

学術ネットワークにおける台湾地震の影響

北村 泰一^{†1} 李 泳錫^{†2}
 崎山 亮^{†3} 岡村 耕二^{†3}

日本時間 2006 年 12 月 26 日 21 時 26 分頃、台湾南方沖を中心とする大規模な地震（台湾南方沖地震）が発生した。この地震は海底を走るケーブルシステムを切断し、アジア太平洋地域のすべての電気通信に障害を与えた。インターネット通信においては、経路情報を複数用意することで冗長性を持たせる運営が行われていたため、迂回経路を選択することが自動的に行われていた。今回の台湾南方沖地震においては、複数のケーブルシステムが近隣に敷設されていたため、台湾を中心として、東西の通信を分断する通信路障害を発生し、地球を周回したり、いったんオセアニア方面を経由するなど大きく南側を回るといった経路が選択された。これらの経路は、通信路長を長いものにしてしまうため、通信品質はきわめて悪いものとなっていた。また、この迂回手段は、回線容量の大きくない回線へのトラフィック集中も発生させた。このため、ネットワーク運用者は手動により BGP の経路情報運営方針の変更を行わざるをえなかった。一方、経路情報の調整による応急措置を行ったにもかかわらず、学術ネットワークへの接続性が得られない研究組織が依然として残り、商用インターネットを学術ネットワークへの接続手法としていた組織もあった。次世代インターネットサービスへの復帰は、これら、研究組織が研究を進めていくうえでは重要であるため、ネットワーク運用者がケーブル接続において、行える範囲で、接続変更と、経路情報の大幅な変更を行い、“孤立”した研究組織を学術ネットワークに接続することも行われた。台湾南方沖地震により、学術ネットワークなどのインターネット通信においては、ネットワークオペレータ同士が常時連絡をとり合い、トラフィック制御を協力して行うことが最も有益であることが分かった。

The Impact of the Taiwan Earthquake against Research and Education Networks

YASUICHI KITAMURA,^{†1} YOUNGSEOK LEE,^{†2} RYO SAIKIYAMA^{†3}
 and KOJI OKAMURA^{†3}

In this paper, we explain how the network failures had been caused by the natural disaster, describe how the restoration methods had been took, and present the lessons learned from the recovery procedures. At 21:26 of December 26th (UTC+9) 2006, a serious undersea earthquake off the coast of Taiwan has occurred, which was measured at 7.1 on the Richter scale. This earthquake caused a big damage to the cable systems under the sea. Because of the fiber cable failures, a lot of communications in the Asia Pacific networks across several countries have been cut off. The first recovery step after earthquake was taken by BGP routers that detoured traffic along the redundant backup paths. However, redundant BGP routes have provided the connectivity of the poor quality. Thus, the operators took the traffic-engineering steps to improve the quality of recovered communication. To avoid filling the narrow-bandwidth link with such detoured traffic, the operators had to change the BGP routing policy. Despite the routing-level first aid, a few institutions could not be directly connected to the R&E network community, because they had only a single link to the network. For these single-link networks, the commodity link was temporarily used for connectivity. Then, cable connection configurations at the switches were changed to provide high bandwidth and the next generation Internet service. What we have learned from the whole restoration procedure are that redundant BGP routing information is useful for recovering connectivity but not intelligent for providing the available bandwidth for the re-routed traffic load, and that the collaboration between operators was valuable to solve traffic engineering issues such as poor-quality re-routing and lost connections of single-link networks.

^{†1} 独立行政法人情報通信研究機構

National Institute of Information and Communications
Technology

^{†2} 国立忠南大学

Chungnam National University

^{†3} 九州大学

Kyushu University

1. はじめに

インターネットを利用した通信容量の増加にともない、ネットワーク構成は、ルータ、スイッチ、異なる線種の伝送路などにより、巨大で複雑になってきた。このような状況下で複数のインターネット通信事業者

(ISP)をまたいだ通信状態管理は非常に難しくなってきた。特に、通信障害対応は、インターネットが重要な通信路となり、研究だけでなく、あらゆる社会活動の伝送にも利用されるようになってきたため、特に短時間での対応が要求されてきている。

文献 1) と 2) では SPRINTTM の Internet Protocol (IP) 基幹部で発生した障害についての報告が行われている。しかし、文献 1) と 2) で記述されている障害は規模が小さく、発生期間も短く、文献の記述も単一 ISP に限定されている。

一般的にインターネットの通信障害は、OSI の階層モデルにおける各階層で発生する。各層における近年の代表的技術の Wavelength Division Multiplexing (WDM), Multi Protocol Label Switching (MPLS) あるいは IP³⁾⁻⁷⁾ が障害発生時の調査対象となる。

各階層で発生する障害に関して、いくつかの手法が提案されてはいるが、現実の運用においては、必ずしもそれらの技術が利用されているわけではない。

現実のネットワークで、提案された手法が利用されない理由の 1 つは、実際のネットワークが研究室で想定しているネットワークよりもはるかに複雑なためである。このため、運用者にとっては、障害が、どのような影響を及ぼすのかを正確には予測できず、実際に障害が発生しても、自分の監視区域での解析が精一杯でそれ以上の分析はできていない。地震、津波などの自然災害は、インターネット通信にも重大な影響を与える。しかも、災害自体の範囲が広範なため、複数の ISP に同時に障害を与える可能性が高い。

本報告では、2006 年 12 月に発生した台湾南方沖地震後に発生した大規模な通信障害に関して報告する。この地震では、海底で発生した地滑りにより、多数のケーブルが同時に切断され、それにともない、通信障害が発生した。この復旧手段として、いくつかの実際に行われた手法に関して述べる。Border Gateway Protocol (BGP)⁸⁾ により自動的に行われた通信経路の変更、BGP 運営方針の変更、スイッチでのケーブルの接続の変更などが実際には行われた。本報告では、大規模な通信障害が、自然災害という予測不能な状態で発生し、その際に、すでに実用化されている自動復旧手法と、その手法では対応できない部分への実際に行われた対応などを述べることで、今後も発生するであろう自然災害による通信障害に関しての対応の手法を示すものである。

21 世紀になってからも、たくさんの災害が発生したが、ネットワーク基幹部が同時に広範囲で使用不能になることはなかった。2006 年 12 月に台湾南方沖で

発生した地震はアジア太平洋地域の学術ネットワークを使用不能とした。日本時間 2006 年 12 月 26 日 21 時 26 分頃、台湾南方沖で大規模な海底地震が発生した。地震の規模を表すマグニチュードは 6.7⁹⁾ であった。この地震は海底で大規模な地滑りを発生させた。このため、この地域に敷設されていた海底ケーブルがすべて損傷を受けた。このケーブルは複数の通信事業者あるいは ISP により共同利用されていたため、多くのインターネットサービスが影響を受けた。この地震は、学術インターネットに関しては、この地域を境に西と東に完全に分離した。この分離された境界部分どうしの通信は、地球を 1 周した経路でしか通信でなくなった。損傷を受けたアジア太平洋地域の学術ネットワークの復旧には、以下に述べるような複数の解法とその修正という段階をふまなければならなかった。その解法には、BGP による自動的な経路修正と、実際の状態にそった経路への修正、さらには、ケーブル接続のスイッチでの変更などがあった。

地震直後、アジア太平洋地域の多くの通信は不通にはならなかった。これは、BGP により、多くのルータで経路が冗長化され、複数の迂回路がすでに用意されていたからであった。通常、BGP による経路制御においては、運営方針に沿った複数の経路が用意されている。したがって、接続断とともに迂回路を用意することができた。しかし、今回に限っては、迂回路はあまり品質の良いものではなかった。たとえば、迂回路の周回遅延 (RTT) が極端に長いなどである。また、用意されていた迂回路が主経路に対して必ずしも帯域が十分ではなく、同時に、緊急時のため、トラフィックが集中し、輻輳が発生した。このため、運用者は手動により通信経路を変更し、品質を向上させなければならなかった。

次に行われたことは、代表的な通信者間の経路の精査であった。BGP による経路制御では、運営方針に沿った、経路制御表を各ルータが持っている。表の上位に書かれている経路ほど、運営方針により合致したものとなっている。この表をもとに、外部に対してのアクセスの経路が BGP で広報されるのであるが、学術ネットワークなどの場合、下位の経路は単なる経路の候補であることが多く、その経路を利用したときの影響は精査されていないことが多かった。このため、特定の回線にトラフィックが集中するようなことが発生した。運用者は、このような事態を避けるため、BGP により広報する経路、受信する経路に手動で制御をかけることを行った。まさに、手動によるトラフィック・エンジニアリングを行うこととなった。

商用 ISP は通信断をさけるため、複数の上流 ISP と契約し、経路を確保していることが多いが、学術ネットワークの場合は、その性格から、複数のケーブルを購入するなどの物理的接続の冗長化は行われていないことが多い。つまり、通信事業者やケーブル会社が、1つの契約回線に対して複数のケーブルを用意し、冗長性を実現しているのが、学術ネットワークの運営で物理的冗長性は考えなくてよかった。ところが、この地震では、すべてのケーブル経路が失われたことから、このような“単一接続”学術ネットワークはその疎通性を、商用 ISP 以外完全に失われた。運営方針の一時的変更、および、各 NOC でのスイッチのケーブル割当てを変更することで、迂回路を作れることが多かった。そのため、スイッチの設定変更、ケーブルの接続変更も行われた。

台湾南方沖地震によるケーブル断から、このような大規模な通信障害に対しては、学術ネットワークにおいては、RTT と通信可能最大帯域を考慮した迂回経路の選択機構が必要であることが分かった。これは、通信において最大効率を目指すトラフィック・エンジニアリングであり、この機構が次世代インターネットサービスを目指す学術ネットワークに必要なものであることが分かった。

本報告は以下のような構成になっている。2章では、アジア太平洋地域の学術ネットワークの活動を紹介する。3章では、通信障害の経過を詳細に記述する。4章では、通信障害復旧の経過を記述する。5章では、通信の障害および復旧において、得られた知見について論じる。6章では、本報告の結論を述べる。

なお、記述は、個別の経路制御情報の変化ごとに記述している。

図1は時系列に展開した場合の図、表、本文の記述の位置を示している。

2. アジア太平洋地域の学術ネットワーク

2.1 Asia Internet Interconnection Initiative (AIII)¹⁰⁾

AIIIは1996年にアジア地域で始まった最初の次世代インターネット学術ネットワークプロジェクトである。AIIIは、ケーブルの敷設を行わず、人工衛星を利用し、インターネットサービスを構築することを目指した。AIIIのメンバは、TH, MY, ID, SG^{*1}であったが、2006年にNPが新たに加わった。AIIIは、衛星を利用するため、帯域の拡張に限界がある。利用可能

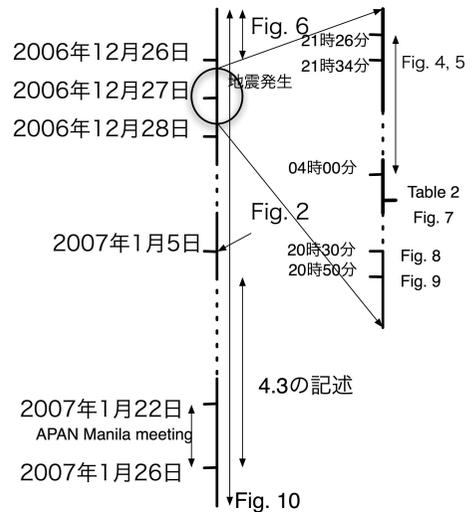


図1 本報告中の図、表、記述の時系列上での位置
Fig.1 Time series of the figures, the tables and the description of this paper.

帯域は、1.5 Mbps から 8 Mbps の間であるが、最近、学術ネットワークなどでさかに行われる実験には必ずしも対応できなくなっている。しかし、AIIIは、先端的通信方式、たとえば、IPv6 unicast および multicast, Uni Directional Link Routing (UDLR), あるいは、次世代 TCP などの研究活動を行っている。

2.2 Asia-Pacific Advanced Network (APAN)¹²⁾

APANは1997年に開始されたアジア太平洋地域の次世代インターネットが集合した研究を目的とした組織である。NOCチームは、JGN2, SINET, KOREN, SingAREN, CERNET, CSTNET, AARNET, TransPAC2, TEIN2などの学術ネットワークの相互接続を提供し、これらの中で、インターネット通信の研究活動が行われている。NOCチームは、アジア太平洋地域の研究開発活動に対して、データの、ワイヤレートに対しての、高効率伝送サービスを提供するような運用を行っている。東京には、APAN Tokyo XPがおかれ、アジア太平洋地域の学術ネットワークの経路交換の中心となっている。

2.3 Trans-Eurasian Information Network 2 (TEIN2)¹³⁾

TEINまたはTEIN1はEUのプロジェクトで、EUとKRを接続した。プロジェクト開始時のTEINの帯域は10Mbpsであった。TEIN2は45Mbpsから1Gbpsの範囲の帯域で構成されたネットワークで、東南アジアの相互接続性を提供すると同時にヨーロッパへのアクセスネットワークとなっており、ヨーロッパ

*1 本報告では、原則として ISO 3166 codes¹¹⁾ を用いる。

表 1 アジア太平洋地域の学術ネットワーク
Table 1 R&E networks in Asia Paicif area.

ネットワーク名	詳述	地域
AIII	Asian Internet Interconnection Initiatives	JP, VN, PH, TH, MY, SG, ID, NP
APAN	Asia-Pacific Advanced Network	JP, KR, CN, TW, HK, VN, TH, MY, SG, AU, NZ, PK, IN, KL, BD
TEIN2	Trans-Eurasia Information Network 2	JP, CN, VN, PH, TH, MY, SG, ID, AU, EU
ASTI	Advanced Science and Technology Institute	PH
CSTNET	China Science & Technology Network	CN
CERNET	China Education and Research Network	CN
GÉANT	AS 登録名	Europe
GÉANT2	GÉANT の後継	Europe
KOREN	Korea Advanced Research Network	KR
SingAREN	Singapore Advanced Research and Education Network	SG
UniNet	Thai Social/Scientific Academic and Research Network	TH
ThaiSARN	Thai Social/Scientific Academic and Research Network	TH
KREONET2	Korea Research Environment Open Network 2 国内ネットワークは KREONET	KR
AARNET	Australia's Research and Education Network	AU
QGPOP	九州 Gigapop	JP
JGN2		JP
SINET		JP
TransPAC2		US
ASNET	TransPAC の後継	TW

の学術ネットワーク GÉANT2 と接続している。

2.4 アジア太平洋地域の学術ネットワーク

表 1 にアジア太平洋地域の学術ネットワークまとめる。

3. 台湾地震による通信障害

3.1 ケーブル切断

日本時間 2006 年 12 月 26 日、強い地震が少なくとも 2 回以上台湾南方沖で発生した。最初の地震は、日本時間で 21 時 26 分⁹⁾、2 回目は 21 時 34 分¹⁴⁾ に発生した。幸いなことに、地震が海底で発生したことで、近くの地盤がきわめて強固であったため、1999 年の台湾中部地震に比べると、死者が出てしまったものの、都市部での被害は、ほとんどなかったといってもよかった。

しかし、2 回の地震はアジアでも有数の海底ケーブル経路の集中する場所で発生し、地震直後の小規模な地滑りと地震発生後 6 時間程経って発生した大規模で広範囲な地滑りによりこれら海底ケーブルに対して限界以上の張力がかかり、断絶した。

日本時間 12 月 27 日午前 4 時頃、アジア地域の学術ネットワークの疎通がいったいに失われた。ケーブル会社および通信事業者の調査の結果、ケーブルシス



図 2 台湾近海のケーブル切断地点¹⁵⁾

Fig. 2 Fiber cut area around Taiwan Island.

テムが破壊されたことが確認された。これが、午前 10 時頃であった。

図 2 の円は今回ケーブルの切断の発生した場所を表している。東アジアのほとんどのケーブルシステムは必ず台湾東南沖を経由している。これは歴史的理由による。さらに、近年の経済活動から、これらのケーブルシステムは複数の通信事業者などにより共同利用されることが一般的となっていた。

3.2 インターネット疎通消失と BGP peering の消失

地震発生後、インターネット通信は商用、学術用ともに疎通を失った。疎通を失った学術ネットワークを以下に列挙すると、MAFFIN (JP, PH 間)、TEIN2 (JP, CN 間)、TEIN2 (JP, SG 間)、CSTNET (CN, US 間)、TEIN2 (HK, KR 間)、ASNET (HK 経由での TW, CN 間)、ASNET (TW, SG 間) などである。これにより、一時的にせよ、香港が、学術および商用の両方のネットワークサービスから孤立するという事態まで発生した。そして、学術ネットワークは、2 つのグループに分かれてしまった。1 つは、JP, KR, TW, US で、もう 1 つのグループは CN, HK, VN, MY, TH, SG, ID, PH であった。インターネット疎通の消失は以下のような順で発生した。

- (1) ケーブルの切断により、リンクレイヤの断絶が発生した。
- (2) 設定順番の第 1 番目の経路での BGP ピアリングが消失した。
- (3) 複数の経路が設定されていた場合、自動的に他の通信可能な経路での経路交換が開始された。

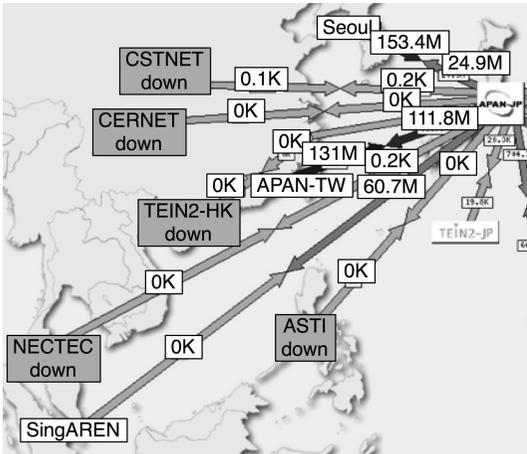


図3 日本時間 2007 年 1 月 5 日 15 時 6 分 14 秒における東京とアジア地域間のトラフィック

Fig.3 Traffic weather map at 15H06M14S (+0900) on January 5th, 2007.

図3は APAN Tokyo XP と国外の間のトラフィックをほぼリアルタイムに表示する“トラフィック・ウェザーマップ”である。このツールは、東京の境界のルータから国外の境界ルータまでの伝送容量と、その逆向きの伝送容量を、数字で表示する。同時に、上記区間をそれぞれ塗り潰された矢印で表示する。矢印の色はトラフィック量が少なきときは青系統の色。トラフィック量が増えるに従って、緑さらに赤へと滑かに色が割り当てられている。

図3は日本時間 2007 年 1 月 5 日 15 時 6 分 14 秒の様子を表している。この時刻は台湾南方沖で地震が発生してから、9 日以上が経過した状態である。このとき、JP と CN の間は、CSTNET および CERNET とともにほぼ 0 Kbps ということから、断絶していることが分かる。また、JP と NECTEC (TH), JP と ASTI (PH) の間も、0 Kbps ということから、通信が断絶していることが分かる。一方、JP と Seoul (KR) の間は、KR から JP 向きのトラフィック量が、153.4 Mbps になっている。3.4 節で記述するとおり、この時点で、KR からは、TW 向きのトラフィックが、CN より流入し、この伝送容量となっている。

リンクレイヤの断絶は BGP セッションの timeout を発生させる。JP と HK との間、HK 経由の JP と CN との間、JP と TH との間、JP と SG との間、JP と PH との間の最短最適経路での BGP ピアリングが失われたため、これらは、可能な限り、迂回経路でのピアリングと置き換わっていった。

一般に、トラフィックが迂回経路に転送された場合、トラフィックが通過する Autonomous System (AS)⁸⁾

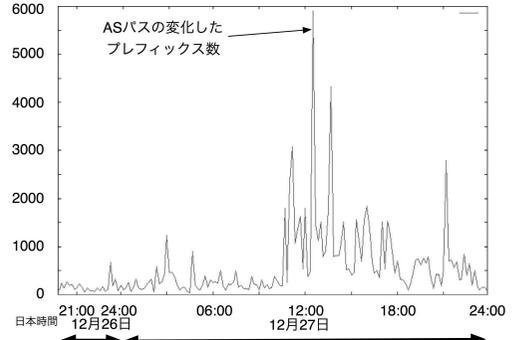


図4 QGPOP における AS パスの変更があった IP プレフィックス数 (2006 年 12 月 26 日から 27 日まで)

Fig.4 Number of IP prefixes that experienced AS path changes (2006.12.26-2006.12.27 at QGPOP).

番号を連続して並べた AS パス長は通常のものより長くなる。なぜなら、特別な設定のない限り、BGP 経路制御においては、最短 AS パス長経路を選ぶようになっているためである。

図4は九州の QGPOP の境界ルータにおける BGP 経路表を 10 分ごとに収集し AS パス長の変化した AS に含まれる IP アドレス群 (IP プレフィックス) 数の様子のうち、日本時間で 2006 年 12 月 26 日と 12 月 27 日の様子を、時系列でグラフ化したものである。

図4より、1,000 以上の IP プレフィックスが、海底ケーブルの切断直後にその通過経路を変化させたことが分かる。

3.3 トラフィック量の変化

CN と KR の間の回線および JP と KR の間の回線は無傷であった。そのため、BGP による経路変更による迂回したトラフィックがこれらの回線に転送されはじめていた。図5と図6は APAN Tokyo XP の国外との境界ルータの伝送容量を、MRTG¹⁶⁾ を利用して、観測および表示したものである。

図5は日本時間での 2006 年 12 月 27 日の JP と KR の間のトラフィックを示したものである。日本時間午前 4 時 30 頃、KR から JP へ向いたトラフィックが大きく変化したことが分かる。一方、図6より、ほぼ同時刻に JP と PH の通信が消失していることが分かる。

図5において KR から JP へのトラフィックが大きく変わったのに対して、JP から KR へのトラフィックが変わっていないのは、経路広報の送受信の運営方針の違いによる。つまり、APAN-KR NOC は、APAN Tokyo XP の広報している経路情報をほぼそのまま外部に広報しているため、CN トラフィックが CN と KR の間の回線を通して伝送される。一方、APAN Tokyo

表 2 2006 年 12 月 27 日に観測された学術ネットワークの経路表¹⁷⁾

Table 2 The monitored routing table for each R&E network on Dec. 27 2006.

宛先	APAN -JP (JP)	CER- NET (CN)	APAN -TW (TW)	KOREN (KR)	ASTI (PH)	CST- NET (CN)	TEIN2	Sing- AREN (SG)	UniNet (TH)	Thai- SARN (TH)	KREO- NET2 (KR)	AAR- NET (AU)
送信元												
APAN -JP		D	直収	直収	U	C	U	直収	D	U	US	US Hawaii
CER- NET	D		D	直収	U	N.A.	TEIN2	TEIN2	TEIN2	TEIN2	TEIN2	TEIN2
APAN -TW	直収	D		APAN -JP	U	U	N.A.	N.A.	N.A.	N.A.	APAN -JP	APAN -JP
KOREN	直収	直収	APAN -JP		U	U	TEIN2	TEIN2	TEIN2	TEIN2	直収	TEIN2
ASTI	U	U	U	U		U	U	U	U	U	U	U
CST- NET	C	N.A.	U	U	U		N.A.	N.A.	N.A.	N.A.	U	N.A.
TEIN2	U	TEIN2	N.A.	TEIN2	U	N.A.		TEIN2	TEIN2	TEIN2	KOREN	TEIN2

D : US を経由した迂回経路による通信
 C : 商用ネットワーク経由での通信
 U : 疎通性を失った状態
 N.A. : 元来通信がない状態

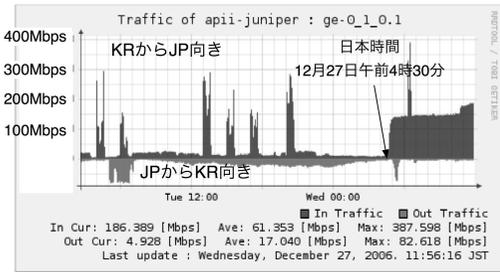


図 5 日本時間 2006 年 12 月 27 日の JP と KR のトラフィック
 Fig. 5 Traffic between JP and KR on December 27th (UTC+9) 2006.

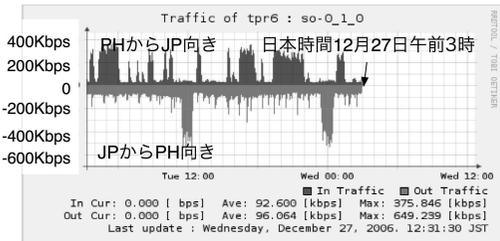


図 6 日本時間 2006 年 12 月 27 日の JP と PH のトラフィック
 Fig. 6 Traffic between JP and PH on December 27th (UTC+9) 2006.

XP から CN 向きには直接回線が存在し, APAN-KR から広報された経路を経路表に加えた場合, CN へのアクセスでは, APAN Tokyo XP には 2 つの経路, つまり, KR 経由の CN への経路と, APAN Tokyo XP から CN へ直接行く経路の 2 つが登録され, どちらの経路を選択するかは AS パス長では判断できないことが発生することから, APAN Tokyo XP では, APAN-KR からの CN の経路を受け取らないというポリシーによるものであった.

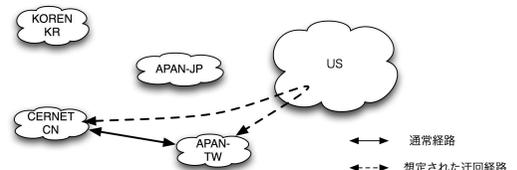


図 7 APAN における地震発生前の経路と想定されていた迂回経路
 Fig. 7 BGP route changes of APAN before earthquake (Dec. 2006).

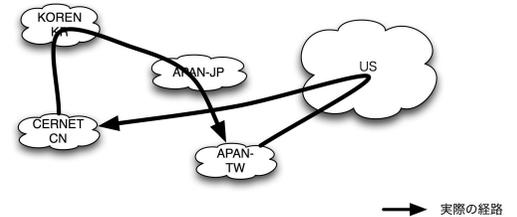


図 8 地震発生後での経路の変化
 Fig. 8 BGP route changes of APAN after earthquake (Dec. 2006).

3.4 BGP 経路表の変更

表 2 はアジア太平洋地域の地震直後の経路制御表である. 表 2 では, ASTI (PH) および CSTNET (CN) が学術ネットワークからの経路を完全に失っていることを灰色で示している. また, APAN-JP (JP) が TEIN2 にまったく通信できなくなっていることも分かる.

また, CERNET (CN) は TW への直接接続経路があり, 図 7 に示す通信が行われていた. また, 万一, この経路に障害が発生した場合は, 図 7 の点線に示す経路が選択されるものと推測されていた. ところが, 図 8 に示すように, KOREN (KR) の経路広報の方針により, TW から CN 向きは US 経由になったが,

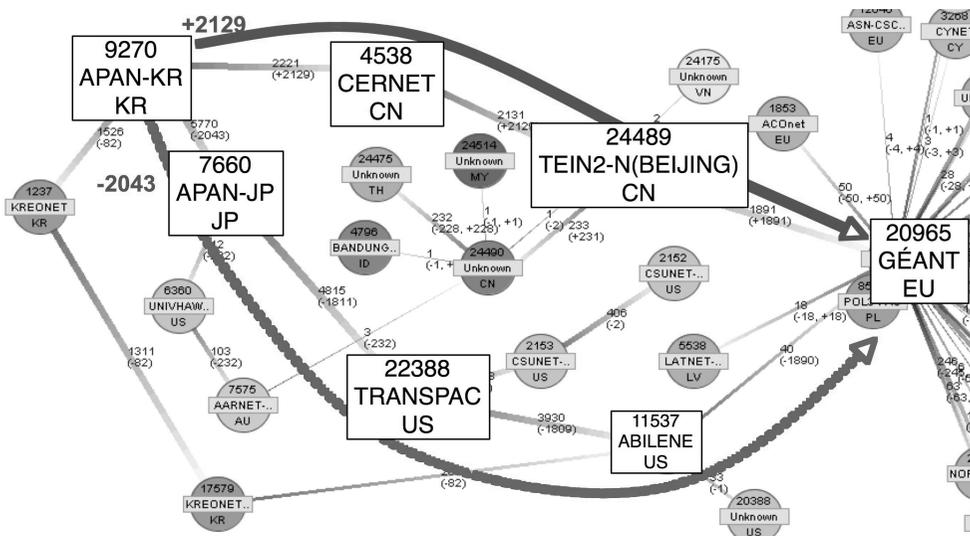


図 9 日本時間 2006 年 12 月 27 日 20 時 30 分における APAN-KR から GÉANT へ修正された経路
 Fig. 9 Reconfigured AS path from APAN-KR to GÉANT (APAN-KR, CERNET, GÉANT) at 20:30, Dec. 27 (UTC+9) 2006.

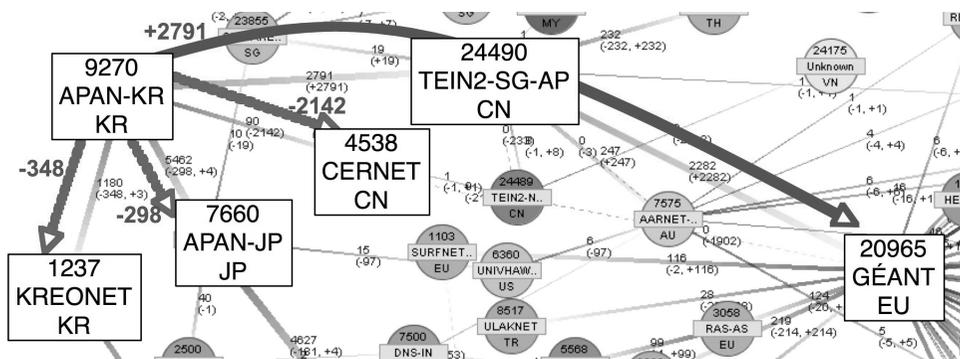


図 10 日本時間 2006 年 12 月 27 日 20 時 50 分における APAN-KR から GÉANT へ修正された経路
 Fig. 10 Reconfigured AS path from APAN-KR to GÉANT through TEIN2-SG at 20:50, Dec. 27 (UTC+9) 2006.

CN から TW 向きは、JP 経由となった。同様に、JP から東南アジアへのアクセスはすべて US 経由となっている。CSTNET (CN) は US への直接回線を持っているが、この回線も、台湾南方沖を通るケーブルを使っていたため、疎通を失った。

3.5 AS レベルの接続形態の変化

BGP による経路の変化を詳細に調べるため、「ABEL2」¹⁸⁾を用いた。データは、QGPOP の境界ルータの BGP 経路表を 10 分ごとに収集をしたもので、各図は 10 分前との経路表の変化を表している。図中の円形領域は AS を表している。AS どうしを接続する線には、観測時のその経路通過プレフィックス数とその増減が示されている。

図 9、図 10 は APAN-KR (AS9270) と GÉANT (AS20965) の間の経路変化を IP プレフィックス数の変化とともに示している。APAN-KR から GÉANT (EU) へは、TEIN2 経路が通常の経路であった。APAN-KR から TEIN2 へのアクセスは、SG へのケーブルを利用していたが、このケーブルが台湾南方沖地震で失われた。APAN-KR がこの状態で EU へのアクセスで利用可能な経路は、JP と US を経由する経路であった。ただし、距離は太平洋、大西洋を超えてのアクセスであるので、遅延が TCP の限界を超えるおそれがあった。このため、関係するネットワークの運用者は、AS パス長を短くするため、BGP のポリシーの変更を行った。つまり、CERNET から APAN-KR に

対して GÉANT の経路情報を広報してもらうこととした。これにより、図 9 に示すように、APAN-KR から GÉANT への経路は CERNET 経由となった。図 9 は日本時間 12 月 27 日 20 時 30 分の様子である。このため、JP と US 経由のプレフィクス数が 2,043 減少し、CERNET 経由での EU 向きの経路のプレフィクス数が 2,129 増加した。

日本時間 20 時 50 分、TEIN2-SG NOC が GÉANT の経路情報の広報を開始した。また、この時点で、KR と SG との間の回線は復旧していた。もちろん、本来のケーブルが復旧したわけではなく、通信事業者が迂回路線を利用したことによって、疎通が復旧しただけのことであった。SG と KR の間の回線帯域は CN と KR の間のそれより大きいことから、KR 運用者が SG から広報される GÉANT の経路情報を優先したことで、KR と SG との間の回線を經由するようになった。図 10 ではその様子を示していて、EU へのアクセスに、CERNET、APAN-JP、KREONET からそれぞれ 2,142、298、348 のプレフィクス数が減少し、TEIN2-SG-AP 向きプレフィクス数が 2,791 増加している。

4. ネットワーク障害復旧手法

地震によるネットワーク障害後、いくつかの手法による通信の復旧が試みられた。それは、とりあえずの通信不能を避けるものであったり、正常だったときに近い効率でデータも伝送できるようにする程度の復旧作業のものであったりした。本章では復旧作業を作業の順番により分類する。

4.1 BGP による自動経路変更

通常、BGP 経路表にある経路の中には、“役に立たない”経路も含まれている。ここで、“役に立たない”とは、その経路が、単なる疎通性であって、RTT が長かったり、十分な帯域がなかったりするという意味である。そのため、ネットワーク運用者は、このような“役に立たない”経路を local preference⁸⁾ の設定、あるいは、経路を受け取らないなどで、除外していた。しかしながら、この台湾南方沖地震は、TW より東の地域と西の地域の多くの経路を使用不能にしたことから、このような“役に立たない”経路がバックアップ経路として有効であった。アジア太平洋の学術ネットワークは、TEIN2 起動後、経路制御が非常に複雑になった。TEIN2 の複雑な経路がときとして、世界中に予想もしない経路を広報したこともあった。しかし、この TEIN2 の複雑な経路により今回の地震では早い時期から疎通性だけは回復させることができた。つま

り、バックアップの AS パスにより、通常使用の経路による疎通性を失った宛先に対して応急措置の疎通性確保を行うことができた。しかし、この BGP による自動経路変更は、運用方針に沿って構築された経路制御表に行われるので、その経路の RTT などのネットワーク的な距離や経路途中の回線の混雑具合は考慮されていない。

4.2 トラフィックエンジニアリングによる BGP 経路制御方針変更

BGP そのものは回線の契約帯域や使用可能帯域という情報を提供しない。さらに、近年、国際間でも VLAN が使われるようになったことから、隣接している AS どうしが実際の距離も近いわけではなくてきている。したがって、今回の事態において、迂回路の QoS 情報は運用者が自分で調べるしかなかった。たとえば、迂回経路が、RTT が長く、輻輳が発生しているなどの通信障害に対して、BGP 経路制御方針を一時的に変更することが必要となった。この様子は図 9、図 10 に示されている。

TEIN2 メンバ (VN, MY, SG, ID, PH) は、ケーブル切断により、APAN Tokyo XP への疎通性を失った。AARNET NOC は TEIN2 メンバの APAN Tokyo XP へのアクセス経路として、AU とハワイを通る経路を提案してきた。この経路は、APAN Tokyo XP ではフィルタリングされていた。なぜなら、実際の距離がとても長く、高効率データ伝送にはまったく向いていなかったためである。この経路を利用することで、TEIN2 メンバの APAN Tokyo XP に対しての疎通は回復したものの、東京とハワイの間の回線が輻輳を起こす事態となった。なぜなら、TEIN2 と東京の間の帯域は通常、622 Mbps あり、東京とハワイの間の帯域は 155 Mbps のためであった。

APAN Tokyo XP は、この時点での、すべての東京にやってくるトラフィックの解析を行った。これには、Netflow¹⁹⁾ による flow データ収集することで行った。

図 11 は Arbor Networks Peakflow SP を用いて、APAN Tokyo XP の境界ルータの flow data を収集し、KR との通信に使われるインタフェースの flow data のみ取り出し、宛先 AS あるいは送信元 AS によって表したものである。

KR から九州経由でやってくる通信については、増加分 (図 5) は、CN からのものであり、さらに解析を行ったところ、TW へ向けてのものであることが分かった。図 11 では、網のかかっている部分、つまり、ERX-CERNET-BKB、TEIN2-SG-AP、TEN2-NORTH-AP、AARNET-AS-AP がそれにあたる。

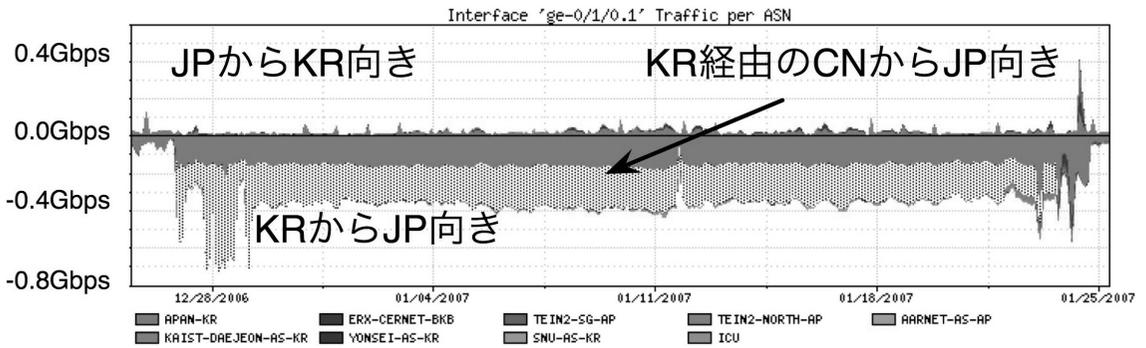


図 11 12月26日から1月26日までのKRからのトラフィック
 Fig. 11 Traffic from KR (December 26–January 26).

JP と KR の間の回線の帯域は 10 Gbps であったことから、すべての TEIN2 トラフィックを KR 経由にする経路制御ポリシー変更を各 NOC に伝え、協力を仰いだ。その結果、東京とハワイの間の回線の輻輳は解消したが、KR と CN の間の回線 (155 Mbps) が輻輳した。そこで、JGN2 の JP と SG の間の回線にさらに協力を依頼し、CN 向け以外のトラフィックを TEIN2-SG NOC から APAN Tokyo XP へ迂回するよう依頼した。これにより、KR-CN の輻輳の回避と、TEIN2 利用者の長い RTT による障害を解消することができた。

4.3 ポートの構成

一部の地域は、学術ネットワークへの接続が 1 回線だけであった。このため、ケーブルが切断されたのち、これらの地域は、商用インターネットサービスを利用して他の学術ネットワークと通信せざるをえなかった。

学術ネットワークの中には境界ルータなどをキャリアホテルと呼ばれるハウジングサービスを利用して設置していることがある。このような施設には他の学術ネットワークが同居していることもあった。さらに、同居している他の学術ネットワークが動作しているケーブルシステムに接続しているという事態もあった。台湾南方沖地震の場合、このような場合において、双方の学術ネットワークが合意のうえで、局内配線により、商用サービス以外のすべてのケーブルシステムが断絶した学術ネットワークを、一時的に他の学術ネットワークの配下にする事で、商用サービスよりは経路や帯域が制御しやすい位置へ変更することも行った。この作業により、接続経路が大幅に変更され、AS パスもまったく異なるものになってしまったものの、学術ネットワークへの直接接続が実現できた。

地震発生 2 日後、KR と SG の間の回線が復旧し、SG は TEIN2 メンバ (ID, SG, TH, MY) の経路

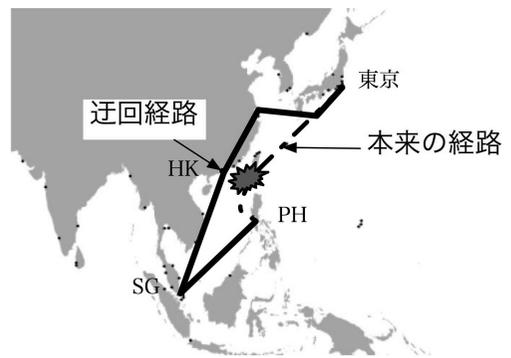


図 12 JP-PH 迂回経路
 Fig. 12 JP-PH detour route circuit.

情報を KR に広報を開始した。ネットワーク運用者が調べたところ、この復旧した SG と KR の間の回線は元のケーブルではなく、迂回路によるものである。これは、RTT が本来のものよりかなり長くなっていたため分かった。同様に、JP と PH の間の回線も迂回経路により、RTT が極端に長くなったものの復旧した。これは、1月22日から26日まで、APAN マニラミーティングが計画されていて、その会議では広帯域を利用したデモも予定されていたため、このような応急措置がとられた。

実際、回線の経路は図 12 に示すように中国大陸内を縦断する経路になっていた。

5. 知 見

地震によるネットワーク障害が、商用 ISP および学術ネットワークで同時に発生した、これを復旧させる手段がネットワーク運用において手探りで行われた。ここでは、回線復旧の過程で得られた知見について述べる。

5.1 障害に対応したケーブル経路の設計

回線利用者や NOC 関係者は一般にケーブル経路には無関心であることが多い。しかし、今回のようなケーブル障害などが発生した場合、障害に対応したケーブル経路設計が重要である。インターネット運用者は、一般には、ケーブルを購入するだけのため、各回線が通常はどの経路を通過しているのかをあらかじめ調査しておくことは、重要である。

5.2 ケーブル障害に対してトラフィックエンジニアリング対応の BGP 経路制御の準備

通常、BGP 経路制御ポリシーは、複数の経路に対して、希望した経路を利用してデータが伝送されることを目指している。そして、通信に利用可能な経路に対して、望ましい順番が用意されている。通常は、AS パス長が最短のものが望ましい経路であるが、AS パス長が同じものが複数ある場合や、AS パス長が長くても、線路長で最短の経路が、望ましい経路であることもあり、それを最初を選択するような重み付けが行われることもある。この際、望ましい順番に用意された経路の中で、下位にならぶ経路は、疎通性を提供する程度だけのこともある。台湾南方沖地震以前、アジア太平洋地域の学術ネットワーク運用担当者の間では、そのような“役に立たない”経路そのものを経路表から削除することさえ行われていた。これは、TEIN2 の起動後、アジア太平洋地域の学術ネットワークの経路が複雑になり、“役に立たない”経路が、AS パス長が見た目短かったり、あるいは、IP プレフィックスの最長一致から、“望ましくない”経路を選択する事象が多数発生したためであった。しかし、地震直後、このような除外していた“役に立たない”経路広報を再度受信することで疎通性を保つことができた。このことから、複雑化した経路は、従来の BGP における経路選択基準の以外に経路情報を整理して保持する基準が必要であることを示している。

このように冗長性を維持した経路制御による経路変更は疎通性を保つという点では成功した。しかし、通信データ容量を維持したり、経路の突然の変更による輻輳の発生を予測したりできていない。なぜなら、BGP においては、回線帯域、回線使用率、あるいは、利用可能帯域という QoS に関わる情報はまったくないので、このような緊急時に変更された経路においては、迂回を繰り返すことによる長遅延が発生することが多い。したがって、BGP 経路制御においては、あわせて、QoS パラメータの検証、トラフィックエンジニアリングによるデータ伝送傾向の検討が特に今回のような事態においては必要であることが分かった。

5.3 ネットワーク運用データの常備

回線復旧において、様々なネットワーク監視ツールを用いた。MRTG¹⁶⁾、ネットワーク・ウェザー・マップ²⁰⁾、BGP 経路表示ツール¹⁷⁾、フロー監視ツールなどである。最初の回線断は、ネットワーク・ウェザー・マップと MRTG における突然の変化によって、把握された。通信復旧作業においては、上記監視ツールによって得られているデータとすでに持っている回線容量や接続地点の知識を総動員する必要があった。そして、通信復旧のための手段を決定していった。iperf²¹⁾ や bwctl²²⁾ を用いることができれば、通信を行う両端での使用可能帯域を推定できる。

例として 4.2 節でのべた迂回路について述べる。東京とシドニーの間で通信する場合、通常、2 つの経路が想定される。1 つは、東京、シアトル、シドニーの経路で、もう 1 つが、東京、ホノルル、シドニーの経路である。前者は、経路すべてにおいて、10 Gbps の帯域を利用することが可能であるが、線路長が長いいため、RTT が長くなる。後者経路は、東京とホノルルの間が 155 Mbps というボトルネックが存在するが、RTT は短い。地震による迂回として AU 経由の経路を用意する場合、上記 2 経路の 1 つを選択するという最終決定をする前に、さらにフローデータを調べた。MRTG¹⁶⁾ や RRDTTool²³⁾ は通信量は観測できても、通信の送信元 AS と宛先 AS ペアを知ることはできない。TEIN2 の通信の多くは US 向きではなく、アジア域内の通信ということが、フローデータの解析から分かり、ボトルネックがあり、輻輳の発生可能性があるにもかかわらず、ハワイ経由の経路を選択したのはこのためであった。また、地震直後に KR から APAN Tokyo XP 向きの通信が急に増加したときもフローデータの解析から理由が判明した。

以上から、ネットワーク監視データや運用データ常備しておくことは、重大な通信障害が発生した場合、それぞれの監視データを収集し、運用データと比べることが、適格な通信復旧にきわめて有効であることが分かった。

5.4 緊急時の運用者間の情報交換

地震はケーブルに損傷を与え、学術ネットワークだけでなく、商用インターネットサービスにも影響をあたえた。特に地震発生直後は、通常の電話にも影響をあたえたが、これは、早急に復旧した。一方、インターネットなどの新しいサービスへの影響は甚大で、VoIP サービスは帯域不足のせいで会話不能の状態であった。さらに、12 月 26 日がクリスマス休暇明けということもあり、あまり多くの運用者が現場にいないという悪

条件も重なった。このため、運用者同士の情報交換は、インスタントメッセージシステムや e-mail が使われた。インスタントメッセージシステムにおいても、帯域不足、遅延のせいで、テキストによる情報交換しかなかった状態であった。しかし、この品質の情報交換で上記の作業を行うことができた。今後、このような広範囲での通信障害が発生する事態が起こるかどうかは分からないが、NOC 間の緊急連絡体制と利用可能帯域に見あった情報共有システムの構築が必要であることが分かった。

6. ま と め

インターネットが発展し、様々な分野において、重要な情報伝送路となった現在、障害対応体制はネットワークの運用上重要な緊急に達成されなければならない事態となった。2006 年 12 月に台湾南方沖で発生した地震によって、その体制が単一ケーブルシステムあるいは単一回線によるもので、広域に発生した場合の影響についてまったく考慮していなかったことが明らかになった。

いくつものケーブルシステムにまたがったバックアップシステムを作っても、広域災害においては意味をなさなかった。しかし、東アジアおよび東南アジア地域という狭い範囲で TEIN2 のような、2 地点間の通信に複数の経路が存在するようなネットワークを構築した場合、いくつかの被害にあわなかった経路を組み合わせることで容易に疎通性を確保できることも分かった。一方、この際の組合せについて、緊急時であることから、ある回線部分にトラフィックが集中する可能性があることからトラフィックエンジニアリングによる運用範囲内での 2 地点間の伝送効率監視システムが通常の NOC で行うトラフィック監視システムのほかに必要であり、それによる BGP の運営方針設定を行う必要があることが分かった。

今回の広範囲の地域とネットワークに影響を及ぼした自然災害による通信障害からの復旧作業を行った結果、ネットワーク管理および緊急事態対処について貴重な経験を得ることができた。今後、回線設計において、インターネットは、回線の経路だけでなく、数ホップ先の AS における帯域使用状況を把握するようなトラフィックエンジニアリングの手法を導入した運用補助システム、および、シミュレータシステムの準備が必要と思われた。

謝辞 地震発生直後より、KDDI OTC(大手町テックニカルセンター)は懸命に作業を行い、APAN Tokyo XP の経路制御の混乱修復に全力をあげた。その働き

は単に業務というよりボランティアに近くその作業に敬意を表す。また、OTC ディレクタの渡辺氏は、図 5、図 6 に示す MRTG のグラフを早急にまとめ、現状把握に奔走した。AARNET NOC は TEIN2 へのアクセスの経路に貴重な助言と作業を行った。KOREN NOC は、休日であるにもかかわらず、CERNET と ASNET の通信を確保するため作業を行った。ハワイ大学 NOC はホノルルでの複雑な経路制御作業および監視を行った。KDDI はケーブルの損傷箇所地点の発見と APAN 地域の対応について助言を与えてくれた。Genkai NOC は、サービスを開始するための準備段階であった回線の運用を前倒して開始し、CERNET と ASNET の通信を確保した。アジア太平洋地域のすべてのネットワーク運用者の協力のおかげで、1 月中旬に疎通性が回復し、1 月末日までに、ケーブルの復旧ができたことに感謝したい。

参 考 文 献

- 1) Iannaccone, G., Chuah, C., Mortier, R., Bhattacharyya, S. and Diot, C.: Analysis of Link Failures in an IP Backbone, *ACM Internet Measurement Conference* (2002).
- 2) Markopoulou, A., Iannaccone, G., Chuah, C., Mortier, R., Bhattacharyya, S. and Diot, C.: Characterization of Failures in an IP Backbone, *IEEE Infocom* (2004).
- 3) Fumagalli, A. and Valcarengi, L.: IP Restoration vs. WDM Protection: Is There an Optimal Choice?, *IEEE Network Magazine*, Vol.14, No.6, pp.34-41 (2000).
- 4) Saharabudde, L., Ramanurthy, S. and Mukherjee, B.: Fault Management in IP-Over-WDM Networks: WDM Protection vs. IP Restoration, *IEEE Journal on Selected Areas in Communications*, Vol.20, No.1, pp.21-33 (2002).
- 5) Sharma, V. and Hellstrand, F.: Framework for Multi Protocol Label Switching (MPLS)-based Recovery, RFC 3469 (2003).
- 6) Ramamurthy, S., Saharabudde, L. and Mukherjee, B.: Survival WDM Mesh Networks, *IEEE/OSA Journal of Lightwave Technology*, Vol.21, No.4, pp.870-883 (2003).
- 7) Gerstel, O. and Ramaswami, R.: Optical Layer Survivability: A Services Perspective, *IEEE Communications Magazine*, Vol.38, No.3, pp.104-113 (2000).
- 8) Lougheed, K. and Rekhter, Y.: Border Gateway Protocol (BGP), RFC 1105 (1989).
- 9) The December 26, 2006, Pingtung, Taiwan earthquake. <http://scman.cwb.gov.tw/rtd/>

- oldeq%5C200612%5C2006122620262170106.gif
- 10) Asian Internet Interconnection Initiative. <http://www.ai3.net>
 - 11) ISO 3166 Code Country. <http://www.iso.org/iso/en/prods-services/iso3166ma/02iso-3166-code-lists/list-en1.html>
 - 12) APAN. <http://www.apan.net>
 - 13) Trans-Eurasian Information Network 2. <http://www.tein2.net>
 - 14) The December 26, 2006, Pingtung, Taiwan earthquake. <http://scman.cwb.gov.tw/rtd/oldeq%5C200612%5C2006122620341570107.gif>
 - 15) Konishi, K.: Cables restoration, *23rd APAN Manila Meeting* (2007).
 - 16) MRTG. <http://oss.oetier.ch/mrtg/>
 - 17) Kurokawa, Y.: Registration of Policy Routing Choices, *22nd APAN Singapore Meeting* (2006).
 - 18) Sakiyama, R., Okamura, K. and Lee, Y.: Visualization of Temporal Difference of BGP Routing Information, *APAN Network Research Workshop*, Singapore (2006).
 - 19) Cisco IOS NetFlow (2007). http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group%_home.html
 - 20) APAN Network Weathermap. <http://mrtg.jp.apan.net/weathermap/index.php>
 - 21) iperf. <http://dast.nlanr.net/Projects/Iperf/>
 - 22) bwctl. <http://e2epi.internet2.edu/bwctl/>
 - 23) RRDTool. <http://oss.oetiker.ch/rrdtool>

(平成 19 年 5 月 11 日受付)

(平成 19 年 11 月 6 日採録)



北村 泰一 (正会員)

1984 年早稲田大学理工学部電気工学科卒業。1986 年早稲田大学大学院電気工学専攻修了。1989 年電波研究所 (後に情報通信研究機構に改組) に入所。APAN 東京 XP のメンバで、広能率データ伝送における研究および運用を行っている。



李 泳錫

1984 年国立ソウル大学計算機工学科を卒業。1997 年国立ソウル大学大学院計算機工学専攻修了。2002 年学位取得。2002 年 10 月から 2003 年 7 月まで、カリフォルニア大学デイビス校ネットワーク研究所で招聘研究員として研究活動を行った。2003 年 7 月、国立忠南大学計算機学科教授に就任。インターネット計測技術、IPv6 およびトラフィックエンジニアリングでの研究を行っている。



崎山 亮

2003 年熊本大学卒業。2007 年九州大学大学院修了。BGP ルータの経路表の解析、計測に興味を持つ。



岡村 耕二 (正会員)

1988 年九州大学卒業。1990 年九州大学大学院情報工学専攻修了。1998 年学位取得。1998 年より九州大学計算機センターにて助教授。先進的インターネット技術およびそれを用いたアプリケーション技術に興味を持つ。