第 XXX 部

M Root DNS サーバの運用

第 30 部
M Root DNS サーバの運用

第 1 章　はじめに

インターネット上の資源は、木構造の名前空間であるドメイン名によって指定される。ドメイン名から、IP アドレスなどの名前に対応した種々の情報を得る操作は名前の解決と呼ばれるが、この名前解決を担当するシステムが DNS——Domain Name System——である。

DNS では、名前空間は Zone と呼ばれる連続した部分空間に分割して管理が行われており、分散的なアルゴリズムによって名前の解決が行われる。木構造の頂点である Root ゾーンの解決を行う DNS サーバは、特に Root DNS サーバと呼ばれており、DNS の名前解決にとって非常に重要である。特に DNS での UDP を用いた場合のメッセージ長の制約から、多数の Root DNS サーバを設定することはできない。DNS ではキャッシュを多用することによって効率を改善するとともに、Root DNS サーバ等の上位ドメインに対応するゾーンを担当するサーバへの問合わせを減らすような努力がなされているが、Root DNS サーバが重要な存在であることには変わりはない。

2009 年は DNSSEC の導入に関する議論が多く行われた。Root DNS サーバの運用者達が集まる会合や、ICANN RSSAC（Root Server System Advisory Committee）/SSAC（Security and Stability Advisory Committee）においても多くの議論が行われ、Root ゾーンに対しての DNSSEC 導入が決定された。実際の導入は 2010 年から段階的に行われていくこととなったが、その事前準備や手順の決定を詳細に行うためのデザインチームが ICANN/VeriSign によって結成された。

第 2 章　M Root DNS サーバの構成

Root DNS サーバは現在 `A.ROOT-SERVERS.NET`〜`M.ROOT-SERVERS.NET` という 13 システムで運用が行われている。このうち、`M.ROOT-SERVERS.NET` は、1997 年 8 月に WIDE Project によって運用が始まった。Root DNS サーバはインターネットにおける分散が制限されている資源の一つであるため、障害等によるサービス中断を最低限に押さえる必要がある。そのため、M Root DNS サーバは、1997 年の運用開始時から、サーバの冗長構成を導入し、主サーバの障害時には副サーバが自動的にサーバ機能を提供するような運用を行っている。

現在は、図 2.1 に示すような基本構成をユニットとし、後述の Anycast を用いてサービスの提供を行っている。各ユニットは 4 台のサーバから構成されており、サーバの OS や DNS ソフトウェアの更新時にもサービスを停止する必要はない。ルータなどの更新時にはサービスを停止せざるを得ないが、サービス停止に先だって経路広告を停止することにより、問合わせは他の active な Anycast サーバによって処理されるため、事実上のサービス停止は発生しない。

なお、2009 年はインターネットにおける経路表の増大に伴い、一部のルータでメモリ不足が発生した。

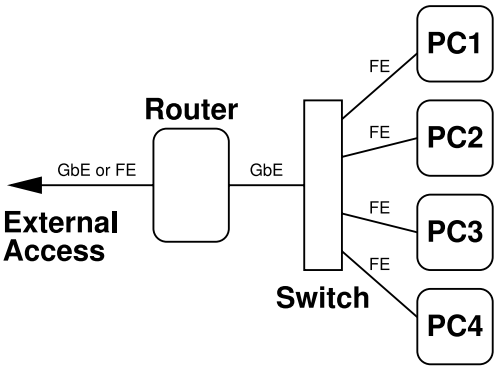


図 2.1.　Anycast 用基本構成

2009 年 12 月現在において、インターネット全体の経路数は 30 万経路を突破し、M Root DNS サーバと IX を接続する一部のルータにメモリ不足によるパフォーマンスの劣化が発生した。そのため、急遽これらのルータを異なる機種に交換、もしくはマネージメントカードの交換を行うことで、経路表の増大に対応した。

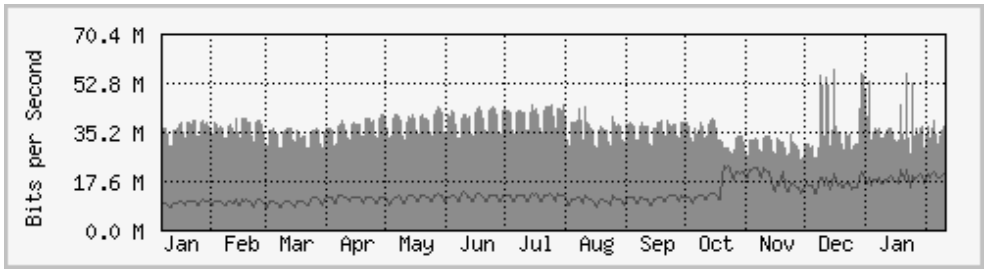
第 3 章　Anycast

Root DNS サーバは 13 台と限られた存在であるため、インターネット上に普く分布させることはできない。そこで、同じデータを供給するサーバを複数インターネット上に設置し、それぞれのサーバは同一サービスアドレスでサービスを提供する様にする。このサービスアドレスを含む経路情報を BGP でアナウンスすることにより、BGP の経路選択ポリシに依存するものの、一つのアドレスで複数台のサーバを運用することができる。この運用方法は RFC3258 “Distributing Authoritative Name Servers via Shared Unicast Addresses”[38] で定義されており、一般的には BGP Anycast と呼ばれている。

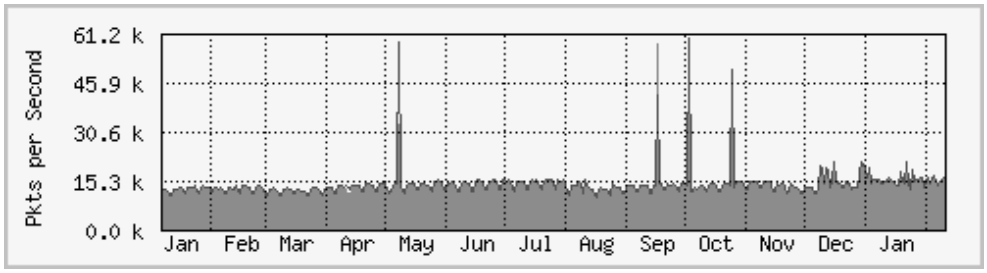
この Anycast に関しては、RFC が出版されたのは 2002 年 4 月であるが、最初の Internet Draft が IETF の DNSOP WG に提案されたのは 1999 年 10 月であり、その間議論が続けられてきた。

M Root DNS サーバでは、2004 年に入り、Seoul (KR) および Paris (FR) での設置を行ない、運用準備を進めてきた。このうち、Seoul に関しては、韓国で唯一の Layer-2 IX である KINX――Korea Internet Neutral Exchange――のご協力を得て、2004 年 7 月 21 日より運用を開始した。経路広告に BGP の `NO_EXPORT` 属性を添付するいわゆる local anycast として運用を行なっているが、学術系のネットワークの収容を目的として NCA――National Computerization Agency――が運用している Layer-3 IX である KIX では、`NO_EXPORT` を外して学術系ネットワークに対して経路の広報を行なっている。しかし、韓国での主要二大 ISP である KT および Daemon への接続性がないため、現在、Seoul で処理されている問合わせは毎秒 50〜100 クエリ程度と多くはない。

一方、Paris は Telehouse Europe、Renater、France Telecom、および Open Transit の協力を得て、Telehouse Voltaire にて 2004 年 9 月 1 日より運用を開始した。ここでは二つの独立な IX である、Renater が運用する SFINX と France Telecom が



(a) トラフィックの推移



(b) パケット数の推移

図 3.1.　2009 年における M-Root DNS 全体の問合わせ数の推移

運用するPARIXに接続している他、2004年10月からはTISCALIが独立にtransitを提供して頂いている。現在は多くのISPに対して`NO_EXPORT`をつけて経路広告を行なっているが、幾つかのISPに対しては`NO_EXPORT`なしに経路広告をしている。ヨーロッパ全域にサービスを提供しているtransit ISPとも多くpeerしているため、そのサービスエリアはフランスに留まっていない。このため、毎秒4000クエリ程度の問合わせがある。

San FranciscoはWIDE San Francisco NOCに設置されており、WIDEとは別なFastEthernetでPAIX/Palo Altoに接続されている。WIDEのLos Angelesでのupstreamであるas701からのトラフィックは東京に送るのではなくSan Franciscoで処理されている。また、アメリカ合衆国の研究教育ネットワークであるInternet2 NetworkとはIPv6によるPAIX上のpeerをしているが、2006年夏にIPv4でのpeerを追加した。これによって、アメリカ合衆国の主な大学からのM-Root DNSサーバへの問合わせはTransPAC等を経由して東京で処理されるのではなく、San Franciscoで処理されるようになり、RTTの改善に貢献している。

図3.1にM-Root全体に対するトラフィックの2009年における推移を示す。2009年10月にoutgoingトラフィックの変化が見て取れるが、パケット数は変化しておらず、増えた理由は不明である。

第4章　他のRoot DNSサーバ

2002年10月22日早朝（日本時間）に発生した13台のRoot DNSサーバをターゲットにしたDDoS攻撃をきっかけに、幾つかのRoot DNSサーバでは、Anycastサーバの設置を図っている。特に、ISCが運用しているF Root DNSサーバでは、APNIC等との協調により、精力的にAnycastサーバの設置を行っている。

2009年12月時点でのRoot DNSサーバの設置状況を表4.1に示す。各サーバの最初の都市が元々運用されていた都市であり、それ以降はAnycastによるものである。Anycastの運用形式も各サーバで異なっており、例えば、`C`ではCogent Communicationsのバックボーンにおける IGPによるAnycastを実施している他、`F`では、Palo Alto, CAとSan Francisco, CAのサーバはグローバルな経路広告を行っているのに対し、その他の`F`サーバは原則として、経路情報に`NO_EXPORT` BGP Communityを添付することによるローカルなAnycastサービスを提供している。

2008年から比較すると、2009年は、A、F、G、I、J、K、Lのサーバにて Anycast拠点が増加している。特にVeriSignが管理するサーバに関しては拠点増加が激しく行われており、さらにAnycast拠点が増え続ける方向と思われる。

表 4.1. Root DNS サーバの設置状況

サーバ	設置都市		
A	Los Angeles, CA	New York, NY	Frankfurt (DE)
	Hong Kong (HK)	Palo Alto, CA	Ashburn, VA
B	Marina Del Rey, CA		
C	Herndon, VA	Los Angeles, CA	New York, NY
	Chicago, IL	Frankfurt (DE)	Madrid (ES)
D	College Park, MD		
E	Mountain View, CA		
F	Ottawa (CA)	Palo Alto, CA	San Jose, CA
	New York, NY	San Francisco, CA	Madrid (ES)
	Hong Kong (HK)	Los Angeles, CA	Rome (IT)
	Auckland (NZ)	Sao Paulo (BR)	Beijing (CN)
	Seoul (KR)	Moscow (RU)	Taipei (TW)
	Dubai (AE)	Paris (FR)	Singapore (SG)
	Brisbane (AU)	Toronto (CA)	Monterrey (MX)
	Lisbon (PT)	Johannesburg (ZA)	Tel Aviv (IL)
	Jakarta (ID)	Munich (DE)	Osaka (JP)
	Prague (CZ)	Amsterdam (NL)	Barcelona (ES)
	Nairobi (KE)	Chennai (IN)	London (UK)
	Santiago de Chile (CL)	Dhaka (BD)	Karachi (PK)
	Torino (IT)	Chicago, IL	Buenos Aires (AR)
	Caracas (VE)	Oslo (NO)	Panama (PA)
	Quito (EC)	Kuala Lumpur (MY)	Suva (Fiji)
	Cairo (EG)	Atlanta, GA	Podgorica (ME)
	Maarten (AN)		
G	Colombus, OH	San Antonio, TX	Honolulu, HI
	Fussa (JP)	Stuttgart-Vaihingen (DE)	Naples (IT)
H	Aberdeen, MD		
I	Stockholm (SE)	Helsinki (FI)	Milan (IT)
	London (UK)	Geneva (CH)	Amsterdam (NL)
	Oslo (NO)	Bangkok (TH)	Hong Kong (HK)
	Brussels (BE)	Frankfurt (DE)	Ankara (TR)
	Bucharest (RO)	Chicago, IL	Washington, DC
	Tokyo (JP)	Kuala Lumpur (MY)	Palo Alto, CA
	Jakarta (ID)	Wellington (NZ)	Johannesburg (ZA)
	Perth (AU)	San Francisco, CA	Singapore (SG)
	Miami, FL	Ashburn, VA	Mumbai (IN)
	Beijing (CN)	Manila (PH)	Doha (QA)
	Colombo (LK)	Vienna (AT)	Paris (FR)
	Taipei (TW)		
J	Dulles, VA (3 sites)	Ashburn, VA	Vienna, VA
	Miami, FL	Atlanta, GA	Seattle, WA
	Chicago, IL	New York, NY	Los Angeles, CA
	Honolulu, HI	Mountain View, CA, (2 sites)	San Francisco, CA (2 sites)
	Dallas, TX	Amsterdam (NL)	London (UK)
	Stockholm (SE) (2 sites)	Tokyo (JP)	Seoul (KR)
	Beijing (CN)	Singapore (SG)	Kaunas (LT)
	Nairobi (KE)	Montreal (CA)	Quebec (CA)
	Dublin (IE)	Sydney (AU)	Cairo (EG)
	Warsaw (PL)	Brasilia (BR)	Sao Paulo (BR)
	Sofia (BG)	Prague (CZ)	Johannesburg (ZA)
	Toronto (CA)	Buenos Aires (AR)	Madrid (ES)
	Vienna (AT)	Fribourg (CH)	Hong Kong (HK)
	Turin(IT)	Mumbai (IN)	Oslo (NO)
	Brussels (BE)	Paris (FR)	Helsinki (FI)
	Frankfurt (DE)	Riga (LV)	Milan (IT)
	Rome (IT)	Lisbon (PT)	San Juan (PR)
	Edinburgh (UK)	Tallin (EE)	Taipei (TW)
	New York, NY	Palo Alto, CA	Anchorage (US)
	Moscow (RU)	Manila (PH)	Kuala Lumpur (MY)
	Luxembourg City (LU)	Guam (US)	Vancouver (CA)
	Wellington (NZ)		
K	London (UK)	Amsterdam (NL)	Frankfurt (DE)
	Athens (GR)	Doha (QA)	Milan (IT)
	Reykjavik (IS)	Helsinki (FI)	Geneva (CH)
	Poznan (PL)	Budapest (HU)	Abu Dhabi (AE)
	Tokyo (JP)	Brisbane (AU)	Miami, FL
	Delhi (IN)	Novosibirsk (RU)	Dar es Salaam (TZ)
L	Los Angeles, CA	Miami, FL	Prague (CZ)
M	Tokyo (JP)	Seoul (KR)	Paris (FR)
	San Francisco, CA		

第5章 DNSSECの導入

2009年における最大の変更点は、DNSSECがRoot zoneに導入されたことである。これは2008年に話題となった、kaminsky attack等のDNSレコード偽装に対する危機感が高まったことが影響している。

Root zoneに対するDNSSEC署名は、ここ数年間Root DNS運用者会議やICAN RSSAC会議にて話題に上がってきた。またその要求も伝えられてきた。2009年にVeriSignとICANNによってRoot DNSSECデザインチームが発足され、Root zone署名に関する文章や`http://www.root-dnssec.org/`といったWebページによってその成果が公開されてきた。また、IETF76においては、Root DNSSECデザインチームによるBoFが開催され、議論が行われた。

実際にRoot zoneの署名が開始されたのは、2009年12月である。VeriSignとICANNによって、Root zoneのKSK、ZSKが作成され、2009年12月1日にDNSSECによって署名されたRoot zoneの公開が開始された。zone署名にはNSECを用いており、NSEC3は導入されなかった。しかし、この段階においてはまだ署名されたRoot zoneが公開されただけであり、実際のRoot DNSサーバにこの署名済みzoneが導入されたわけではない。

実運用されているRoot DNSサーバへのDNSSEC署名済みzoneの導入は、2010年1月から段階的に行われていく予定である。まず2010年1月にL Root DNSサーバが署名済みzoneを導入し、次にA Root DNSサーバに導入される。M Root DNSサーバは2010年3月3日に導入される予定であり、切り替え後のトラフィック遷移やサーバへの負荷監視を注意して行う必要がある。

なお、導入されるDNSSEC署名済みzoneは、DURZ（deliberately unvalidatable root zone）と呼ばれるものであり、署名を検証することのできない鍵とともに配布されている。つまり、本当に署名を検証できるzoneと鍵のセットが公開されるのはまだ先であり、2010年7月にRoot zoneのtrust anchorとともに公開される予定である。

第6章 まとめ

M Root DNSサーバは、12年以上に渡り安定的にサービスを提供してきた。特に多階層の冗長構成の導入により、サービスの停止を伴わずにサーバやサーバソフトウェアの保守作業が可能になったことは、サービス停止を伴う保守作業は72時間前に他のRoot DNSサーバオペレータに連絡することが要請されていることを考えると、運用面で大きなメリットがある。また、数多くのISPやIXの協力により、サーバそのものの安定運用に留まらず、インターネットの広い範囲に対して安定なサービスを提供できたことも特筆すべきである。また、Root DNSサーバのIPv6によるサービスも始まり、IPv6の展開に新たなtriggerになったと言うことができる。さらに、DNSSECによる署名もいよいよ導入され、2010年はRoot DNSサーバにとって大きな変更が発生する年となる。M Root DNSサーバでは、WIDE Projectの監督責任のもと、JPRSと共同で管理運用を行い、DNSSECの本格運用に備えた機材更新も含め、今後の対応を行っていく所存である。