Übung 2

Florian Hofer, Werner Nutt
Thomas Tschager

2. Wireshark I: Protokoll-Stack und Ethernet

Lernziele

Verstehen, wie Protokolle und Schichten in Paketen repräsentiert sind. Untersuchen, wie Frames der Bitübertragungsschicht im Detail aufgebaut sind.

Anforderungen

Wireshark: Dieses Labor verwendet das Software-Tool Wireshark, um Netzwerkverkehr zu erfassen und zu untersuchen. Wenn Wireshark nicht bereits auf Ihrem Computer installiert ist, können Sie es von www.wireshark.org herunterladen.

Schritt 1: Erfassen eines Trace

Erzeugen Sie einen Trace, der die Protokollstruktur von Paketen zeigt. Rufen Sie in Ihrem Browser eine URL Ihrer Wahl auf und laden Sie die entsprechende Webseite herunter. Der Datenverkehr zwischen dem Webserver und dem Client, liefert den Trace.

1. *Schließen Sie unnötige Browser-Tabs und -Fenster.* Durch die Minimierung der Browseraktivität verhindern Sie, dass Ihr Computer unnötige Webinhalte abruft, und Sie vermeiden zufälligen Datenverkehr im Trace.
2. *Starten Sie Wireshark und beginnen Sie ein Capture mit einem Filter von „*`tcp port 80`*“ oder „*`tcp.port == 80`*“.* Aktivieren Sie „`enable network name resolution`“. Dieser Filter zeichnet nur den Standard-Webverkehr auf und nicht andere Arten von Paketen, die Ihr Computer senden kann. Durch die Aktivierung übersetzt Wireshark die Adressen der Computer, die Pakete senden und empfangen, in Namen, die Ihnen zu erkennen helfen, ob die Pakete zu oder von Ihrem Computer gehen.
3. *Wenn das Capture gestartet wird, öffnen Sie eine einfache Webseite Ihrer Wahl in Ihrem Browser.*
4. *Nachdem das Laden erfolgreich war, kehren Sie zu Wireshark zurück und verwenden Sie die Menüs oder Schaltflächen, um die Aufzeichnung zu stoppen.* Wenn Sie es geschafft haben, wird das obere Wireshark-Fenster mehrere Pakete anzeigen, und höchstwahrscheinlich wird es voll sein. Wie viele Pakete erfasst werden, hängt von der Größe der Webseite ab. Es sollten mindestens 8 Pakete im Trace sein, typischerweise sind es 20-100. Glückwunsch, Sie haben einen Trace aufgenommen!

2: Untersuchen Sie den Trace

Wireshark erlaubt uns, ein Paket (aus dem oberen Panel) auszuwählen, und uns dann die Protokollschichten in Bezug auf die Header-Felder (im mittleren Panel) und die Bytes, aus denen sich das Paket zusammensetzt (im unteren Panel), anzeigen zu lassen. Beachten Sie, dass wir hier „Paket“ als Oberbegriff verwenden. Streng genommen wird eine Informationseinheit auf der Verbindungsebene als Frame bezeichnet. Auf der Netzwerkschicht wird es ein Paket, auf der Transportschicht ein Segment und auf der Anwendungsschicht eine Nachricht genannt. Wireshark sammelt Frames und stellt uns die übergeordneten Paket-, Segment- und Nachrichtenstrukturen vor, die innerhalb der Frames zu erkennen sind. Wir werden oft den Terminus "Paket" verwenden, da jeder Rahmen ein Paket enthält und es oft das Paket oder die Details auf höherer Ebene sind, die von Interesse sind.

Wählen Sie ein Paket aus, für das die Spalte Protokoll „HTTP“ und die Spalte Info ein GET ist. Es ist das Paket, das die von Ihrem Computer an den Server gesendete Web-(HTTP-)Anfrage enthält. (Sie können auf die Spaltenüberschriften klicken, und nach den jeweiligen Werten zu sortieren. Aber es sollte nicht schwierig sein, ein HTTP-Paket mit einfachem Nachschauen zu finden. Sehen wir uns genauer an, wie die Paketstruktur die verwendeten Protokolle widerspiegelt.

Da wir eine Webseite abrufen, wissen wir, dass die verwendeten Protokollschichten die unten gezeigten niedrig. Das heißt, HTTP ist das Webprotokoll auf der Anwendungsschicht, das zum Herunterladen von URLs verwendet wird. Wie viele Internet-Anwendungen läuft es auf den Protokollen der TCP/IP-Transport- und Vermittlungsschicht. Die Protokolle für die Sicherungs- und die Bitübertragungsschicht hängen von Ihrem Netzwerk ab, sind jedoch in der Regel kombiniert, entweder in Form von Ethernet (dargestellt), wenn Ihr Computer verkabelt ist, oder 802.11 (nicht dargestellt), wenn Ihr Computer über WiFi verbunden ist.

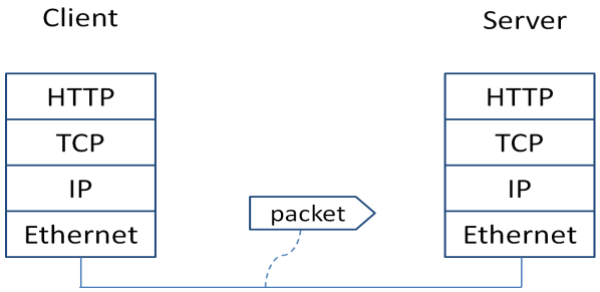

Abbildung 1: Protokoll-Stack beim Herunterladen einer Webseite

Wenn Sie das HTTP-GET-Paket ausgewählt haben, schauen Sie genau hin, um die Ähnlichkeiten und Unterschiede zu sehen zwischen dem Paket und unserem Protokoll-Stack. Die Protokollblöcke sind im mittleren Panel des Wireshark-Fensters aufgelistet. Sie können jeden Block expandieren (indem Sie auf das Dreiecks-Icon klicken) und seine Details zu sehen.

- Der erste Wireshark-Block ist „Frame“. Dies ist kein Protokoll, es ist ein Datensatz, der allgemeine Informationen über das Paket enthält, einschließlich wann es erfasst wurde und wie viele Bits es lang ist.
- Der zweite Block ist „Ethernet“. Das passt zu unserem Diagramm! Beachten Sie, dass Sie möglicherweise einen Trace auf einem Computer mit 802.11 (WiFi) erstellt haben, aber immer noch einen Ethernet-Block anstelle eines 802.11-Blocks sehen. Warum? Standardmäßig wandelt Wireshark den 802.11-Header in einen Pseudo-Ethernet-Header um.
- Dann kommen IP, TCP und HTTP, die genauso sind, wie wir es erwartet haben. Beachten Sie, dass die Reihenfolge von unten nach oben verläuft. Dies liegt daran, dass die Header-Informationen des Protokolls der unteren Schicht, wie in der Abbildung auf Folie 22 in Kapitel 2, an die Vorderseite der Informationen des Protokolls der oberen Schicht angefügt werden. Das heißt, die Protokolle der untersten Schicht stehen im Paket „auf dem Draht“ an erster Stelle.

Suchen Sie nun ein anderes HTTP-Paket, die Antwort des Servers an Ihren Computer, schauen Sie sich die Struktur dieses Pakets an und finden Sie die Unterschiede zum HTTP-GET-Paket heraus. Dieses Paket sollte „200 OK“ im Info-Feld haben, was einen erfolgreichen Abruf anzeigt. In unserem Trace sollten sich zwei zusätzliche Blöcke im Detail-Panel befinden:

- *Der erste zusätzliche Block sagt etwas wie* „[3 Reassembled TCP Segments ....]“. Die Details in Ihrem Capture werden variieren, aber dieser Block beschreibt mehr als das Paket selbst. Höchstwahrscheinlich wurde die Web-Antwort als eine Reihe von Paketen über das Netzwerk gesendet, die zusammengesetzt wurden, nachdem sie am Computer angekommen waren. Das Paket mit der Bezeichnung HTTP ist das letzte Paket in der Webantwort, und der Block listet Pakete auf, die zusammengefügt wurden, um die komplette Webantwort zu erhalten. Jedes dieser Pakete wird mit dem Protokoll TCP angezeigt, obwohl die Pakete einen Teil einer HTTP-Antwort enthalten. Nur das letzte Paket wird mit dem Protokoll HTTP angezeigt, wenn die komplette HTTP-Nachricht erkannt werden kann, und Wireshark listet die Pakete auf, die zusammengefügt wurden, um die HTTP-Antwort zu erhalten.
- *Der zweite zusätzliche Block beginnt mit* "Line-based text data ....". Die Details in Ihrem Capture werden variieren, aber dieser Block beschreibt den Inhalt der abgerufenen Webseite. In unserem Fall ist er vom Typ text/html, obwohl es sich auch um text/xml, image/jpeg oder viele andere Typen handeln könnte. Wie beim Frame-Block handelt es sich hier um kein echtes Protokoll. Stattdessen haben wir hier eine Beschreibung der Paketinhalte, die Wireshark produziert, um uns zu helfen, den Netzwerkverkehr zu verstehen.

Step 3: Paketstruktur

Um die Paketstruktur genau zu verstehen, geben Sie eine grafische Darstellung eines HTTP-GET-Pakets, das die Position und Größe der HTTP-, TCP-, IP- und Ethernet-Protokoll-Header in Bytes anzeigt. Ihre Abbildung kann das Gesamtpaket einfach als langes, dünnes Rechteck darstellen.

Um die jeweilige Größe zu berechnen, beachten Sie, dass Wireshark zu jedem Paket Informationen über die Größe der Nutzlast und der Header angibt. Zum Beispiel zeigt uns ein Klick auf den IP-Version 4-Header eines Pakets in unserem Trace, dass die Länge 20 Bytes beträgt. Sie können auch die gesamte Paketgröße verwenden, die in der Spalte Länge oder im Rahmendetailblock angezeigt wird.

Step 4: Protokoll-Overhead

Schätzen Sie den Download-Protokoll-Overhead oder den prozentualen Anteil der Download-Bytes, der für den Protokoll-Overhead verwendet wird. Betrachten Sie dazu die HTTP-Daten (Header und Message) als Nutzlast für das übertragende Netzwerk und Header der unteren Schichten (TCP, IP und Ethernet) als Overhead. Man möchte, dass dieser Overhead klein ist, so dass die meisten Bits verwendet werden, um Inhalte zu transportieren, die für Anwendungen wichtig sind. Um dies herauszufinden, schauen Sie sich zunächst nur die Pakete in der Download-Richtung für einen einzelnen Webabruf an. Sie können nach der Spalte Destination sortieren, um sie zu finden. Das erste Pakete sollte ein kurzes TCP-Paket sein, das als SYN ACK aufgelistet ist. Dies Paket leitet eine Verbindung ein. Es folgen meist längere Pakete in der Mitte (von etwa 1 bis 1,5KB), von denen das letzte ein HTTP-Paket ist. Dies ist der Hauptteil des Downloads. Das letzte wird ein kurzes TCP-Paket sein, das Teil der Beendigung der Verbindung ist. Überprüfen Sie für jedes Paket, wie viel Overhead es in Form von Ethernet/IP/TCP-Headern hat und wie viele nützliche HTTP-Daten es in der TCP-Nutzlast trägt. Schauen Sie sich auch das HTTP-Paket in Wireshark an, um zu erfahren, wie viele Daten in den TCP-Nutzlasten aller Download-Pakete enthalten sind.

Schritt 5: Demultiplex-Schlüssel

Wenn ein Ethernet-Frame auf einem Computer eintrifft, muss die Ethernet-Schicht das darin enthaltene Paket an die nächsthöhere zu verarbeitende Schicht übergeben. Der Vorgang, die richtige höhere Schicht zu finden, um empfangene Pakete zu verarbeiten, wird Demultiplexing genannt. Wir wissen, dass in unserem Fall die höhere Schicht IP ist. Aber woher weiß das Ethernet-Protokoll das? Schließlich hätte die höhere Schicht ein anderes Protokoll sein können (z.B. ARP). Wir haben das gleiche Problem auf der IP-Schicht - IP muss in der Lage sein festzustellen, dass der Inhalt der IP-Nachricht ein TCP-Paket ist, damit es an das TCP-Protokoll zur Verarbeitung übergeben werden kann. Die Antwort ist, dass Protokolle, um die höhere Schicht zu bestimmen, in ihrem Header Informationen verwenden, die als "Demultiplex-Schlüssel" bezeichnet werden.

Sehen Sie sich die Ethernet- und IP-Header eines Download-Pakets im Detail an, um die folgenden Fragen zu beantworten:

1. Welches Feld im Ethernet-Header ist der Demultiplex-Schlüssel, der ihm sagt, dass die nächsthöhere Schicht IP ist? Welcher Wert wird in diesem Feld verwendet, um "IP" anzuzeigen?
2. Welches Feld im IP-Header ist der Demultiplex-Schlüssel, der ihm sagt, dass die nächsthöhere Schicht TCP ist? Welcher Wert wird in diesem Feld verwendet, um "TCP" anzuzeigen?

Schritt 5: Gültigkeitsbereich der Ethernet-Adressen

Jeder Ethernet-Frame trägt eine Quell- und Zieladresse. Eine dieser Adressen ist die Ihres Computers. Sie ist die Quelle für die gesendeten Frames und das Ziel für empfangene Frames. Aber wie lautet die andere Adresse? Angenommen, Sie haben einen entfernten Internet-Server gepingt, dann kann dies nicht die Ethernet-Adresse des entfernten Servers sein, da ein Ethernet-Frame sich nur innerhalb eines LANs bewegt. Stattdessen wird es die Ethernet-Adresse des Routers oder des Standard-Gateways sein, wie z.B. Ihr Access Point im Falle von 802.11. Dies ist das Gerät, das Ihr LAN mit dem Rest des Internets verbindet. Im Gegensatz dazu geben die IP-Adressen im IP-Block jedes Pakets den Quell- und Zielpunkt der gesamten Route an. Dies sind Ihr Computer und der entfernte Server.

Fertigen Sie eine Abbildung an, die die relativen Positionen Ihres Computers, des Routers und des entfernten Servers zeigt. Markieren Sie Ihren Computer und den Router mit ihren Ethernet-Adressen. Markieren Sie Ihren Computer und den entfernten Server mit ihren IP-Adressen. Zeigen Sie auf der Zeichnung, bis wohin Ihr LAN geht und wo der Rest des Internets beginnt.

Finden Sie für sich selbst heraus …

Wir empfehlen Ihnen, Protokolle und Schichten zu explorieren, nachdem Sie diese Labor beendet haben. Einige Ideen:

- Sehen Sie sich ein kurzes TCP-Paket an, das keine Daten höherer Ebene enthält. Für welche Entität ist dieses Paket bestimmt? Wenn es keine Daten für höhere Schichten trägt, dann scheint es für ein höheres Protokoll wie HTTP nicht sehr nützlich zu sein!
- In einem klassischen Schichtmodell wird an eine Nachricht von einer höheren Ebene ein Header der unteren Ebene angehängt und sie wird zu einer neuen Nachricht. Aber das ist nicht immer der Fall. Oben sahen wir einen Trace, in dem die Web-Antwort (eine HTTP-Nachricht bestehend aus einem HTTP-Header und einer HTTP-Nutzlast) in mehrere Nachrichten der unteren Schicht (mehrere TCP-Pakete) umgewandelt wurde. Stellen Sie sich vor, Sie haben (wie in Schritt 2) die Paketstruktur für das erste und letzte TCP-Paket mit der Web-Antwort gezeichnet. Wie werden sich die Zeichnungen unterscheiden?
- In dem oben beschriebenen klassischen Schichtenmodell hängen die unteren Schichten Header-Zeilen an die von den höheren Schichten übergebenen Nachrichten an. Wie wird sich dieses Modell ändern, wenn eine untere Schicht Verschlüsselung/Komprimierung hinzufügt?