

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΕΑΡΙΝΟ ΕΞΑΜΗΝΟ 2021-2022

ΜΑΘΗΜΑ «ΔΙΚΤΥΑ ΥΠΟΛΟΓΙΣΤΩΝ»

ΔΙΔΑΣΚΩΝ: ΓΕΩΡΓΙΟΣ Δ. ΣΤΑΜΟΥΛΗΣ, ΚΑΘΗΓΗΤΗΣ

ΒΟΗΘΟΙ: ΔΙΟΝΥΣΗΣ ΔΑΜΑΣΙΩΤΗΣ, ΙΑΚΩΒΟΣ ΠΙΤΤΑΡΑΣ, ΘΕΟΔΟΣΗΣ ΓΙΑΝΝΟΠΟΥΛΟΣ

Σειρά Εργαστηριακών Ασκήσεων 3

Παράδοση: 2/6/2022

Σκοπός

Χρήση του εργαλείου Wireshark για ανάλυση πρωτοκόλλων σε γραφικό περιβάλλον για την άντληση πληροφοριών για τα πρωτόκολλα TCP, UDP, HTTP καθώς και γενικές ερωτήσεις/ασκήσεις ως προς τα παραπάνω πρωτόκολλα καθώς και το πρωτόκολλο IP.

Γενικές πληροφορίες

Η εργασία θα πρέπει να εκπονηθεί σε ομάδες τριών (3) ατόμων. Για την παράδοση της εργασίας θα πρέπει να ετοιμάσετε σύντομη αναφορά με τις απαντήσεις σας και με τεκμηρίωση (screen shots), την οποία θα υποβάλλετε ηλεκτρονικά μέσω **e-class**. Για απορίες μπορείτε να απευθυνθείτε στον βοηθό κ. Δαμασιώτη, στο email diondam@gmail.com.

1) Πρωτόκολλο TCP

Το TCP είναι το κύριο πρωτόκολλο επιπέδου μεταφοράς που χρησιμοποιείται στο Διαδίκτυο. Το πρωτόκολλο TCP εγγυάται την αξιόπιστη και στην ορθή σειρά μετάδοση των δεδομένων. Περιλαμβάνει έναν μηχανισμό ελέγχου ροής δεδομένων (flow control), ο οποίος επιτρέπει στον παραλήπτη να περιορίσει την ποσότητα των bytes που θα στείλει ο αποστολέας σε κάθε δεδομένη χρονική στιγμή. Επίσης, το TCP υλοποιεί κι έναν μηχανισμό ελέγχου συμφόρησης. Η ιδέα αυτού του μηχανισμού είναι να ελέγχει πόσο γρήγορα θα στέλνει το TCP δεδομένα ώστε να αποτρέψει τον αποστολέα από το να υπερφορτώσει το δίκτυο.

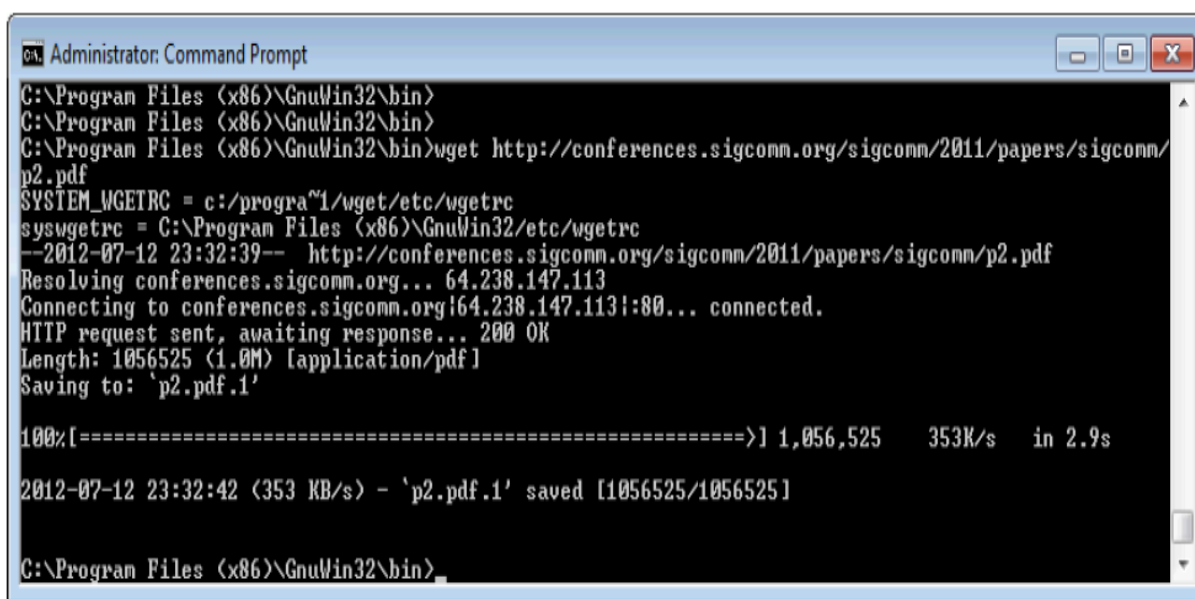
Ο σκοπός της χρήσης του παραθύρου συμφόρησης είναι να καθορίσει για κάθε πηγή πόση χωρητικότητα είναι διαθέσιμη στο δίκτυο ώστε να γνωρίζει τον όγκο των δεδομένων που μπορούν να μεταφερθούν. Διατηρεί μία μεταβλητή κατάσταση για κάθε σύνδεση, η οποία λέγεται congestion window (παράθυρο συμφόρησης) και χρησιμοποιείται από την πηγή για να περιορίσει τον όγκο των δεδομένων που επιτρέπεται να σταλούν σε μία χρονική στιγμή. Το TCP χρησιμοποιεί το μηχανισμό "Additive Increase/Multiplicative Decrease" και μειώνει δραστικά το παράθυρο συμφόρησης όταν θεωρείται ότι το επίπεδο συμφόρησης είναι υψηλό, ενώ αυξάνει το παράθυρο όταν θεωρείται ότι το επίπεδο συμφόρησης είναι χαμηλό. Το TCP αντιλαμβάνεται τα timeouts σαν μια ένδειξη συμφόρησης. Κάθε φορά που συμβαίνει timeout, η πηγή μειώνει το παράθυρο συμφόρησης στο μισό της προηγούμενης τιμής του (στην πραγματικότητα η τιμή του παραθύρου συμφόρησης μειώνεται στο 1, δηλ. στο TCP maximum segment size, και μετά αυξάνει γρήγορα στο μισό της προηγούμενης τιμής του). Η μείωση αυτή οφείλεται στο τμήμα "multiplicative decrease" του μηχανισμού. Κάθε φορά που η πηγή στέλνει επιτυχώς έναν αριθμό τμημάτων, προστίθεται η αξία ενός maximum segment size στο παράθυρο συμφόρησης. Η αύξηση αυτή οφείλεται στο τμήμα "additive increase" του μηχανισμού.

Το TCP χρησιμοποιεί επίσης τον μηχανισμό "Slow Start" για να αυξήσει το παράθυρο συμφόρησης γρήγορα κατά την εκκίνηση μιας συνδέσεως TCP, οπότε το παράθυρο συμφόρησης αυξάνει εκθετικά παρά γραμμικά.

Στην παρούσα άσκηση θα καταγράψετε το αποτύπωμα μιας σύνδεσης TCP μέσω της οποίας θα μεταφερθεί μια ποσότητα των δεδομένων. Ο όγκος των δεδομένων θα πρέπει να είναι ενδιαμέσος (δηλ. της τάξεως των 500 KB) και όλα τα δεδομένα να μεταφερθούν μέσω της ίδιας σύνδεσης. Πολλές εφαρμογές χρησιμοποιούν το TCP για μεταφορά, συμπεριλαμβανομένων των web browsers. Να εκτελέσετε απλά μια λήψη web για να εξασκηθείτε με τα στοιχεία σε μια σύνδεση TCP. Επισημαίνεται ότι το TCP είναι σε θέση να μεταφέρει δεδομένα και στις δύο κατευθύνσεις ταυτόχρονα. Ωστόσο, στη συγκεκριμένη άσκηση αποστέλλεται μόνο ένα περιεχόμενο και μόνο από τον απομακρυσμένο server στον τοπικό υπολογιστή (μετά την αρχική αίτηση).

1. Να βρείτε μια διεύθυνση URL (βλέπε Άσκηση 3), όπου το αντίστοιχο αρχείο είναι μέτριου μεγέθους, και να το κατεβάσετε χρησιμοποιώντας το πρωτόκολλο HTTP (αντί HTTPS). Μπορείτε να χρησιμοποιήσετε το πρόγραμμα περιήγησής σας για να ψάξετε, ίσως για μια εικόνα (.jpg) ή ένα αρχείο PDF (.pdf). Είναι απαραίτητο να εξασφαλίσετε ότι η ανωτέρω διεύθυνση URL είναι ένα μοναδικό αντικείμενο δεδομένων και όχι μια ιστοσελίδα (π.χ. ένα .html) με πολλαπλούς πόρους.
2. Να ανακτήσετε το ανωτέρω URL με wget ή curl για να βεβαιωθείτε ότι είστε σε θέση να ανακτήσετε τουλάχιστον 500 KB περιεχομένου εντός λίγων δευτερολέπτων. Για παράδειγμα, να μεταβείτε στον browser σας σε μία HTTP ιστοσελίδα που περιέχει λίστα pdf αρχείων ή jpg εικόνων και να επιλέξετε ένα αρχείο το οποίο θα κατεβάσετε, χρησιμοποιώντας την εντολή wget.

Τα επιτυχημένα παραδείγματα λήψης αρχείων φαίνονται στο παρακάτω σχήμα.



```
Administrator: Command Prompt
C:\Program Files (x86)\GnuWin32\bin>
C:\Program Files (x86)\GnuWin32\bin>
C:\Program Files (x86)\GnuWin32\bin>wget http://conferences.sigcomm.org/sigcomm/2011/papers/sigcomm/p2.pdf
SYSTEM_WGETRC = c:/progra~1/wget/etc/wgetrc
syswgetrc = C:\Program Files (x86)\GnuWin32\etc\wgetrc
--2012-07-12 23:32:39-- http://conferences.sigcomm.org/sigcomm/2011/papers/sigcomm/p2.pdf
Resolving conferences.sigcomm.org... 64.238.147.113
Connecting to conferences.sigcomm.org|64.238.147.113|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 1056525 (1.0M) [application/pdf]
Saving to: 'p2.pdf.1'

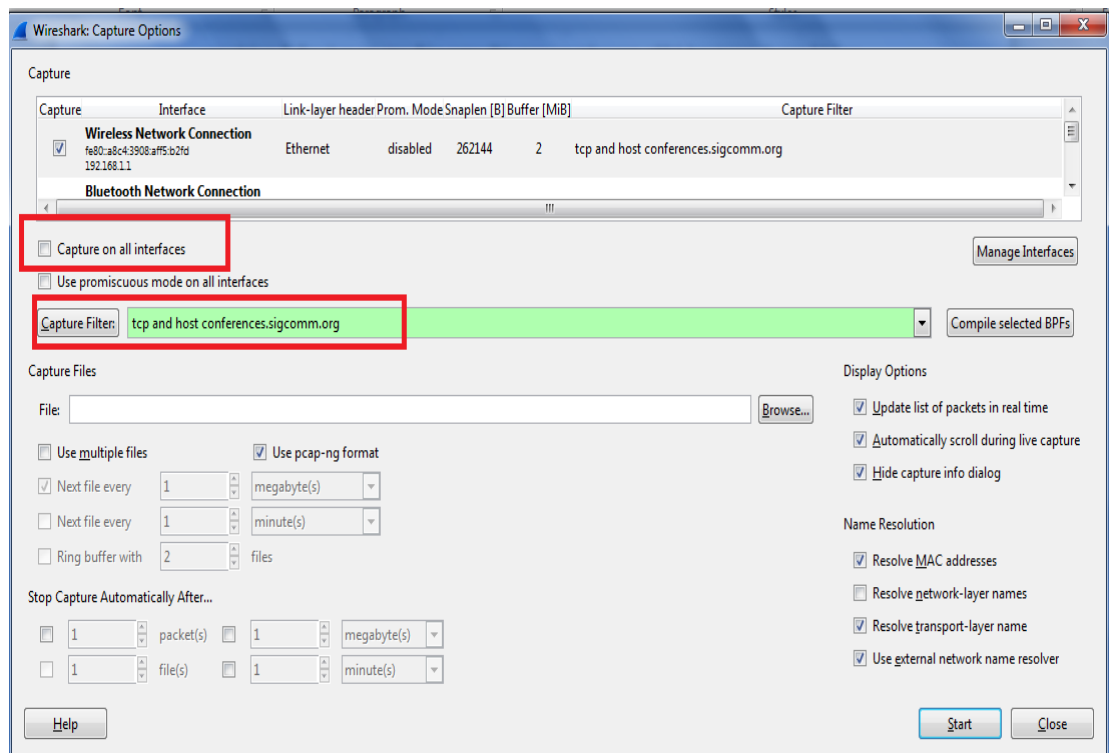
100%[=====>] 1,056,525 353K/s in 2.9s

2012-07-12 23:32:42 (353 KB/s) - 'p2.pdf.1' saved [1056525/1056525]

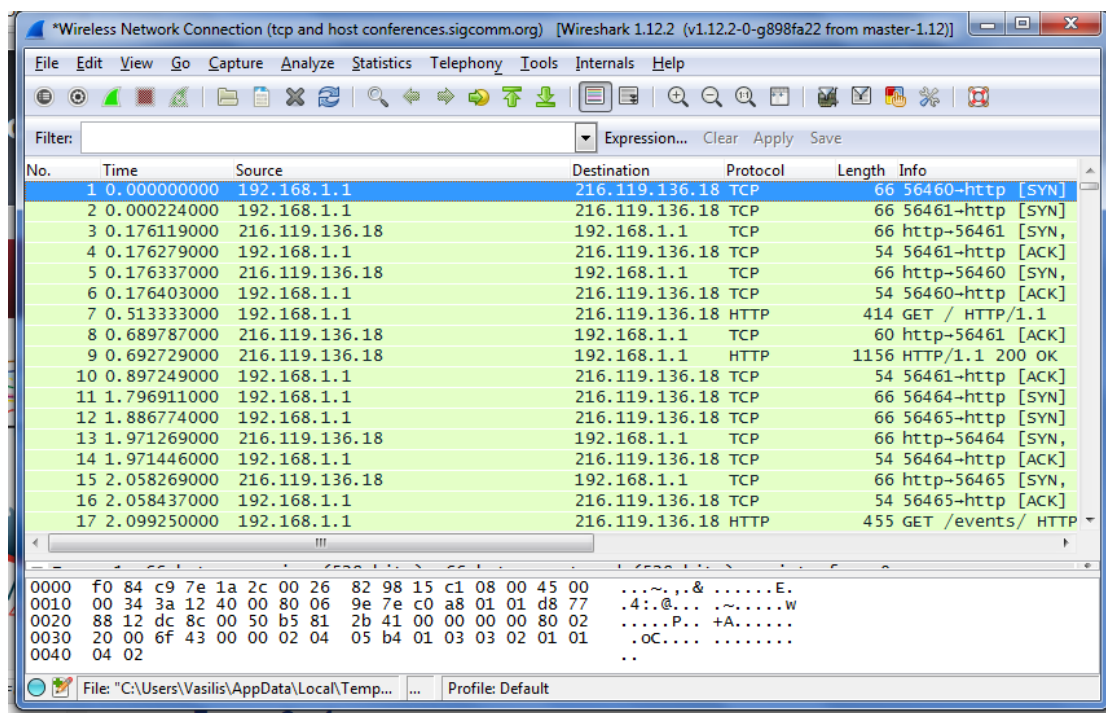
C:\Program Files (x86)\GnuWin32\bin>
```

Figure 1. Μία επιτυχημένη λήψη με τη χρήση της εντολής wget (Windows)

3. Να ξεκινήσετε μία καταγραφή με το Wireshark με φίλτρο "tcp and host xx.xx.xx", όπου xx.xx.xx είναι το όνομα του απομακρυσμένου Server από τον οποίο θα μεταφέρετε το περιεχόμενο, π.χ., "conferences.sigcomm.org" στο παράδειγμα που δείχνει το σχήμα παρακάτω. Η ιδέα του φίλτρου είναι να καταγράψει την κίνηση TCP μόνο μεταξύ του υπολογιστή σας και του διακομιστή (server). Το παράθυρο λήψης πρέπει να είναι παρόμοιο με αυτό που εμφανίζεται παρακάτω. Να επιλέξετε τη διεπαφή από την οποία θα γίνει η σύλληψη. Η διεπαφή αντιστοιχεί στην κύρια σύνδεση που χρησιμοποιείτε από τον υπολογιστή σας (ενσύρματη ή ασύρματη σύνδεση) για να συνδεθείτε στο Internet. Να ελέγξετε αν είναι επιλεγμένη "capture packets in promiscuous mode" και αν ναι, να την από-επιλέξετε. Είναι επιθυμητό μόνο να καταγραφούν τα πακέτα που αποστέλλονται προς/από τον υπολογιστή σας. Να αφήσετε τις υπόλοιπες επιλογές στις προεπιλεγμένες τιμές τους. Το φίλτρο σύλληψης, εάν υπάρχει, χρησιμοποιείται για να αποτρέψει την καταγραφή της υπόλοιπης κυκλοφορίας του υπολογιστή σας.



4. Μόλις η καταγραφή ξεκινήσει, να επαναλάβετε την παραπάνω εντολή `wget/curl`. Αυτή τη φορά, τα πακέτα θα καταγράφονται από το Wireshark.
5. Όταν η εντολή ολοκληρωθεί, να σταματήσετε την καταγραφή στο Wireshark. Πρέπει τώρα να έχετε ένα ίχνος παρόμοιο με αυτό που φαίνεται στην παρακάτω εικόνα:



Το «μεσαίο» χρονικά τμήμα μιας σύνδεσης TCP είναι αυτό όπου γίνεται η μεταφορά και η λήψη των δεδομένων. Αυτό είναι το κύριο μέρος της σύνδεσης. Στη συνέχεια θα εξετάσετε την εξέλιξη του ρυθμού λήψης δεδομένων στην πάροδο του χρόνου.

6. Κάτω από το μενού Statistics, να επιλέξετε το "IO Graph". Από προεπιλογή, το γράφημα αυτό δείχνει το ρυθμό των πακέτων στη πάροδο του χρόνου. Να το αλλάξετε ώστε να δείχνει το ρυθμό λήψης (download rate) όπως περιγράφεται παρακάτω. Αντιθέτως, το ίχνος πρέπει να λαμβάνεται κοντά στον υπολογιστή που λαμβάνει τα δεδομένα, και αυτό επιτυγχάνεται με χρήση του Wireshark.
- Στον x-axis, να προσαρμόσετε το tick interval και pixels per tick. Το tick interval πρέπει να είναι αρκετά μικρό για να εξετάσετε τη συμπεριφορά του πάνω από το ίχνος. Τα pixels per tick μπορεί να ρυθμιστούν για να δώσουν μεγαλύτερο ή μικρότερο εύρος στο γράφημα.
 - Στον y-axis, να αλλάξετε τη μονάδα σε Bits/Tick. Η προεπιλογή είναι Packet / Tick.
 - Να προσθέσετε ένα φίλτρο για να εμφανίζονται μόνο τα πακέτα λήψης. Υποθέτοντας ότι η λήψη είναι από το συνηθισμένο port 80 του web server, μπορείτε να εισάγετε ένα φίλτρο "tcp.srcport == 80". Πρέπει επίσης να πατήσετε Enter, και ίσως χρειαστεί να κάνετε κλικ στο κουμπί "Graph" για να το εμφανίσετε εκ νέου.
 - Για να δείτε το αντίστοιχο γράφημα για την κίνηση Upload, να εισάγετε ένα δεύτερο φίλτρο στο επόμενο πλαίσιο. Και πάλι αν υποθέσουμε ότι η συνηθισμένη θύρα του web server είναι 80, το φίλτρο είναι "tcp.dstport == 80". Αφού πατήσετε το πλήκτρο Enter και κάνετε κλικ στο κουμπί Graph, θα πρέπει να έχετε δύο γραμμές στο γράφημα, όπως φαίνεται παρακάτω.

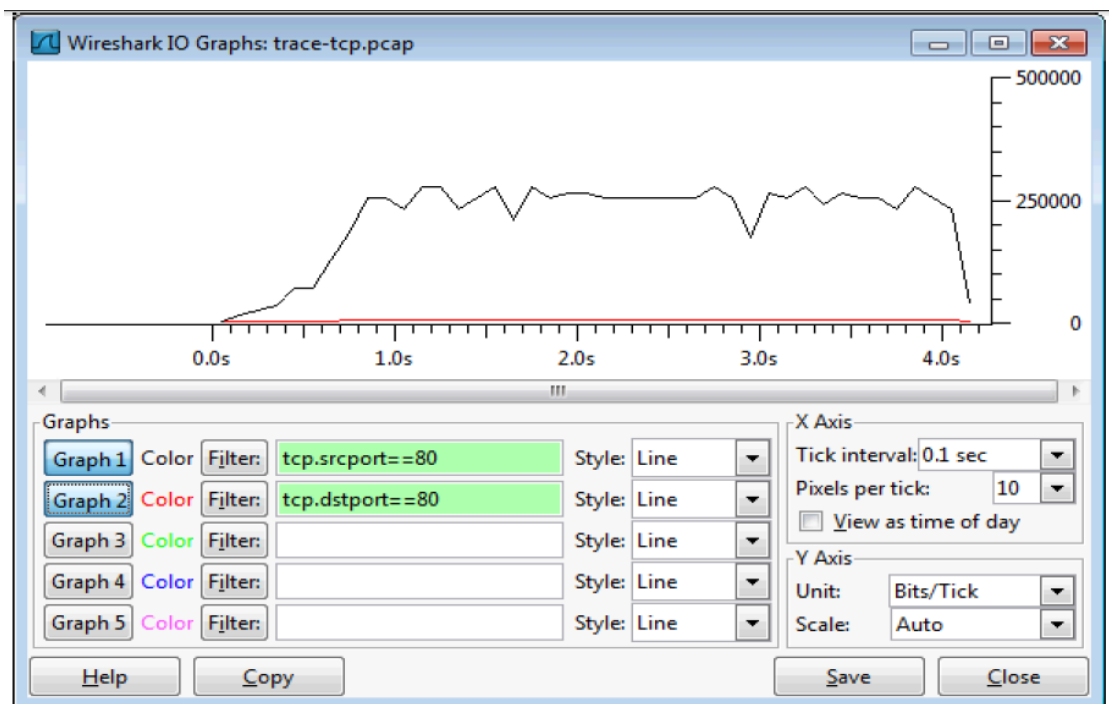


Figure 4. TCP download rate over time via an IO graph

Από αυτό μπορούμε να δούμε το ρυθμό λήψης του δείγματος να αυξάνεται εκθετικά από το μηδέν έως ένα μέγιστο επίπεδο ρυθμού, το οποίο είναι κατά προσέγγιση το ίδιο κάθε φορά. Αυτό αντιστοιχεί στη φάση αργής εκκίνησης. Ο ρυθμός λήψης στο συγκεκριμένο παράδειγμα ανωτέρω όταν η σύνδεση είναι σε λειτουργία είναι περίπου 2,5 Mbps. Ο ρυθμός αποστολής είναι σταθερός και αποτελεί μικρό ποσοστό της συνολικής κυκλοφορίας. Στο παράδειγμα του σχήματος, η λήψη συνεχίζεται αρκετά σταθερά μέχρι να ολοκληρωθεί. Στην πράξη μπορεί να εμφανισθούν διαφορετικές συμπεριφορές αυξομειώσης του ρυθμού λήψης. Για παράδειγμα, εάν το διαθέσιμο εύρος ζώνης επηρεάζεται σημαντικά από ανταγωνιστικά downloads από τον ίδιο server, τότε ο ρυθμός λήψης ορίζεται από τον server και όχι από το δίκτυο, ενώ εάν χαθούν αρκετά πακέτα στο δίκτυο τότε ο ρυθμός λήψης επηρεάζεται σημαντικά από αυτό.

Να απαντήσετε στις ακόλουθες ερωτήσεις :

- 1) Ποιος είναι ο ρυθμός λήψης δεδομένων σε packets/second και bits/second όταν η σύνδεση TCP λειτουργεί καλά;

- 2) Ποιο ποσοστό αυτού του ρυθμού λήψης αφορά τη λήψη περιεχομένου; Να το υπολογίσετε περιγράφοντας όλα τα βήματα που ακολουθήσατε. Υπόδειξη: Να παρατηρήσετε ένα τυπικό πακέτο λήψης IP. Μπορείτε να δείτε το μέγεθος του πακέτου και πόσα bytes του TCP (ωφέλιμο φορτίο) περιέχει;
- 3) Ποιος είναι ο ρυθμός αποστολής δεδομένων σε packets/second και bits/second λόγω των πακέτων που περιέχουν τα ACK;

Να παρατηρήσετε (συμπεριλαμβάνοντας στην απάντησή σας και πάλι ορισμένα σχετικά screenshots και κάποιο σύντομο σχολιασμό) τα πακέτα IP κατά την λήψη στη μέση του ίχνους σας ως προς τα παρακάτω χαρακτηριστικά:

- a) Θα πρέπει να δείτε ένα μοτίβο των τμημάτων TCP που λάβατε και τα οποία μεταφέρουν δεδομένα και ACKs που αποστέλλονται πίσω στον server. Συνήθως θα υπάρχει ένα ACK για κάθε δύο τμήματα, γιατί σκόπιμα στέλνεται μια αθροιστική επιβεβαίωση ανά δύο διαδοχικά νέα τμήματα. Αυτά τα ACK ονομάζονται Delayed ACKs.
- b) Δεδομένου ότι παρακολουθείται η διαδικασία λήψης, ο αριθμός ακολουθίας των ληφθέντων τμημάτων αυξάνει συνεχώς. Κατά συνέπεια, ο αριθμός των ACK, με τα οποία διαβιβάζονται μεταγενέστερα τμήματα πακέτων θα αυξηθεί αντίστοιχα.
- c) Δεδομένου ότι παρακολουθείται η διαδικασία λήψης, ο αριθμός ακολουθίας των μεταδιδόμενων τμημάτων δεν θα αυξηθεί (μετά το αρχικό GET). Ομοίως και ο αριθμός ACK για τα εισερχόμενα πακέτα δεν θα αυξηθεί.
- d) Κάθε τμήμα μεταφέρει πληροφορίες παραθύρου για να ενημερώσει το άλλο τελικό σημείο επικοινωνίας το μέγεθος του χώρου που παραμένει ελεύθερος στο buffer για να χρησιμοποιηθεί για αποστολή δεδομένων.

Να απαντήσετε στην παρακάτω ερώτηση:

4. Εάν το πιο πρόσφατα ληφθέν τμήμα TCP από τον server έχει αριθμό ακολουθίας X, τότε τι αριθμό ACK θα πρέπει να έχει το επόμενο TCP τμήμα που μεταδίδεται;

2) Πρωτόκολλο UDP

Το πρωτόκολλο επιπέδου μεταφοράς UDP παρέχει μια υπηρεσία “βέλτιστης προσπάθειας” (best effort) χωρίς σύνδεση (connectionless). Είναι μια «μινιμαλιστική» επέκταση στο επίπεδο μεταφοράς της υπηρεσίας “best-effort” του πρωτοκόλλου IP. Τα τμήματα UDP μπορεί να χαθούν (μη αξιόπιστη μετάδοση) ή να παραδοθούν εκτός σειράς στο ανώτερο στρώμα. Κάθε τμήμα UDP αντιμετωπίζεται ανεξάρτητα από τα άλλα. Το UDP είναι ένα λιτό πρωτόκολλο μεταφοράς για να στέλνεται πληροφορία με όσο πιο γρήγορο ρυθμό γίνεται, δηλ. χωρίς επιβράδυνση από μηχανισμούς ελέγχου, αλλά και χωρίς εγγύηση αξιοπιστίας.

Με τη βοήθεια του Wireshark να καταγράψετε την κίνηση όταν κάνετε χρήση της υπηρεσίας DNS. Να εφαρμόσετε φίλτρο σύλληψης για να παρατηρείτε μόνο την κίνηση που σχετίζεται με τη διεύθυνση IP του υπολογιστή σας και να ξεκινήσετε την καταγραφή. Να ανοίξετε ένα παράθυρο εντολών και να καθαρίσετε την προσωρινή μνήμη DNS (DNS cache) που διατηρεί ο υπολογιστής χρησιμοποιώντας κατάλληλη εντολή. Στη συνέχεια να εκτελέσετε την εντολή nslookup <http://grad.cs.aueb.gr/> ακολουθούμενη από <Enter> ώστε να αποστείλετε ερώτημα στον τοπικό εξυπηρετητή DNS για τη διεύθυνση IP του υπολογιστή <http://grad.cs.aueb.gr/> και μετά να τερματίσετε την καταγραφή στο Wireshark. Στη συνέχεια να εφαρμόσετε φίλτρο απεικόνισης ώστε να παραμείνουν μόνο πακέτα IP που μεταφέρουν τμήματα UDP.

- a. Ποια είναι η σύνταξη του φίλτρου σύλληψης που χρησιμοποιήσατε;
- b. Ποια είναι η σύνταξη του φίλτρου απεικόνισης που χρησιμοποιήσατε;

Να παρατηρήσετε το πρώτο τμήμα UDP που αποστάλθηκε από τον υπολογιστή σας.

- c. Να καταγράψετε τα ονόματα και το μήκος των πεδίων της επικεφαλίδας του τμήματος UDP.
- d. Ποιο είναι το συνολικό μέγεθος της επικεφαλίδας UDP;
- e. Ποιος είναι ο αριθμός πρωτοκόλλου για το UDP στην επικεφαλίδα του πακέτου IP εντός του οποίου ενθυλακώνεται;
- f. Ποιο είναι το μήκος του τμήματος βάσει του μεγέθους του πακέτου αυτού;
- g. Τι εκφράζει το πεδίο μήκος (Length) της επικεφαλίδας UDP;
- h. Ποιο είναι το μέγιστο μέγεθος τμήματος UDP που μπορεί να μεταφερθεί από ένα πακέτο IP; Να αιτιολογήσετε την απάντησή σας.

Να παρατηρήσετε τα μηνύματα DNS που ανταλλάχθηκαν.

- i. Να επιβεβαιώσετε ότι το πρωτόκολλο μεταφοράς που χρησιμοποιήθηκε για την επικοινωνία με τον εξυπηρετητή DNS είναι το UDP.
- j. Ποια είναι η διεύθυνση IP του εξυπηρετητή DNS;
- k. Να καταγράψετε τις θύρες (προέλευσης και προορισμού) του πρωτοκόλλου μεταφοράς που χρησιμοποιήθηκαν για μια ερώτηση και απάντηση για εγγραφή DNS τύπου A.
- l. Ποια από τις παραπάνω θύρες αντιστοιχεί στο πρωτόκολλο εφαρμογής DNS;

3) Πρωτόκολλο HTTP

Ο παγκόσμιος ιστός (World Wide Web) είναι μια απέραντη συλλογή εγγράφων που εμφανίζεται με τη μορφή ιστοσελίδων. Η ιστοσελίδα είναι μια συλλογή αντικειμένων, όπως αρχείο HTML, εικόνες JPEG, αρχεία ήχου, Java applet, κλπ, και μπορεί να περιέχει παραπομπές προς άλλες ιστοσελίδες. Η ανάγνωση των ιστοσελίδων μπορεί να γίνει με τη βοήθεια ενός προγράμματος-πελάτη (client) γνωστού ως περιηγητή ή φυλλομετρητή (web browser). Οι χρήστες μπορούν να ακολουθήσουν τους υπερσυνδέσμους με ένα κλικ του ποντικιού, διαδικασία γνωστή, ως πλοήγηση (navigation). Ο περιηγητής ή πλοηγός ιστού προσκομίζει την ιστοσελίδα και την απεικονίζει στην οθόνη του υπολογιστή, μεταφράζοντας το κείμενο και τις εντολές μορφοποίησης. Οι υπερζεύξεις απεικονίζονται έντονα, είτε με υπογράμμιση είτε με τη χρήση ενός συγκεκριμένου χρώματος, είτε και με τα δύο.

Σε κάθε ιστοσελίδα έχει αντιστοιχηθεί ένα μονοσήμαντο όνομα, ώστε να μην υπάρχει καμία αμφιβολία όσον αφορά την αναγνώρισή της. Συγκεκριμένα, σε κάθε σελίδα εκχωρείται μια ταυτότητα, ένας Ενιαίος Εντοπιστής Πόρων URL (Uniform Resource Locator) που στην ουσία παίζει το ρόλο του παγκόσμιου ονόματος της σελίδας. Τα URL αποτελούνται από τρία μέρη: το πρωτόκολλο, το όνομα DNS ή τη διεύθυνση IP της μηχανής στην οποία βρίσκεται η σελίδα και ένα τοπικό όνομα που προσδιορίζει μονοσήμαντα τη συγκεκριμένη σελίδα (συνήθως απλώς ένα όνομα καταλόγου ή αρχείου που υπάρχει στη μηχανή όπου βρίσκεται η σελίδα). Για παράδειγμα, το URL της κεντρικής ιστοσελίδας του ΟΠΑ είναι: www.aueb.gr/index.php. Το URL αυτό αποτελείται από τρία μέρη: το πρωτόκολλο (http), το όνομα DNS της μηχανής στην οποία εντοπίστηκε η σελίδα (www.aueb.gr) και το όνομα του αρχείου που περιέχει τη σελίδα (index.php), με κάποια σημεία στίξης που διαχωρίζουν τα κομμάτια.

Ανάκτηση HTML σελίδας

Για να επισκεφθεί ο χρήστης μια ιστοσελίδα αρκεί να συμπληρώσει σωστά τη διεύθυνση της ιστοσελίδας στη γραμμή διευθύνσεων του πλοηγού ιστού. Για τη μεταφορά των πληροφοριών που περιέχει η ιστοσελίδα χρησιμοποιείται το πρωτόκολλο επιπέδου εφαρμογής HTTP (hypertext transfer protocol). Το HTTP μεταφέρει πόρους (resources). Πόρος είναι μια μονάδα πληροφορίας που προσδιορίζεται μοναδικά από ένα URL (το R στο ακρωνύμιο URL). Η πιο κοινή περίπτωση πόρου είναι το αρχείο. Όμως, ένας πόρος μπορεί είτε να δημιουργείται δυναμικά ως αποτέλεσμα εντολών (π.χ. PHP script), είτε να είναι έγγραφο

διαθέσιμο σε πολλές γλώσσες ή οτιδήποτε άλλο. Σε κάθε ιστότοπο (ιστοσελίδα, web site) υπάρχει μία διεργασία εξυπηρετητή (web server), η οποία παρακολουθεί («ακούει») την TCP θύρα (port) 80 περιμένοντας εισερχόμενες συνδέσεις από πελάτες πλοηγούς (web clients). Μόλις εγκατασταθεί η σύνδεση, ο πελάτης στέλνει μία αίτηση (HTTP request) και ο εξυπηρετητής στέλνει μία απάντηση (HTTP response) που περιέχει διάφορα αντικείμενα. Μετά η σύνδεση απελευθερώνεται, δηλαδή, το HTTP δε διατηρεί πληροφορία κατάστασης (stateless) μεταξύ διαδοχικών δοσοληψιών (transactions).

Όπως έχει ήδη αναφερθεί, το HTTP είναι πρωτόκολλο επιπέδου εφαρμογής και ακολουθεί το υπόδειγμα πελάτη/εξυπηρετητή (client/server model). Οι μορφές της αίτησης και απόκρισης HTTP είναι ταυτόσημες και χρησιμοποιούν αγγλικές λέξεις. Και στις δύο περιπτώσεις τα μηνύματα περιλαμβάνουν:

- Μια αρχική γραμμή,
- Καμία ή πολλές γραμμές επικεφαλίδων,
- Μια κενή γραμμή (blank line), δηλαδή, τους χαρακτήρες <CR><LF> , και
- Προαιρετικά το κυρίως σώμα του μηνύματος HTTP (π.χ. ένα αρχείο, τα δεδομένα μιας ερώτησης, κλπ.).

Η αρχική γραμμή είναι διαφορετική για τις αιτήσεις και αποκρίσεις. Η αρχική γραμμή των ερωτήσεων περιέχει τρία μέρη που χωρίζονται με κενά: το όνομα της μεθόδου (πάντα με κεφαλαία γράμματα), την τοπική διαδρομή (local path) του αιτούμενου πόρου και τέλος την έκδοση του πρωτοκόλλου HTTP που χρησιμοποιείται. Η αρχική γραμμή των αποκρίσεων αποκαλείται γραμμή κατάστασης. Περιέχει και αυτή τρία μέρη που χωρίζονται με κενά: την έκδοση του HTTP, έναν κωδικό κατάστασης που υποδεικνύει το αποτέλεσμα της αίτησης και μια Αγγλική λέξη που περιγράφει τον κωδικό κατάστασης. Οι γραμμές επικεφαλίδων παρέχουν πληροφορίες για την αίτηση ή την απόκριση ή για το αντικείμενο που περιέχει το κυρίως σώμα. Εμφανίζονται όπως το σύνηθες κείμενο: μια επικεφαλίδα ανά γραμμή της μορφής "Header-Name: value" και καταλήγουν σε <CR><LF>. Το κυρίως σώμα του μηνύματος HTTP ακολουθεί τις επικεφαλίδες. Στις αποκρίσεις, το κυρίως σώμα επιστρέφει τον αιτούμενο πόρο στον πελάτη ή κάποιο επεξηγηματικό κείμενο σε περίπτωση λάθους. Στις αιτήσεις, αυτό είναι το μέρος όπου αρχεία ή δεδομένα που εισάγει ο χρήστης αποστέλλονται στον εξυπηρετητή. Όταν το μήνυμα HTTP περιέχει σώμα, συνήθως, οι επικεφαλίδες περιγράφουν το σώμα (τον τύπο και το μήκος του).

Μέχρι σήμερα έχουν αναπτυχθεί 2 εκδόσεις του πρωτοκόλλου, η HTTP 1.0 (περιγράφεται στο RFC 1945) και η HTTP 1.1 (περιγράφεται στο RFC 2068). Οι συνδέσεις HTTP διακρίνονται σε μη επίμονες (Non-persistent HTTP) και σε επίμονες (Persistent HTTP). Στις μη επίμονες συνδέσεις HTTP, μέσω μίας σύνδεσης TCP μπορεί να αποσταλεί το πολύ ένα αντικείμενο κάθε φορά. Το πρωτόκολλο HTTP/1.0 χρησιμοποιεί μη επίμονες συνδέσεις. Στην περίπτωση των επίμονων συνδέσεων HTTP, μέσω της ίδιας σύνδεσης TCP μπορούν να αποσταλούν πολλαπλά αντικείμενα μεταξύ πελάτη-εξυπηρετητή. Το πρωτόκολλο HTTP/1.1 χρησιμοποιεί επίμονες συνδέσεις. Περισσότερες πληροφορίες για το HTTP μπορείτε να βρείτε στην ιστοσελίδα <http://www.w3.org/Protocols/>

Σε αυτή την άσκηση θα καταγραφούν τα μηνύματα HTTP που παράγονται κατά την επίσκεψη μιας ιστοσελίδας με χρήση του Internet Explorer. Προτού αρχίσετε την καταγραφή να φροντίσετε να αδειάσετε την προσωρινή/κρυφή μνήμη (cache) του πλοηγού. Να επιλέξετε Tools -> Internet Options, στην πινακίδα (tab) General να πιέσετε το κουμπί Delete, στην οθόνη που θα εμφανισθεί να επιλέξετε το Delete Files, να επιβεβαιώσετε την πρόθεσή σας, να περιμένετε να ολοκληρωθεί η διαγραφή και μετά να κλείσετε τα παράθυρα διαλόγου.

Επειδή τα μηνύματα HTTP μεταφέρονται σε περισσότερα από ένα τμήματα TCP (και τελικά πακέτα IP) αφού ξεκινήσετε το Wireshark, να ακολουθήσετε από το μενού του κεντρικού παραθύρου τη διαδρομή Edit -> Preferences..., να κάνετε κλικ στο σύμβολο δίπλα στο Protocols στην αριστερή πλευρά του παραθύρου, κατόπιν να εντοπίσετε και να κάνετε κλικ στο πρωτόκολλο HTTP και να βεβαιωθείτε ότι όλες οι επιλογές περί ανασύνθεσης και αποσυμπίεσης είναι επιλεγμένες. Στη συνέχεια, στο πρωτόκολλο TCP να βεβαιωθείτε ότι το «Allow subdissector to reassemble TCP streams» είναι επιλεγμένο και να φροντίσετε το «Validate the TCP check-sum if possible» να μην είναι επιλεγμένο. Τέλος, να πιέσετε OK για να κλείσει το παράθυρο και να εφαρμοσθούν οι αλλαγές σας. Με τις επιλογές αυτές μηνύματα που μεταφέρονται σε

περισσότερα από ένα πακέτα θα αποκωδικοποιηθούν από το Wireshark ως πλήρη μηνύματα HTTP και όχι αποσπασματικά.

Να ξεκινήσετε μια καταγραφή να επισκεφτείτε μία ιστοσελίδα με πρωτόκολλο http (όχι https, π.χ. την <http://grad.cs.aueb.gr/>) και να σταματήσετε την καταγραφή όταν φορτωθεί πλήρως η σελίδα. Να εφαρμόσετε κατάλληλο φίλτρο απεικόνισης ώστε να παραμείνουν μόνο μηνύματα του πρωτοκόλλου HTTP.

- a. Ποιο είναι το φίλτρο απεικόνισης που εφαρμόσατε;
- b. Ποια είναι η έκδοση του πρωτοκόλλου HTTP που χρησιμοποιεί ο πλοηγός ιστού σας και ποια η έκδοση του πρωτοκόλλου HTTP που χρησιμοποιεί ο εξυπηρετητής ιστού;
- c. Ποιες γλώσσες εκτός της Αγγλικής δηλώνει ενδεχομένως ο πλοηγός ιστού σας ότι μπορεί να δεχτεί από τον εξυπηρετητή;
- d. Να εφαρμόσετε κατάλληλο φίλτρο ώστε να παραμείνουν μόνο τα πρώτα τμήματα TCP των τριπλών χειραψιών που διεξήχθησαν με τον εξυπηρετητή www.aueb.gr. Ποια είναι η σύνταξη του φίλτρου; Πόσες συνδέσεις TCP έγιναν και ποιες οι αντίστοιχες θύρες πηγής;
- e. Αφού απενεργοποιήσετε το φίλτρο απεικόνισης της ερώτησης 2.4, να εφαρμόσετε νέο φίλτρο ώστε να παραμείνουν μόνο οι εντολές HTTP προς τον εξυπηρετητή ιστού (HTTP requests). Ποια είναι η σύνταξη του; Πόσες εντολές HTTP απέστειλε ο υπολογιστής σας προς τον εξυπηρετητή ιστού;
- f. Να περιγράψετε τα πεδία των τμημάτων με τα οποία αποστέλλεται το περιεχόμενο της σελίδας.
- g. Να εξηγήσετε τη λογική των sequence και acknowledgement numbers των τμημάτων TCP που έχουν γίνει captured.
- h. Εκτός από την ίδια την ιστοσελίδα, ο πλοηγός ιστού ζήτησε και κάποιες εικόνες. Πόσες εικόνες κατέβασε ο πλοηγός ιστού; Οι εικόνες επιστρέφονται στον υπολογιστή σας από την ίδια διεύθυνση IP?

Επανάκτηση HTML σελίδας

Η ανάγκη για μείωση του χρόνου εξυπηρέτησης των αιτήσεων των πελατών και η ανάγκη για μείωση του φορτίου στις ζεύξεις πρόσβασης των τοπικών δικτύων οδήγησαν στην επινόηση της τεχνικής της ενδιάμεσης αποθήκευσης (Web caching). Στόχος ήταν η ταχύτερη ικανοποίηση του αιτήματος του πελάτη ακόμη και χωρίς την ανάμιξη του αρχικού εξυπηρετητή πηγής. Ο χρήστης ρυθμίζει τον πλοηγό του ώστε η περιήγηση στον ιστό να γίνεται μέσω ενδιάμεσου εξυπηρετητή (proxy server). Εάν τα ζητούμενα αντικείμενα περιέχονται εκεί (π.χ, αν τα έχει ήδη αναζητήσει κάποιος άλλος χρήστης του ιδίου δικτύου προηγουμένως) επιστρέφονται από τον ενδιάμεσο εξυπηρετητή. Διαφορετικά, ο ενδιάμεσος εξυπηρετητής ζητά τα αντικείμενα από τον εξυπηρετητή πηγής, κρατά ένα αντίγραφο και τα προωθεί στον πελάτη.

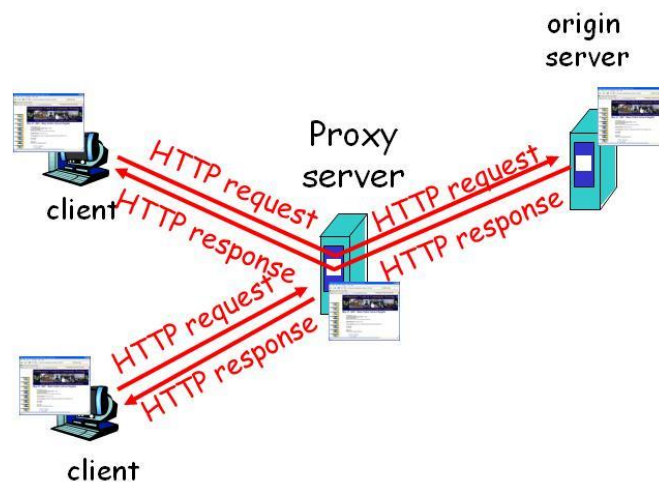


Figure 5. HTTP

Μια συμπληρωματική λύση στο ίδιο πρόβλημα δόθηκε μέσω των πλοηγών ιστού. Συγκεκριμένα, οι περισσότεροι πλοηγοί σήμερα κάνουν τοπική αποθήκευση, δηλαδή, κρατούν στο δίσκο του υπολογιστή διάφορα αντικείμενα από τις ιστοσελίδες που επισκέπτονται. Όταν λοιπόν ο χρήστης επισκεφτεί ξανά την ίδια σελίδα, τότε αντί για την HTTP μέθοδο GET εκτελείται η conditional GET. Σύμφωνα με την τελευταία, το ζητούμενο αντικείμενο μεταφέρεται μόνο εάν ισχύουν οι συνθήκες που περιγράφονται στην επικεφαλίδα της εντολής.

Σε αυτό το μέρος της άσκησης θα καταγραφούν τα μηνύματα HTTP που παράγονται κατά την επίσκεψη μιας ιστοσελίδας με χρήση του Internet Explorer. Πριν ξεκινήσετε την καταγραφή να φροντίσετε να αδειάσετε την προσωρινή/κρυφή μνήμη (cache) του πλοηγού. Αφού ξεκινήσετε μια νέα καταγραφή με το Wireshark, να επισκεφθείτε τη σελίδα <http://grad.cs.aueb.gr/>. Αμέσως να ανανεώσετε τη σελίδα πατώντας το κουμπί Reload. Κατόπιν να σταματήσετε την καταγραφή και να εφαρμόσετε κατάλληλο φίλτρο απεικόνισης ώστε να παραμείνουν μόνο μηνύματα του πρωτοκόλλου HTTP.

- i. Να παρατηρήσετε τις επικεφαλίδες πρωτοκόλλου HTTP του πρώτου μηνύματος τύπου GET. Υπάρχει γραμμή IF-MODIFIED-SINCE; Ποιος είναι ο κωδικός κατάστασης (status code) που επιστρέφει ο εξυπηρετητής ως απόκριση στο μήνυμα αυτό;
- j. Όπως είδαμε παραπάνω το ζητούμενο αντικείμενο δεν είχε τροποποιηθεί και γι' αυτό το λόγο ο πλοηγός ιστού απεικόνισε στο χρήστη το ήδη αποθηκευμένο αντικείμενο. Στην αντίθετη περίπτωση, δηλαδή, εάν το ζητούμενο αντικείμενο είχε τροποποιηθεί, τότε ο πλοηγός ιστού θα κατέβαζε το τροποποιημένο αντικείμενο από τον εξυπηρετητή, θα το αποθήκευε στη μνήμη cache και θα το απεικόνιζε στο χρήστη. Συνεπώς, να κάνετε τις ανωτέρω παρατηρήσεις και για την μεταφορά μιας άλλης, πιο δυναμικά μεταβαλλόμενης ιστοσελίδας, και να επαληθεύσετε ότι όντως ο πλοηγός ιστού θα κατέβαζε εκ νέου το τροποποιημένο αντικείμενο.

Στη συνέχεια να ξεκινήσετε μια νέα καταγραφή με το Wireshark. Με τον Internet Explorer να επισκεφθείτε τη σελίδα <http://grad.cs.aueb.gr/>. Να περιμένετε λίγο περισσότερο από ένα λεπτό και στη συνέχεια να ανανεώσετε τη σελίδα πατώντας το κουμπί Refresh. Κατόπιν να σταματήσετε την καταγραφή και να εφαρμόσετε κατάλληλο φίλτρο απεικόνισης ώστε να παραμείνουν μόνο μηνύματα του πρωτοκόλλου HTTP.

- k. Ποιος είναι ο κωδικός κατάστασης (status code) που επιστρέφει ο εξυπηρετητής ως απόκριση στο πρώτο μήνυμα HTTP τύπου GET του πλοηγού ιστού για το κατέβασμα της σελίδας; Πότε τροποποιήθηκε για τελευταία φορά το περιεχόμενο αυτό, σύμφωνα με την απάντηση στο πρώτο μήνυμα HTTP τύπου GET;
- l. Να παρατηρήσετε τις επικεφαλίδες του πρωτοκόλλου HTTP του πρώτου μηνύματος απόκρισης από τον εξυπηρετητή. Υπάρχει γραμμή με πεδίο Set-Cookie; Ποια είναι η τιμή της;
- m. Να παρατηρήσετε τα πεδία του πρωτοκόλλου HTTP του επόμενου μηνύματος τύπου GET. Υπάρχει γραμμή με πεδίο Cookie; Ποια είναι η τιμή της;

4) Γενικές ερωτήσεις και ασκήσεις

Άσκηση 1

Έστω ένας υπολογιστής ο οποίος χρησιμοποιεί στην διεπαφή του με το δίκτυο το μηχανισμό του τρύπιου κουβά για την μορφοποίηση της μετάδοσης δεδομένων. Ο μέγιστος επιτρεπτός ρυθμός εισαγωγής δεδομένων στο δίκτυο είναι 2.5MByte/s και ο ρυθμός μετάδοσης στο σύνδεσμο από τον υπολογιστή προς τον κουβά 3.5MByte/s.

- a) Έστω ότι ο υπολογιστής επιθυμεί να στείλει 350MB στο δίκτυο και τα στέλνει με μία ριπή (burst). Ποια πρέπει να είναι η ελάχιστη χωρητικότητα που θα έχει ο κουβάς προκειμένου να μην συμβεί απώλεια δεδομένων;
- b) Έστω ότι η χωρητικότητα του κουβά είναι 200MB. Να παρουσιάσετε το προφίλ της κίνησης εξόδου όταν η κίνηση εισόδου είναι η ανωτέρω ριπή των 350MB.
- c) Για χωρητικότητα του κουβά ίση με 200MB, ποια είναι η μεγαλύτερη χρονική διάρκεια της ριπής (burst) από τον υπολογιστή, ώστε να μην συμβεί απώλεια δεδομένων;

Άσκηση 2

Έστω ότι σε ένα δίκτυο η εισαγόμενη κίνηση ρυθμίζεται με τη βοήθεια ενός κουβά κουπονιών. Ο ρυθμός παραγωγής κουπονιών του κουβά αυτού είναι 10 Mbytes/sec, η χωρητικότητα του είναι 1 Mbyte και ο μέγιστος ρυθμός μετάδοσης δεδομένων είναι 50 Mbytes/sec.

Ποια είναι η μέγιστη διάρκεια της ριπής εξόδου όταν η αποστολή των δεδομένων γίνεται με την προαναφερθείσα μέγιστη ταχύτητα αποστολής και ο κουβάς είναι αρχικά γεμάτος; Ποιος είναι ο συνολικός χρόνος εξόδου μιας ριπής εισόδου συνολικής διάρκειας 40 msec? Να παρουσιάσετε το προφίλ της κίνησης εξόδου και του περιεχομένου του κουβά στην περίπτωση αυτή.

Εφόσον έχουμε κουβά κουπονιών ισχύει:

Ρυθμός εξόδου κίνησης = μέγιστος ρυθμός μετάδοσης (εισόδου), όταν στον κουβά υπάρχουν κουπόνια προς κατανάλωση.

Ρυθμός εξόδου κίνησης = ρυθμός τροφοδοσίας του κουβά με κουπόνια, όταν ο κουβάς είναι άδειος.