

第6部

特集6 M-Root DNSサーバの運用

加藤 朗、関谷 勇司、石原 知洋、遠峰 隆史

第1章 M Root DNSサーバの概要

Root-DNSサーバは、DNSによる名前解決の要である。Cacheを多用するため、毎回のDNSの名前解決の度に参照されるわけではないが、基本的には全ての名前解決に関与しているといっても過言ではなく、DNSによる名前解決を行う上で重要な役割を担っている。

1995年頃まではRoot-DNSサーバはA-Iの9つのサーバで運用されていた。I-Root以外は全てU.S.で稼働していたが、ヨーロッパやアジア太平洋地域へのインターネットの展開により、それらの地区でのRoot-DNSサーバの必要性がIEPGなどで議論されるようになった。このため、1997年に入るとJ-Mの4つが追加された。DNSのUDPを用いる際のメッセージ長は、IP/UDPヘッダを除くと512byteが上限であったため[43]、事実上13がRoot-DNSサーバ数の上限となっている。同年5月にはI-RootがRIPE/NCCの運用で英国のLondonのIXであるLinxに運用が移され、M-Rootは同年8月にWIDE Projectの運用で東京に移動した[44]。この移動は、前もってサーバ基盤を準備しておき、それぞれのIPv4アドレスを新しいものに変更するという形で実施されている。

M-Root DNS Serverは1997年8月に運用を開始してから23年以上、大きなトラブルなく運用を継続している。当初はルータとしてCisco 4500 1台、サーバとしてPentiumPro 200MHz PC 2台(primary/backup)という小規模なハードウェアで運用していた。当初の問い合わせは600qps程であり、現在の80,000 qps程度に比べると1%にも満たない量であった。

2002年に東京拠点で“Anycast in a Rack”として、Anycast

の予備的な運用を、他のRoot DNSサーバに先駆けて開始して以来、2004年に掛けて、Seoul、San Francisco、Parisの各拠点で運用を開始した。当初は、202.12.27.33というアドレスを使用しIPv4のみのサービスを提供していたが、2008年2月からは13あるRootサーバの最初の4つの一つとして、2001:dc3::35というアドレスでIPv6サービスを、後述のSeoul拠点以外で開始した。2013年には全拠点でのIPv6サービスを開始し、現在に至っている。

M-Rootは運用開始当初からWIDE Projectが責任をもって運用しているが、2005年12月から日本レジストリサービス(JPRS)との共同運用を実施している。

第2章 運用拠点

東京拠点では、2004年から、Anycast [45]を用いてDIX-IE、JPIX、JPNAPのそれぞれに別々なクラスタを割り当てて運用しており、他のサービスとはネットワークやサーバを共有していない。

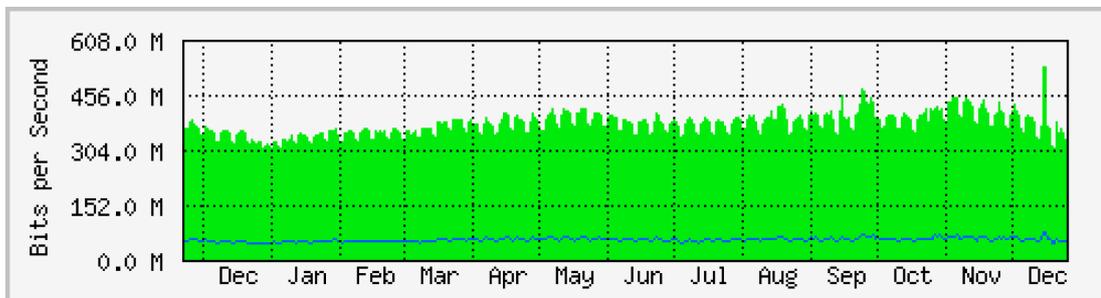
2002年から運用を始めた大阪拠点は、当初は東京のバックアップとして、東京拠点がダウンした場合のバックアップとしての運用であった。これは、BGP [46]でAS-Pathを数個prependすることにより実現していた。しかし、これだけではトラフィックは皆無だったため、同じルータ配下に別なサーバを設置し、RFC1918 [47] (いわゆるプライベートアドレス)等の逆引きに対してMXDOMAINエラーを返すAS112サービス[48]の運用を開始した。また2013年からはAS-Path prependを廃し、定常的なサービスを提供している。

Seoul拠点は、2004年7月に、韓国のIXであるKINXに接続する形態で運用を開始した。韓国国内のISP事情で、規模の大きなISPとpeeringできないため、トラフィックは多くないが、最初の海外でのAnycastの運用として、同年3月にサーバ基盤を設置し、運用経験を積んできた。ここは、国際transitを提供して頂いていないため、主に韓国国内に対するサービスを提供しているが、IXに接続する東南アジアのISPもあるため、必ずしも韓国限定というわけではない。また当初はIPv4のみの運用であったが、2013年には、実質的なpeer数は極めて少ないが、IPv6のサービスを提供している。また、KINXの協力で、KINXに接続していない韓国国内のISPへのトランジットの提供を受け、一定数の問い合わせが処理されるようになった。

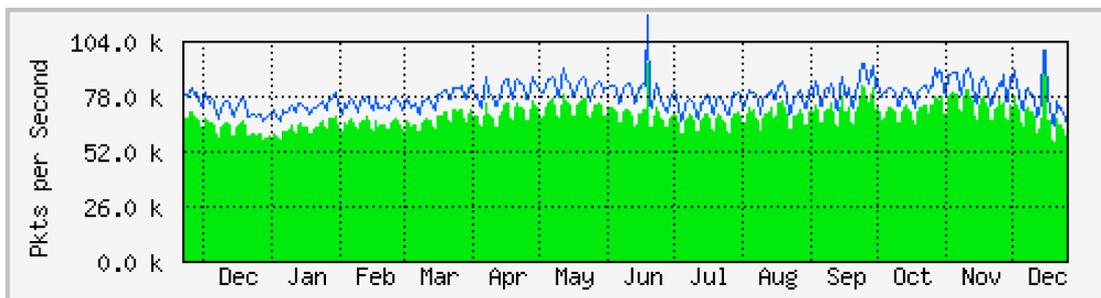
Paris拠点は、2004年9月に運用を開始した。当初はデータセンター内部にスイッチを持っていた、フランスの研究開発ネットワークであるRenaterが運用するSFINXと、France Telecomが運用するPariXの2つのIXに接続していたが、接続するIXも徐々に増え、Equinix Parisを始め、

NL-IXおよびFrance-IXが加わった。M-Rootとしてのサービスは、主にヨーロッパが中心になっているが、アフリカの都市からのアクセスも少なくない。各RootサーバへのRTTの調査では、最もRTTが短いのがM-Root (Paris)という都市もあることがしばしば報告されている。Paris拠点では、サーバは独立に設置しているが、JPのサーバであるe.dns.jpも運用されている。なお、Paris拠点はクラスタ単位での問い合わせ密度が一番高かったため、2016年11月より別なクラスタが追加され、処理の分散が図られている。

San Francisco拠点は、従来から存在したWIDE San Francisco NOC (現在は廃止)に併設する形態で2004年10月から運用を開始している。Paris拠点と同様な形態でe.dns.jpサーバも運用されている。諸般の事情で、2017年12月にSan Francisco市内のデータセンターからSan Jose近郊のデータセンターに移設され、またParis拠点と同様に処理の分散化を図るべく翌2018年1月からは2クラスタ体制で運用が行われている。



(a) トラフィック



(b) パケット数

図1 M-Root DNS 全体の問い合わせ

第3章 2020年のトラフィック傾向

図1にM-Root全体に対するトラフィックの2020年における推移を示す。前年に比べて大きな変動も無く、トラフィック量ならびにクエリ量とも安定した推移を示している。12月に入り若干トラフィックが減少しているのは、Chromeブラウザの実装の変化であると言われている。

第4章 2020年の主な変更点

2020年は、東京拠点の1クラスタ分の機材更新を実施した。また、2020年8月にParis拠点の1クラスタ分の機材更新を実施する予定であったが、COVID-19のため作業することができず、2021年3月に延期している。これに関しても、作業のため渡仏することは容易ではないため、遠隔操作で更新作業ができないかどうか検討している。

第5章 APNICとの協力関係

2002年頃からAPNICでは、APNIC地域でのRootDNSサーバの運用に資金的なサポートをするプログラムを設けている[49]。このプログラムでは、サーバ等の調達をAPNICの資金で行い、それらを貸与することで、経済的に余裕がない各国のNICあるいはIXがRoot DNSサーバをホストできるようになっている。Internet Systems Consortiumが運用するF-Rootを初め、I-Root、K-Rootを含む幾つかのRoot DNSサーバがこのプログラムによって、APNIC地域でのAnycastインスタンスの増強を図ってきた。

近年のM-Rootは、JPRSの予算でサーバの更新および拡張を行っている。予算の制約もあって、インスタンス数の増強は必ずしも容易ではなかった。拠点数を増やすことは、管理コストの増加はあるものの、総合的なサーバの処理能力は向上し、また顧客からみた応答時間の減少も期待できる利点がある。そのため、2017年頃から、M-RootでもこのAPNICのプログラムによって、特にアジア太平洋地域でのサーバ増強を行い、footprintの拡大ができないかどうかの検討が始まった。

最近ではサーバ能力が向上し、仮想化技術などの進展も併せると、1Uサーバ1台に小型スイッチ、ソフトウェアルータがあれば、特定コミュニティに対してサービスを提供する、いわゆるLocal Anycastサーバを設置・運用することが可能になってきている。小型スイッチを用いない方式も考えられるが、近年ではSFPなどを用いることによって、多くのIXが要求する1000GBase-LRを直接サーバに収容できるようにするNICの入手が困難になっている問題を解決することができる。また、多くのサーバは、サーバ本体を管理するインターフェースがあり、これを用いることによってOSなどのインストールも遠隔地から実施することができる。しかし、インターネットに直接接続するのは問題があり、ACLでアクセスを制御したい。そのため、外部にL3スイッチを設置し、メディア変換を行うとともに、ACLによってアクセス制御を実施することができる。

2019年になると、WIDE Project、JPRS、APNICの三者で具体的な議論が始まり、2020年8月にWIDE Project、JPRS、APNICの三者でMoUに署名されるに至った。この枠組みによって、

1. Anycast拠点に対する機材やソフトウェアを、基本的な保守込みでAPNICが貸与
2. 運用は、今までと変わらず、WIDE ProjectとJPRSで実施
3. データセンタのスペース、IXへのアクセス、およびサーバクラスタの管理に関する接続性はホストが提供
4. 以上の点を、ホスト、APNIC、JPRS、WIDEProjectの4者によるMoUとして署名

ということで、拠点の増設を促進することが可能となった。特にホストは3.に示された各要件を提供しなければならないが、ほとんどのIX運用者は既に持っている資源であり、必要な空間も5U程度であることから、これらの準備には大きな問題はないことが多い。機材調達の予算は必要だったが、1.によってAPNICから機材等が貸与されるため、容易にRoot DNSサーバを誘致することができるようになる。

このMoUに基づいた最初のケースは、APNIC自身がホストとなったケースで、オーストラリアのBrisbaneにあるデータセンタに設置されたものである。最初の実装であることから、APNICオフィス内に暫定的に設置され、最小限の設定をAPNICの技術者に依頼したあと、遠隔でサーバのOSの導入などの作業を実施した。これは、今後の運用導入を円滑に実施するためであり、これによってサーバ調達に関する要件も一部変更されている。当初IPアドレスの変更が必要になると想定されたが、接続環境を設定することにより、アドレスの変更作業なしに実環境に投入できるように計らって頂いた。その後、データセンタ内にある2つのIXに対して接続され、2020年12月に運用を開始した[50]。

遠隔地での運用に関して、従来であれば誰かが訪問し、詳しい説明を行うこともできた。アジア太平洋地域では、通信の自由化は必ずしも十分に進んでいるとはいえ、機材受け入れに関して監督官庁の承認が必要になる場合や、運用に際して政府からの免許を受ける必要がある場合もあり、必ずしも簡単に話が進むとは限らない。暫定版である、と断った上で送った技術的要件に基づいて発注が行われてしまったケースもあり、一筋縄ではいかないこともあることが分かっている。

APNICの資金による機材提供は、発注プロセスや納品後のライセンス登録などをAPNICの担当者が制御できる上、一部の機能に関してはM-Rootオペレータもアクセスできるようになっている。ライセンスの定期的な確認や更新作業も必要であり、また将来の話であるが、導入した機材の保守が簡単に延長できるのが5年間であり、それを過ぎないように機材も更新も必要となっている。そのため、APNICの担当者と連絡を密にし、定期的にビデオ会議を開催している。実際、既に複数の拠点に関して話を進めており、APNICの新しい会計年度が始まる2021年1月になってから、これらの拠点に必要な機材等の調達プロセスを開始することになっている。以上のように、APNICの協力によって、あるいはAPNICとは独立した誘致にも対応することによって、今後の新しい拠点への展開を行い、Root DNSサービスの安定的な提供に努めたい。

謝辞

機材更新に関して、機材の提供を頂いたM-Rootの共同運用をしている日本レジストリサービスおよび同社の関係技術者諸氏に感謝します。

M-Rootの運用は、東京拠点ではDIX-IE、JPIX、JPNAP、大阪拠点ではNSPIX3、JPNAP-Osaka、JPIX-Osakaの、Seoul拠点ではKINXおよびKREONET2、Paris拠点ではTelehouse Paris、Renater、SFINX、PariX、FranceTelecom、Equinix Paris、NL-IX/OpenTransit、France-IX、TINET、およびAFNICの、San Francisco拠点ではSFMIXおよびFCIX、Brisbane拠点に関してはAPNIC、この他数多くのISPや諸氏のご協力のもと運用されており、これらの団体および関係者に感謝します。