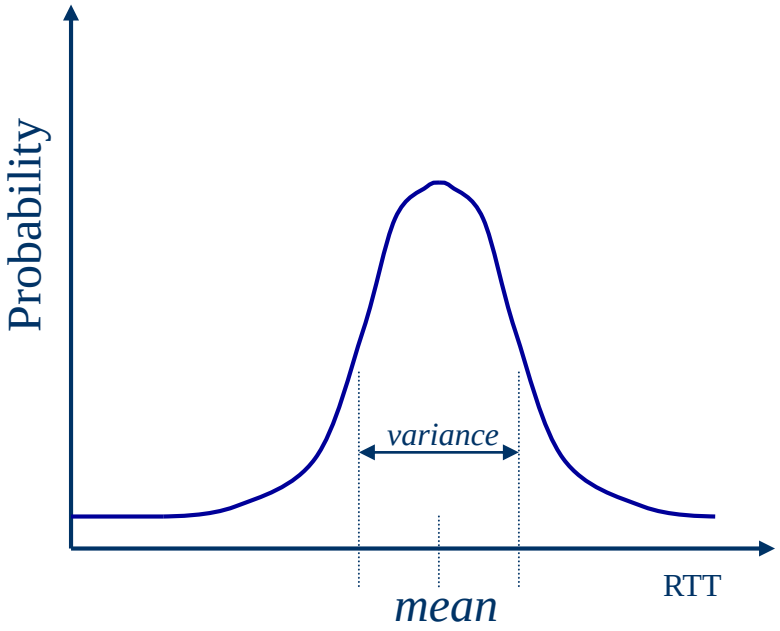
Πώς προσδιορίζεται το RTT;

- SampleRTT: ο χρόνος από τη μετάδοση του τεμαχίου μέχρι τη λήψη της ACK.
- Επειδή το SampleRTT μεταβάλλεται, είναι επιθυμητή μια εξομαλυμένη τιμή για το RTT, όχι το τρέχον SampleRTT.

TCP: Αξιόπιστη μετάδοση

Χρόνοι Round Trip και Timeout

- Υπάρχει κάποια (άγνωστη) κατανομή των RTT.
- Προσπαθούμε να εκτιμήσουμε ένα RTO για να ελαχιστοποιήσουμε την πιθανότητα μιας εσφαλμένης λήξης χρόνου.



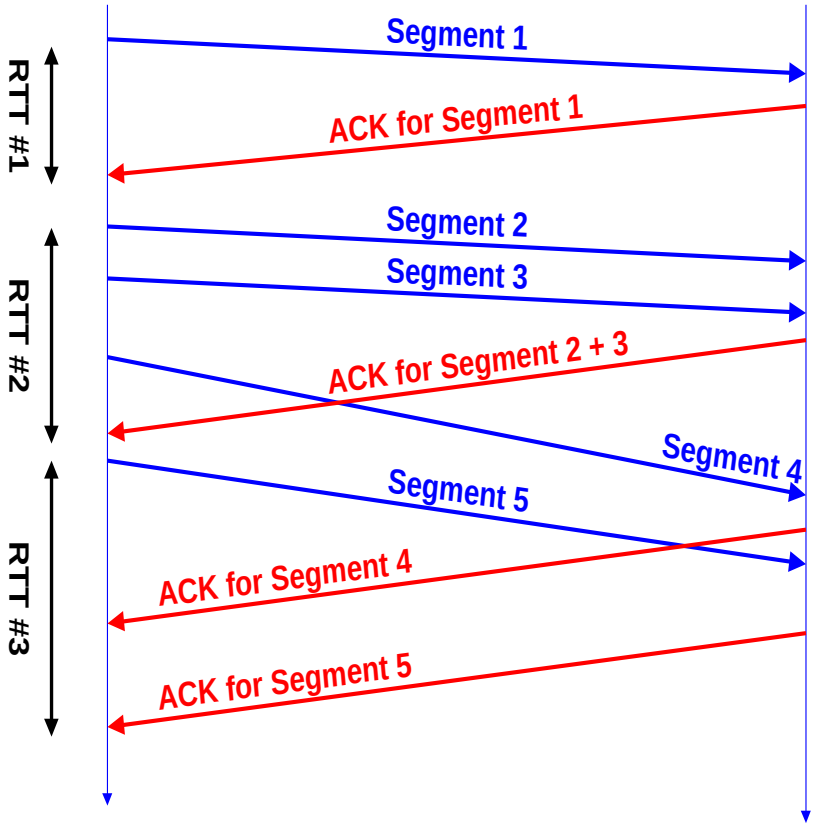
- Οι ουρές στους δρομολογητές μεγαλώνουν όταν υπάρχει περισσότερη κίνηση, μέχρι να γίνουν ασταθείς.
- Καθώς αυξάνει το φορτίο, η variance της καθυστέρησης αυξάνει απότομα.

TCP: Αξιόπιστη μετάδοση

Ρύθμιση της τιμής του Timeout

Η τιμή του RTO τίθεται βάσει των μετρήσεων του RTT που πραγματοποιεί το TCP.

- Κάθε σύνδεση TCP μετράει τη χρονική διαφορά μεταξύ της αποστολής ενός τεμαχίου και της λήψης της αντίστοιχης ACK.
- Υπάρχει μόνο μια μέτρηση σε ισχύ κάθε φορά (δηλ., οι μετρήσεις δεν επικαλύπτονται).
- Στο διπλανό σχήμα φαίνονται τρεις μετρήσεις RTT.



TCP: Αξιόπιστη μετάδοση

Ρύθμιση της τιμής του Timeout

- Η RTO υπολογίζεται βάσει των μετρήσεων του *RTT.*
  - Χρησιμοποιείται εκθετικός σταθμισμένος κινούμενος μέσος όρος (*srtt*) για την εκτιμώμενη καθυστέρηση και τη variance (*rttvar*) της καθυστέρησης.

- Οι μετρήσεις RTT εξομαλύνονται ως εξής:

$$srtt_{n+1} = \alpha \ SampleRTT + (1-\alpha)\, srtt_n$$

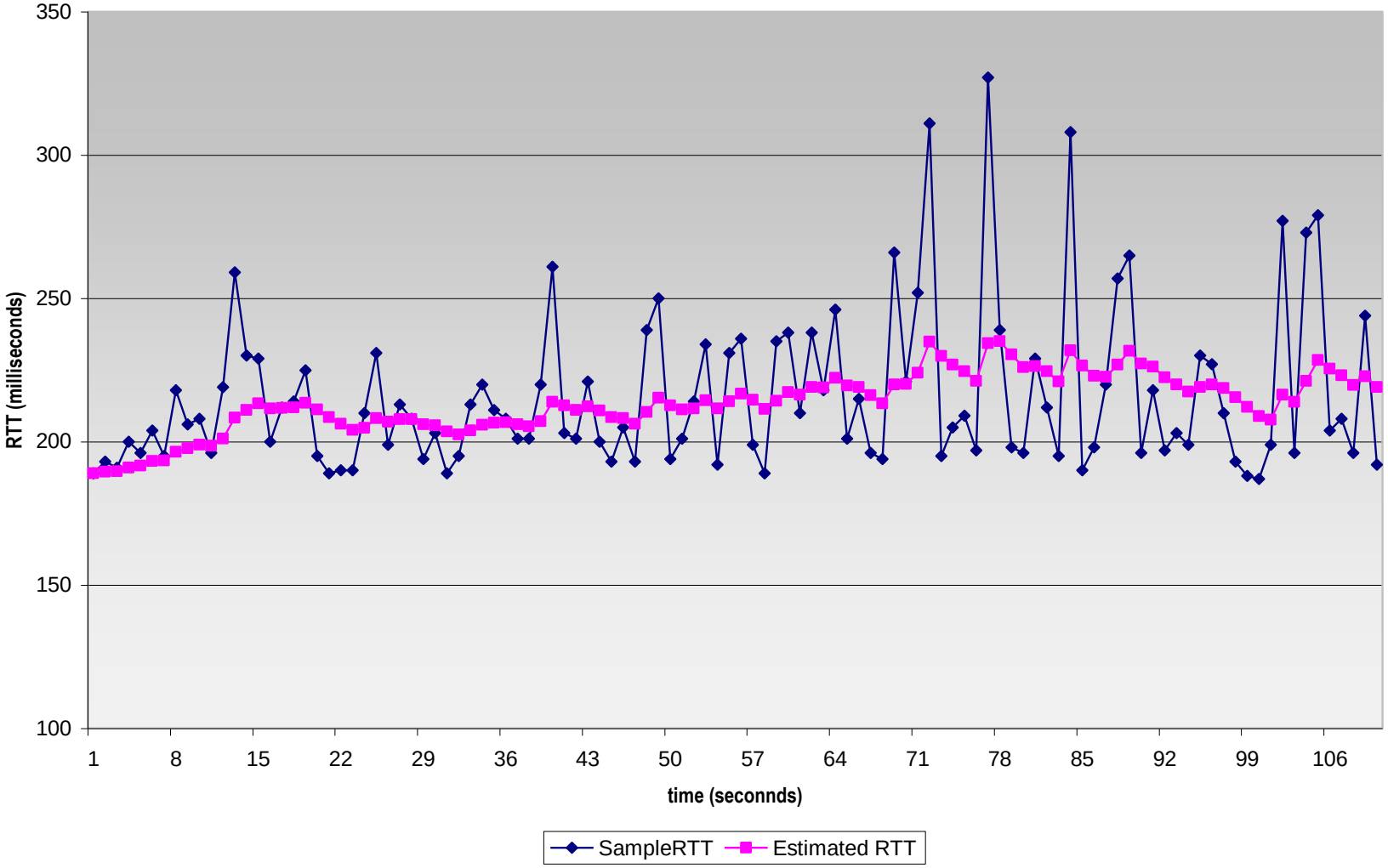
$$rttvar_{n+1} = \beta\,(\,|\,SampleRTT - srtt_n\,|\,) + (1-\beta)\, rttvar_n$$

- Οι τιμές του α =1/8 και β =1/4

$$RTO_{n+1} = srtt_{n+1} + 4\, rttvar_{n+1}$$

TCP: Αξιόπιστη μετάδοση

Ρύθμιση της τιμής του Timeout



TCP: Αξιόπιστη μετάδοση

Ρύθμιση της τιμής του Timeout

- Αρχική τιμή του RTO:
  - Ο πομπός θέτει την αρχική τιμή του RTO:

$$RTO_0 = 3\ sec$$

- Υπολογισμός του RTO μετά την πρώτη μέτρηση του RTT:

$$srtt_1 = SampleRTT$$

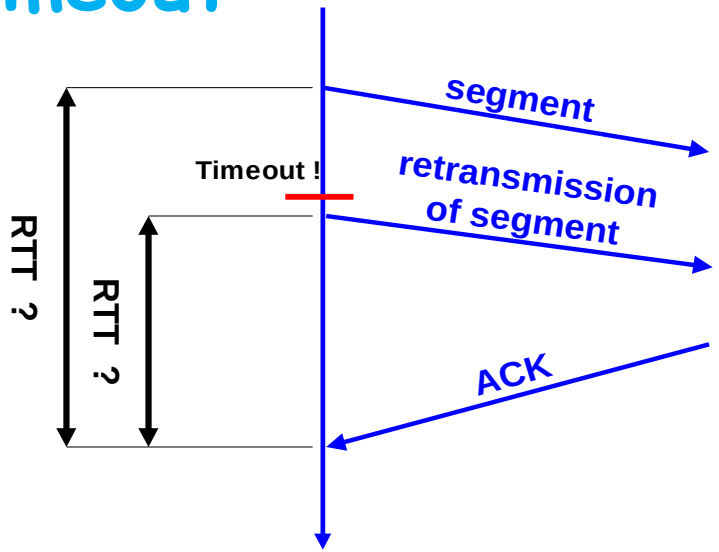
$$rttvar_1 = SampleRTT / 2$$

$$RTO_1 = srtt_1 + 4\ rttvar_1$$

TCP: Αξιόπιστη μετάδοση

Ρύθμιση της τιμής του Timeout

Αν ληφθεί ACK για τεμάχιο που επαναμεταδόθηκε, ο πομπός δεν μπορεί να ξέρει αν η ACK ανήκει στο αρχικό ή στο τεμάχιο που επαναμεταδόθηκε.



- **Αλγόριθμος του Karn**:
  - Μην ενημερώνεις την *RTT* για τεμάχια που επαναμεταδόθηκαν.
  - Ξαναξεκίνα τις μετρήσεις RTT μόνο μετά τη λήψη ACK που αφορά κανονικό τεμάχιο.
  - Όταν εμφανιστεί ένα timeout, η τιμή του RTO διπλασιάζεται (*εκθετική οπισθοχώρηση*)

$$RTO_{n+1} = min\,(\,2\,RTO_n\,,\,64)\ seconds\,.$$

TCP: Διαχείριση συνδέσεων

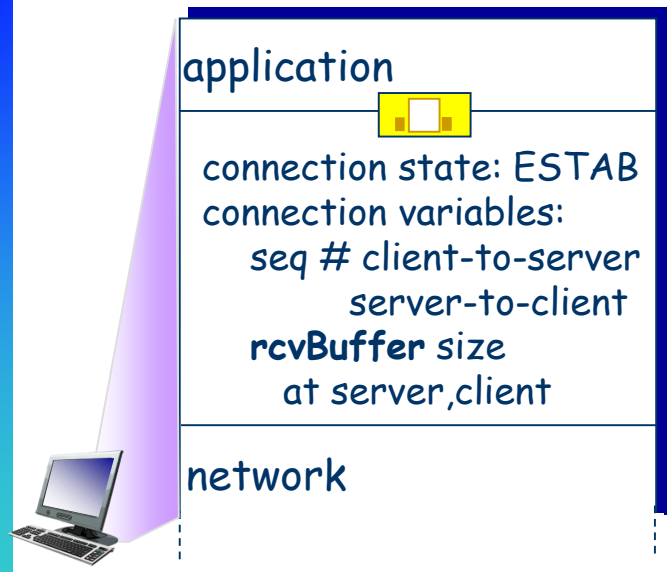
- Εγκατάσταση σύνδεσης
- Απόλυση σύνδεσης
- Ειδικά σενάρια
- Διαγράμματα καταστάσεων

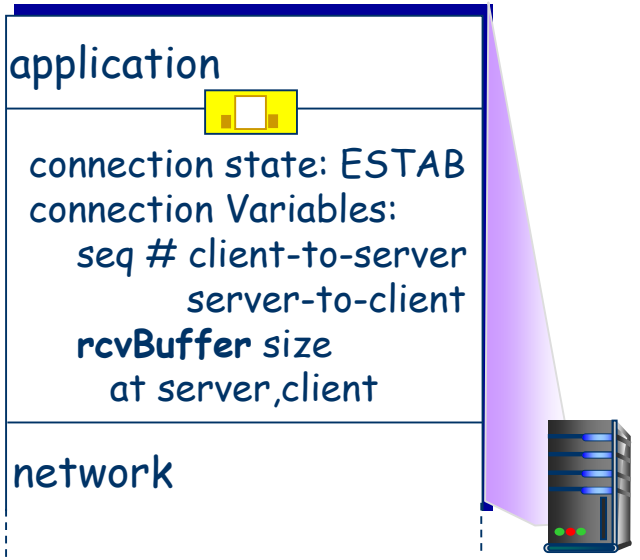
TCP: Διαχείριση συνδέσεων

Εγκατάσταση σύνδεσης

Πριν ανταλλάξουν δεδομένα, ο πομπός και ο δέκτης πραγματοποιούν τριμερή χειραψία.

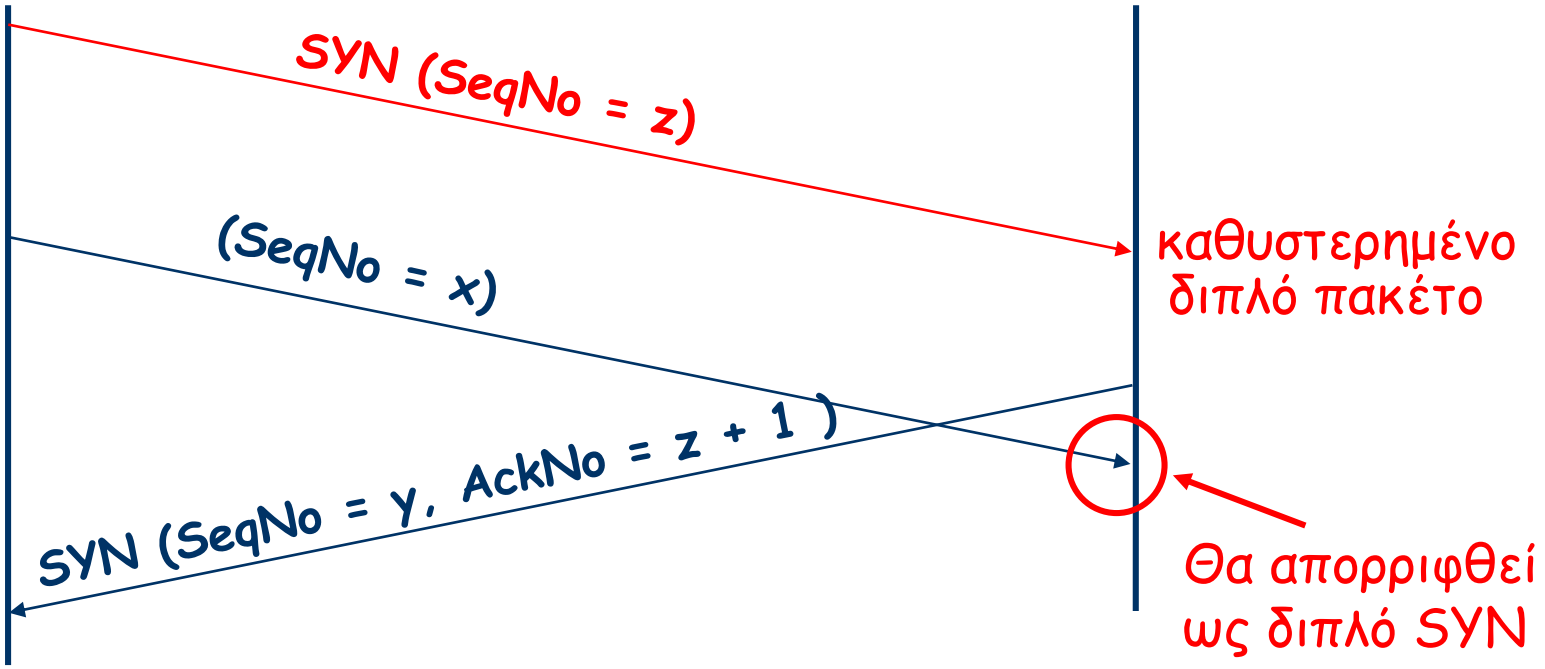
- Συμφωνούν για την εγκατάσταση της σύνδεσης (καθένας γνωρίζει ότι ο άλλος επιθυμεί την εγκατάσταση σύνδεσης).
- Συμφωνούν για τις παραμέτρους σύνδεσης.





TCP: Διαχείριση συνδέσεων

Γιατί δεν αρκεί η διμερής χειραψία



Όταν ο client αρχίζει τη μετάδοση δεδομένων (ξεκινώντας με SeqNo = x+1), ο server θα απορρίψει όλα τα δεδομένα.

TCP: Διαχείριση συνδέσεων

Τριμερής χειραψία

Κατάσταση client

LISTEN

επιλογή αύξ. αριθμ. x
αποστολή μηνύματος TCP SYN

SYNSENT

SYNbit=1, Seq=x

επιλογή αύξ. αριθμ. y
αποστολή μηνύματος
TCP SYNACK

SYNbit=1, Seq=y
ACKbit=1; ACKnum=x+1

αποστολή ACK για SYNACK;
αυτό το τεμάχιο μπορεί να έχει
και δεδομένα προς τον server

ESTAB

ACKbit=1, ACKnum=y+1

λήψη ACK(y)

Κατάσταση server

LISTEN

SYN RCVD

ESTAB

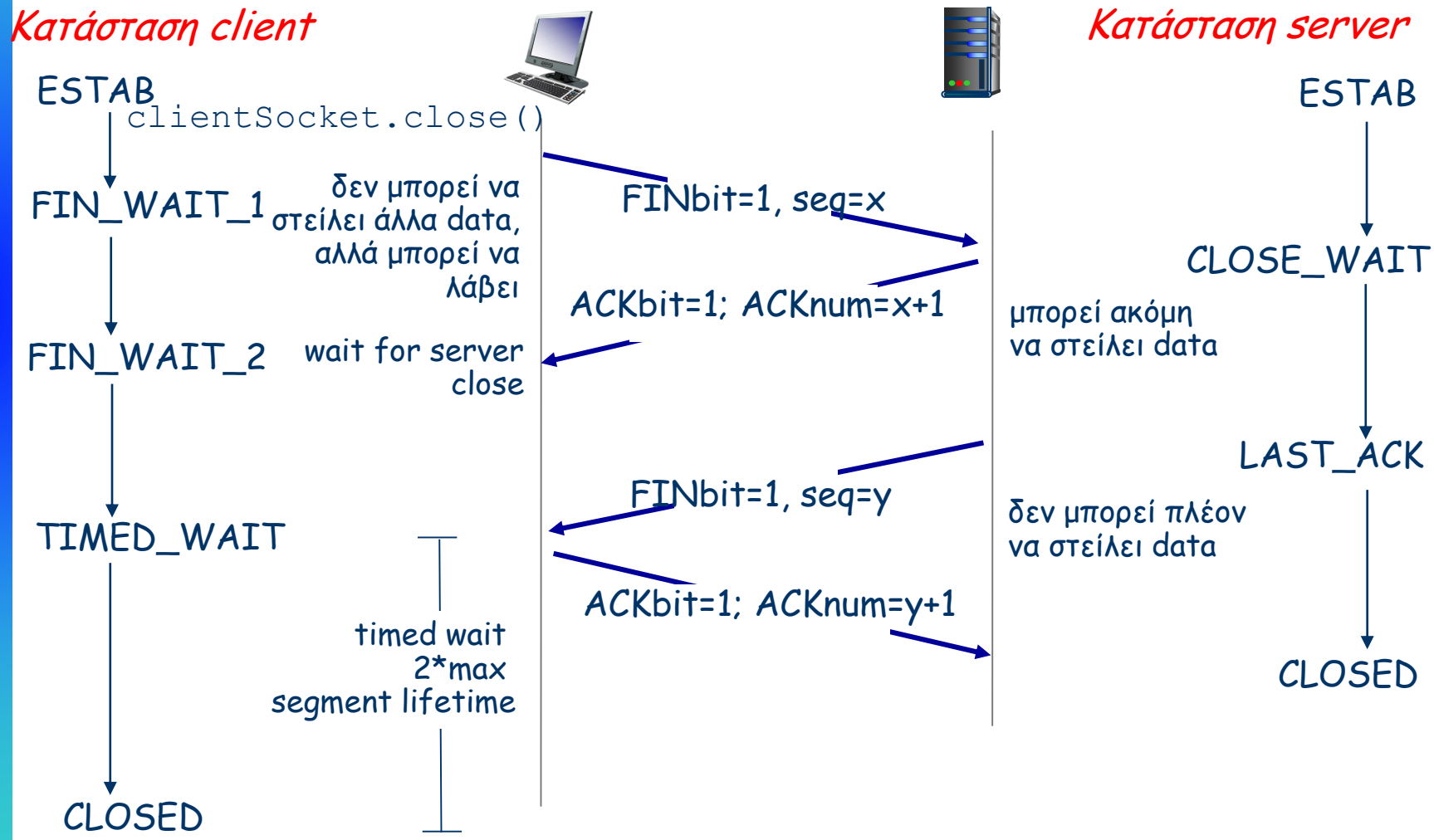
TCP: Διαχείριση συνδέσεων

Απόλυση σύνδεσης

- Κάθε άκρο της ροής δεδομένων πρέπει να τερματίσει ανεξάρτητα (**"half-close"**):
  - Αποστολή τεμαχίου TCP με το FIN bit = 1.
- Απάντηση στο λαμβανόμενο FIN με ACK.
  - Μετά τη λήψη FIN, το αποστελλόμενο ACK μπορεί να συνδυάζεται με αποστολή FIN.
- Οι ταυτόχρονες ανταλλαγές FIN πρέπει να αντιμετωπίζονται.

TCP: Διαχείριση συνδέσεων

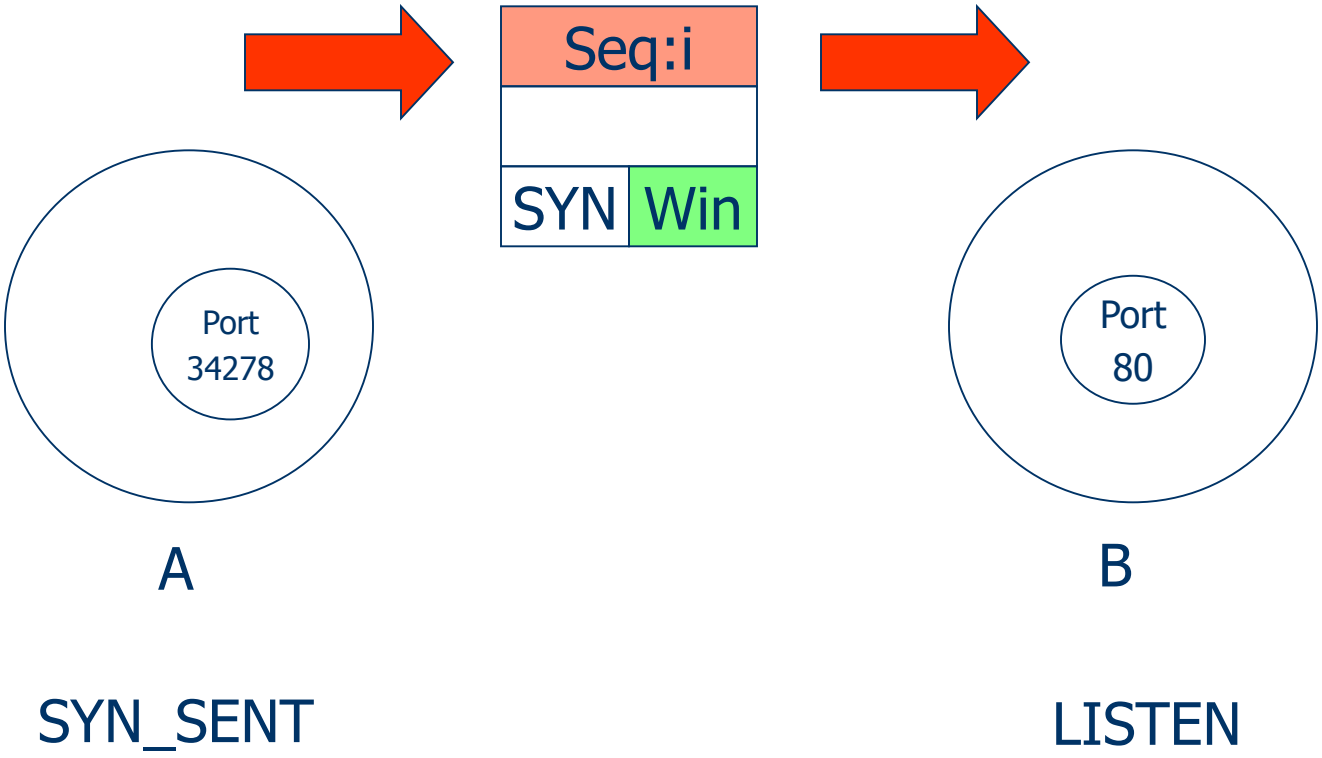
Απόλυση σύνδεσης



TCP: Διαχείριση συνδέσεων

Παράδειγμα εγκατάστασης σύνδεσης

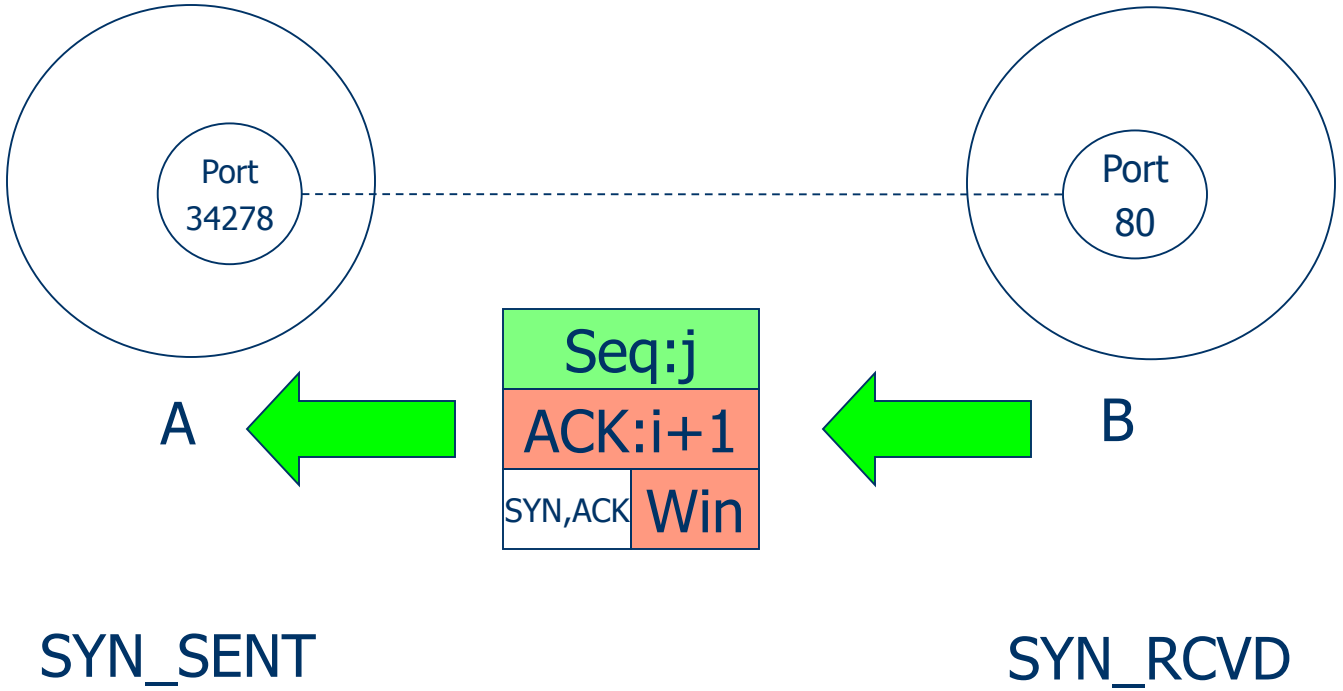
- Εγκατάσταση σύνδεσης: Βήμα 1



TCP: Διαχείριση συνδέσεων

Παράδειγμα εγκατάστασης σύνδεσης

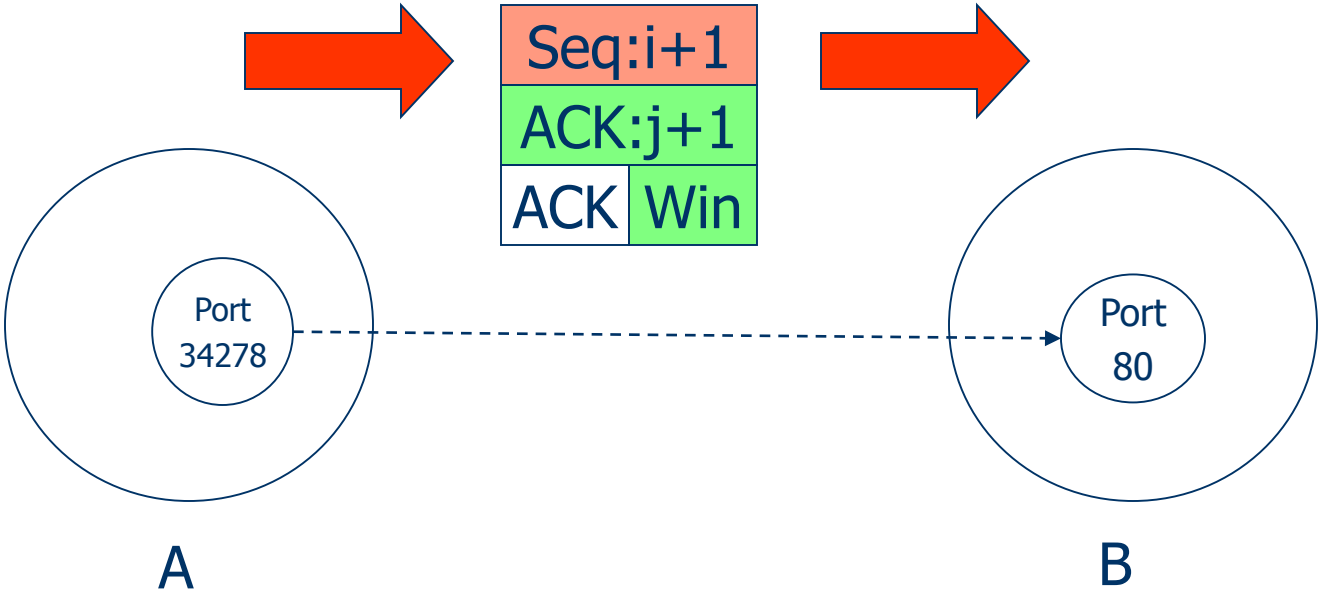
- Εγκατάσταση σύνδεσης: Βήμα 2



TCP: Διαχείριση συνδέσεων

Παράδειγμα εγκατάστασης σύνδεσης

- Εγκατάσταση σύνδεσης: Βήμα 3



ESTABLISHED

SYN_RCVD

TCP: Διαχείριση συνδέσεων

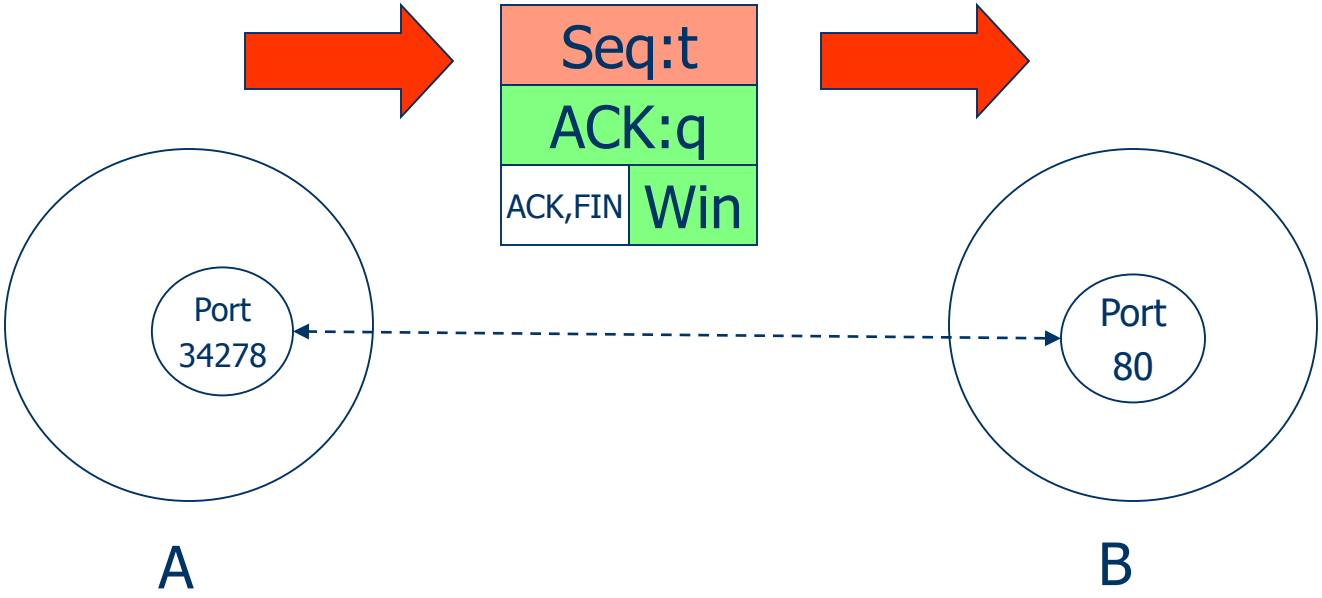
Παράδειγμα εγκατάστασης σύνδεσης

- Και οι δύο στην κατάσταση ESTABLISHED, ανταλλαγή δεδομένων...

TCP: Διαχείριση συνδέσεων

Παράδειγμα απόλυσης σύνδεσης

- Απόλυση σύνδεσης: Βήμα 1



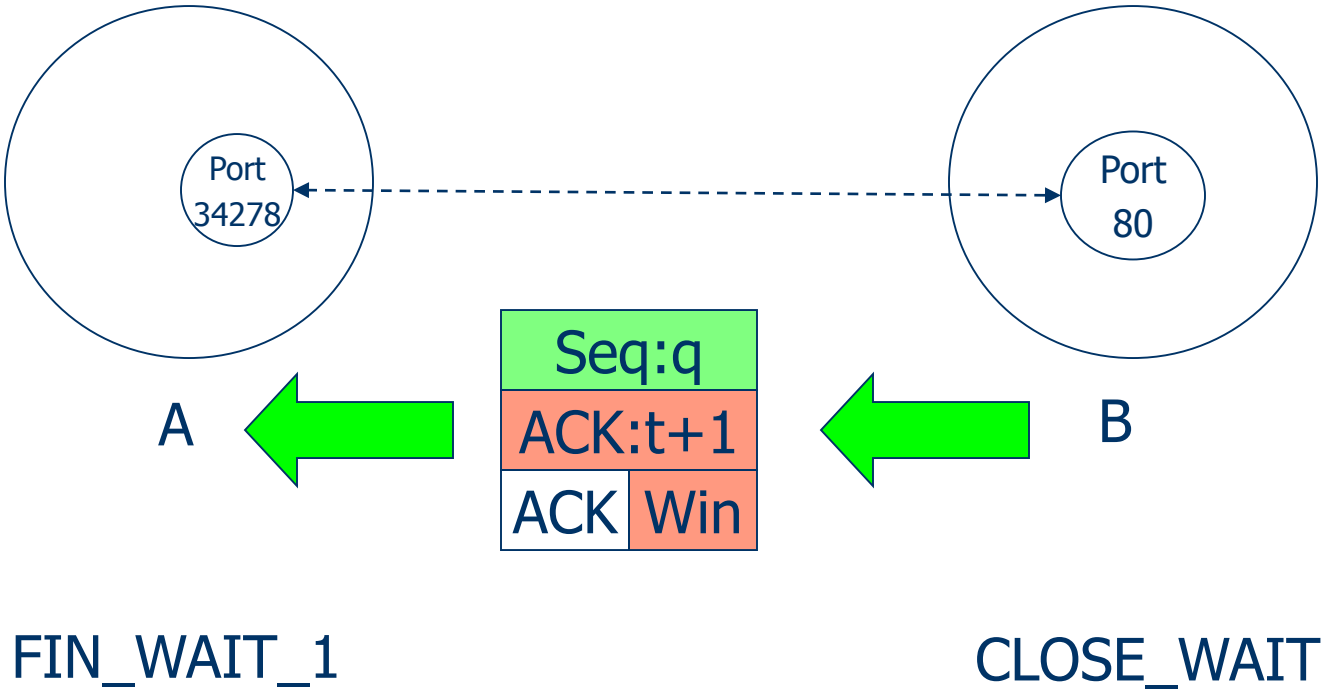
FIN_WAIT_1

ESTABLISHED

TCP: Διαχείριση συνδέσεων

Παράδειγμα απόλυσης σύνδεσης

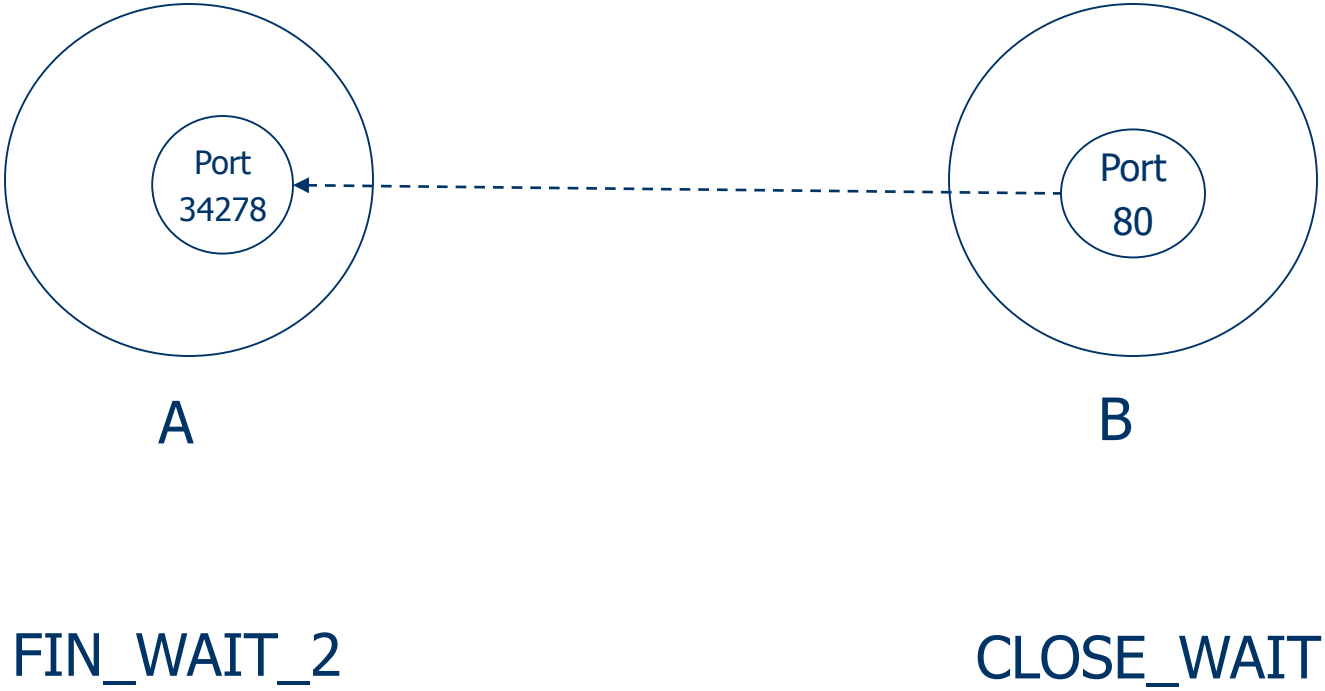
- Απόλυση σύνδεσης: Βήμα 2



TCP: Διαχείριση συνδέσεων

Παράδειγμα απόλυσης σύνδεσης

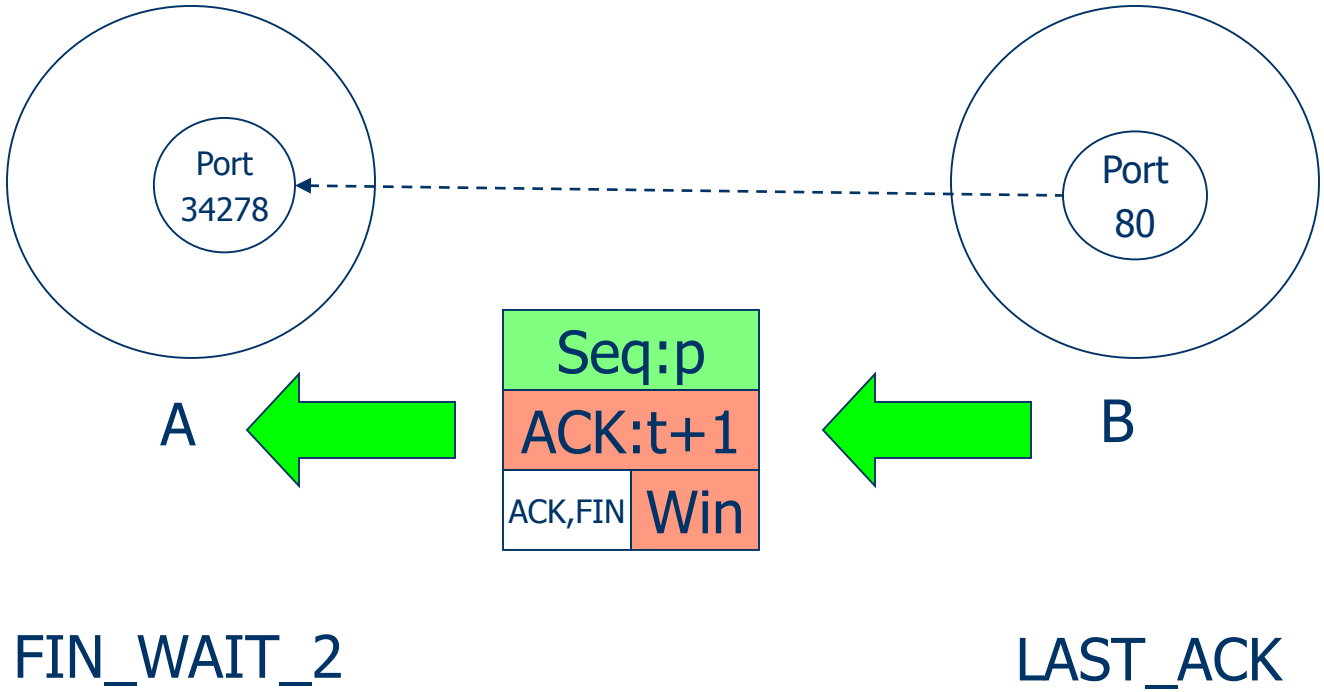
- Η εφαρμογή στον Β πρέπει να κλείσει



TCP: Διαχείριση συνδέσεων

Παράδειγμα απόλυσης σύνδεσης

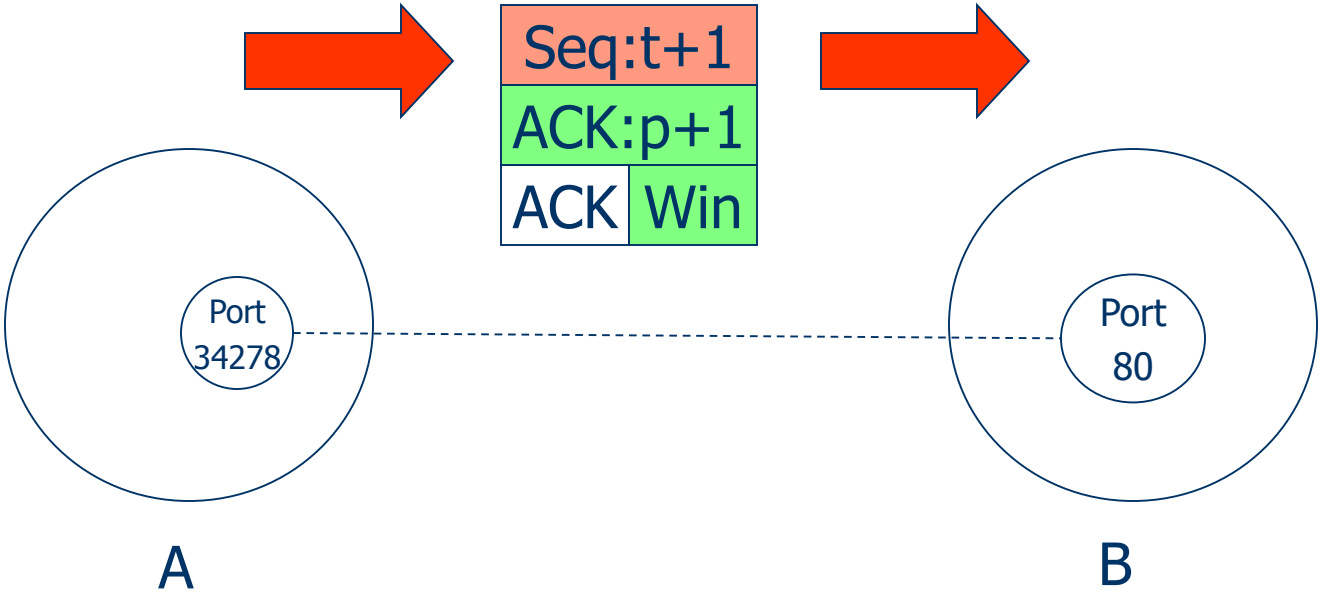
- Απόλυση σύνδεσης: Βήμα 3



TCP: Διαχείριση συνδέσεων

Παράδειγμα απόλυσης σύνδεσης

- Απόλυση σύνδεσης: Βήμα 4



TIME_WAIT

LAST_ACK

TCP: Διαχείριση συνδέσεων

Παράδειγμα απόλυσης σύνδεσης

- Ο Client περιμένει 2*MSL πριν μεταβεί στην κατάσταση CLOSED (ο Β μπορεί να ξαναστείλει ένα FIN).

TCP: Διαχείριση συνδέσεων

Κατάσταση αναμονής 2MSL = TIME_WAIT

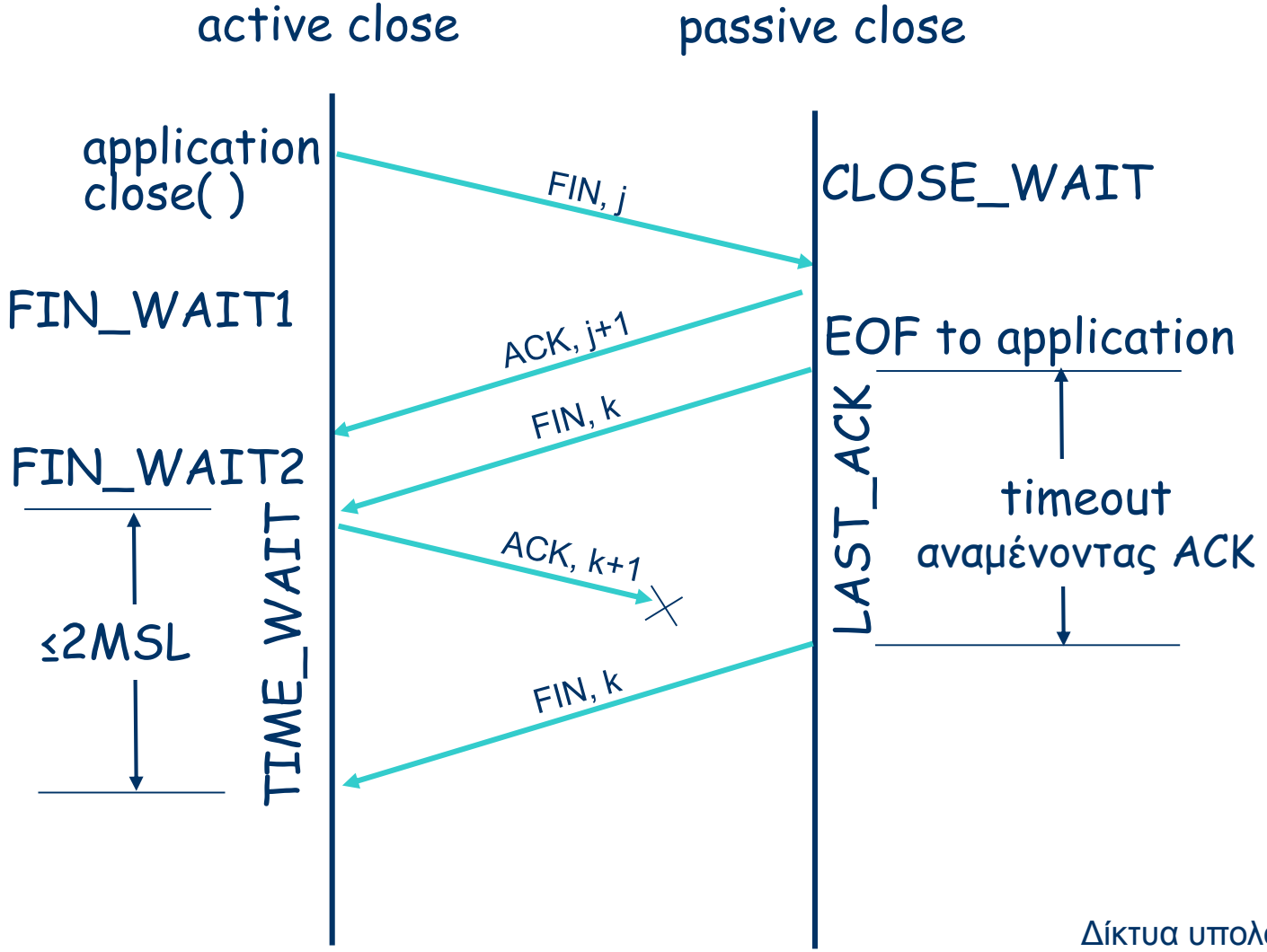
- Όταν το TCP κάνει active close, και στέλνει την τελική ACK, η σύνδεση πρέπει να παραμείνει στην κατάσταση **TIME_WAIT** για διπλάσιο χρόνο από τη μέγιστη ζωή τεμαχίων (maximum segment lifetime, MSL).

  2MSL= 2 * Maximum Segment Lifetime

- Γιατί;
  Δίνεται στον TCP client μια ευκαιρία επαναποστολής της τελικής ACK. (Ο server θα κάνει timeout αφού στείλει το τεμάχιο FIN και θα ξαναστείλει το FIN).
- Το MSL τίθεται στα 2 min ή 1 min ή 30 sec.

TCP: Διαχείριση συνδέσεων

TIME_WAIT



TCP: Διαχείριση συνδέσεων

Επαναφορά συνδέσεων

- Η επαναφορά (reset) συνδέσεων γίνεται με την ενεργοποίηση της σημαίας RST .
- **Πότε ενεργοποιείται η σημαία RST;**
  - Όταν φθάνει η αίτηση σύνδεσης και δεν αναμένει καμία διαδικασία server στο σημείο προορισμού.
  - Όταν αποσταλεί τεμάχιο που δεν αναμένεται καθόλου σε υπάρχουσα σύνδεση, π.χ. ο αύξων αριθμός είναι εκτός περιοχής.
- Η επαναφορά μιας σύνδεσης αναγκάζει τον δέκτη του RST να πετάξει τα αποθηκευμένα δεδομένα. Ο δέκτης δεν επαληθεύει το τεμάχιο RST.

TCP: Μεταφορά δεδομένων

Η μεταφορά δεδομένων μέσω TCP μπορεί να χαρακτηριστεί ως:

αλληλοδραστική μεταφορά - telnet, rlogin
μαζική μεταφορά - ftp, mail, http

Το TCP έχει ευριστικές μεθόδους για να χειρίζεται αυτούς τους τύπους μεταφοράς δεδομένων.

Για αλληλοδραστική μεταφορά δεδομένων:
- Επιχειρεί να περιορίσει τον αριθμό των πακέτων.

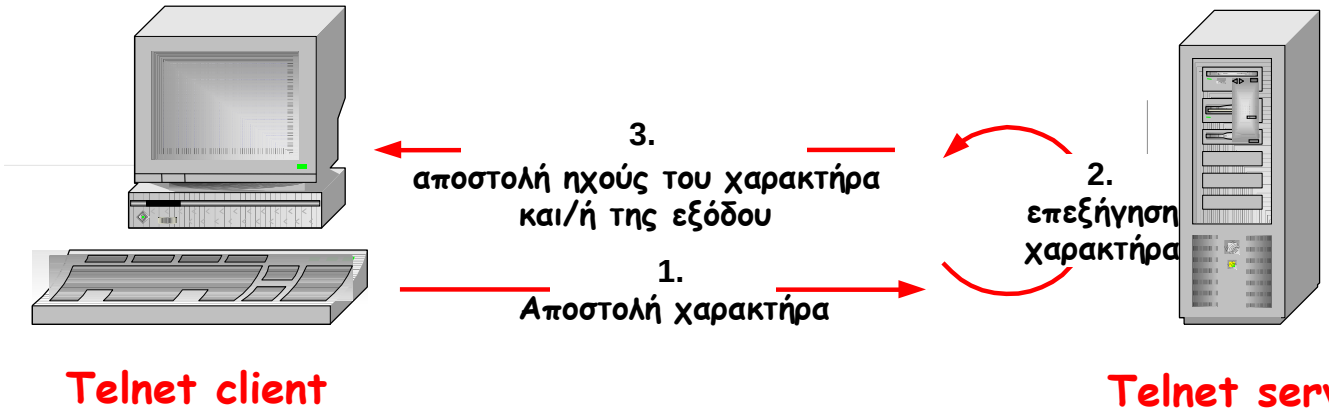
Για μαζική μεταφορά δεδομένων:
- Επιχειρεί να ελέγξει τη ροή δεδομένων.

TCP Data Flow

- Interactive Data: small sized packets (Telnet, Rlogin etc.)
- Bulk Data: full sized packets (File Transfer Protocol, Electronic Mail etc.)
- Usually in the internet, on a packet count basis, about half of TCP segments contain bulk data and the other half contain interactive data
- On a byte count basis the ratio is around 90% bulk data and 10% interactive data

TCP: Μεταφορά δεδομένων

Αλληλοδραστικές εφαρμογές: Telnet



Οι εφαρμογές απομακρυσμένου τερματικού (π.χ., Telnet) στέλνουν χαρακτήρες σε έναν server. Ο server επεξεργάζεται τον χαρακτήρα και στέλνει την έξοδό του στον client.

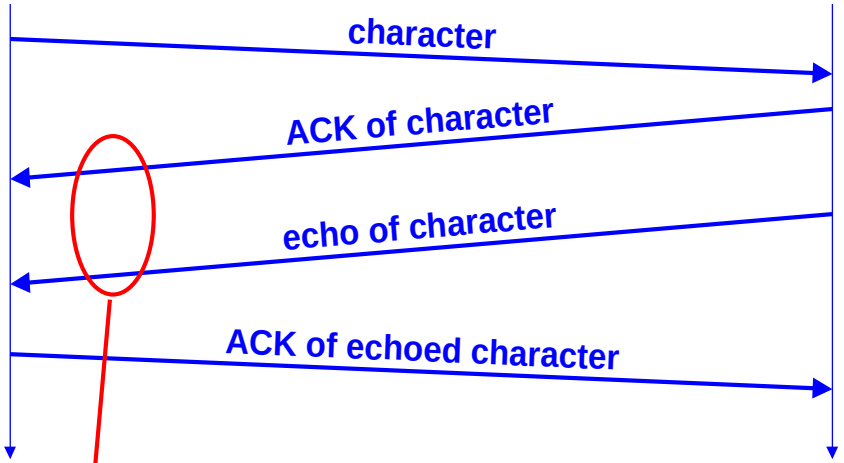
Για κάθε πληκτρολογούμενο χαρακτήρα, βλέπουμε 3 πακέτα:

1. **Client → Server:** Αποστολή πληκτρολογούμενου χαρακτήρα
2. **Server → Client:** Επαλήθευση του πρώτου πακέτου και ηχώ του χαρακτήρα
3. **Client → Server:** Επαλήθευση του δεύτερου πακέτου

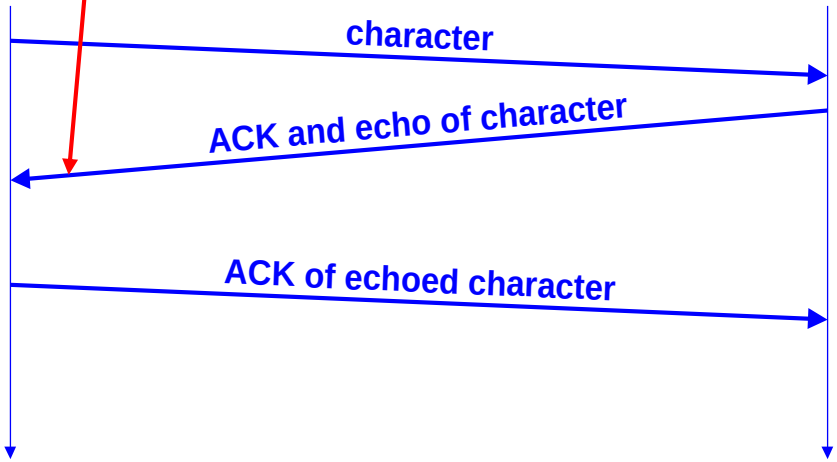
TCP: Μεταφορά δεδομένων

Αλληλοδραστικές εφαρμογές: Telnet

- Θα αναμέναμε 4 πακέτα ανά χαρακτήρα:



- Ωστόσο, εμφανίζεται η παραπλεύρως απεικονιζόμενη μορφή:



- Το TCP έχει καθυστερήσει την αποστολή μιας ACK

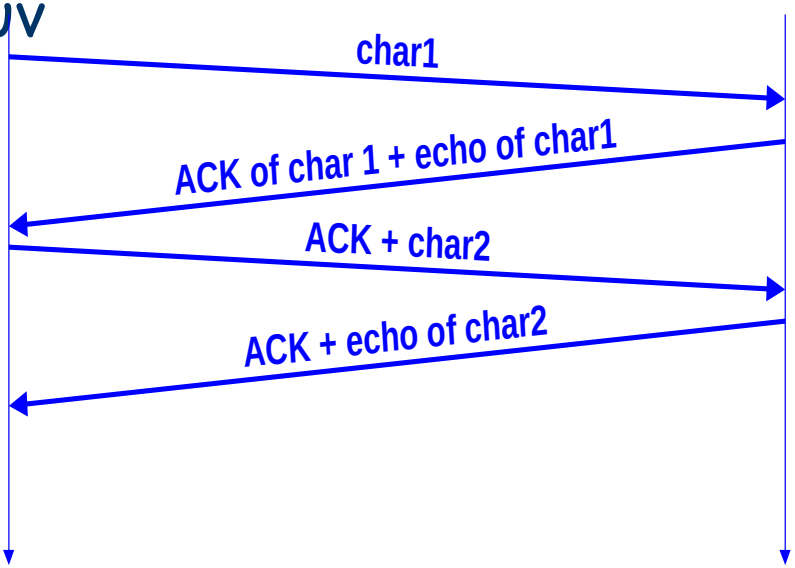
TCP: Μεταφορά δεδομένων

- The total number of bytes transmitted for a single keystroke (let us assume 1 byte of data) is: TCP header=20 bytes; IP header=20 bytes; Ethernet Header/trailer=18 bytes.
  - For Data segment: (1+20+20+18) bytes = 59 bytes
  - For Ack: (20+20+18) bytes = 58 bytes
  - For echo: (1+20+20+18) bytes = 59 bytes
  - For ack of echo: (20+20+18) bytes = 58 bytes
  - Total= 234 bytes

TCP: Μεταφορά δεδομένων

Σύνοδος telnet σε απομακρυσμένον host

- **Παρατήρηση:** Η μετάδοση των τεμαχίων ακολουθεί διαφορετικό τρόπο, δηλ., υπάρχουν μόνο δύο πακέτα ανά πληκτρολογούμενο χαρακτήρα.



- Η καθυστερημένη επαλήθευση δεν εμφανίζεται.
- Ο λόγος είναι ότι υπάρχουν πάντα δεδομένα έτοιμα προς αποστολή, όταν φθάνει η ACK.

TCP: Μεταφορά δεδομένων

Αλγόριθμος του Nagle

- Έχει ως στόχο την αποφυγή της μη αποτελεσματικής χρήσης του εύρους ζώνης.
- Πομπός:
  - Αποθηκεύει προσωρινά όλα τα δεδομένα χρήστη, αν εκκρεμούν ανεπαλήθευτα δεδομένα.
  - Στέλνει, όταν όλα έχουν επαληθευτεί ή έχει δεδομένα που συμπληρώνουν ένα τεμάχιο MSS.
- Δέκτης:
  - Δίνει εντολή αποστολής, μόνο όταν μπορεί να αυξήσει επαρκώς το παράθυρο λήψης.

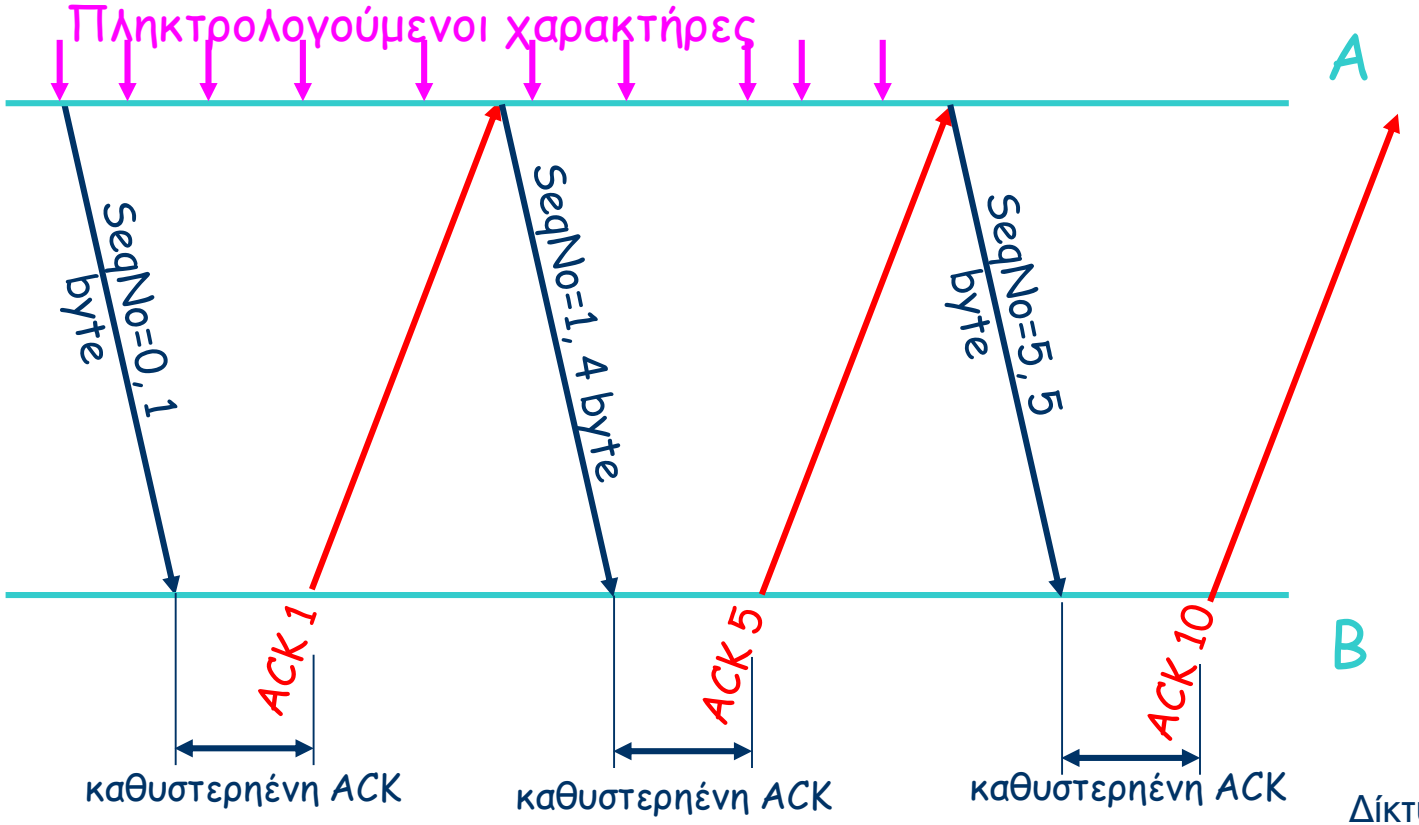
Nagle Algorithm

- Problem: The existence of a large number of small packets can generate congestion on WANs (in Rlogin 1 byte of data is transmitted as a datagram of 41 bytes, 20 bytes for IP header and 20 bytes for TCP header)
- Solution (Nagle Algorithm): a TCP connection can have only one outstanding small segment (not yet acknowledged). No additional small packets can be sent until ack is received. Instead small amounts of data are collected by TCP and sent in a single segment when the ack comes. The algorithm is self-clocking: the faster the ack comes the faster the data is sent
- Sometimes we may need to disable Nagle's algorithm (i.e mouse movements)

TCP: Μεταφορά δεδομένων

Αλγόριθμος του Nagle

- Μόνο ένα τεμάχιο ενός byte μπορεί να μεταδίδεται (Επειδή δεν υπάρχουν δεδομένα προς απόστολή από τον B προς A έχουμε καθυστερημένες ACK).



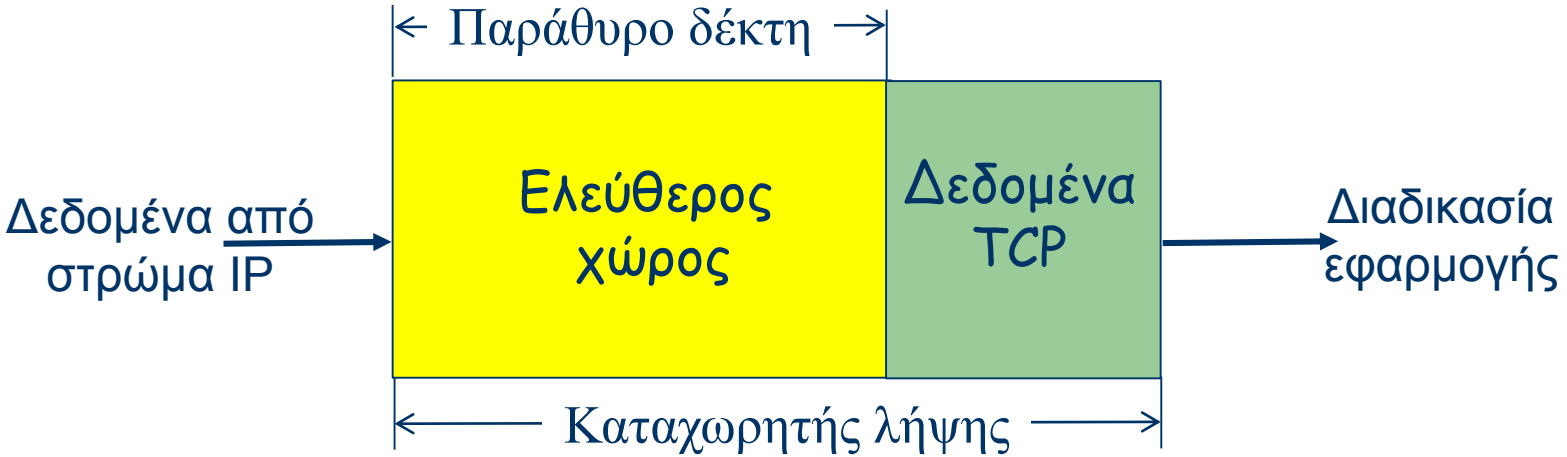
TCP: Μεταφορά δεδομένων

Ιδιότητες του αλγορίθμου του Nagle

- Εφαρμόζεται μόνο σε μικρά πακέτα. Στις μεταφορές μεγάλων αρχείων υπάρχουν πάντα πλήρη MSS για αποστολή.
- Ο αλγόριθμος είναι αυτοχρονιζόμενος:
  - Βασικά εφαρμόζει Stop & Wait για μικρά τεμάχια.
  - Σε LAN, το μικρό RTT δεν εισάγει μεγάλη αναμονή, οπότε ο αλγόριθμος δεν είναι αποτελεσματικός.
  - Σε WAN, το μεγάλο RTT εισάγει περισσότερη αναμονή, αλλά ο αλγόριθμος είναι πιο αποτελεσματικός σε μακριές ζεύξεις.
- Όταν απαιτείται μικρή καθυστέρηση, ο αλγόριθμος προκαλεί ανεπιθύμητες καθυστερήσεις.
- Οι εφαρμογές μπορεί να απενεργοποιήσουν τον αλγόριθμο.

TCP: Έλεγχος ροής

- Η πλευρά λήψης της σύνδεσης TCP έχει έναν καταχωρητή λήψης:



- Η διαδικασία εφαρμογής μπορεί να αργεί να διαβάσει από τον καταχωρητή λήψης.
- Έλεγχος ροής: ο πομπός δεν πρέπει να υπερχειλίσει τον καταχωρητή του δέκτη μεταδίδοντας πολλά, πολύ γρήγορα.
- Προσαρμογή του ρυθμού αποστολής δεδομένων στον ρυθμό ανάγνωσης της λαμβάνουσας εφαρμογής.

TCP: Έλεγχος ροής

- Το TCP χρησιμοποιεί έλεγχο ροής με ολισθαίνον παράθυρο:
  - Δεν χρησιμοποιεί NACK
  - Μόνο συσσωρευτικές ACK
- Οι επαληθεύσεις δεν προκαλούν αυτόματα αλλαγές στο μέγεθος παραθύρου του πομπού.
- Ο δέκτης επιστρέφει δύο παραμέτρους στον πομπό.

AckNo	window size (win)
32 bits	16 bits

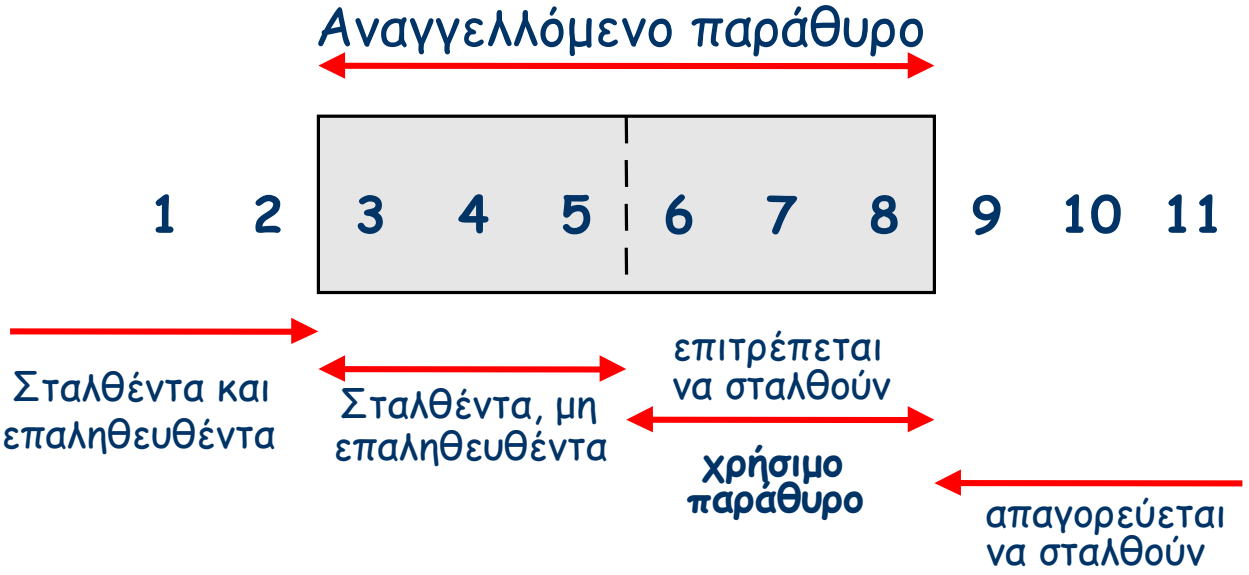
- Ο πομπός μπορεί να στείλει δεδομένα μέχρι το διαφημιζόμενο παράθυρο, δηλαδή, τα byte:

  AckNo, AckNo+1, …, AckNo + win -1
- Ο δέκτης μπορεί να επαληθεύσει χωρίς να αλλάξει το παράθυρο.
- Ο δέκτης μπορεί να αλλάξει το παράθυρο χωρίς να επαληθεύσει.

TCP: Έλεγχος ροής

Ολισθαίνον παράθυρο

Το πρωτόκολλο ολισθαίνοντος παραθύρου λειτουργεί σε επίπεδο byte:

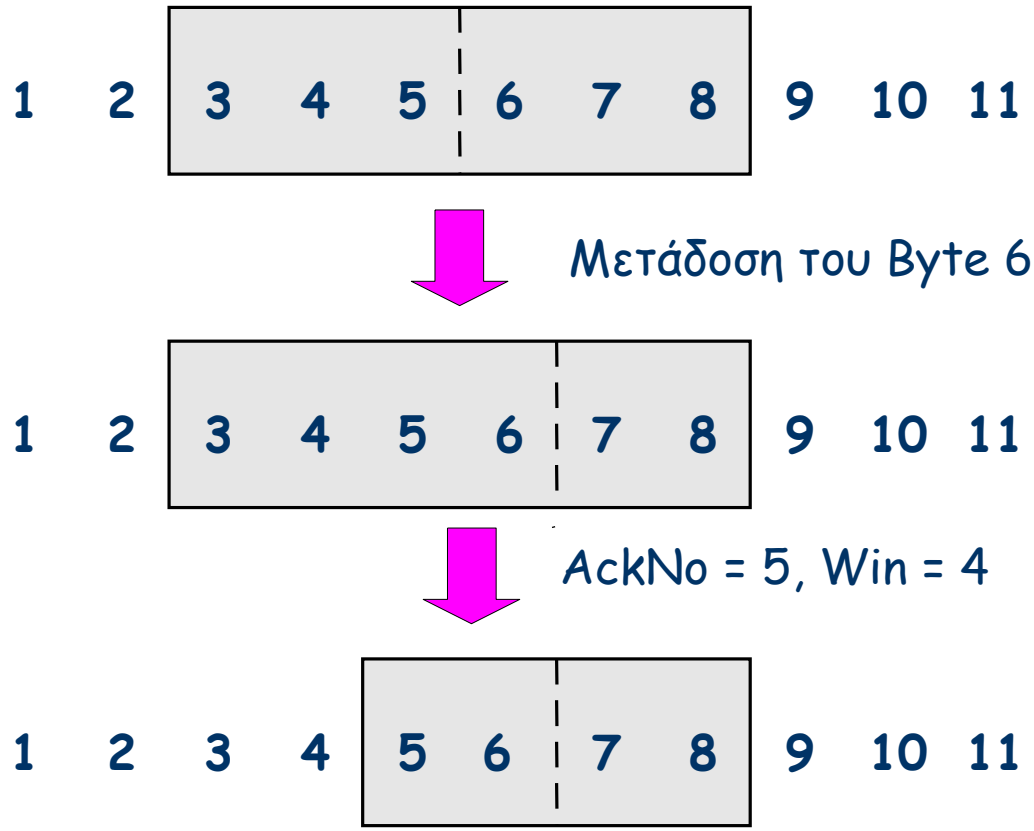


Ο πομπός μπορεί να στείλει μόνο τους αύξοντες αριθμούς 6,7,8.

TCP: Έλεγχος ροής

Ολισθαίνον παράθυρο: κλείσιμο

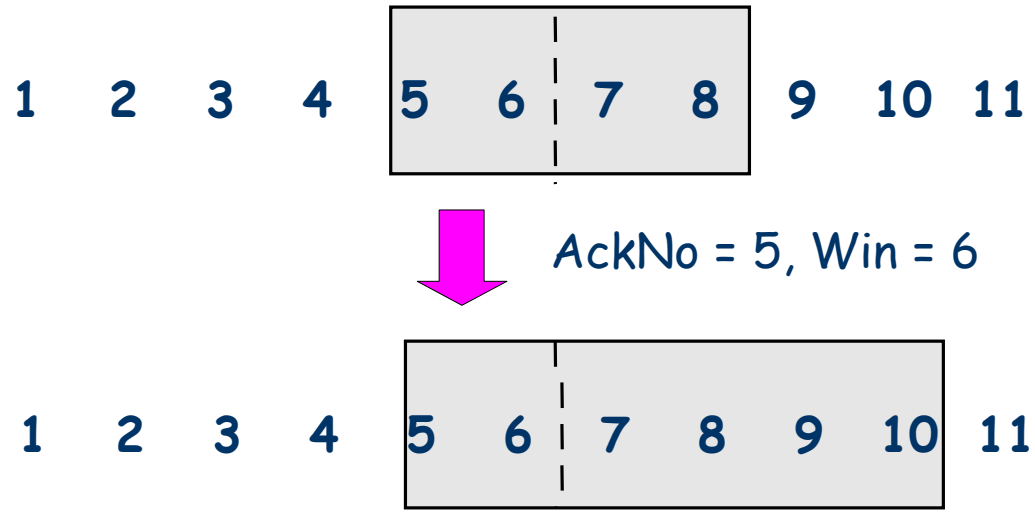
Αποστολή ενός byte (με SeqNo = 6) και λήψη της επαλήθευσης (AckNo = 5, Win=4):



TCP: Έλεγχος ροής

Ολισθαίνον παράθυρο: άνοιγμα

Λήψη επαλήθευσης που μεγαλώνει το παράθυρο προς τα δεξιά (AckNo = 5, Win=6):

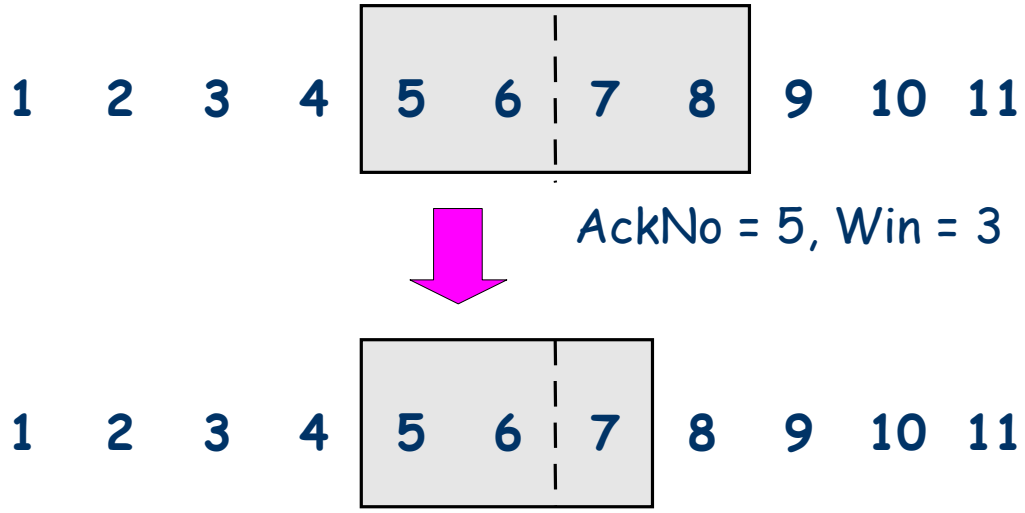


Ο δέκτης ανοίγει το παράθυρο όταν ο καταχωρητής TCP αδειάζει (εννοώντας ότι τα δεδομένα παραδίδονται στην εφαρμογή).

TCP: Έλεγχος ροής

Ολισθαίνον παράθυρο: συρρίκνωση

Λήψη επαλήθευσης που περιορίζει το παράθυρο από δεξιά (AckNo = 5, Win=3):



Η συρρίκνωση παραθύρου δεν πρέπει να χρησιμοποιείται.

TCP: Έλεγχος ροής

Ολισθαίνον παράθυρο: παράδειγμα