

インターネット

増原 英彦

1 情報ネットワーク

今日の計算機の多くが接続されているインターネットは、情報ネットワークの一つである。情報ネットワークとは、計算機をはじめとする機器や人間などの複数の参加者がお互いに情報をやりとりするための仕組みであり、郵便制度、電信、電話、ニュースの配信、為替取引、銀行オンラインなど様々なものがある。

計算機どうしを接続するネットワークは、一つの部屋、建物、敷地の中のように狭い範囲の計算機どうしをつなぐものと、より遠方の計算機をつなぐものに分けられる。前者は LAN (local area network)、後者は WAN (wide area network) と呼ばれる。

計算機ネットワークは主に電気信号を使って情報を伝えているため、通信容量(一定時間に送受信できるデータ量)と、通信距離は両立しない。従って歴史的には、LAN としては大容量の通信をするネットワークが、WAN としては小容量の通信をするネットワークが別々に運用されてきた。

2 インターネット

インターネット(the Internet) は LAN のような小さなネットワークどうしを接続して作られた1つのネットワークである。インターネットでは、各計算機はインターネット・プロトコル(IP) と呼ばれる手順で通信をする。以降で説明するいくつかの特徴から、非常に多くの計算機が接続された、世界中に広がったネットワークに成長している。なお、広い意味でのインターネット (internet) は、「ネットワークどうしをつなぐネットワーク」のことを指し、IP に依らないものも含む。また、IP を用いているが、いわゆる「インターネット」には接続されていない独立したネットワークも含まれる。

2.1 IP による通信の仕組み

インターネットの通信の仕組みを理解するために、ある計算機 A と別の計算機 B との間の通信の際の手順を概観しよう。(例えば、A と B はそれぞれ異なる組織のメールサーバであった場合を考えてみればよい。)

まず、前提として A,B を始めとするインターネット上の全ての計算機には IP アドレスと呼ばれる番号が付けられている。IP アドレスは 32 ビット¹の整数であるが、8 ビットの数字を 4 つ並べた形で書かれることが多い。例えば、10011101010100100010001100100011 という 2 進数は IP アドレスの一例であるが、8 ビットずつに分けて 10011101 01010010 00100011 00100011 とし、さらにそれぞれを 10 進数にして 157.82.35.35 のように書く。計算機の間で通信が行われるときには、必ずこの IP アドレスを使って「何番の計算機から何番の計算機へ」という形で行われる。

¹現在普及している IP バージョン 4 の場合である。IP バージョン 6 では 128 ビットの整数が使われる。

1つのIPアドレスの上位ビットはネットワークアドレスと呼ばれ、その計算機が接続されているLANなどのネットワークに割り当てられている番号になっている。つまり、同じLANに接続されている機器のIPアドレスの上位何ビットかは同じ値になっている。上の例に出たIPアドレスであれば、先頭の16ビットがネットワークアドレスになっている。つまり、157.82.xx.xxというIPアドレスは物理的に同じネットワークに接続された機器だけが持つことになる。

計算機間の通信は、パケット交換と呼ばれる方式で行われる：

1. 計算機AがBにメッセージを送る際には、計算機Aはメッセージの内容を数百バイトの大きさのパケット(小包)と呼ばれる小さなデータに分ける。
2. 1つ1つのパケットには、送り主AのIPアドレス、送り先BのIPアドレスや時刻、パケットの番号、中身などが含まれる。(電子メールのメッセージにも似ているが、送り主や送り先は計算機であって、利用者ではないことに注意。)
3. 計算機Aは1つ1つのパケットを、自身が接続されているLANへ順に送る。
4. LANにはネットワークスイッチあるいはインターネットルータと呼ばれるLANどうしをつなぐ機器が接続されており、各パケットを別のネットワークへ転送する。この際、ネットワークルータは送られて来たパケットをどこへ転送すべきかは、パケットに書き込まれている宛先のIPアドレスのネットワークアドレスを使って決める。
5. パケットは、ネットワークスイッチやインターネットルータを何個か通過して計算機Bが接続されているLANまで転送される。
6. 最後に、計算機Bが送られて来たパケットを受け取る。
7. 各パケットは転送途中のネットワークの混雑や故障などにより順番が変わってしまったり届かなかったりするものもある。そのため、計算機Bのソフトウェアが、パケットの中に書き込まれた番号などを使って順序を戻したり、届かなかったメッセージを計算機Aに送り直すように要求する(この要求もネットワークを通じて送られる)などして、元のメッセージを復元した上で処理される。

送られたメッセージに応じて計算機Bから計算機Aに返信が行われる際には、上と同様の手順でパケットが送られる。

2.2 特徴

インターネットでは接続されている全ての機器²が対等に通信できる。旧来の銀行オンラインシステムや鉄道発券オンラインシステムでは、サーバと呼ばれる中心的な計算機と、端末の役割を果たす計算機から構成されており、そこではサーバと端末の間でしか通信ができないのが普通であった。そういったものと比べると、より一般的なネットワークだと言える。

メッセージをパケットに分割して送る方法はまた、インターネットの通信コストや品質を決める大きな特徴の1つである。パケットに分割して送る方では、2地点をつなぐ1本の通信線の上に、様々な宛先のパケットを順に送ることができる。そのため、インターネットでは大抵の場合には高速な通

²計算機に限らずプリンタのような機器でもインターネットに接続されていることがある。

信が安価に行える。しかし同時に、どの程度の容量の通信ができるかの最低値は保証されていない。この性質はベストエフォートと呼ばれている。

この性質は、回線交換型と呼ばれるネットワークとの対比でよく説明できる。例えば電話ネットワークは回線交換型であるが、このようなネットワークで A から B へ通信を行うときは、まず A と B の間の経路 (回線) を確保し、それから実際のデータを送り、通信が終了すると確保していた経路を解放する。

例えば 2 地点を結ぶ通信線を用意する際に、その 2 地点では最大 100 組の計算機のペアが同時に通信することが見込まれていたとする。そのとき、

- 回線交換型では、1 Mbps の通信線 (回線) を 100 本用意する、一方、
- パケット交換型では、10 Mbps の通信線を 1 本だけ用意する。

(Mbps は通信容量の単位で 1 Mbps は 1 秒間に百万ビットのデータを送る能力を表わす。) 前者は合計で 100 Mbps の通信容量となるので、後者よりもコストが高い。このとき回線交換型では、2 地点間の計算機間の通信は常に 1 Mbps の容量で行われる。一方、パケット交換型では、同時に通信している計算機のペアの数によって、各計算機にとっての通信容量が変化する。例えば 1 組の計算機しか通信をしていないときは、その計算機が 10 Mbps の容量を全て使うことができるので、回線交換型よりも高速な通信ができる。一方、100 組の計算機が全て同時に通信したとすると、一組の計算機の通信は 0.1 Mbps 程度になるので、回線交換型よりも遅くなる。インターネットでは、非常に多数の計算機が接続されているけれども、同時に多数の計算機が大量の通信をすることはほとんどないため、パケット交換型のメリットが大きい。

インターネットはまた制御が分散している。インターネットで最も重要な点は、ネットワークルータが送られて来たパケットを適切なネットワークへ転送する機能であるが、この機能は各 LAN に接続されたネットワークルータが担っており、「全体を制御している 1 つの機械」というものが存在しない。そのため 1 つの機器が故障したとしても、残りのネットワーク全体は動き続けることができるため、故障に強いネットワークになっている。また、もし全体を制御する機械が存在していた場合、ネットワークが大きくなるにつれその機械が処理すべきデータが膨大なものになってしまう。制御が分散していることは、インターネットが今日のような非常に大きな規模のネットワークに急速に成長できた理由の 1 つだと言える。

問題 5-1: (パケット交換型対回線交換型) 1 節に挙げたインターネット以外の情報ネットワークを、パケット交換型と回線交換型に分類せよ。

問題 5-2: (ネットワークの方式とマルチメディア) 電話やビデオ放送のように、音声や動画のようなデータを実時間で送受信することを考えた場合、パケット交換型と回線交換型ではどのような利害得失があるかを考えよ。

問題 5-3: (ネットワークの方式と料金体系) 電話は通話時間に応じた料金がかかる。一方、専用の回線を使って企業や大学などがインターネットに接続した場合、料金は一定額であることが普通である。この違いをパケット交換型と回線交換型の違いによって説明できるだろうか?

問題 5-4: (携帯電話のデータ通信) 電話線を通してインターネットに接続する場合には、接続時間に応じて料金が徴収される。また、携帯電話を通してインターネットに接続する場合に、送受信したパケットの量に応じて課金を行うことが行われている。一方で、

3 ドメイン名システム

別の計算機と通信をするためには、その計算機の IP アドレスを知っている必要がある。しかし、数字の羅列をそのまま人間が使うには不便である。(cf. 電話番号) また、IP アドレスは、ネットワークの接続の仕方によって変更されることもしばしばあるため、長期間にわたって使うのにも不便である。

これを解決するため、インターネットではドメイン名システム (domain name system) を用いている。これは、人間にとって意味のある名前を計算機につけて、その名前から IP アドレスを調べる仕組みである。(cf. 電話帳)

インターネットには世界中の計算機が接続されているため、計算機に名前をつける際には、世界中で重複がないようにする必要がある。そのためドメイン名システム (domain name system) では、組織の構造を反映した階層的 (hierarchical) な名前をつけることにしている。

3.1 階層的な名前

階層的な名前のつけ方は、多数の要素があるときに、それぞれに固有の名前をつけるための方法である。これはドメイン名システムに限らず、(次回以降で説明する) ファイルシステムや、人間社会における住所の表記など多くの場面で用いられている。

その方法を少し形式的に説明すると次のようなものになる:

- まず、全ての要素を以下のようにして分類する
 - 要素をいくつかの集団あるいは要素に分ける
 - それぞれの集団あるいは要素に異なる名前をつける (*)
 - それぞれの集団に含まれる要素について、上の手順を繰り返して分類する
- ある要素または集団 x が集団 y に含まれていた要素だったとする。要素または集団 x の「階層的な名前」は、 y における x の名前 (上の (*) 部でつけたもの) に y の「階層的な名前」を加えたものとする

この方法の利点は、(*) 部で名前をつける際に、その集団の中だけで異なる名前をつければよい点である。逆の言い方をすれば、別の集団で同じ名前が使われていても問題がないということである。また、名前のつけ方に関する方針も、集団ごとに自由に決めて構わない。名前からその名前が表わすものへの関数を考えると、このような集団は関数の定義域 (ドメイン) と見ることができる。「ドメイン名システム」と名付けられているのはそのためである。

【例】東京大学教養学部の住所 住所を考える際には世界中の住所を分類することからはじめる。まずは《国》で分けられる。《国》に固有の名前をつけたものが「日本」である。次に日本中の住所は《都道府県》に分けられ、その名前が「東京都」である。別の国では別のやり方 (例えば《州》) で分けているかも知れないが、この方針は「日本」というドメインで決めればよい。さらに東京都は《区》に分けられ、その名前は「目黒」となる。東京都には「目黒」という名前の区は1つだけしかないが、もし他の都道府県に「目黒」という名前の区があっても構わない。目黒区は《町》³に分けられ、その名前は「駒場」となる。さらに《丁目》《番地》《号》と続いている。

³町という呼び方は適切でないかも知れない。

このようにして分類した場合、駒場の階層的な名前は、目黒区における名前の「駒場」に目黒区の階層的な名前「日本国東京都目黒区」を加えたものなので、「日本国東京都目黒区駒場」となる。普段住所を扱うときは日本というドメインを対象としていることが明らかなので、「日本国」というドメインにおける階層的な住所「東京都目黒区駒場」とだけ書くことも多い。

3.2 実際のドメイン名

インターネットで実際に使われているドメイン名は、次のようなものである。

- 一番大きな集団(トップドメインと呼ぶ)は、
 - 国名をアルファベット2文字で表わしたもの — 日本であれば「jp」、ツバルという国であれば「tv」など。
 - 米国では⁴、組織の性格をアルファベット3文字で表わしたもの — 商業(com)・教育機関(edu)・政府(gov)・任意団体(org)・ネットワークプロバイダ(net)など。

のどちらかを使う。米国が特別扱いになっているのは、歴史的な理由である。

- それ以下の階層は、それぞれの国の方針に従う。例えば日本では、2番目の階層には
 - 組織の性格をアルファベット2文字で表わしたもの — 商業(co)・学術機関(ac)・政府(go)・任意団体(or)・ネットワークプロバイダ(ne)など
 - 都道府県名をローマ字で書いたもの

のどちらか⁵を使う。また、ツバルのように、2番目以下の階層の名前を決める権利を私企業に売却している国もある。

- 階層的な名前を書く際には、

《集団内の名前》 《ピリオド》 《集団の階層的な名前》

の順に書く。つまり、より上位の集団が後に来る。

3.3 ドメイン名サーバ

ドメイン名から IP アドレスを決めるためには、どこかに電話帳のようなものがある必要がある。ドメイン名システムでは、名前から IP アドレスを調べるためのドメイン名サーバと呼ばれる計算機を使うことでこの「電話帳」を実現している。

およそ次のような仕組みである:

- 各ドメインには、そのドメインを担当するドメイン名サーバが存在する

⁴トップドメインによっては米国以外でも使えるものもある。

⁵最近はこれに加えて、任意の団体名を日本字とアルファベットにしたものを使えるようにする動きがある。

- ドメイン名サーバは、担当するドメインの中にある計算機の IP アドレスを答える (例えば、ecc.u-tokyo.ac.jp を担当するドメイン名サーバは、その中にある smtp という計算機の IP アドレスを答える)
- ドメイン名サーバは、担当するドメインの中に含まれるドメインを担当するサーバの IP アドレスを答える (例えば、u-tokyo.ac.jp を担当するドメイン名サーバは、その中にある ecc というドメインのドメイン名サーバを答える)
- ドメイン名サーバは、担当するドメイン外のドメイン名について、答えを返すことのできるドメイン名サーバの IP アドレスを答える (例えば、u-tokyo.ac.jp を担当するドメイン名サーバが u-tokyo.ac.jp 以外のドメイン名を聞かれたときには、トップドメインのドメイン名サーバを答える)
- トップドメインは、世界中に散在する数十台のサーバによって担当されている

一方、ネットワークに接続された各計算機は、その組織のドメイン名サーバの IP アドレスだけを知っている。その計算機が、あるドメイン名の IP アドレスを調べる際には、

1. まず、自組織のドメイン名サーバに問い合わせる
2. そのドメイン名サーバが IP アドレスを教えてくれれば完了
3. そのドメイン名サーバが、問い合わせるべきドメイン名サーバの IP アドレスを教えてくれれば、そのドメイン名サーバにもう一度問い合わせ、2に戻る

という手順を踏む。各ドメイン名サーバは、自組織外のドメイン名に対してはトップドメインを担当するサーバを⁶、そうでない場合には下位組織のドメイン名サーバを教えてくれる。結果として、ドメイン名の階層化を上位から下位に向かって問い合わせながら目的の情報を見つけることができる。

問題 5-5: (トップレベルドメイン) 国名以外のトップレベルドメインには com や edu などがあるが、それ以外にどのようなものがあるか。

問題 5-6: (jp ドメイン) 日本のトップドメインは jp で、その次の階層には co や ac などがある。それ以外にどのようなものがあるか。

4 実習: WWW と情報検索

練習 5-7: (WWW) HWB「10. ウェブブラウザ」に従い、WWW ブラウザの基本的な使用方法を理解せよ。以降の課題を行う際には⁷、見つけた WWW ページの場所や内容を後から参照したり、ダウンロードした WWW ページを元に自分のページを作成するので、ページの保存方法や栞 (ブックマーク) などの使い方を覚えておくとよい。

⁶実際には過去の問い合わせの結果を覚えておくなどの工夫がなされているため、直接 IP アドレスを答える場合もある。

⁷この授業の課題はもちろんだが、それ以外の授業や研究などの多くの場面で必要となることである。

練習 5-8: (検索エンジン) HWB「11. 情報検索」に従い、計算機ネットワークを通して情報を収集する方法の基本を理解せよ。

練習 5-9: (情報の検索) 以下の問いについて調べ、各問いについての答えとその根拠を示せ。WWW 検索エンジンによって調べた場合には、検索に用いたキーワードと、回答に用いた WWW ページが全体の何番目に提示されたかについても記すこと。

ただし、以下の点に注意せよ。(1) 調べる手段は WWW 検索エンジンに限らず、任意のものを用いてよい。(2) 答えが1つとは限らないような問いもある。(3) WWW ページには根拠の無い情報も多いので、明確な根拠があるかどうか、2つ以上の出典から確認できる内容であるかどうかを気をつけよ。

1. イラクに現在派兵している国のリスト
2. 1997年時点での EU 加盟国
3. マイクロソフトが過去5年間に提案した革新的な技術の中で、最も重要なもの3つ
4. ウェブログ (weblog) の意味と、その呼び方を最初に提案した人物の名前
5. 受動喫煙が胎児に与える主な影響
6. マイナスイオンが体に良い影響を与える原理