

APAN 東京 XP の構築と運用について

北辻 佳憲* 小林 克志** 北村 泰一** 加藤 朗*** 小西 和憲*
*(株)KDD 研究所 **郵政省通信総合研究所 ***東京大学

概要

1997年12月に運用を開始した APAN (Asia-Pacific Advanced Network) は、接続地点、帯域、接続性、サービスを拡充し、アジア・太平洋と米国を相互に接続する国際実験ネットワークとして世界的に認識されつつある。これまでの国際実験ネットワークは一对向の国と国をポイントツーポイント回線で接続する接続形態が多かったのに対し、APAN はポイントツーポイントの回線を各国 XP に終端して複数の国をマルチポイントに接続する国際実験環境を提供する。本稿では、APAN 東京 XP の構築と運用について、特にネットワーク構成、経路制御、今後の展開について述べる。

Construction and Operation of APAN Tokyo XP

Yoshinori KITATSUJI* Katsushi KOBAYASHI** Yasuichi KITAMURA**
Akira KATO*** Kazunori KONISHI*

*KDD R&D Laboratories Inc. **Communications Research Laboratories ***University of Tokyo

abstract

APAN (Asia-Pacific Advanced Network), which has initiated its provision of the high performance research and educational network environment in Dec, 1997, is widely used with increasing exchange points, expanding bandwidth on the international links and providing stable reachabilities and expanded services. APAN provides a multi-point testbed network to Asia-Pacific and US project collaborators while many existing international testbed networks are limited to point-to-point environment. This paper describes the network construction and operation of APAN Tokyo XP, focusing on network configurations, routing, and future plans.

1. はじめに

APAN^[1] は各国の研究機関がアジア・太平洋に持つ国際回線を相互接続した多地点実験ネットワークで、APAN 東京 XP (eXchange Point, 以下東京 XP とする) を中心に、APAN ソウル XP (韓国)、SingAREN^[2] (シンガポール)、MSC^[3] (マレーシア)、PHnet^[4] (フィリピン)、ACSys^[5] (オーストラリア)、StarTAP2^[6] (米国) と相互接続を行っている。国内においては IMnet^[7]、WIDE、CRL^[8]、大阪大学と相互接続を行っている。

APAN はプロジェクト指向の利用規約を IPv4 のユニキャスト通信に設けており、APAN Allocation Committee^[9] に承認されたプロジェクト参加組織の間の通信を許可している。また、IPv6、マルチキャストなどに関しては通信を提供すること自身を実験と位置付け、利用者を制限していない。

本稿では、東京 XP の接続構成、利用規約にもとづく経路制御、ネットワーク運用、今後のネットワークの拡大および実験等について述べる。

2. 接続構成

東京 XP は 6 カ国 7 XP と主に ATM 回線で相互接続を行い、国内においても主要な学術情報ネットワークと相互接続を行っている。本節では、東京 XP と各国 XP との相互接続構成および東京 XP の物理的な接続構成について述べる。

2.1 相互接続構成

APAN の国際回線の接続構成を図 1 に、国際回線の帯域

表 1: APAN の国際回線と帯域

接続組織	国	帯域
ACSys	オーストラリア	768Kbit/s
AIT	マレーシア	1.5Mbit/s
MSC	マレーシア	192Kbit/s
PHnet	フィリピン	768Kbit/s
SingAREN	シンガポール	2Mbit/s
ソウル XP	韓国	8Mbit/s
StarTAP2	米国	100Mbit/s

を表 1 に示す。東京 XP はソウル XP、MSC、SingAREN、StarTAP2 に ATM 回線で相互接続し、ACSys および AIT^[10] へ、それぞれ RWCP^[11] および WIDE を経由して接続している。国内においては、IMnet と FDDI および ATM 回線で、WIDE、CRL、大阪大学と ATM 回線で相互接続を行っている。SingAREN は StarTAP2 およびソウル XP と ATM 回線でそれぞれ 14Mbit/s、2Mbit/s で相互接続を行っている。

2.2 東京 XP の物理的な接続構成

東京 XP を構成する主な機器とその用途を表 2 に、主な機器の物理的な接続を図 2 に示す。東京 XP と接続組織間のトラフィックおよびルータ間のトラフィックは主に TPS-Tokyo (Fore 社製 ASX-1000) を経由して交換され、DNS、WWW、メール等のサーバ(図中省略)は MS-Tokyo (Foundry Networks 社製 BigIron4000) を経由して TPR-

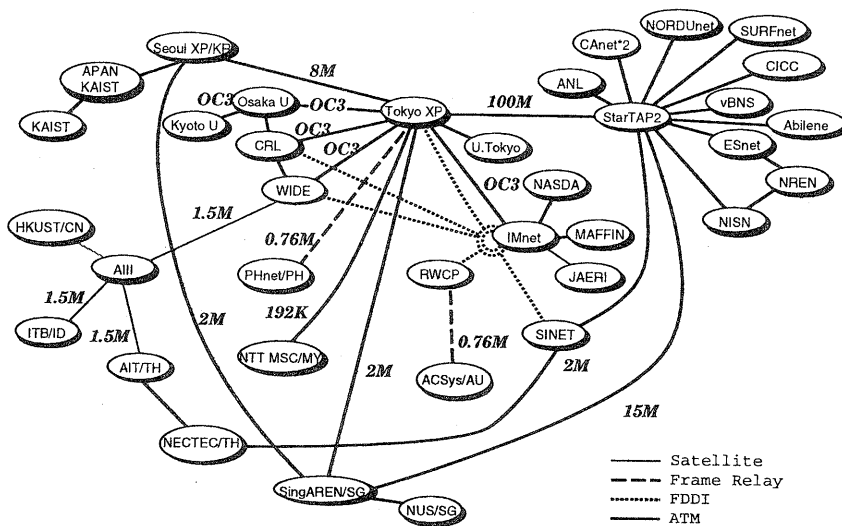


図 1: 国際回線の接続図

表 2: 東京 XP を構成する主な機器

機器の種類	機器名	用途
ルータ	TPR-Tokyo	IPv4 トラフィックの転送。 東京 XP の主ルータ。
ルータ	TPPR-Tokyo	IPv4 トラフィックのソース を検証。
ルータ	TP6R-Tokyo	IPv6 トラフィックの転送。
ルータ	TPM2R-Tokyo	JP Mbone との接続。
ルータ	DS-Tokyo	DiffServ 実験。
ATM スイッチ	TPS-Tokyo	接続回線およびルータのトラフィックの交換。
L3 スイッチ	MS-Tokyo	サーバ向けトラフィックおよび保守用。
サーバ	oc3mon	OC3mon (米国 ATM 回線用)。
サーバ	oc3mon2	OC3mon (韓国、シンガポール ATM 回線用)。
サーバ	surveyor	国際回線の一方方向遅延の検証。
サーバ	skitter	RTT および接続性の検証。

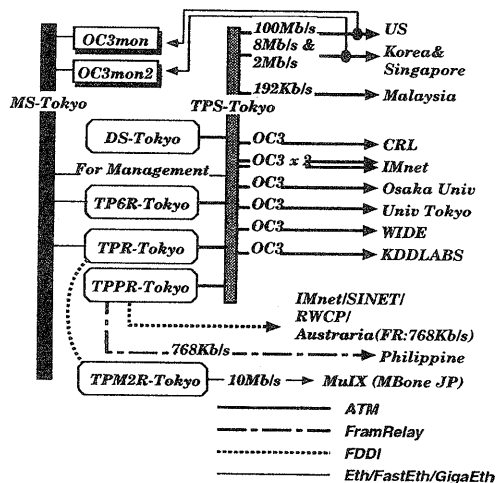


図 2: 東京 XP の主な機器の接続図

Tokyo (Cisco 社製 7505) に接続される。IMnet, SINET, RWCP に接続する FDDI 回線およびフィリピン向けフレームリレー回線は TPPR-Tokyo (Cisco 社製 7505) に接続され、JP Mbone 向けの FDDI 回線は TPM2R-Tokyo (Cisco 社製 4500) を経由して MuIX (Multicast Inter eXchange) に接続される。

東京 XP には OC3mon^[12]、Skitter^[13]、Surveyor^[14]等の計測機器が設置してある。OC3mon は ATM 回線を通るセルをキャプチャしてキャプチャデータを保存するプログラムで、東京 XP では米国回線と韓国/シンガポール回線をキャプチャできるように 2台が設置されている。Surveyor は KAIST^[15] および SingAREN に設置されているものと共に一方方向遅延を計測する。Skitter は世界中の約 20000 台のホストへの到達性と RTT を計測する。

3. 経路制御

3.1 ユニキャストの経路制御

ユニキャスト通信において認定されていない組織の利用を制限するために、東京 XP では図 3 のように TPR-

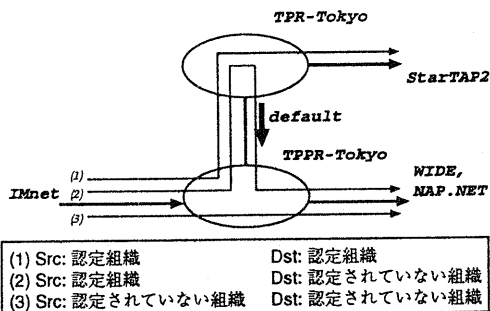


図 3: 2 台のルータによる経路制御

Tokyo と TPPR-Tokyo の 2 台のルータを使用している。TPPR-Tokyo は認定組織の経路を持っており、認定組織のバケットしか送り込まれないことを想定している。このため、組織全体が認定されている組織と、認定組織のバケットしか送り込まない組織は、直接 TPPR-Tokyo に接続される。

一方、認定組織とそうではない組織の両方のバケットを送り込む組織は、TPPR-Tokyo に接続する。TPPR-Tokyo は接続インタフェースでポリシーベースドルーティングの設定を行っている。TPPR-Tokyo のポリシーベースドルーティングはバケットのソースアドレスを検証し、認定組織からのバケットであれば TPR-Tokyo へ転送し、認定されていない組織からのバケットであれば通常の経路制御を行う。TPPR-Tokyo で行われる通常の経路制御が、認定されていない組織からのバケットに限られるために経路表に認定組織の経路は保持されない。さらに、TPR-Tokyo では、デフォルト経路を TPPR-Tokyo へ向け、認定組織宛ではないバケットを TPPR-Tokyo へ転送する。

TPPR-Tokyo へ転送されるバケットはそれぞれ、次の経路を通して接続組織へ送り出される。

- (1) バケットのソースとデスティネーションの両方が認定組織である場合
バケットはソースが検証されて TPR-Tokyo へ転送される。TPR-Tokyo ではバケットのデスティネーションをもとに、接続組織へ送り出される。
- (2) バケットのソースが認定組織でデスティネーションが認定組織ではない場合
バケットはソースが検証されて TPR-Tokyo へ転送される。TPR-Tokyo ではバケットのデスティネーションに一致する経路が無い場合、デフォルト経路にしたがって TPPR-Tokyo へ転送される。TPPR-Tokyo では TPR-Tokyo から転送されたバケットは通常の経路制御が行われ商用プロバイダ等へ送り出される。
- (3) バケットのソースが認定組織ではない場合
TPPR-Tokyo でソースが認定組織とマッチせず通常の経路制御が行われ、商用プロバイダへ送り出される。

経路情報の交換に関しては、どちらのルータに接続する組織も認定組織の経路を東京 XP から受けなければな

表 3: マルチキャスト接続組織

接続組織	接続方法	プロトコル
ACSys	DVMRP トンネル	DVMRP
SingAREN	直接接続	PIM-DM
ソウル XP	直接接続	MBGP/PIM-SM/MSDP
StarTAP2	直接接続	MBGP/PIM-SM/MSDP
CRL	直接接続	MBGP/PIM-SM/MSDP
JP Mbone	DVMRP トンネル	DVMRP
KDDLABS	直接接続	MBGP/PIM-SM/MSDP
NASDA	DVMRP トンネル	DVMRP

らない。そのため、全ての BGP のセッションは TPR-Tokyo に終端し、TPPR-Tokyo に接続する組織は BGP multihop の設定を行う。

3.2 マルチキャストの経路制御

東京 XP では PIM-SM^[16] と MSDP^[17] を導入し、マルチキャストの経路制御を行っている。マルチキャストはユニキャストと異なり、利用規約がないため RPF (Reverse Path Forwarding) チェックに認定組織とそうではない組織の両方の経路が必要となる。しかし、ユニキャストの経路は TPR-Tokyo と TPPR-Tokyo に分けて保持されるため、東京 XP ではどちらかのルータにマルチキャスト用の経路を別に持たなければならない。以上のことから、東京 XP は TPR-Tokyo に MBGP^[18] を導入している。東京 XP とマルチキャストで接続する組織を表 3 に示す。

APAN におけるマルチキャスト通信は、ネットワーク構築の初期の段階から試験的に運用され、当初は PIM-DM (PIM-Dense Mode) を利用してきた。1999 年 7 月に APAN 各国 XP で MSDP/PIM-SM/MSDP への移行を順次行い、現在に至っている。

ACSys は RWCP ネットワークを経由して東京 XP と接続し、マルチキャスト用のルータが PIM をサポートしていないために、DVMRP トンネルを確立してトンネル上で経路制御を行っている。

SingAREN については初期の段階から PIM-DM を使用してきたが、2000 年 3 月にルータの OS のバージョンアップが行われ、東京 XP との接続に MBGP/PIM-SM/MSDP を導入する予定となっている。

JP Mbone と東京 XP の接続は、1999 年 12 月に東京 XP が MuIX (WIDE) に接続することで相互接続を開始した。

東京 XP は APAN 各国 XP と同様に、ルータが持つマルチキャストセッションを Web に可視化するツールを導入し、<http://www.jp.apan.net/multicast/sd.shtml> で公開している。通常のセッション数は 50~60 である。

4. ポリシーベースドルーティングの影響

TPPR-Tokyo で行われているポリシーベースドルーティングはルータにおける通常の経路制御と異なるため、ルータの性能に影響する負荷を発生させる可能性がある。

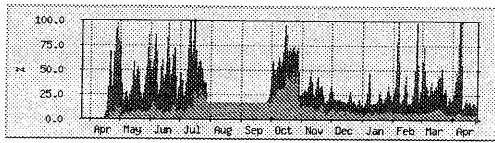


図 4: TPR-Tokyo の CPU 使用率の変化

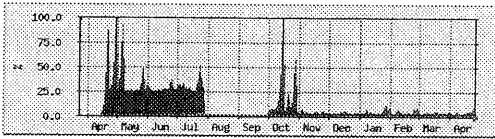


図 5: TPPR-Tokyo の CPU 使用率の変化

図 4,5 に TPR-Tokyo および TPPR-Tokyo の CPU 使用率のグラフを示す。

両図は 1999 年 4 月から 2000 年 4 月の両ルータの CPU 使用率を示しており、MRTG^[19] によるリアルタイム計測によって得られたグラフである。計測では 5 分間毎に平均 CPU 使用率を計算している。CPU グラフにおいては、淡色は 2 日間全体の平均値を示しており、濃色は 2 日間の中での最大の 5 分間の平均値を示している。なお、1999 年 8 月から 9 月にかけて計測を停止していた。

図 4 の TPR-Tokyo の CPU 使用率は 1999 年 10 月からほぼ半減しているが、これはルータのルーティングスイッチプロセッサの高速化 (CISCO RSP2 から RSP4 へ変更) を行ったためである。TPR-Tokyo は ATM インタフェースおよび FastEthernet インタフェースに CX-AIP および CX-EIP2 を装備し、パケットの送出インタフェースの決定処理を RSP4 で行っているために TPR-Tokyo に比べて高い値を示している。

図 5 の TPPR-Tokyo の CPU (RSP4) 使用率は、複数のインタフェースを統合する VIP2-50 をサポートして以来 1-2% にまで低下した。ポリシーベースドルーティング PBR は VIP2-15 に挿入されている FDDI インタフェースで行われているため、ルータの CPU がパケットの転送に影響を受ける程の負荷が観測されていない。

5. トラフィック計測

APAN の主な国際回線として米国回線、韓国回線、シンガポール回線の日常のユニキャストトラフィックの変化を図 6~8 に示すトラフィックの計測は MRTG で行い、1999 年 4 月から 2000 年 4 月のトラフィックを示している。これらの図は、前述の図と同様に、2 日間の平均値とその間の 5 分平均の最大値を、送り出されるトラフィックについては細線で、入ってくるトラフィックについては塗面でそれぞれ示している。

米国回線 (70Mbit/s) は StarTAP2 から東京 XP へ向かうトラフィックが多く、韓国回線 (8Mbit/s) は東京 XP からソウル XP へ向かうトラフィックが若干多い。両回線のトラフィックとも順調な増加を示している。

APII-SG 回線 (2Mbit/s) はマルチキャストトラフィックが 3 月まで定期的に双方向に流れていたが 2000 年 3 月からは減少した。この時期に SingAREN のボーダルータの OS をアップグレードを行っており、新しい OS へ変

更後、マルチキャストブルーニングが正常に行われていると想像される。

6. ネットワーク運用

東京 XP の運用には、障害対策、障害対応、新規回線または新規プロジェクトによる接続調整等がある。以下にそれぞれについての現状を述べる。

6.1 接続調整

ユーザは国際実験またはデモで APAN のネットワークを利用するために、APAN Allocation Committee へ利用申請を提出する。ユーザは共同実験者のネットワークまでの ATM による PVC 接続か、IPv4 または IPv6 の接続性を選択できるが、ATM によるエンド・エンドの PVC が割り当てられることは少ない。これは、バス上の複数の組織がすべて全て ATM をサポートしていることが少ない、接続組織がダイレクト PVC の割当を容認しないなどの理由のためである。

APAN Allocation Committee は、利用者が希望するバスの回線のオーナーおよび APAN の接続組織と調整を行い、申請者へ利用の可否を伝える。利用が承認されると、APAN Allocation Committee は NOC (Network Operations Center: 各国 XP 運用者) へ接続調整の連絡を行う。接続調整の多くは新規回線またはトンネル回線の設定と BGP による経路の広報/受け入れで、設定状況によって、利用者、接続組織に接続性の確認の協力を要請する。

NOC における接続調整が完了すると、NOC は APAN Allocation Committee および利用者へ連絡する。最後に NOC は障害対策として、新規利用の PVC および IP address の WEB 登録や、必要であればトラフィックの収集等を行う。

現在の認定組織はアジア・太平洋地域で 55 組織が、米国・ヨーロッパ地域に 2000 余りが承認されている。

6.2 障害対策

東京 XP は APAN 各国 XP と同様に、障害対策として、Web による運用ページの開設とルータ/スイッチのログおよび設定のバックアップを行っている。Web はパスワード認証付きで接続組織の運用者 (必要な場合には利用者) に公開し、東京 XP の状況把握に利用されている。

障害対策として、東京 XP では回線の監視とログの収集および設定の定期的な保存を行っている。東京 XP の運用開始から 1 年余りにかけて障害の大部分が回線障害であったため、東京 XP では回線の監視に重点をおいている。監視の方法は、接続組織および東京 XP のルータまたはサーバへ ping を定期的に行ってメールによる接続性の自動報告を行う。また、ATM 回線においては OC3mon のプログラムを一部変更して PVC 毎のパケットの増加を定期的に監視し、増加が観測されない PVC をメールで自動的に報告するようにしている。なお、現在はネットワークの状態が安定しており、1 年以上回線障害が生じていない。

ルータ/スイッチのログは常時サーバへ転送され、設定に関しても一日に一度自動的にサーバへ保存される。また、サーバに保存される全てのログと設定は一日に一度別のサーバへバックアップされる。

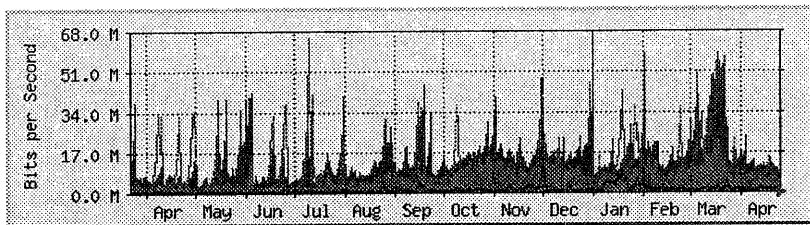


図 6: TransPAC 回線 (70Mbit/s) のトラフィックの変化

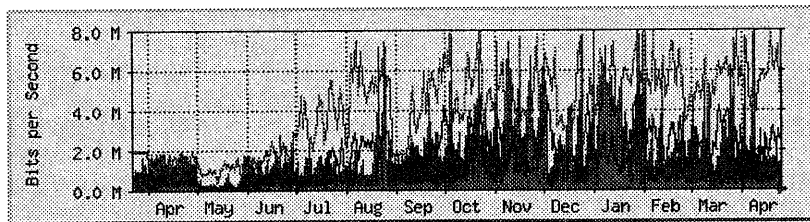


図 7: APII-KR 回線 (7Mbit/s) のトラフィックの変化

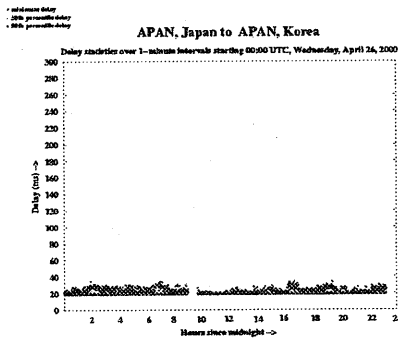


図 9: 東京 XP から KAIST への一方向遅延

ネットワークの調整等で止むをえず発生するサービスの停止および不意の障害等に対してはトラブルチケットを発行し、接続組織等の関係者へメールで連絡を行って復旧活動を行う。

6.3 ツール

東京 XP では運用上有用な運用ツールおよび性能評価ツールをサーバへインストールし、これを運用者および利用希望者間で調整をとりながら使用している。東京 XP で利用されているツールを表 4 に示す。

Surveyor の計測例として、東京 XP と KAIST の間の一方向遅延を図 9, 10 に示す。グラフは 2000 年 4 月 26 日の計測を示しており、東京 XP から KAIST 方向の遅延に定常的なばらつきが観測されている。

7. 展望

これまで、APAN のネットワークはアジア・太平洋の接続地点の拡大、回線の増速を行ってきた。ここでは、今後予定されているデモの支援、回線接続計画、実証実験について説明する。

7.1 デモの支援

マルチポイントの実験環境を提供する APAN のネットワークは、これまでに多くのデモを支援してきた。代表的なデモの支援として、iGrid, Super Computing,

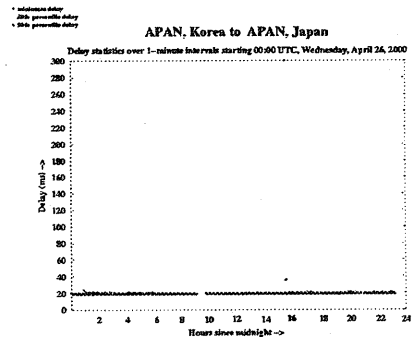


図 10: KAIST から東京 XP への一方向遅延

表 4: 東京 XP のツール

項目	用途
netperf	サーバクライアント型の帯域測定
mping	複数の ICMP パケットを送出する単体での帯域測定
surveyor	一方向遅延測定
skitter	接続性および RTT 計測
oc3mon	ATM トラフィックキャプチャシステム
mrtg	SNMP によるカウンタ値の取得とグラフ化
pchar	ICMP の応答間隔を計測する単体での帯域計測
BGP	ルータの BGP 経路の可視化
route view	
multicast session view	マルチキャストセッションの可視化

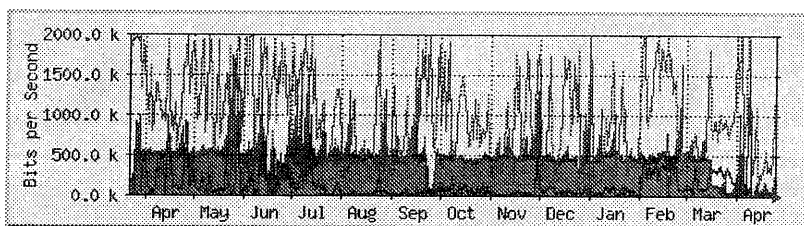


図 8: APII-SG 回線 (2Mbit/s) のトラフィックの変化

Internet Workshop, Global Observation Information Network 等の国際会議があり、今後 INET2000, NLANR/Internet2/APAN Joint Meeting のデモが計画されている。

特に 2000 年 6 月に行われる INET2000 のデモにおいては、東京 XP とハワイ大学の間にならな米国回線 (45Mbit/s) を一時的に接続する。この東京 XP-ハワイ大学回線は、すばる望遠鏡で撮影される映像を INET2000 会場へ伝送するために使用される計画が立てられており、その他に、QBone Network (東京 XP-シカゴ回線を經由) を利用して MPEG 伝送のデモも予定されている。

7.2 新規回線の接続

APAN KR はソウル XP から米国への新規回線計画を、APAN CH が AIII と衛星による新規回線計画をそれぞれ立てている。ソウル XP-米国の回線が確立されれば、アジアと米国の間の冗長性がさらに強化される。また、2000 年 12 月に行われる NLANR/Internet2/APAN Joint Meeting に既存の東京 XP-シカゴ回線の増速とともにハワイ大学との接続が予定されている。

7.3 QoS ネットワークの構築

APAN QoS WG の活動の一つに、QBone ネットワークの設計/構築/試験運用がある。東京 XP および Seoul XP は KDD 研究所と KT の研究所の間のダイレクトな PVC 割当を行い、両機関の DiffServ 実験を支援してきた。2000 年 2 月の APAN QoS WG 会合において、この実験ネットワークを拡大することが承認され、マルチポイント実験環境を構築することとなった。現在、KT の研究所、東京 XP、KDD 研究所、StarTAP2、ANL^[20]、EMERGE^[21] の DiffServ ルータの相互接続が行われている。QBone ネットワークは既存のネットワークとは独立して構築されており、接続は全て ATM PVC で行われ、接続性を確保することからルータは一社に統一されている。

ANL は GARNET^[22] と呼ばれる実験ネットワークを既に構築しており、EMERGE と相互接続を行っている。これらの実験ネットワークはこれまでに、シェーピング技術およびキューイング技術の実証実験、リソース管理およびリアルタイムデータ解析システムの開発と試験利用等を行ってきた。

QBone ネットワークで行われる実験は、現在利用可能な技術をもとにどの程度の品質が得られるかを実証的に試験し、リソース管理手法や運用技術の確立を目指している。また、INET2000 で行われるデモの一部のトラフィックを本ネットワークで伝送する計画とともに、新規のデ

モの調整を通して QBone ネットワークにおける運用ノウハウの習得が期待される。

QBone ネットワークは実験の進展に併せて随時接続を拡大し、将来はこれらの実験で得られた技術を APAN の既存のネットワークに導入することを目標としている。

8. おわりに

本稿では、APAN 東京 XP の接続構成、ポリシーベースルーティングと MBGP を用いた経路制御、代表的な国際回線のトラフィック、障害対策および回線監視等に関するネットワーク運用、QBone 実験ネットワークおよび新規回線計画等について述べた。

参考文献/URL

- [1] Asia-Pacific Advanced Network, <http://www.apan.net/>
- [2] Singapore Advanced Research & Education Network <http://www.singaren.net.sg/>
- [3] Multimedia Super Corridor, <http://www.mdc.com.my/>
- [4] Philippine Network Foundation, <http://www.ph.net/>
- [5] Australian Commonwealth Government's Cooperative Research Centers, <http://acsys.anu.edu.au/>
- [6] Science, Technology, And Research Transit Access Point <http://www.startup.net/>
- [7] Inter-Ministry Research Information network, <http://www.imnet.ad.jp/>
- [8] 郵政省通信総合研究所, <http://www.crl.go.jp/>
- [9] APAN Allocation Committee, <http://alloc.apan.net/>
- [10] Asian Institute of Technology, <http://www.ait.ac.th/>
- [11] Real World Computing Partnership, <http://www.rwcp.or.jp/>
- [12] National Laboratory for Applied Network Research, <http://www.nlanr.net/NA/Oc3mon/>
- [13] <http://www.caida.org/tools/measurement/skitter/>
- [14] Surveyor Project, <http://www.advanced.org/surveyor/>
- [15] Korea Advanced Institution of Science and Technology, <http://www.kaist.ac.kr/>
- [16] Estrin, D., et al., "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification," RFC 2362, June 1998.
- [17] Farinacci, D., "Multicast Source Discovery Protocol (MS-DP)," draft-ietf-msdp-spec-05.txt, February 2000.
- [18] Bates, T., et al., "Multiprotocol Extensions for BGP-4," RFC 2283, February 1998.
- [19] Multi Router Traffic Grapher, <http://ee-staff.ethz.ch/oetiker/webtools/mrtg/mrtg.html>
- [20] Argonne National Laboratories, <http://www.anl.gov/>
- [21] ESnet/MREN Regional Grid Experimental NGI Testbed, <http://www.evl.uic.edu/cavern/EMERGE>
- [22] GARNET, <http://www-fp.mcs.anl.gov/qos/garnet.htm>