

新米エンジニアのための

初歩の

インターネット技術

《第4回 NSPをめぐるルーティング》

浅羽 登志也

asaba@ij.ad.jp

株式会社インターネットイニシアティブ

今回は、自分の組織の内部のネットワークをちょっと離れて、外の世界に目を向けてみます。いわゆるプロバイダーとは何か、プロバイダーの向こうには何があるのかを考えてみましょう。そのうえで、インターネット全体のルーティングがどのように行われているのかのイメージを、マクロ的な視点からとらえていきたいと思います。

はじめに

お客様に聞かれて答えに困る質問というのがいくつかある。

質問その1:「インターネットはどこ会社がやっているのですか?」

回答その1:「え? ええっと、その、あの、日本では私共や某アメリカ大手電話会社系 特別第二種電気通信事業者さんとか、某国内大手コンピュータメーカーの特別第二種電気通信事業者さんですとか、いろいろございます。アメリカには、UUNETやSprintなどがサービスをしていますか……」

質問その2:「つまり、インターネットはいくつもあるということですか?」

回答その2:「いえ、あの、全体では1つなんですけど、いくつものネットワークの集合体みたいなものなんです。つまり、インター、ネットなわけですね、ハイ。」

とまあこんな感じで分かったような分からないような説明になってしまう。では、その全体で1つという“インターネット”は、いったいどのようにして構成されているのだろうか?

NSPとIX

図1に、今日のインターネットがどのように構成されているのかのイメージを示す。NSPというのは、Network Service Providerの略で、インターネットへの接続サービスをエンドユーザーに提供するネットワーク事業者のことを指す。NSPには特定の目的のためにインターネット接続を提供しているものと、純粹に通信サービスとしてインターネット接続を提供しているものがある。

前者の例としてはもともとはNSF（全米科学財団）が資金を提供している研究をサポートするために作られたNSFNET（この号が書店に並ぶ頃にはなくなってしまっているだろうが）であったり、宇宙科学関係の研究者のためのNSN（NASA Science Network）などがある。日本では、ネットワークを研究するためのネットワークであるWIDEインターネットや、理科学系の大学や研究室をつなぐためのTISN、省庁をつなぐ省際ネットなどがある。

また後者としては、Alternet、ANS、PSI、Sprintlink、InternetMCIなどアメリカだけでも星の数ほど存在するし、日本でも、IIJやSpin、InfoWeb、Meshなど、この1年ほどの間にたくさんのNSPがサービスを開始している。

IXはInternet eXchangeの略で、複数のNSP間の相互接続ポイントである。異なるNSP間での相互接続のやり方としては、たとえば図1のNSP-AとNSP-Dのように直接リンクを張って相互接続するやり方もある。しかし、NSPの数が増え、それらの間の相互接続を行うためにいちいち直接リンクを張るのはコスト的に考えても無駄である。そこで、IXの考え方が登場した。複数のNSPが1か所にリンクを張り、そこを経由してそれぞれの間のデータのやりとりをしようというものである。

IXとしては、古くはFIX（Federal Internet eXchange）やCIX（Commercial Internet eXchange）などがあり、最近ではMAE（Metropolitan Area Ethernet）や、NAP（Network Access Point）などが

ある。

まとめると、NSPの1つひとつがインターネットのある一部分を運営していて、それらが、直接リンクを張り合ったり、またはIXを経由するなどして相互に接続されて全体として1つの大きなネットワークができあがっている。これが今日のインターネット像なのである。各NSPがその業績を伸ばし、どんどんネットワークを拡大していけばいくほど、全体としてのインターネットも大きくなるのである。もちろん、それに付随していろいろな技術的な問題点も出てくるのではあるが。

NSPを選ぶということ

さて、ではエンドユーザーから見るとこれはどういうことになるのだろうか？ともかくどれか1つのNSPからサービスを受ければ、大体の場合はインターネット全体にアクセスが可能となる。どのNSPを選ぶかにはいろいろなファクターがあるだろう。接続料金の安いところという選び方もあるだろうし、直接接続されている組織数が多いところという選び方もある。

先日出張先のボストンから成田に帰って来るときにこんなことがあった。同じくボストンから成田に帰って来るTさんとほぼ同じ時間のフライトだったのだが、それぞれ航空会社が違って、私はUnited Airlineで、TさんはAmerican Airlineであった。フライトスケジュールをよく聞いてみると、Tさんのほうが私より数分後の出発であるにもかかわらず、成田に着くのは数分早いのである。やられた！思わず心の中でそう叫んでしまった。

途中の経路も異なっていて、UAのほうはシカゴ経由、AAのほうはサンノゼ経由である。ううむ。であれば、AAでサンノゼでの乗り継ぎの時間が長めになるようにスケ

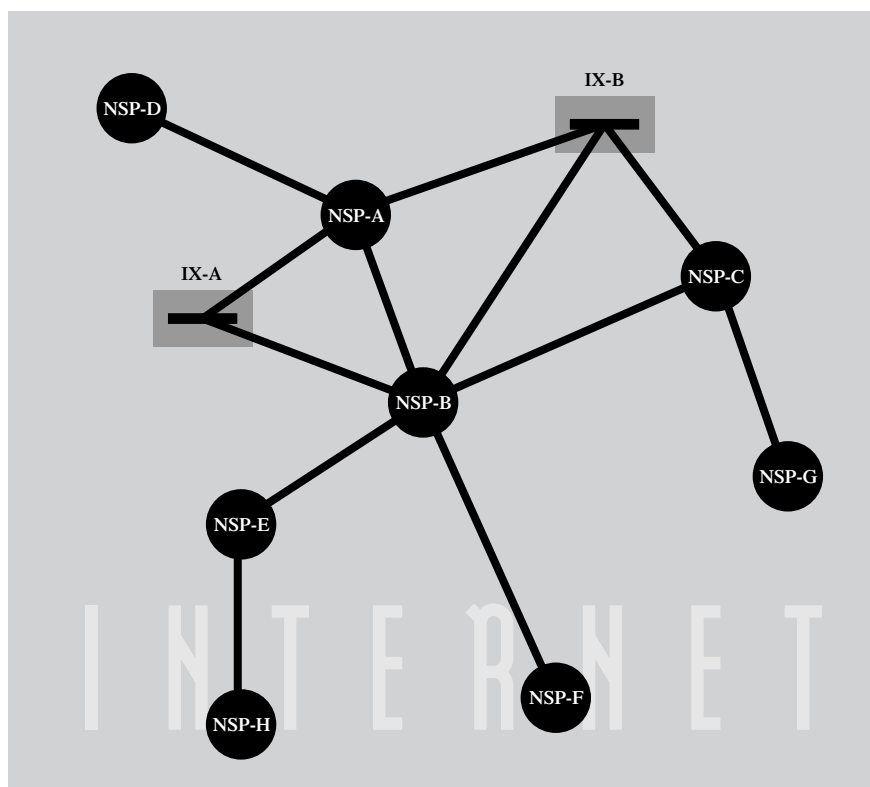


図1 インターネットの構成

ジュールすれば、そこで車を借りてひとっばしりすればPCショッピングができたのかあ。なるほど、なるほど。AAだとそういう楽しみもあるわけだな。でもでも、UAを使えばマイレージも稼げるし、レッドカーペットクラブだってあるじゃないか。そう考えるとやはりUAも捨てがたい。うむ。などと、ぶつぶつぶやきながら、それじゃあねとボストンでTさんと別れて成田に帰って来たわけだが、成田空港で、さてTさんはもうとっくに行ってしまったのかいな、とふと到着便の掲示板を見上げてみると、「AAXX 便delayed」とあるではないか！

なあんだ、結局私のほうが早く着いたのねと、ちょっぴり勝ち誇ったような気分になって足取りも軽く家路についたのである。

つまりNSPを選ぶということは、こういうことなのである。UAやAAがいればNSPである。空港はIXに相当する。同じ相手と通信する場合でも、NSPによってその経路が違ったり、かかる時間が違ったりするわけだ。予定外にディレイが大きくなったりすることもあるが、その予定外の頻度も違ったりする。また、レッドカーペットクラブのような付加的なサービスの内容も違っていたりするのである（決してAAが悪いと言っているわけではない。単なる例ですよ！）。どちらを選ぶかはそれぞれの事情に照らし合わせて判断すればよいのである。

今回のように単一の航空会社のみで旅行できない場合だってあるだろう。そのような場合はIXであるどこかの空港で違う航空会社の飛行機に乗り換えることになる。この観点から考えると、より多くの空港にダイレクトで行ける路線を持っている航空会社を選んでおけば、いろいろなところで、より近い経路で旅行できる可能性が大きくなるだろう。これも、NSPを選ぶときの1つのファクターになりうる。

traceroute の使用例 (IJJ ISOC)

```
asaba:ijj> traceroute info.isoc.org
traceroute to info.isoc.org (198.6.250.9), 30 hops max, 40 byte packets
 1 otemachi.iij.net (192.244.176.49) 8 ms 3 ms 3 ms
 2 iij-gate.iij.net (192.244.176.26) 4 ms 3 ms 4 ms
 3 stockton.iij.net (192.244.176.210) 109 ms 108 ms 108 ms
 4 San-Jose1.CA.ALTER.NET (137.39.224.65) 133 ms 120 ms 122 ms
 5 San-Jose3.CA.ALTER.NET (137.39.101.97) 122 ms 159 ms 130 ms
 6 Vienna1.VA.ALTER.NET (137.39.12.1) 222 ms 229 ms 209 ms
 7 Falls-Church4.VA.ALTER.NET (137.39.100.33) 221 ms 221 ms 206 ms
 8 Falls-Church1.VA.ALTER.NET (137.39.8.2) 310 ms 216 ms 223 ms
 9 IS-gw.ALTER.NET (137.39.238.162) 241 ms 240 ms 223 ms
10 info.isoc.org (198.6.250.9) 207 ms 223 ms 220 ms
```

traceroute の使用例 (インプレス ISOC)

```
%traceroute info.isoc.org      traceroute コマンドでリモートシステム名を指定
traceroute to info.isoc.org (198.6.250.9), 30 hops max, 40 byte packets
 1 impgate (192.218.90.250) 0 ms 0 ms 10 ms      トレースの起点
 2 iij-imp (192.218.90.5) 0 ms 0 ms 0 ms
 3 192.244.180.121 (192.244.180.121) 370 ms 360 ms 390 ms
 4 iij-gate.iij.net (192.244.176.26) 380 ms 470 ms 370 ms
 5 stockton.iij.net (192.244.176.210) 470 ms 570 ms 460 ms
 6 San-Jose1.CA.ALTER.NET (137.39.224.65) 680 ms 990 ms 990 ms
 7 San-Jose3.CA.ALTER.NET (137.39.101.97) 800 ms 910 ms *  応答のないパケット
 8 Vienna1.VA.ALTER.NET (137.39.12.1) 840 ms 940 ms 1040 ms
 9 Falls-Church4.VA.ALTER.NET (137.39.100.33) 880 ms 740 ms 610 ms
10 Falls-Church1.VA.ALTER.NET (137.39.8.2) 600 ms 290 ms 400 ms
11 IS-gw.ALTER.NET (137.39.238.162) 250 ms 270 ms 250 ms
12 info.isoc.org (198.6.250.9) 270 ms 390 ms 420 ms      終点のホスト
```

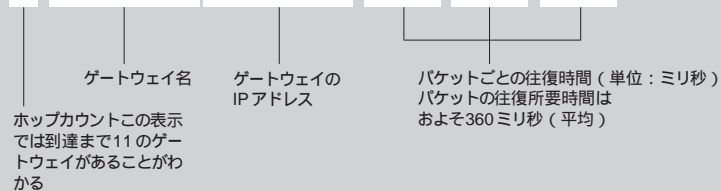


図2 traceroute コマンドとその出力例

```

asaba:wide>traceroute info.isoc.org
traceroute to info.isoc.org (198.6.250.9), 30 hops max, 40 byte packets
 1 jp-gate (133.4.11.1) 3 ms 2 ms 3 ms
 2 jp-entry (133.4.1.2) 4 ms 4 ms 5 ms
 3 us-entry (133.4.43.2) 132 ms 134 ms 132 ms
 4 border1-serial2-1.SanFrancisco.mci.net (204.70.32.13) 138 ms 136ms 144 ms
 5 border3-fddi0-0.SanFrancisco.mci.net (204.70.2.163) 139 ms 163ms 135 ms
 6 Palo-Alto1.CA.ALTER.NET (137.39.29.3) 144 ms 146 ms 152 ms
 7 San-Jose3.CA.ALTER.NET (137.39.29.1) 140 ms 142 ms 141 ms
 8 Vienna1.VA.ALTER.NET (137.39.12.1) 242 ms 216 ms 226 ms
 9 Falls-Church4.VA.ALTER.NET (137.39.100.33) 221 ms 221 ms 225 ms
10 Falls-Church1.VA.ALTER.NET (137.39.8.2) 251 ms 231 ms 217 ms
11 IS-gw.ALTER.NET (137.39.238.162) 245 ms 225 ms 241 ms
12 info.isoc.org (198.6.250.9) 222 ms 217 ms 224 ms
    
```

図3 tracerouteの使用例 (WIDE経由 ISOC)

traceroute による経路調査

そこで、あるNSPのサービスに加入して通信を行う場合、他のいくつかのNSPの中ではどういう経路を通っているのかが気になるところだ。これを調べるためにtracerouteというツールがある。tracerouteは、目的のホストまでに経由するすべてのルータと、それぞれのルータまでのRTT (Round Trip Time : 往復所要時間) を表示してくれる便利なツールだ。

たとえばIIJのオフィスのマシンからinfo.isoc.org (インターネット学会 : 本部は米国バージニア州レストン) までの経路を調べてみよう。図2の左端の数字は、tracerouteを実行したマシンからいくつめのルータかを示している。その右のホスト名はルータの名前であり、括弧の中はそのルータのIPアドレスである。さらに右の3

つの数字は、そのルータまでのRTTである。tracerouteは、一度に3つのパケットを飛ばして、経路上のルータまでのRTTを調べているため、数字が3つ出るのである。

経路上のルータのホスト名から推測すると大体どのあたりを通っているのかが分かるだろう。すなわち、IIJからはまずAlternetのSan Joseにあるルータに入り、そこからアメリカ大陸を横断してVirginia州のVienna市にあるルータに行き、さらにFalls Church市にあるルータからisocに到達している。info.isoc.orgにたどり着くまでに結局9つのルータを経由していることになる。

さて、ではこれを他のNSP経由でやるとどうなるだろう？ 図3は、WIDEのマシンから同じinfo.isoc.orgにtracerouteを実行した結果である。こちらも途中のルータのマシン名から推測すると、まずMCIのSan Franciscoのルータに入り、そこからAlter-

netのPalo Altoのルータを経由してAlternetのSan Joseのルータに入っている。そこからあとは、IIJからの場合と同じだ。かかるRTTも大体同じくらいであることが分かる。この2つのtracerouteの結果から、AlternetとMCIがアメリカ西海岸で相互に接続されていることが分かる。

NSP内でのルーティング

さて、これでともかくインターネット上のさまざまなサービスを受けるためには、いづれか1つのNSPに接続を持つ必要があることが分かった。

特定のNSPに接続されると、前回解説したように、接続された組織が用いているネットワークアドレスに対する経路情報がNSP内部のルータ上にセットされる。また、その組織の内部では、内部のネットワークの情報以外には、インターネットへの出口が分かればよいので、デフォルトの経路がNSP側のルータに向けてセットされればよい。

同じNSPに接続されている他の組織に関しても同様であり、これによって、そのNSPに接続されている組織同士の間での通信が可能になる。

図4の例で、組織AがNSPに接続されると、組織Aからはデフォルトの経路がNSP内部のルータAに向かって設定され、ルータAでは、組織Aに対する経路は組織内部に向けられる。NSP内部の他のルータでは、組織Aに対する経路はルータAにより近づく方向に設定される。同様に、組織Bが接続された場合には、組織Bからはデフォルトの経路がNSP内部のルータBに向かって設定され、ルータBでは組織Bに対する経路は組織B内部に向かって設定される。また、組織Bに対する経路はルータBにより近づく方向に設定される。

このような経路の設定の上で、組織Aから組織Bにデータを送りたいときには、組織Aの内部では、デフォルトの経路に従って、IPデータグラムはルータAに転送される。いったんNSPの中に入ると、組織Bの経路がルータBにどんどん近づいていくように設定されているので、これにより、ルータAからルータBまでデータグラムは転送されていく。最終的にルータBから組織Bの内部へと送り届けられるのである。

このように、NSPの内部のルータでは、そこに接続されている組織の経路情報が、その組織に向かうように設定されており、同じNSPに接続されている組織同士の通信が可能となるのである。

では、同一NSPにつながれていない組織との間でのルーティングはどうなるのだろうか？

NSP間のルーティング

デフォルト経路の利用

ここで再び図1に戻ろう。たとえば、

NSP-D、NSP-F、NSP-G、NSP-Hの場合には、NSP内部に接続されていない組織に至る道は1つしかない。したがって、ここでもデフォルト経路を用いることができる。すなわち、NSP-DからはNSP-Aに、NSP-FからはNSP-Bに、NSP-GからはNSP-Cに、NSP-HからはNSP-Eに、それぞれデフォルトの経路を向けておけば、あとはデフォルト経路の先のNSPが何とかしてくれそうだ。

では、NSP-Eの場合はどうだろうか？ NSP-Eは、他の2つのNSPとの接続を持っており、1つはNSP-Hのみにつながっており、もう1つはNSP-Bを経由して他のNSPにつながる接続となっている。この場合には、NSP-Eの内部では、もちろん、NSP-E自身に接続されている組織に対する経路情報を持っている。また、NSP-Hの内部の組織に対する経路情報は、NSP-Hに対して向くように持てばよいだろう。それ以外のNSPに接続されている組織に対しては、すべてNSP-B経由になるので、ここでデフォルト経路を使えばよさそうだ。

デフォルト経路が使えない場合

それでは、残りのNSP-A、NSP-B、NSP-Cではどうだろうか？ このような場合には、もはやデフォルト経路は使えそうもない。というのは、デフォルト経路をどちらに向ければよいのか明らかではないからだ。さて、ではデフォルト経路を使わないということはどういうことなのだろうか？ デフォルト経路は、いわば“自分自身が個別に知っているもの以外すべて”的な意味合いを持つ経路情報である。これを使わないということはすなわち、すべてを個別に知らないといけないうことになる。したがって、これらのNSPでは、インターネット上のすべての組織に至る経路情報を持っていないといけないうことになる。

すべての組織に至る経路情報を持っていないといけないう一言で言っても、そんなに簡単なものではない。たとえば、NSP-Aでは、NSP-Bに接続されている組織に対する経路をIX-A経由にすればいいのだが、はたまた、IX-B経由にすればいいのか頭を悩ませなければいけないう。

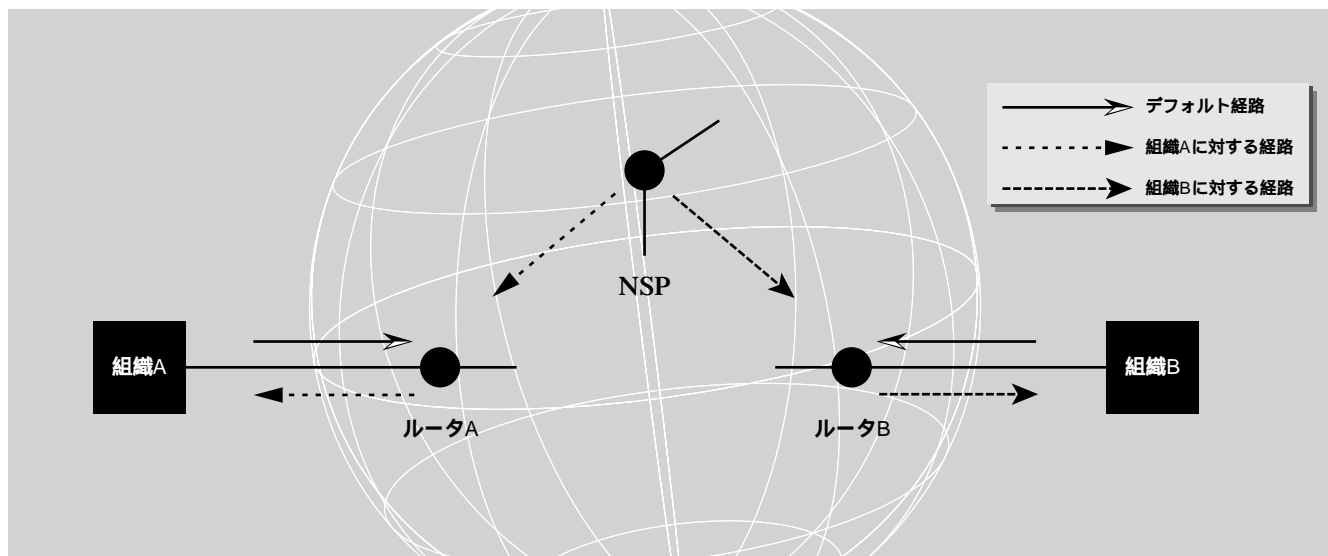


図4 NSP内部でのルーティング

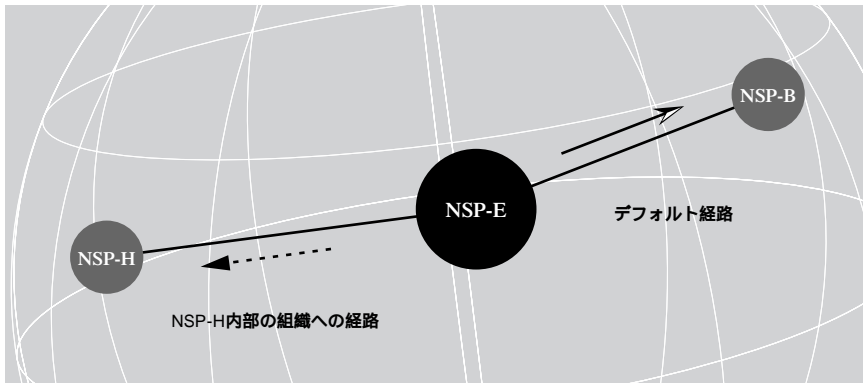


図5 NSP-Eの場合

おや、でも待てよ。NSP-Bと一言で言ってしまうてよいのだろうか？たとえばNSP-Bに接続されている組織のうち、よりIX-Aに近い組織に関してはIX-A経由で、また、よりIX-Bに近い組織に関してはIX-B経由にしてしまえばよいではないか？

実際のインターネット上の経路制御においては、こういうことは実はあまりやらない。なぜならば、それをやろうとすると、扱わなければいけない情報が一気に増えてしまうからである。近さというものを表現するための指標が必要になるし、それに基づいた経路選択をする際の計算量もぐっと増えてしまう。

したがって、他のNSPに接続されている組織に関する経路情報はひとまとめにして取り扱うのが普通である。つまり、ある組織に対する経路情報は、その組織が接続されているNSPに代表されるということになる。そして、いったんそのNSPを離れると、経路選択の判断は個別の組織ではなく、その組織が接続されているNSPを基準に行われるということである。

そうすると、たとえばNSP-AからNSP-Gに至る経路を考えた場合には、可能な経路としては以下の4つがある。

1. NSP-A (IX-A) NSP-B (IX-B)
NSP-C NSP-G
2. NSP-A (IX-A) NSP-B NSP-C NSP-G

3. NSP-A (IX-B) NSP-B NSP-C NSP-G
4. NSP-A (IX-B) NSP-C NSP-G

経路1はあまり意味がなさそうだ。1を使うぐらいなら3で十分だからである。そうすると問題は、2、3、4のどの経路を選ぶかということになる。

ポリシーに基づく経路選択

これは一概にどれがよいとは言えない。経由するNSPの数をできるだけ減らすという観点から考えれば、4がいちばんよいことになるが、何らかの理由でNSP-Aは、NSP-CとIX-Bの間のリンクを使いたくないかもしれない。その場合には4よりもむしろ2か3の経路のほうが好ましい経路ということになる。これに対する一般的な回答はもはや存在せず、それぞれのNSPにおける判断に委ねるしかない。

このように、NSP間での経路選択を複数候補の中から行う場合には、なんらかのポリシーが基準になって行われることが多い。

この、ポリシーに基づく経路選択を行うための概念として、AS (Autonomous System) というものがある。ASとは、単一の管理主体によって運営されているひとまとまりのネットワークのことをさす概念である。すなわち、1つのASは1つのポリシーを表していると考えるのである。通常は、1つのNSPが1つのASを構成している。

同一のAS内部の経路選択では、ポリシーに基づく判断というものはあまり必要ない。1つのASは1つのポリシーだからである。したがって1つのAS内部での経路選択は、純粋に目的の組織に至る経路のコストが最小になるように選ぶなど、計算で求めることができる。しかし、AS間の経路選択はそうはいかず、なんらかのポリシーに基づいた判断が必要になってくるのである。

おわりに

うむ。ポリシーだってえ？ なんだかキナ臭い話だよなあ。たとえば「XX主義に属するASは経由したくない！」とか、「主義のASからの経路は受け入れられない！」などと言いつけるのだろうか？ かんべんしてくれよう。

しかし、それも仕方がない。インターネットとは、いわば社会の縮図なのである。いや、これからは社会はインターネットの縮図と言っても過言ではないかもしれない(?)。

アドレスが枯渇するとか、経路情報が爆発するので何とかしないとイケないとかいう議論を聞いていると、フロンガスによるオゾン層の破壊問題とか砂漠化の問題とか、そういった現実の社会問題、環境問題に関する議論がふと頭をよぎる。冗談ではなく、今後は地球的視野に立ったインターネットのあり方、利用の仕方というものを考えて、そのうえで、それを支える技術のあり方を真剣に考えていかなければいけないのだろうな。などと、いつになくシリアスになりつつ……

しかし……とりあえずは、月刊誌における連載のあり方というものを今後は真剣に考えていかなければいけないのだろうな、という気持ちのほうが実は強かったりする今日この頃ではある。



[インターネットマガジン バックナンバーアーカイブ] ご利用上の注意

このPDFファイルは、株式会社インプレスR&D(株式会社インプレスから分割)が1994年～2006年まで発行した月刊誌『インターネットマガジン』の誌面をPDF化し、「インターネットマガジン バックナンバーアーカイブ」として以下のウェブサイト「All-in-One INTERNET magazine 2.0」で公開しているものです。

<http://i.impressRD.jp/bn>

このファイルをご利用いただくにあたり、下記の注意事項を必ずお読みください。

- 記載されている内容(技術解説、URL、団体・企業名、商品名、価格、プレゼント募集、アンケートなど)は発行当時のものです。
- 収録されている内容は著作権法上の保護を受けています。著作権はそれぞれの記事の著作者(執筆者、写真の撮影者、イラストの作成者、編集部など)が保持しています。
- 著作者から許諾が得られなかった著作物は収録されていない場合があります。
- このファイルやその内容を改変したり、商用を目的として再利用することはできません。あくまで個人や企業の非商用利用での閲覧、複製、送信に限られます。
- 収録されている内容を何らかの媒体に引用としてご利用する際は、出典として媒体名および月号、該当ページ番号、発行元(株式会社インプレス R&D)、コピーライトなどの情報をご明記ください。
- オリジナルの雑誌の発行時点では、株式会社インプレス R&D(当時は株式会社インプレス)と著作権者は内容が正確なものであるように最大限に努めましたが、すべての情報が完全に正確であることは保証できません。このファイルの内容に起因する直接のおよび間接的な損害に対して、一切の責任を負いません。お客様個人の責任においてご利用ください。

このファイルに関するお問い合わせ先

株式会社インプレスR&D

All-in-One INTERNET magazine 編集部

im-info@impress.co.jp