

## ***HTTP lab 2023***

Σε αυτό το εργαστήριο, θα εξερευνήσουμε διάφορες πτυχές του πρωτοκόλλου HTTP: τη βασική αλληλεπίδραση GET/απόκρισης, μορφές μηνυμάτων HTTP, ανάκτηση μεγάλων αρχείων HTML, ανάκτηση αρχείων HTML με ενσωματωμένα αντικείμενα και έλεγχο ταυτότητας και ασφάλεια HTTP. Πριν ξεκινήσετε αυτά τα εργαστήρια, ίσως θελήσετε να αναθεωρήσετε την Ενότητα 2.2 του βιβλίου.

### **HTTP GET/response**

Ας ξεκινήσουμε την εξερεύνηση του HTTP κατεβάζοντας ένα πολύ απλό αρχείο HTML - ένα αρχείο που είναι πολύ σύντομο και δεν περιέχει ενσωματωμένα αντικείμενα. Κάντε τα εξής:

- Εκκινήστε το πρόγραμμα περιήγησής σας.
- Εκκινήστε το Wireshark packet sniffer, όπως περιγράφεται στο Εισαγωγικό εργαστήριο (αλλά μην ξεκινήσετε ακόμα τη λήψη πακέτων). Εισαγάγετε "http" (μόνο τα γράμματα, όχι τα εισαγωγικά και με πεζά γράμματα) στο παράθυρο προδιαγραφών φίλτρου οθόνης, έτσι ώστε μόνο τα μηνύματα HTTP που έχουν καταγραφεί θα εμφανίζονται αργότερα στο παράθυρο λίστας πακέτων. (Εδώ μας ενδιαφέρει μόνο το πρωτόκολλο HTTP και δεν θέλουμε να δούμε την ακαταστασία όλων των πακέτων που έχουν καταγραφεί).
- Περιμένετε λίγο περισσότερο από ένα λεπτό (θα δούμε γιατί σύντομα) και μετά ξεκινήστε τη λήψη πακέτων Wireshark.
- Εισαγάγετε τα ακόλουθα στο πρόγραμμα περιήγησής σας
- <http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file1.html>
- Το πρόγραμμα περιήγησής σας θα πρέπει να εμφανίζει το πολύ απλό αρχείο HTML μιας γραμμής.
- Σταματήστε τη σύλληψη πακέτων Wireshark.

1. Το πρόγραμμα περιήγησής σας εκτελεί HTTP έκδοση 1.0, 1.1 ή 2; Ποια έκδοση του HTTP τρέχει ο διακομιστής;
2. Ποιες γλώσσες (αν υπάρχουν) δείχνει ότι το πρόγραμμα περιήγησής σας μπορεί να αποδεχτεί στον διακομιστή;
3. Ποια είναι η διεύθυνση IP του υπολογιστή σας; Ποια είναι η διεύθυνση IP του διακομιστή [gaia.cs.umass.edu](http://gaia.cs.umass.edu);
4. Ποιος είναι ο κωδικός κατάστασης που επιστρέφεται από τον διακομιστή στο πρόγραμμα περιήγησής σας;
5. Πότε τροποποιήθηκε τελευταία φορά το αρχείο HTML που ανακτάτε στον διακομιστή;
6. Πόσα bytes περιεχομένου επιστρέφονται στο πρόγραμμα περιήγησής σας;

## HTTP CONDITIONAL GET/response

Θυμηθείτε από την Ενότητα 2.2.5 του κειμένου, ότι τα περισσότερα προγράμματα περιήγησης ιστού εκτελούν προσωρινή αποθήκευση αντικειμένων και επομένως συχνά εκτελούν GET υπό όρους κατά την ανάκτηση ενός αντικειμένου HTTP. Πριν εκτελέσετε τα παρακάτω βήματα, βεβαιωθείτε ότι η προσωρινή μνήμη του προγράμματος περιήγησής σας είναι άδεια . Τώρα κάντε τα εξής:

- Εκκινήστε το πρόγραμμα περιήγησής σας και βεβαιωθείτε ότι η προσωρινή μνήμη του προγράμματος περιήγησής σας έχει διαγραφεί, όπως συζητήθηκε παραπάνω.
- Εκκινήστε τον ανιχνευτή πακέτων Wireshark
- Εισαγάγετε την ακόλουθη διεύθυνση URL στο πρόγραμμα περιήγησής σας
- <http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file2.html>
- Το πρόγραμμα περιήγησής σας θα πρέπει να εμφανίζει ένα πολύ απλό αρχείο HTML πέντε γραμμών.
- Εισαγάγετε ξανά γρήγορα την ίδια διεύθυνση URL στο πρόγραμμα περιήγησής σας (ή απλώς επιλέξτε το κουμπί ανανέωσης στο πρόγραμμα περιήγησής σας)
- Σταματήστε τη λήψη πακέτων Wireshark και πληκτρολογήστε "http" (και πάλι, με πεζά χωρίς τα εισαγωγικά) στο παράθυρο προδιαγραφών φίλτρου οθόνης, έτσι ώστε μόνο τα μηνύματα HTTP που έχουν καταγραφεί να εμφανίζονται αργότερα στο παράθυρο λίστας πακέτων.

Απάντησε τις παρακάτω ερωτήσεις:

8. Επιθεωρήστε τα περιεχόμενα του πρώτου αιτήματος HTTP GET από το πρόγραμμα περιήγησής σας στον διακομιστή. Βλέπετε μια γραμμή "IF-MODIFIED-SINCE" στο HTTP GET;

9. Επιθεωρήστε τα περιεχόμενα της απόκρισης διακομιστή. Ο διακομιστής επέστρεψε τα περιεχόμενα του αρχείου; Από που φαίνεται αυτό;

10. Τώρα επιθεωρήστε τα περιεχόμενα του δεύτερου αιτήματος HTTP GET από το πρόγραμμα περιήγησής σας στον διακομιστή. Βλέπετε μια γραμμή "IF-MODIFIED-SINCE:" στο HTTP GET ; Εάν ναι, ποιες πληροφορίες ακολουθούν την κεφαλίδα "IF-MODIFIED-SINCE:";

11. Τι είναι ο κωδικός και η φράση κατάστασης HTTP που επιστράφηκε από τον διακομιστή ως απόκριση σε αυτό το δεύτερο HTTP GET; Ο διακομιστής επέστρεψε τα περιεχόμενα του αρχείου; Εξηγείστε.

## Ανάκτηση μεγάλων εγγράφων

Στα μέχρι τώρα παραδείγματά μας, τα έγγραφα που ανακτήθηκαν ήταν απλά και σύντομα αρχεία HTML. Ας δούμε στη συνέχεια τι συμβαίνει όταν κατεβάζουμε ένα μεγάλο αρχείο HTML. Κάντε τα εξής:

- Εκκινήστε το πρόγραμμα περιήγησής σας και βεβαιωθείτε ότι η προσωρινή μνήμη του προγράμματος περιήγησής σας έχει διαγραφεί, όπως συζητήθηκε παραπάνω.
- Εκκινήστε τον ανιχνευτή πακέτων Wireshark
- Εισαγάγετε την ακόλουθη διεύθυνση URL στο πρόγραμμα περιήγησής σας <http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html>
- Το πρόγραμμα περιήγησής σας θα πρέπει να εμφανίζει τη μάλλον μακροσκελή Διακήρυξη Δικαιωμάτων των ΗΠΑ.
- Σταματήστε τη λήψη πακέτων Wireshark και πληκτρολογήστε "http" στο παράθυρο προδιαγραφών φίλτρου οθόνης, έτσι ώστε να εμφανίζονται μόνο μηνύματα HTTP που έχουν καταγραφεί.

Απάντησε τις παρακάτω ερωτήσεις :

12. Πόσα μηνύματα αιτήματος HTTP GET έστειλε το πρόγραμμα περιήγησής σας; Ποιος αριθμός πακέτου περιέχει το μήνυμα GET message το κείμενο της Διακήρυξης των Δικαιωμάτων (Bill of Rights);
13. Ποιος αριθμός πακέτου περιέχει τον κωδικό κατάστασης και τη φράση που σχετίζονται με την απάντηση στο αίτημα HTTP GET;
14. Ποιος είναι ο κωδικός κατάστασης και η φράση στην απάντηση;
15. Πόσα τμήματα TCP που περιέχουν δεδομένα χρειάζονταν για να φέρουν την ενιαία απόκριση HTTP και το κείμενο της Διακήρυξης των Δικαιωμάτων;

## Έγγραφο HTML με ενσωματωμένα αντικείμενα

Τώρα που είδαμε πώς το Wireshark εμφανίζει την κυκλοφορία πακέτων που έχει συλληφθεί για μεγάλα αρχεία HTML, μπορούμε να δούμε τι συμβαίνει όταν το πρόγραμμα περιήγησής σας κατεβάζει ένα αρχείο με ενσωματωμένα αντικείμενα, δηλαδή ένα αρχείο που περιλαμβάνει άλλα αντικείμενα (στο παρακάτω παράδειγμα, αρχεία εικόνας ) που είναι αποθηκευμένα σε άλλους διακομιστές.

Κάντε τα εξής:

- Εκκινήστε το πρόγραμμα περιήγησής σας και βεβαιωθείτε ότι η προσωρινή μνήμη του προγράμματος περιήγησής σας έχει διαγραφεί, όπως συζητήθηκε παραπάνω.
- Εκκινήστε τον ανιχνευτή πακέτων Wireshark
- Εισαγάγετε την ακόλουθη διεύθυνση URL στο πρόγραμμα περιήγησής σας <http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html>

Το πρόγραμμα περιήγησής σας θα πρέπει να εμφανίζει ένα σύντομο αρχείο HTML με δύο εικόνες. Αυτές οι δύο εικόνες αναφέρονται στο βασικό αρχείο HTML. Δηλαδή, οι ίδιες οι εικόνες δεν περιέχονται στο HTML. Αντίθετα, οι διευθύνσεις URL για τις εικόνες περιέχονται στο αρχείο HTML που έχετε λάβει. Όπως συζητήθηκε στο εγχειρίδιο, το πρόγραμμα περιήγησής σας θα πρέπει να ανακτήσει αυτά τα λογότυπα από τις υποδεικνυόμενες τοποθεσίες web. Το λογότυπο του εκδότη μας ανακτήθηκε από τον ιστότοπο [gaia.cs.umass.edu](http://gaia.cs.umass.edu). Η εικόνα του εξωφύλλου της 8ης έκδοσης (ένα από τα αγαπημένα μας εξώφυλλα) αποθηκεύεται σε έναν διακομιστή στη Γαλλία.

- Σταματήστε τη λήψη πακέτων Wireshark και πληκτρολογήστε "http" στο παράθυρο προδιαγραφών φίλτρου οθόνης, έτσι ώστε να εμφανίζονται μόνο μηνύματα HTTP που έχουν καταγραφεί.

Απάντησε τις παρακάτω ερωτήσεις :

16. Πόσα μηνύματα αιτήματος HTTP GET έστειλε το πρόγραμμα περιήγησής σας; Σε ποιες διευθύνσεις Διαδικτύου στάλθηκαν αυτά τα αιτήματα GET;

17. Μπορείτε να πείτε εάν το πρόγραμμα περιήγησής σας κατέβασε τις δύο εικόνες σειριακά ή αν λήφθηκαν από τις δύο ιστοσελίδες παράλληλα; Εξηγείστε.

## Έλεγχος ταυτότητας HTTP

Τέλος, ας δοκιμάσουμε να επισκεφτούμε έναν ιστότοπο που προστατεύεται με κωδικό πρόσβασης και να εξετάσουμε τη σειρά των μηνυμάτων HTTP που ανταλλάσσονται για έναν τέτοιο ιστότοπο. Το URL [http://gaia.cs.umass.edu/wireshark-labs/protected\\_pages/HTTP-wireshark-file5.html](http://gaia.cs.umass.edu/wireshark-labs/protected_pages/HTTP-wireshark-file5.html) προστατεύεται με κωδικό πρόσβασης. Το όνομα χρήστη είναι "wireshark-students" (χωρίς τα εισαγωγικά) και ο κωδικός πρόσβασης είναι "network" (και πάλι, χωρίς τα εισαγωγικά). Ας αποκτήσουμε, λοιπόν, πρόσβαση σε αυτόν τον «ασφαλή» ιστότοπο που προστατεύεται με κωδικό πρόσβασης. Κάντε τα εξής:

- Βεβαιωθείτε ότι η προσωρινή μνήμη του προγράμματος περιήγησής σας έχει διαγραφεί, όπως αναφέρθηκε παραπάνω, και κλείστε το πρόγραμμα περιήγησής σας. Στη συνέχεια, ξεκινήστε το πρόγραμμα περιήγησής σας
- Εκκινήστε τον ανιχνευτή πακέτων Wireshark
- Εισαγάγετε την ακόλουθη διεύθυνση URL στο πρόγραμμα περιήγησής σας [http://gaia.cs.umass.edu/wireshark-labs/protected\\_pages/HTTP-wireshark-file5.html](http://gaia.cs.umass.edu/wireshark-labs/protected_pages/HTTP-wireshark-file5.html)
- Πληκτρολογήστε το όνομα χρήστη και τον κωδικό πρόσβασης που ζητήσατε στο αναδυόμενο πλαίσιο.
- Σταματήστε τη λήψη πακέτων Wireshark και πληκτρολογήστε "http" στο παράθυρο προδιαγραφών φίλτρου οθόνης, έτσι ώστε μόνο τα μηνύματα HTTP που έχουν καταγραφεί θα εμφανίζονται αργότερα στο παράθυρο λίστας πακέτων.
- Σημείωση: Εάν δεν μπορείτε να εκτελέσετε το Wireshark σε ζωντανή σύνδεση δικτύου, μπορείτε να χρησιμοποιήσετε το "κλασικό" ίχνος πακέτων http-ethereal-trace-5 ή άλλα πρόσθετα ίχνη, όπως σημειώνονται στην υποσημείωση 2, για να απαντήσετε στις παρακάτω ερωτήσεις.

Απάντησε τις παρακάτω ερωτήσεις :

18. Ποια είναι η απάντηση του διακομιστή (κωδικός κατάστασης και φράση) ως απόκριση στο αρχικό μήνυμα HTTP GET από το πρόγραμμα περιήγησής σας;

19. Όταν το πρόγραμμα περιήγησής σας στέλνει το μήνυμα HTTP GET για δεύτερη φορά, ποιο νέο πεδίο περιλαμβάνεται στο μήνυμα HTTP GET;