

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[15]、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K-Rootが英国のLondonのIXであるLinxにおけるRIPE/NCCの運用に移管され、M-Rootは同年8月にWIDE Projectの運用で東京に移動した[16]。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24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Cisco4500 一台、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80,000qps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[17]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つのうちのひとつとして、2001:dc3::35というアドレスの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。さらに、後述のように、APNICサポートを受けて、主にアジア太平洋地域に、小型の拠点の運用を2020年末から実施している。

第2章 運用拠点

東京拠点では、2004年から、Anycast[17]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[18]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[19]（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[20]の運用を開始した。また2013年からはAS-Pathprepend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がないため、主に韓国国内に対するサービスを提供していた。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大手ISPへのトランジットの提供を受け、またKINXにアクセスしている韓国外のISPも増加してきたことから、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最も高かったため、2016年11月より別なクラスタを追加し、処理の分散を図っている。

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。

2020年に予定されていたParis拠点の1クラスタ分の機材の更新は、COVID-19で渡航ができないため、一年間延期し、2021年に実施することになった。しかし、翌年になってもこの状況は大きく改善することはなかったため、通常どおり我々が現地に向かって作業することは困難であった。そのため、機材の販社のParis支店に場所を用意して頂き、Linuxサーバを踏み台として、その奥にサーバやネットワーク機器を仮組みして頂いた。遠隔操作によってネットワーク機器の設定やサーバへのOSのインストールなどの作業を実施した。その後、それらの機材をデータセンターに搬入し、連絡を取りながら交換作業を依頼した。

従来の拠点のハードウェア更新作業は、複数台あるサーバは一台づつ交換しておき、ルータや一部のスイッチの交換時のみサービスを停止するという方法で、ダウンタイムは1時間程度で済んでいた。今回は、作業手順の煩雑化やそれによる問題発生を防ぐため、すべてを停止し、交換作業を依頼した。そして電源投入し、接続等を確認したのち、サービスを再開した。そのため、27時間程度のダウンタイムになった。この期間中もM-RootのParisにあるもう一つの別のクラスタは平常運転しており、問題は報告されていない。

第3章 2022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2022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。

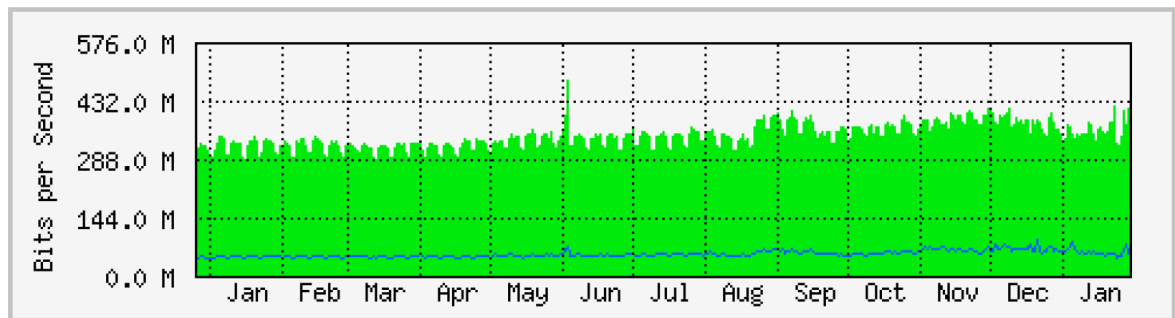
第4章 2022年の主な変更点

2021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、前述のように、Paris拠点の1クラスタ分の機材更新は、COVID-19の影響、およびそれに伴う機材の納期遅れから、2021年6月に延期して実施した。

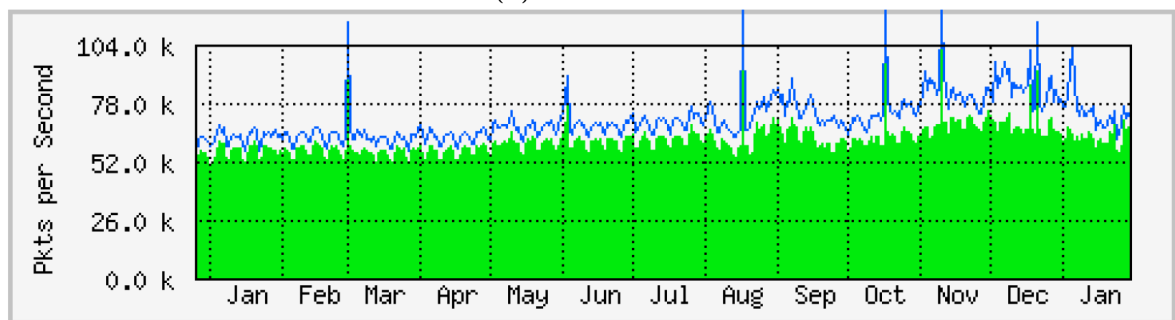
また、後述のようにベトナムのハノイにおいて、主に

ベトナム国内へのサービス提供を目的とした、いわゆる“Small Anycast”ノードが2021年10月に、ベトナムの

Local Internet RegistryであるVNNICのホストのもとに運用を開始している。



(a) トラフィック



(b) パケット数

図1 M-Root DNS全体の問合わせ

表1 APNICサポートのM-Root 一覧

略称	国	都市	種別	接続 IX 等	運用開始	備考
BNE	AU	Brisbane	Small	MegaIX, QLD-IX, IX Australia	Dec 2020	
HAN	VN	Hanoi	Small	VNIX Hanoi	Oct 2021	(*1)
GUM	GU	Mangilao	Small	MARIIX (GOREX)	Feb 2022	
KUL	ML	Kuana Lumpur	Small	MyIX	Mar 2022	
BKK	TH	Bangkok	Small	BKNIX	Nov 2022	
SIN	SG	Singapore	Large	SingAREN, Equinix	Nov 2022	
(MNL)	PH	Quizon	Small	PhOpenIX-Manila	—	(*2)
(JKT)	ID	Jakarta	Small	IIX, OpenIXP	—	(*2)
(KHH)	TW	Kaohsiung	Small	TWIX	—	(*2)
(DAC)	BD	Dahka	Small	BDIX	—	(*2)
(UBN)	MN	Ulaarnbaatar	Small	MISPA	—	(*2)
(KTM)	NP	Kathmandu	Small	NPIX	—	(*2)

*1 ハードウェアはVNNIC 提供

*2 運用調整中

第5章 APNICの協力によるアジア太平洋地区への展開

APNIC Foundationからの資金をベースに、WIDEProject、JPRS、APNICの三者で2020年8月にMoUが締結され、これに基づいて、アジア太平洋地域でのM-Rootサーバの展開が可能になった。このMoUに基づいた最初のケースは、APNIC自身がホストとなり、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。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。

この枠組みを受け、WIDE Project、JPRS、APNIC毎月2回程度ビデオ会議を開催し、サーバ設置の可能性について調整している。機材等はAPNICが提供するが、データセンター内にサーバを設置するスペース、IXのポートおよびそれに至るCross Conenctファイバ、さらにM-Rootサービスを実施するためのIPv4/IPv6アドレスおよび接続性、リモートハンドサービスをIXなどのホスト側に提供して頂くことになる。もちろん、それぞれの実状にあった修正をしながら、上記3者とM-Rootのホストとなる現地事業者を加えた4者でMoUを締結している。

しかしながら、最初に相談をしてからMoUの締結に至るまでの時間が長いケースが多いこと、およびCOVID-19の影響によるベンダーからの機材の納入遅れがあり、必ずしも順調に作業が進んでいるわけではない。2021年中に運用を開始できたのは、10月のベトナムのハノイだけであった。この他、機材は到着し、インストール作業は遠隔で実施しているものの、ACLなどの問題で現在調整中な拠点が一つ、一部の機材は納品されたものの全てが揃っていない拠点が一つある。さらに、MoUには署名し機材は発注したものの、まだ納品されいない、あるいは納期すら不明な拠点が6拠点を数えている。さらに、関係者の合意は得られ、MoUに署名する手続きに直ぐにも進めそうな拠点が一つ、具体的な話が進行している拠点が一つある。

これ以外に、アジア太平洋地域ではわが国は北東にあり、地理的中心ではないことから、いろいろなところからの問い合わせのRound Trip Timeという観点からは中

心ではない。APNIC担当地域の中心であり、上陸している海底ケーブルの数などを考えると、Singaporeにも拠点が必要である。これは、主に拠点設置国あるいはその周辺国へのサービスを限定している、いわゆる“*Small Anycast*”では十分とは言えず、Globalに経路を広告する、“*Global Anycast Node*”とすべきである。この場合、機材の量やそれに必要なスペースや資金もSmall Anycastとは異なるため、現在その実現に向けてAPNICと調整が行われている。

第6章 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。MRoot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、EquinixParis、NL-IX/OpenTransit、France-IX、TINET、およびAFNICの、San Francisco拠点ではSFMIXおよびFCIX、Brisbane拠点およびアジア太平洋地域のSmall Anycast拠点(準備中も含む)に関してはAPNIC、Hanoi拠点に関してはVNNIC、Guam拠点に関してはUniversity of Guam、Kuala Lumpur拠点に関してはMyIX、Bangkok拠点に関してはBKNIX、Singapore拠点に関してはAPNIC、その他数多くの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。