ネットワークを流れるパケットを覗いたり作ったりしてみよう

サイエンスワークショップ 2018 @多摩科学技術高校 2018/03/15-16

1. はじめに

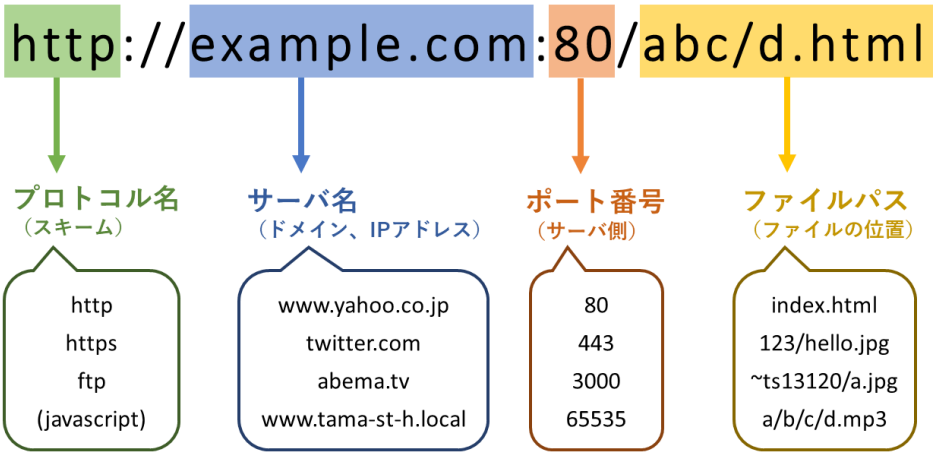
私たちが何気なく使っているネットワークはデータのまとまりであるパケットによってやり取りされています。今回はネットワークの基礎とパケットについて実際に手を動かしながら学習していきます。

2. Web の仕組みを理解しよう

この章では、身近で触れる機会の多い Web を切り口として、ネットワークの基礎と Web 技術について理解することを目指します。

2.1. URL について

URL(Uniform Resource Locator)とは、ネットワーク上に存在するデータを特定するために使われる統一された書き方です。国際規格の RFC3986 で定められています。 例えば Yahoo Japan のトップページの URL は https://www.yahoo.co.jp/です。



Web ブラウザでは URL が http://から始まる場合は、ポート番号を示す:80 を省略できます。同様に https://から始まる場合は、ポート番号:443 を省略できます。これは、80 番(http)と 443 番(https)がそれぞれのプロトコルの標準ポートであるためです。

演習

Web ブラウザを起動して、ポート番号を指定して Web サイトにアクセスしてみよう。URL が http から始まるサイトと https から始まるサイトの両方で試してみよう。

2.2. HTML について

私たちが目にする Web ページ（ホームページ）の多くは、HTML(HyperText Markup Language)という言語で記述されています。 以下は HTML で簡単な Web ページ作成した例です。

```
<!doctype html>
<html>
   <head>
      <title>こんにちは</title>
   </head>
   <body>
      <h1>ホームページです</h1>
      <p>日々向上の　熱き思い　響け世界に　多摩科学技術高校</p>
   </body>
</html>
```

HTML ではタグと呼ばれる記法によって記述されます。タグには開始タグと終了タグがあります。開始タグと終了タグの間に文字列やコンテンツを含めることで Web ページを作り上げます。

```
<開始タグ> 適当な文字列 <終了タグ>
```

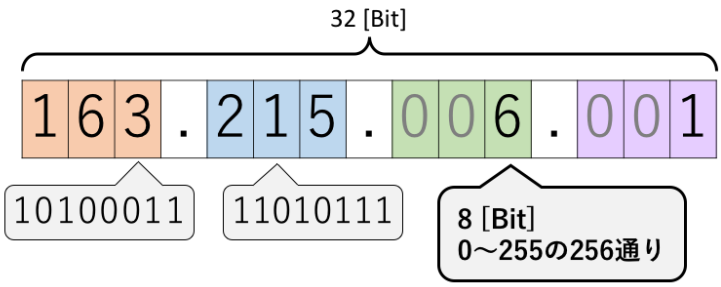
演習

テキストエディタを起動して上記の HTML を打ち込んだファイル hello.html を作成します。これを Web ブラウザで開き、どのように表示されるか確認してみよう。

2.3. IP アドレスについて

IP アドレスとはネットワーク上における住所です。ネットワーク内でコンピュータやスマートフォンなど個々の機器を識別するために利用されます。広く普及している IPv4(バージョン 4)の IP アドレスは 32 ビットで表現されます。

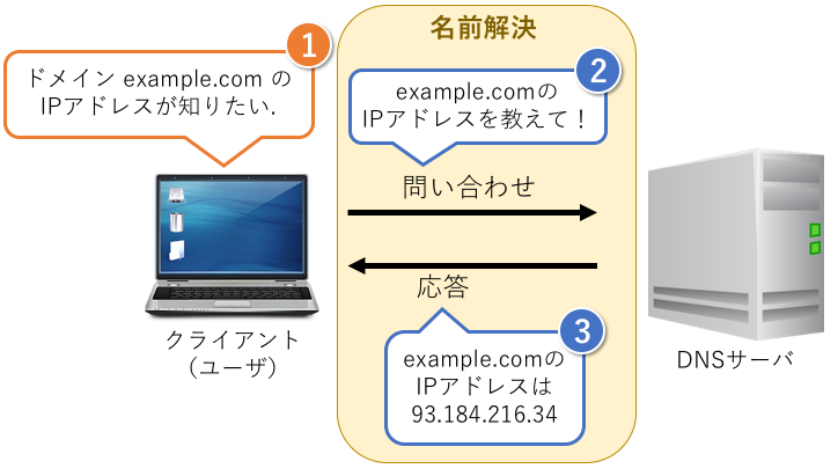
情報技術基礎で学んだように、1 ビットで表現できる情報は 0 と 1 の 2 通りです。すなわち、8 ビットで表現できる情報は $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^8 = 256$ 通りです。また、2 進数の桁数とその数の表現に必要なビット数は一致しています。例えば $(215)_{10}$ は $(11010111)_2$ であり 8 ビットで表現できます。



- 0と255は予約アドレスのため、実際には1〜254の254通り
- 32ビットで表現できるアドレスは4,294,967,296通り

2.4.　DNS について

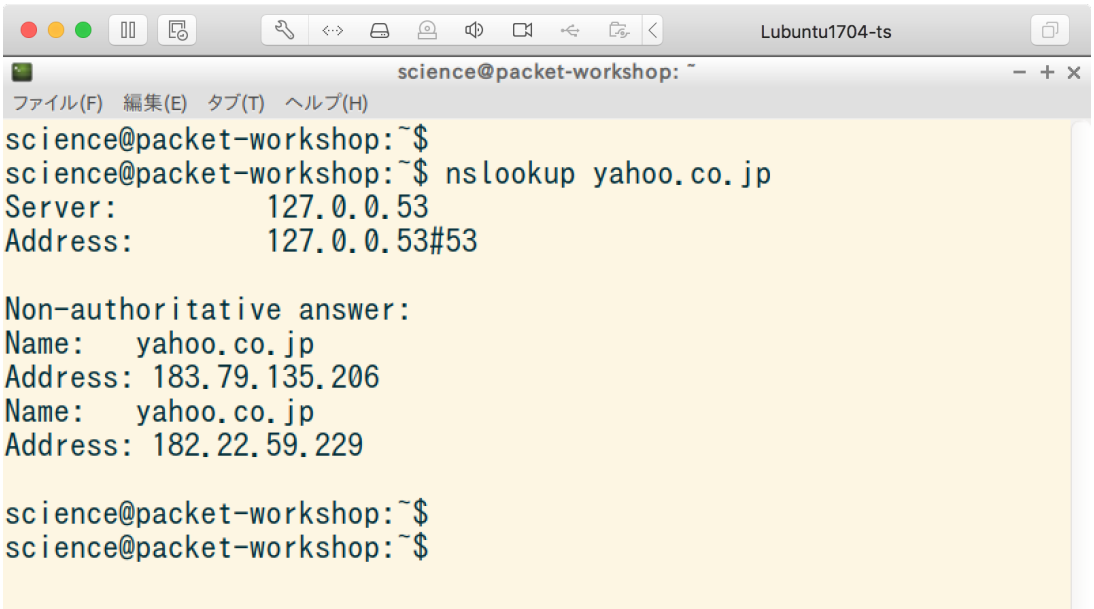
DNS(Domain Name System)とは、IP アドレスとドメインを結びつける仕組みです。ドメインとは英数字とハイフン、ドットから構成されるサイト毎の識別子です。ドメインの例には yahoo.co.jp や abema.tv などがあります。一般にドメインから IP アドレスを調べることは名前解決と呼ばれます。例えば、google.co.jp は名前解決をすると IP アドレスが 172.217.26.3 であることが分かります。



コマンドラインで名前解決する

以下では、Web サイトの URL からドメインを取り出してみます。つまり、名前解決をコマンドを打ち込むことによって行います。

①コマンドラインに nslookup yahoo.co.jp と入力します。



コマンドの実行結果から yahoo.co.jp の IP アドレスは 183.79.135.206 と 182.22.59.229 であることが分かります。

②次に IP アドレス 183.79.135.206 を Web ブラウザのアドレスバーに入力してアクセスしてみます。
これで DNS が IP アドレスとドメインを結びつけていることが確認できました。

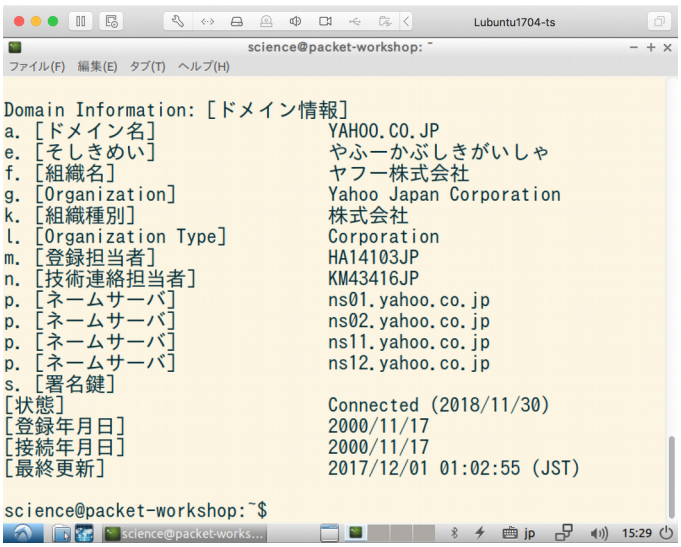
演習

http://50.60.70.80/のような IP アドレスによるアクセスではなく、DNS というドメイン http://example.com/によるアクセスが一般的である理由を考えてみよう。ポイントは「見た目」と「覚えやすさ」です。余裕のある人は、技術的な視点で「サーバの IP アドレスが変わった時」の DNS を利用することによるメリットを考えてみよう。

2.5. ドメインの所有者を調べる

それぞれのドメインには所有者情報が登録されています。これにより「どこにいる誰が所有しているか」を知ることが出来ます。インターネットのサービスで調べることも出来ますが、ここではコマンドラインで調べてみます。

①ターミナルを起動してコマンド whois yahoo.co.jp を入力して実行します。



②コマンドの実行結果を確認していきます。Organization と Registered Date と Last Update などを確認してみます。

演習

ドメイン tama-st-h.ed.jp の所有者、登録日、最終更新について調べてみよう。

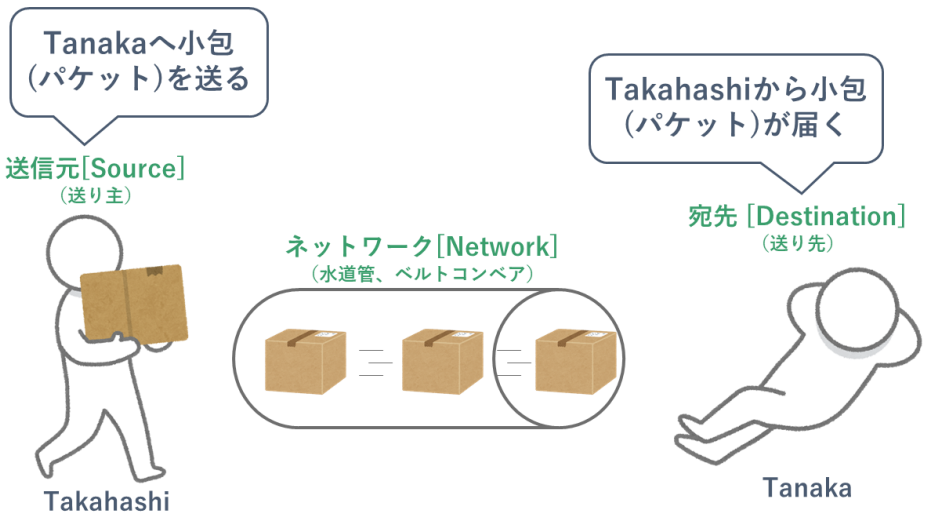
所有者	
登録日	
最終更新	

3. パケットに触れてみよう

この章では Web ブラウザで Web サイトにアクセスする際にどのようなパケットが送信されているか確かめます。また、パケットのキャプチャに関する基本操作の習得を目指します。

3.1. パケットとは

ネットワーク上におけるデータは、パケットとよばれる一定サイズのデータのまとまりによってやり取りされます。パケットには必ず送信元と送信先、タイプなどのあらかじめ定められた情報を付与する必要があります。この仕組みは、小包を離れた相手へ郵送することに似ています。

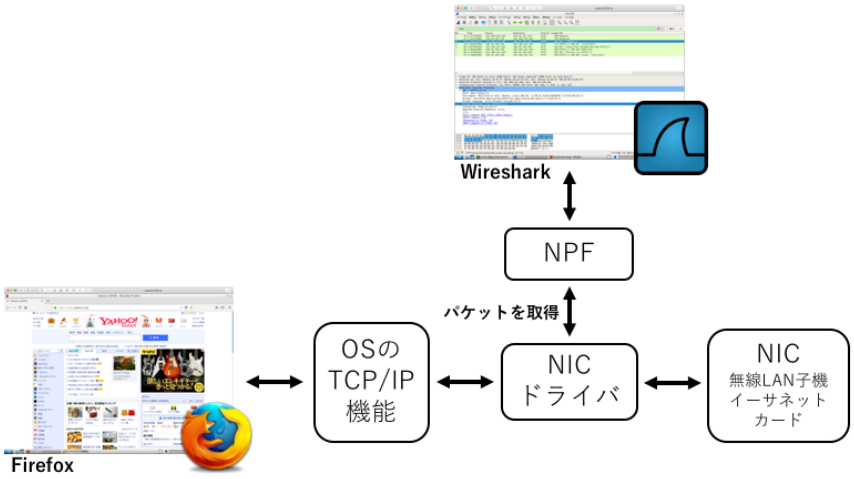


3.2. パケットのキャプチャ

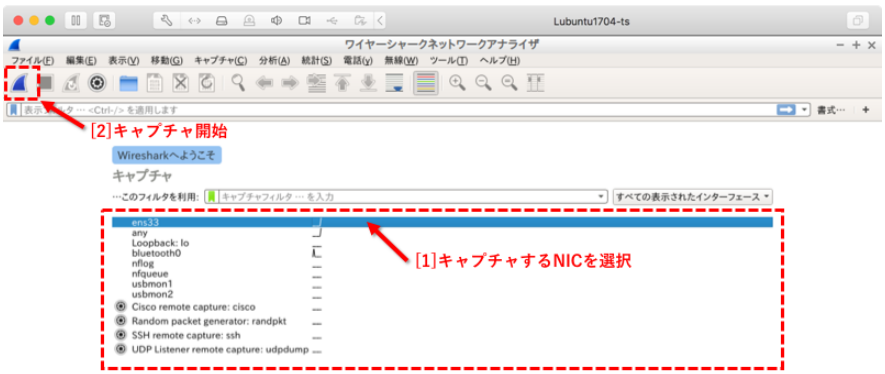
ネットワークを流れるパケットは専用のソフトウェアを使用することで、取得することが出来ます。一般にパケットを取得することは、パケットをキャプチャするとよばれます。今回は広く知られている Wireshark というソフトウェアを使用してパケットを見てみます。

3.3. Wireshark の使い方

Wireshark はコンピュータの NIC(ネット接続用のハードウェア)でやり取りされるパケットを取得してウィンドウに表示します。

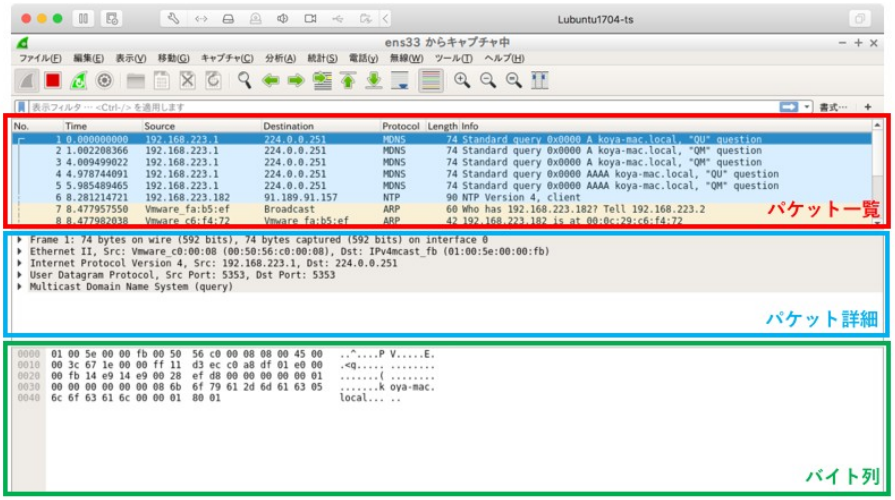


①Wireshark を起動する下図のウィンドウが表示されます。Wireshark の起動には管理者権限(Linux: root)が必要なため、一般ユーザからは「sudo wireshark &」で起動します。



画面中央で、キャプチャに使用する NIC「enp33」を選択します。ウィンドウの左上にある青いアイコンをクリックするとキャプチャが開始されます。

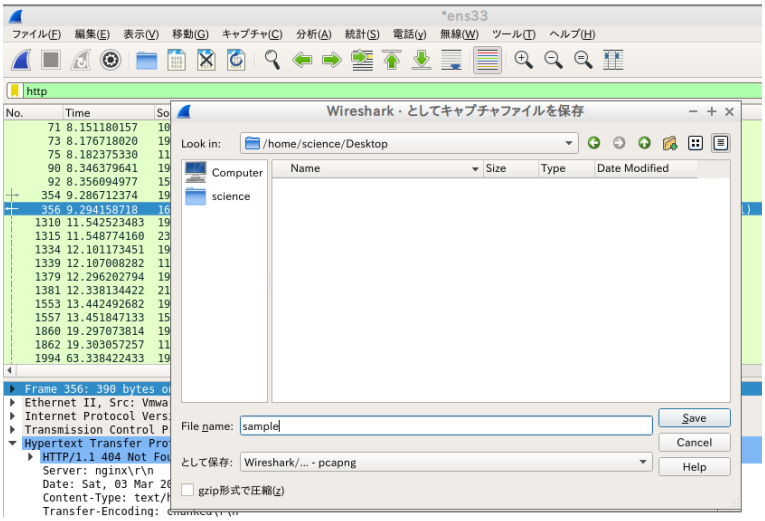
②キャプチャを開始すると下図のウィンドウが表示されます。



「パケット一覧」にはキャプチャしたパケットがリスト形式で表示されます。「パケット詳細」には「パケット一覧」で選択されたパケットの詳細が表示されます。「バイト列」には選択されたパケットの 16 進数表記が表示されます。

③キャプチャを終了するには左上にある赤い四角のアイコンをクリックします。

④キャプチャしたパケットの保存は、メニューバーの「ファイル」から「保存(S)」を選び、ファイル名と保存先を指定することで行なえます。



演習

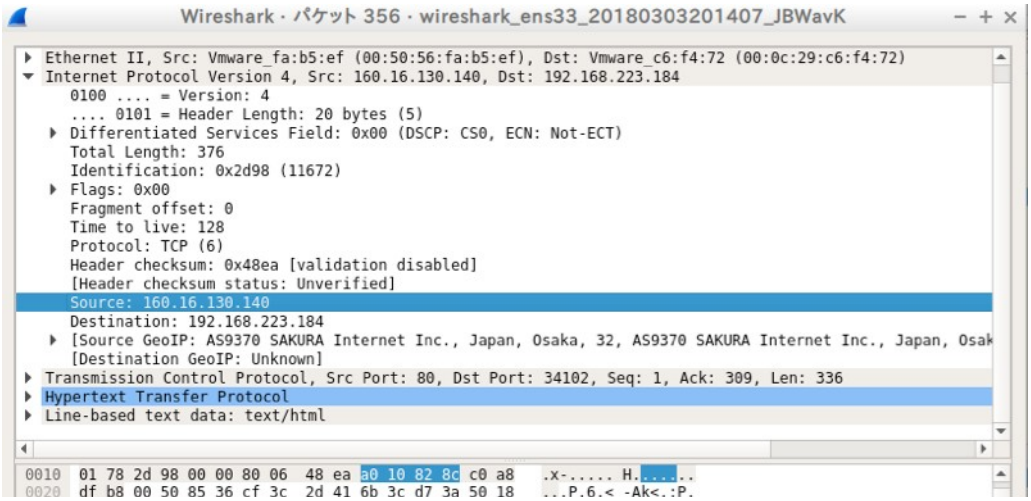
Wireshark でパケットキャプチャをしてみよう。キャプチャしたデータを sample.pcapng としてデスクトップに保存してみよう。

3.4. パケットの構造

パケットは下図のような階層構造になっています。上にいくほどレイヤーが低くハードウェアに近くなります。逆に下にいくほどソフトウェアに近くなります。

Wireshark	役割	OSI 参照モデル(参考)
–	0 と 1 を電圧の高低や光の点滅で表現, ケーブルやコネクタの形状の規定	物理層
Frame	通信媒体で直接接続された機器間でのパケットの転送と識別	データリンク層
Ethernet II		
IPv4 [Internet Protocol Version 4]	アドレスの管理と経路の選択	ネットワーク層
TCP [Transmission Control Protocol]	データ転送の管理と信頼性向上	トランスポート層
HTTP [Hypertext Transfer Protocol]	通信の状態管理	セッション層
	データ・フォーマットの交換	プレゼンテーション層
	特定のアプリケーションに特化した通信	アプリケーション層

パケット一覧から適当なパケットを選びダブルクリックします。左側に表示される▶をクリックすることで各レイヤーの情報を見ることが出来ます。例えば、IP(Internet Protocol)を開くと 12 行目に「Souce: 160.16.130.140」と表示されています。つまり、このパケットは IP アドレスが 160.16.130.140 のサーバから送信されたものであると分かります。



演習

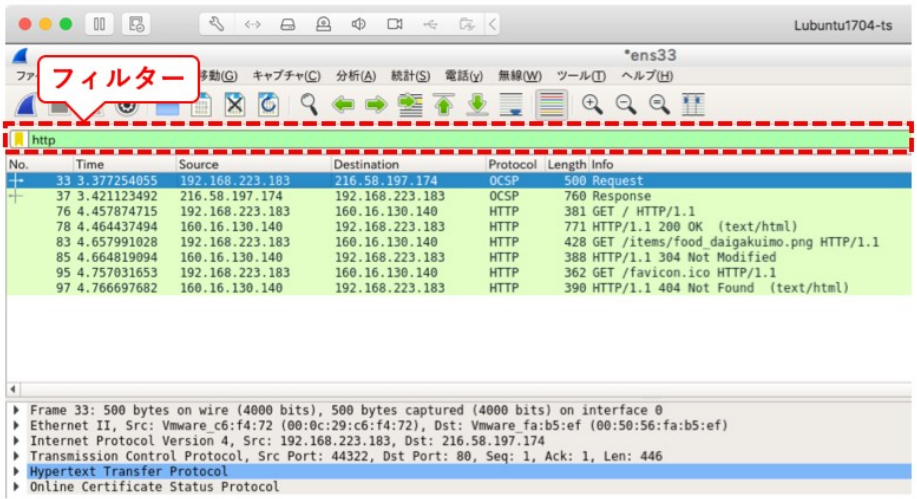
Wireshark でキャプチャしながら多摩科学技術高校の Web サイトへアクセスして、Web サーバの IP アドレスをパケットから調べてみよう。余裕のある人は、別の方法でも確認してみよう。

3.5. HTTP パケットをキャプチャ して解析

Wireshark を使って HTTP パケットの解析を行ってみます。パケットを解析することで、Web ページのデータがどのように通信されているか理解することを目指します。

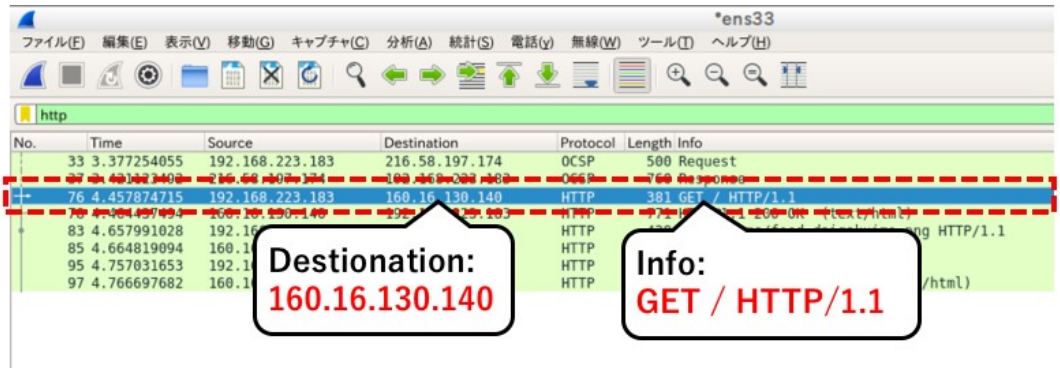
練習

①Wireshark を起動し、パケットのキャプチャを開始します。
②Firefox を起動して http://ddos.tokyo/にアクセスしてみます。
③Wireshark でパケットのキャプチャを停止します。
④Wireshark のウィンドウ上部にあるフィルタでキャプチャしたパケットを絞り込みます。



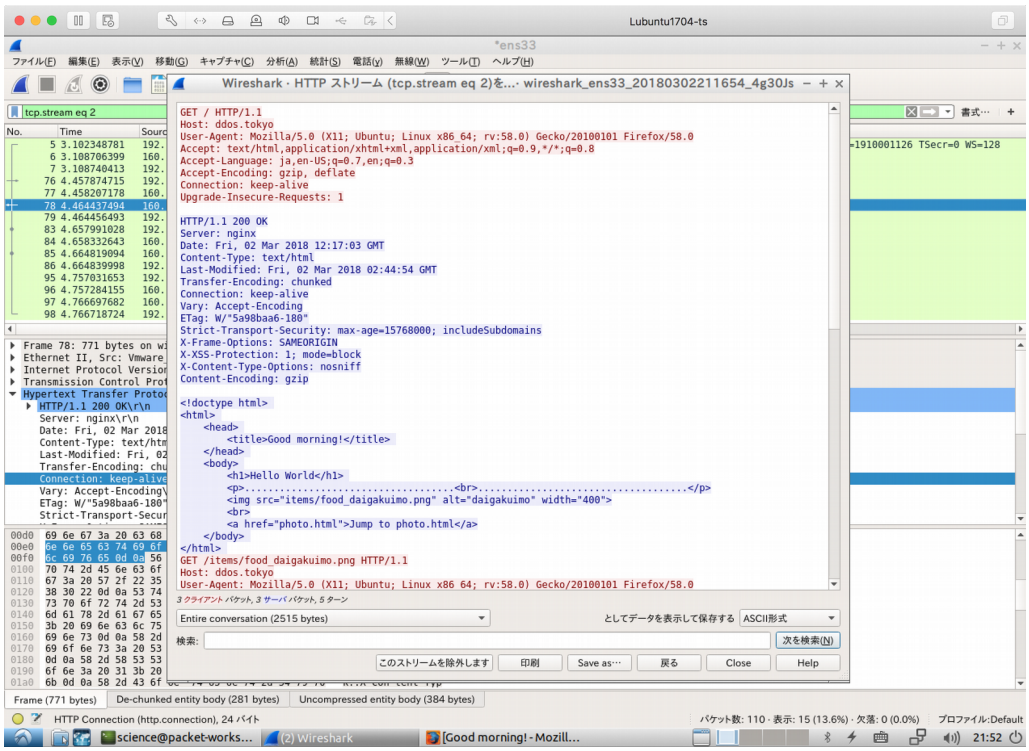
フィルターで「http」を指定して Enter を押します。これによって全てのパケットの中から HTTP のみが表示されます。パケットが大量にある場合、条件でパケットを絞り込むことで効率的に解析が行えます。

⑤表示されたパケットの中から Destination(宛先)が 160.16.130.140 かつ Info が「GET / HTTP/1.1」のパケットを探して選択します。



⑥選択したパケットのマウスカーソルをあわせて右クリックします。メニューから「追跡」→「HTTP ストリーム」の順でクリックします。この操作によって、Web サーバと手元のコンピュータでやり取りされたデータを順序ごとに見ることが出来ます。

⑦表示されたウィンドウに表示された文字列のうち、赤色の箇所は「手元のコンピュータから Web サーバへ」送信したデータです。逆に、青色の箇所は「Web サーバから手元のコンピュータへ」送信されたデータ（手元のコンピュータから見れば受信）です。



演習 1

上記の手順から考えられたことや気づいたことをまとめてみよう。 ポイントは「HTML」や「通信の内容が読めるか」などです。送信元(Source)と送信先(Destination)にも注目してみよう。

演習 2

ここでは、先ほど閲覧した Web ページに生年月日や氏名、クレジットカード番号などの個人情報が含まれていた場合を考えてみます。私たちがそうした Web ページへアクセスした時に、悪意のある誰かがパケットをキャプチャしていました。この時、どのような危険性があるか考えてみよう。ポイントは演習 1 と同じように「通信の内容が読めるか」です。

演習 3

この演習では、いつも私たちが安全にオンラインショッピングや Web メールが利用できる理由を確かめてみます。これによって通信の安全性と暗号化について理解します。

① Wireshark を起動してキャプチャを開始します。
② Firefox を起動して、「**https**://ddos.tokyo/form.html」にアクセスします。
③ Wireshark のキャプチャを停止します。
④先ほどと同様にフィルターを使って HTTP パケットだけを表示してみます。

練習と演習 3 で異なる点について考えてみよう。ポイントは「HTTP パケットの数」です。余裕がある人は HTTP パケット以外のパケットの中で増えたパケットを探してみよう。

4. おわりに

このワークショップを通じて、通信の仕組みやネットワークについて少しでも興味を持ってもらえれば幸いです。学校の授業よりも難易度は高めに設定し、演習など自分で考える箇所も多くしました。この資料の内容について、さらに調べたり先生に質問したりすることでスキルを磨いてください。

現在、世の中で動作するソフトウェアのほとんどがネットワークを利用しているものです。ネットワークを理解することは、ネットワークやインフラのエンジニアのみならず、プログラマにも必要な技術（スキル）です。ぜひ、基礎的な部分だけでも学んでおくことをオススメします。