インターネットシミュレータの構築報告

鈴　木　常　彦†

約 20 台のサーバと 50-60 台規模の仮想マシン上にオープンソースソフトウェアを主体とした小規模なインターネットシミュレータを構築した。DNS と BGP ルーティングを稼働させたインターネットの実践的な実験、教育環境をキャンパス内で提供することを目的としている。仮想マシンには現状 VMWare を用いているが、ホスト OS には Debian GNU/Linux、仮想マシンの OS には FreeBSD、ルーティングには Zebra を用いており、誰もが構築、利用、拡張できるオープンな設計をとっている。今日、インターネットは実験ネットワークとしてではなくインフラとして扱われるようになり、研究者や学生はファイアウォールの内側に入ってしまった。結果として障壁なくインターネットの研究を行うには、このような実験環境が必要となってきている。

Development of an Internet Simulator

Tsunehiko Suzuki†

This paper introduces our small size internet simulator with open source software on 50-60 virtual machines on 20 servers. This simulator has practical DNS systems and BGP4 based networks. For everyone, the simulater is built using free softwares except VMWare for virtual machine, such as Debian GNU/Linux for host OS, FreeBSD for guest OS, and Zebra for routing daemons. Today, the Internet came to be treated as infrastructure , not as an experiment network, and the researcher and the student have entered the inside of the firewall. Therefore, such an experiment environment is necessary to study the Internet without a wall.

1. はじめに

実験ネットワークであった The Internet (以下インターネット) は、いつの頃からかインフラと呼ばれるようになり、研究者や学生たちが自由に実験のできるネットワークではなくなってしまった。セキュリティ確保の名の下に、企業も大学もファイアーウォールの内側に入ってしまい、多くの研究者や学生たちは障壁のない裸のインターネットに触れられなくなってしまった。

ファイアウォールの内側からはかつてのインターネットにはつながらない。外部へ向かった HTTP のセッションは L4 スイッチに囚われ、プロキシサーバと接続させられる。80 番ポート以外の Web サーバがあっても、そのサイトとつながるすべはない。

外部へ向かった SMTP セッションは同様に、リレーサーバと会話をすることになる。相手のサーバにセッションを張ったつもりが、SMTP の Greeting メッセージには自サイトのサーバ名が現れる。

HTTP と SMTP 以外のプロトコルはその存在を否定されているかに見える。

サーバも設置には許可を必要とし、必要最小限のポートだけが通信を許される。

このような状況がそこかしこでインターネットを分断し、それはかつて End to End ネットワークと呼ばれたインターネットとよべるネットワークではなくなってしまっている。

こうした状況下において、自由にインターネットプロトコルを弄び、ネットワークの研究、実験を障壁なく行うためには、インターネットをインフラ視しているユーザのいない、別なインターネットが必要となる。

そこで、筆者の研究室では小規模なインターネットを独自に構築することにした。着手したのは 2003 年だった。

同様の趣旨で構築された大規模なシステムとして北陸先端科学技術大学院大学の StarBED[1)2)3)] があるが、筆者は手近に利用できるシステムを目指し、オープンソースソフトウェア主体による小規模構成のインターネットシミュレータを独自に構築してきた。

† 中京大学情報理工学部情報システム工学科
Chukyo University

本システムは ns[8] や CISCO 社の PacketTracer のようにアプリケーション上でパケットをシミュレートするものではなく、実際に TCP/IP プロトコルを流す小規模な実ネットワークである。インターネットシミュレータと称する所以は、インターネットを小規模に模したトポロジーでノードを結び、The Internet で用いられている BGP4 でルーティングを行い、DNS によるオルタナティブなドメイン空間を構築、運用しているところにある。

筆者の主要な研究テーマは、インターネットを構成する基盤技術の実装や運用の脆弱性にあり、WWW サーバや DNS サーバ、あるいはルータの実装を実運用実験ができるネットワークが必要であり、それらを抽象化したシミュレーションモデルでは目的に適さない。実際にサーバやクライアントの各種実装を TCP/IP で相互接続し、インターネットと類似した実験ができる本シミュレータはそうした研究に最適なものとなっている。

ただし、実際のインターネットは大規模な複雑系システムをなしているとともに、大勢の人間がつながっている有機的なネットワークであり、そのオートポエティックな性質までシミュレーションすることは残念ながら不可能である。

本システムは、現在、筆者の所属する中京大学における教育用、研究用のネットワークとして実稼働している。

2. システムの構成

本システムの物理構成図を図 1 に示す。一つ一つの箱は複数の仮想マシン (IA パソコンを BIOS からシミュレートしている VMWare 上のゲスト OS として構築) で動作する仮想ノードをホストしている 1U サーバ (以下物理ノード) である。

それぞれの物理ノードが内部に 2,3 個の仮想ノードを持ち、1 つの物理ノードが 1 つの AS(Autonomous System) ネットワークを構成している。またノードは大きく 7 つのグループに別れており、それぞれが IX(Internet eXchange) を模したスイッチングハブで相互接続されている。

この仮想インターネットを構成する表 LAN の他に、それぞれの物理ノードは管理用の裏 LAN で接続されており、すべての物理ノードの OS はファイルサーバからネットワークブートで立ち上がるようになっており、効率的な集中管理が可能となっている。

物理ノードの OS は Debian GNU/Linux[4] を用い、ファイルサーバと仮想ノードの OS は FreeBSD[5] である。仮想ノードを動作させるのには 現状 VMWare[6] を用いているが、これも Xen などフリーな環境への移行を計画している。

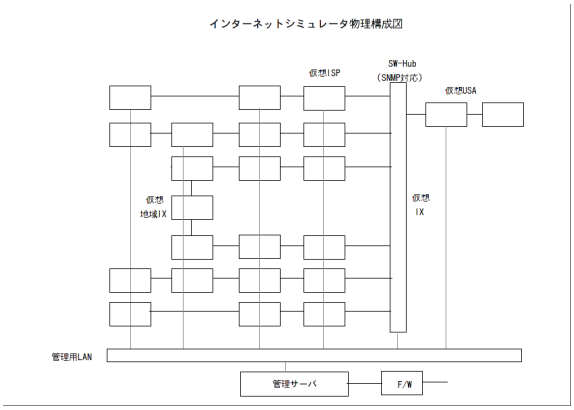

図 1 シミュレータ物理構成図

仮想ノードに FreeBSD を用いたのは、パケットロスや遅延等をシミュレートできる dummynet の機能を用いたかったためである。仮想環境に VMWare を選んだのは、実験の目的や必要に応じて Windows や Linux なども混在させることも想定したためであり、設計として FreeBSD に特化しているわけではない。物理ノードが GNU/Linux になっているのは単に VMWare を動作させるための制約によるものである。

3. ルーティングの構成

今日のインターネットにおいて、一般の研究者や学生が運用する機会が少ないのが、AS 間で用いられているルーティングプロトコルである BGP4 の運用である。実運用の BGP4 で事故を起こすと、それは経路ハイジャックとなり世界に混乱をきたすこともありえる[9]。

本シミュレータでは、仮想ノードの構成する AS 間のルーティングにインターネットと同じ BGP4 を用いており、AS 間ルーティングの運用や実験を誰もが自由に行うことができる。

本来、大学はそれぞれが AS として運用されてしかるべきであるが、それができていないのはひとえにルーティング技術者の不足であり、本システムで学んだ技術者が、各大学、各地域を支えてくれるようになることを目指している。

BGP4 を動作させるルーティングデーモンには、オープンソースのルーティングソフトウェアである Zebra[7] を用いた。Zebra はルーティングデーモンとしての機能において実ルータと遜色なく動作する上に、ISP においてシェアの高い CISCO 社の ISO に似た設定書式と CUI を持つため、実務を学びたい学生にも良く適している。

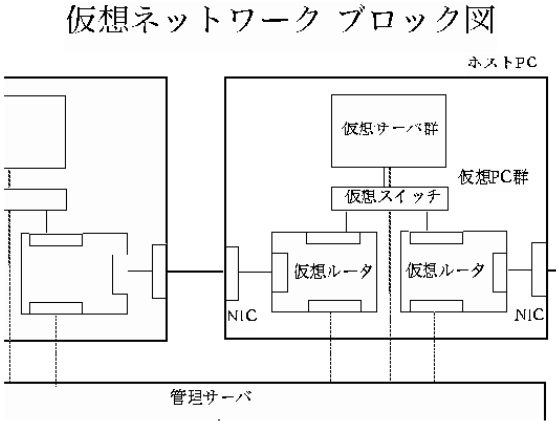

図 2 仮想ノード構成図

図 2 で示すように、仮想ノードの中で動作する Zebra による仮想ルータが物理ノードが持つ 2 つの物理インターフェイス間をルーティングするとともに、隣接ノードの Zebra との間で BGP peer を張る。これらのノードはブロックのように連結し仮想インターネットを組み立てることができるようにモデル化してある。

本システムにおける仮想インターネットでは、RFC1918 のプライベートアドレス空間のうち、10.0.0.0/8 と 172.16.0.0/12 をグローバルアドレス空間と想定し、192.168.0.0/16 を仮想インターネットのプライベートアドレス空間として、例えば管理用裏 LAN に用いている。10.0.0.0/8 は適宜 CIDR ブロック化して割り当てを行っている。

(ex31)		(ex11)		(ex71)
AS51009		AS51003		AS51001
		10.0.0.2	$\leftarrow IX \rightarrow$	10.0.0.70
10.1.1.2	$\longleftrightarrow$	10.1.1.1		
10.7.0.0/16		10.1.0.0/16		10.128.0.0/11
10.7.0.0/24		10.1.0.0/24		10.129.0.0/16
etc.		etc.		etc.

表 1 仮想ネットワークの一部

表 1 はシミュレーションしているネットワークの一部を示す。ex11,ex31,ex71 はそれぞれ独立した仮想 AS の境界ルータである。

ex11 と ex71 は IX で相互接続する複数の Tier1 ISP 群のなかの 1 対の AS ピアである。また、ex31 は ex11 をトランジットして IX の先の他の ISP たちとつながる地域 ISP を模している。

ex71 の 10.0.0.70 と ex11 の 10.0.0.2 が一つの BGP4 peer、ex31 の 10.1.1.2 と 10.1.1.1 がもうひとつの BGP4 peer を構成している。それぞれ内部では、static に記述した複数の prefix (例えば ex71 での 10.128.0.0/11, 10.129.0.0/16) を BGP4 に redistribute することにより、複数のネットワークを抱えた現実の AS を模すようにしている。 以下の表 2 は表 1 の各境界ルータの経路表である。

それぞれの経路が AS PATH を伴って現れていることがわかる。

表 3 に境界ルータのルーティングデーモンである Zebra の設定 (コンフィグ) を示す。

Zebra の設定はファイルで記述できる一方、Zebra デーモンへの telnet によって行うこともできるので、シミュレータの利用者はあたかもルータにログインする感覚でオペレーションを行うことができる。

4. DNS の構成

仮想 USA と称したノードに DNS のルートサーバを立て、インターネットを模した DNS 名前空間を使用できるようになっている。現在、.com をもじった .nom という TLD (Top Level Domain) も動いている。ルートサーバのソフトウェアには 現在 tinydns を用いている。

新しいドメインを立ち上げたい実験者は、仮想ノードで DNS サーバを動作させ、ルートサーバ、あるいは TLD サーバから委譲を受けることにより、シミュレータ上で自分のドメインを運用することができる。

5. 管 理 環 境

本システムは telnet あるいは ssh でリモートログインし管理する部分のほか、VMWare の GUI やその上の仮想マシンの GUI 環境にアクセスするために、VNC[11] を用いることができる。また 20 台ほどあるホストサーバの設定はファイルサーバ (NFS) 上で一括管理することができる。

すべての物理ノードは裏 LAN を介して NFS サーバに接続されており、各サーバの設定はファイルサーバ上で作業ができるようになっている。また、BGP ルーティングの基本設定を施した仮想ノードが VMWare のイメージファイルとしてテンプレート化されており、これをコピーして新しい仮想ノードを接続していける

```
zebra71# sh ip bgp
BGP table version is 0, local router ID is 10.0.0.70

   Network          Next Hop       Metric LocPrf Weight Path
*> 10.0.0.0/24      0.0.0.0             0         32768 ?
*> 10.1.0.0/16      10.0.0.2            0             0 51003 i
*> 10.1.1.0/24      10.0.0.2                          0 51003 i
*> 10.7.0.0/16      10.0.0.2                          0 51003 51009 i
*  10.128.0.0/11    0.0.0.0                       32768 i
*>                  0.0.0.0             0         32768 i
*> 10.128.1.0/24    0.0.0.0             0         32768 ?
*> 10.129.0.0/16    0.0.0.0             0         32768 i
*> 172.16.0.0/24    0.0.0.0             0         32768 ?

zebra31# sh ip bgp
BGP table version is 0, local router ID is 10.1.1.2

   Network          Next Hop       Metric LocPrf Weight Path
*> 10.0.0.0/24      10.1.1.1                          0 51003 51001 ?
*> 10.1.0.0/16      10.1.1.1                          0 51003 i
*> 10.1.1.0/24      10.1.1.1            0             0 51003 i
*> 10.7.0.0/16      0.0.0.0             0         32768 i
*> 10.128.0.0/11    10.1.1.1                          0 51003 51001 i
*> 10.128.1.0/24    10.1.1.1                          0 51003 51001 ?
*> 10.129.0.0/16    10.1.1.1                          0 51003 51001 i
*> 172.16.0.0/24    10.1.1.1                          0 51003 51001 ?
```

表 2 経路表

ようになっている。

またファイルサーバは FreeBSD の CVSUP サーバにもなっており、ファイアウォールサーバを介して、インターネットから OS や各種ソフトウェアをシミュレータ側へ取り込むことができる。

基本的に外部ネットワークからの利用はセキュリティを考慮して、ssh 等での VPN 接続を想定しており、装置と外部をルーティングすることは考えていない。

なお、仮想サーバの基本設定部分が VNC を介した GUI 環境になっていることは、初心者への受けはよいのであるが、実のところエキスパートにとっては不便

```
router bgp 51001
 bgp router-id 10.0.0.70
 aggregate-address 10.128.0.0/11
 redistribute connected
 neighbor 10.0.0.2 remote-as 51003
 neighbor 10.0.0.2 next-hop-self


router bgp 51003
 bgp router-id 10.0.0.2
 network 10.1.0.0/16
 neighbor 10.0.0.70 remote-as 51001
 neighbor 10.0.0.70 next-hop-self
 neighbor 10.1.0.2 remote-as 51003
 neighbor 10.1.0.2 next-hop-self


router bgp 51009
 bgp router-id 10.1.1.2
 network 10.7.0.0/16
 neighbor 10.1.1.1 remote-as 51003
 neighbor 10.1.1.1 next-hop-self
```

表 3 ルータの設定

であることは否めない。開発開始当初は GUI しか実現できなかったためであるが、現在、CUI 環境への移行を図っているところである。

6. 適　　用

6.1 地域 IX の研究への適用

本シミュレータには中央の IX の他に、一部の枝の先を相互接続して地域 IX を模擬したスイッチがある。これにより、部分的 (no export) な経路アナウンスやパンチングホール等を用いた地域 IX の運用に関する研究が行えるようになっている。

筆者の活動拠点である東海地域には、UCAN[10] という地域ネットワークの活動コミュニティがあり、2004 年 6 月から 2007 年 3 月まで学術ネットワークと地域 ISP の間の相互接続実験を行った。

この際にも、地域にルーティング技術者が稀少であることが問題として浮かびあがった。そこで、技術者のいない非 BGP サイトを実験用 AS に収容し、プライベート AS で BGP ルーティングを提供するというようなことを行ったが、この際にもそうしたルーティングの事前確認には本シミュレータが大いに役立った。

6.2 DNS の研究への適用

筆者の研究テーマの一つは DNS の運用上の脆弱性であり、本シミュレータでは、それぞれの仮想ノードで、tinydns のほか、いくつかのバージョンの BIND や NSD を動作させることにより、その脆弱性を確認することができる。

現在、DNS amplification による DDoS 攻撃のシミュレーションや、DNS キャッシュサーバへのポイズニングのシミュレーションなどを試行中である。また、IPv6 環境下での DNS の運用上の問題点の洗い出しにも用いていく予定である。

こうした実験には DNS サーバの数が多いほど有意なデータをとることができる。DNS サーバの数は、仮想ノードに複数の IP アドレスを付与することにより、未実証ではあるが装置全体で数百から 1 千台程度まで稼働させることが可能と考えている。

また、特に誤った設定や DNS サーバソフトウェア実装上の問題などは、抽象化されたシミュレーションでは不十分であり、実際の実装を稼働させることのできる本シミュレータが有効であると考えられる。

7. おわりに

現在のシミュレータの外観を図 6.2 に示す。4 年ほどかけて構築してきたものであるが、まだまだ改良の余地は大きい。

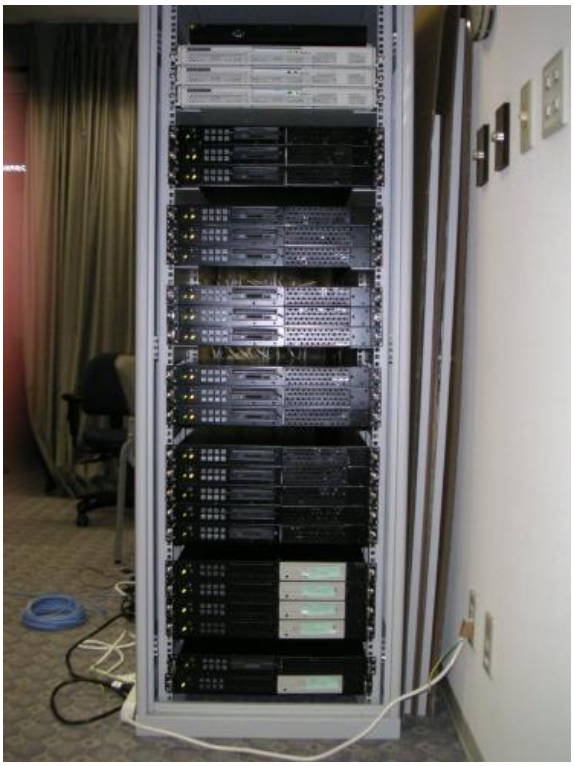

図 3　シミュレータ外観

本システムの仮想ノードを構成する重要な基盤となっている VMWare は当初商用のものを用いたが、拡張コストやライセンス管理の面、オープンな利用の面などで扱いづらいものとなっており、現在はフリーバージョンの VMWare Server のほか、VirtualBox など他の選択肢も視野に入れつつ、主として Xen への移行による再構築を試行中である。

また、従来 1 つの OS あたりルーティングテーブルが 1 つという制約があったが、現在の技術では 1 つの OS で複数のルーティングテーブルを持たせることが可能なため、仮想インターフェイスを増やすことにより、現状規模の物理リソースのままで、数倍から数十倍の規模のノードを持った仮想インターネットが構築できる目途もついている。

機能の上での課題としては、ネットワークの動的な変化をシミュレーションすることが必要と考えている。たとえば、一部の peer をシミュレーション中にダウンさせ、いわゆるフラッピングを発生さて、時系列上なネットワークの変化を生じさせることにより、より実際のネットワークに近い動きをシミュレーションしたい。

これを行うには、現在はルータのコンフィギュレーションを手作業で書き換える必要があるため、プログラマブルな動的設定変更の機能を盛り込む必要がある。

さて、今後の教育、研究現場においては、シミュレータで技術を学び、シミュレータで運用実験を行い、あらたなネットワークアプリケーションはシミュレータで検証できたものだけが実ネットワークで接続を許されるという形になっていくのかもしれない。

果たしてそれでインターネットが発展できるかどうかはわからないが、少なくとも基盤となる技術とインターネットの精神をかろうじて伝えていくことができれば、本システムを構築した意味もあるだろう。

本システムは残念ながら StarBed に規模は劣るが、インターネットに手が届かなくなった現在の教育現場において手軽に利用、拡張可能なシミュレータの意義は大きいといえる。

謝　　辞

本システムの主要部分の開発は中京大学特定研究助成によるものです。研究助成共同提案者の伊藤誠先生、開発を支援いただいた情報理工学部の先生方、構築を手伝ってくれた株式会社リフレクションの山崎浄君、伊藤剛志君に感謝いたします。

参 考 文 献

1) http://www.starbed.org/
2) Toshiyuki Miyachi, Ken-ichi Chinen and Yoichi Shinoda: StarBED and SpringOS: Large-scale General Purpose Network Testbed and Supporting Software, Valuetools 2006, Pisa, Italy, ISBN 1-59593-504-5, Oct, 2006.
3) Toshiyuki Miyachi, Junya Nakata, Razvan Beuran, Ken-ichi Chinen, Kenji Masui, Satoshi Uda, Yasuo Tan and Yoichi Shinoda: Realistic Simulation of Internet, ASC 2006, Tokyo, Japan, ISBN 4-431-49021-3, pp386-390, Oct. 2006.
4) http://www.debian.org/
5) http://www.freebsd.org/
6) http://www.vmware.com/
7) http://www.zebra.org/
8) http://www.isi.edu/nsnam/ns/
9) http://d.hatena.ne.jp/memecomputing/20080227
10) http://www.ucan.initiative.jp/
11) http://www.realvnc.com/