

3. ブロードバンドトラフィックレポート

マクロレベルな視点で見た、震災によるトラフィックへの影響

2011年3月に発生した東日本大震災では通信インフラも大きな被害を受けました。しかし、マクロレベルな視点で見ると、震災によるトラフィックへの影響は限定的なものでしかありませんでした。ここでは、東日本大震災によるトラフィックへの影響を報告した後、2011年5月30日から1週間のトラフィック量やポート使用量を用いて2010年からのトラフィックの変化を探っていきます。

3.1 はじめに

本レポートでは、IJJが運用しているブロードバンド接続サービスのトラフィックを分析して、その結果を報告します。昨年、IIR Vol.8で、以前からあったP2Pファイル共有からWebサービスへの移行の流れが改正著作権法、いわゆるダウンロード違法化を契機にヘビーユーザにも浸透してきたと報告しました。本レポートでは、まず、ブロードバンドトラフィックへの東日本大震災の影響について報告します。その後、これまでと同様に利用者の1日のトラフィック量やポート別使用量等を元に、この1年間のトラフィックの変化を報告します。

3.2 震災の影響について

2011年3月11日に発生した東日本大震災では、回線の切断、設備と機器の損壊等、通信インフラも大きな被害を受けました。しかし、国内インターネット全体への影響は限定的なものでした。震災直後に固定電話や携帯電話が繋がらなくなったため、インターネットが情報交換に大きな役割を果たしました。

今回のレポートでは、まず、震災によるブロードバンドトラフィックへの影響を見ていきます。図-1に、IJJが

運用するブロードバンド接続サービスでの宮城県と全国における2011年3月1ヵ月間のトラフィックを示します。ここでは、Y軸のトラフィック量が絶対値で開示できないため、ピーク値を1として正規化しています。

東北地方でのIJJの拠点である仙台データセンターの設備に被害はありませんでしたが、東京と仙台の間を結ぶ冗長構成のバックボーン回線がともに切断されました。しかし、翌日3月12日午前中にはバックボーン回線は復旧しました。この時点では、宮城県内のほとんどの地域が停電中で、仙台データセンターは自家発電で運用を行っていましたが、トラフィックはほとんどありませんでした。その後の電気と回線の復旧に伴い徐々にトラフィックが回復していき、震災発生後10日ほどで以前の85%くらいまで回復し、その後は横ばい状態になりました。東北電力からの送電、NTT東日本の通信サービスも、震災発生後10日ほどで90%近くまで急速に回復し、一部を除いた地域は、4月末くらいまでに回復しています。

全国レベルのトラフィックを見ると、震災発生直後に20%くらい減少しています。夜間には少し戻りますが、震災当日は停電に加えて首都圏で多くの人々が帰宅難民となり、トラフィックもその影響を受けました。しかし、翌日の土曜日は、前週の土曜日の85%くらいま

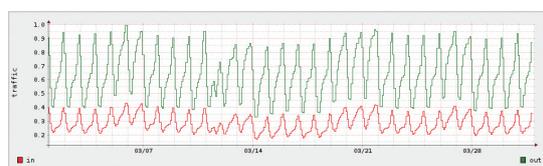
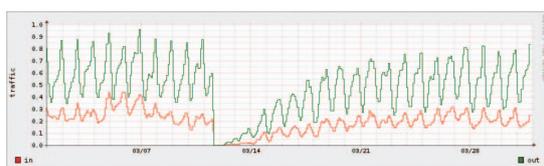


図-1 2011年3月でのブロードバンドトラフィック (左側:宮城県、右側:全国)

で回復しています。その後の1週間ほどは計画停電の影響もあって数%減となりますが、以降はほぼ元のトラフィック量に回復しています。このような状況は、IIJのブロードバンド接続サービスに限ったものではなく、他ISPも同様の状況であったと聞いています。

図-2に、過去4年間でブロードバンド全体の月平均トラフィックを示します。ここからも、震災の影響が限定的なものであったことがわかります。昨年報告した^{*1}、2010年1月に施行された改正著作権法、いわゆるダウンロード違法化のときに比べて、その影響が小さかったと言えます。また、2010年1月以降、OUT(ダウンロード)が増える一方でIN(アップロード)は横ばいであり、ファイル共有の利用が減っていることがうかがえます。ただし、トラフィックに対する影響がどの程度あるのかは不明ですが、障害に強いP2Pの特性を利用して震災情報を共有するユーザが増えたという話もあります。

トラフィックへの影響が限定的であった理由はいくつか考えられます。まず、地震と津波によって壊滅的な被害を受けた地域に主要設備がなかったため、基幹サービスへの影響が小さくて済んだことです。これは、あくまでマクロレベルな視点で見たときの影響です。震災発生から4ヵ月経つ本稿執筆時点でも、まだサービスを復旧ができていない被災地が数多く存在しています。また、3月11日以降も、度重なる余震によって回線断が多発しましたが、現場の迅速な復旧作業のおかげで、

大きな障害には至りませんでした。今回の震災の影響がこの程度でおさまっているのは、復旧に尽力された多くの方々のおかげです。

また、震災後の電力不足と節電によるトラフィックへの影響もあまり見られません。当初、電力不足が問題になると考えられた地域は、東北電力及び東京電力管内に限られていました。とはいえ、これらの地域におけるブロードバンドトラフィック量の合計は全体の半分以上を占めています。しかし、実際には、節電によって利用が減った分を、情報収集等のためにインターネット利用が増えた分で相殺しているように見えます。そもそも、節電の為にインターネットの利用を控えた利用者は少ないと思われる。企業等では、社内のサーバやPCをできるかぎり停止する等の措置が取られました。しかし、そのような場合にも、社内トラフィックは大きく減ったとしても、我々通信事業者から見える組織間トラフィックは相手先との調整が必要なため削減が難しいと考えられます。さらに、震災と計画停電の実施により、社内や家庭内のサーバをクラウドに移行したり、遠隔地へバックアップしたりする傾向が加速し、それがトラフィックの押し上げ要因となっている可能性もあります。

このようにマクロレベルな視点で見ると、震災によってトラフィックが大きく変わったようには見えません。次節以降では、2010年5月末からのデータを使い、より詳細なトラフィック傾向の変化を分析します。

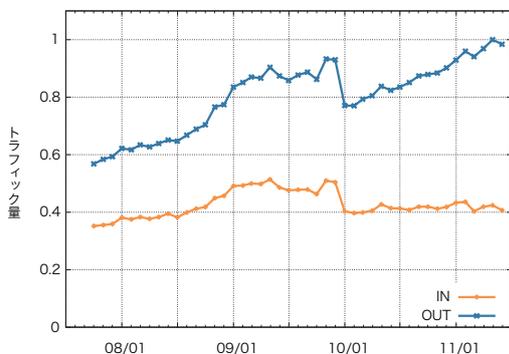


図-2 過去4年間のブロードバンドトラフィック量の推移 (2011年5月のOUT値を1として正規化)

*1 「ブロードバンドトラフィックレポート:P2Pファイル共有からWebサービスへシフト傾向にあるトラフィック」(長健二郎著、Internet Infrastructure Review Vol.8 25 ~ 30ページ) http://www.iiij.ad.jp/development/iir/pdf/iir_vol08_report.pdf

3.3 データについて

今回のレポートも、前回までと同様に個人及び法人向けのブロードバンド接続サービスでファイバーとDSLの利用者を収容しているルータから、Sampled NetFlowによって収集したデータを利用しています。ブロードバンドトラフィックは平日と休日でその傾向が異なるため、1週間分のデータを解析の対象にしています。今回は、2011年5月30日から6月5日までの1週間分のデータを、前回解析した2010年5月24日から30日の1週間分のデータと比較します。

各利用者の使用量は、利用者に割り当てられたIPアドレスと、観測したIPアドレスを照合して求めました。また、Sampled NetFlowでパケットをサンプリングして統計情報を取得しています。サンプリングレートは、ルータの性能や負荷を考慮して1/8192に設定しました。観測した使用量にサンプリングレートの逆数を掛けることで、全体の使用量を推定しています。このようなサンプリング方法を採用することで、使用量の少ない利用者のデータには誤差が生じる恐れがありますが、使用量がある程度以上である利用者のデータでは統計的に意味のある数字が得られます。

なお、ここ数年、DSLからファイバーへの移行が進み、2011年に観測したユーザ数の88%がファイバー利用者であり、トラフィック量全体の93%を占めるまでになっています。また、本レポート内でのトラフィックのIN/OUTは、ISP側から見た方向を表しています。INは利用者からのアップロード、OUTは利用者へのダウンロードになります。

3.4 利用者の1日の使用量

まずは、ブロードバンド利用者の1日の使用量をいくつかの視点から見ていきます。ここでの1日の使用量は、各利用者の1週間分のデータを7で割った1日平均です。

図-3に、利用者の1日の平均使用量の分布(確率密度関数)を示します。ここでは、IN(アップロード)とOUT(ダウンロード)に分け、X軸に利用者のトラフィック量、Y軸にその出現確率をそれぞれ示しています。また、図-3の左側では2010年と2011年を、右側では2005年と2011年をそれぞれ比較しています。X軸は対数表示で、その範囲は 10^4 (10KB) から 10^{11} (100GB) です。最も使用量が多い利用者のトラフィックは1.2TBにも及び、一部の利用者がグラフの範囲外になりますが、おおむね 10^{11} (100G) までの範囲に分布しています。なお、グラフ左側に現れている突起は、サンプリングレートの影響によるノイズです。

IN(アップロード)とOUT(ダウンロード)の各分布は、片対数グラフ上で正規分布となる、対数正規分布に近い形をしています。これは、線形グラフにおいて左端近くにピークがあり右方向になだらかに減少する、いわゆるロングテールな分布になります。OUTの分布がINの分布より右にあり、ダウンロード量がアップロード量よりも1桁ほど大きくなっています。

IN(アップロード)の分布の右端に注目してみると、もう1つの小さな分布の山があることに気がきます。実際にはOUT(ダウンロード)の分布にも、メインの分布に重なっていますが同様の分布の山があります。これら

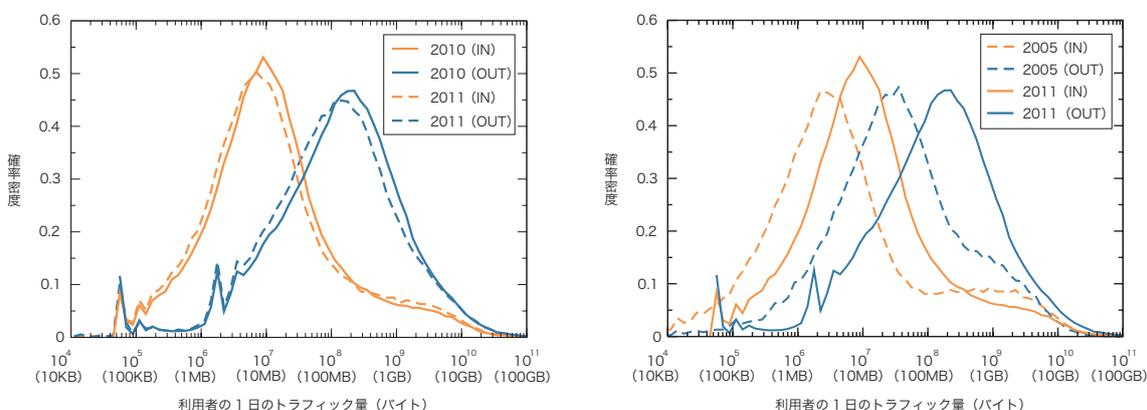


図-3 利用者の1日のトラフィック量分布(左側: 2010年と2011年の比較、右側: 2005年と2011年の比較)

の分布は、INとOUTでほぼ同じ位置にあり、IN/OUTのトラフィックが対称であるヘビーユーザの存在を示しています。そこで、ここでは便宜上、大多数を占めIN/OUTのトラフィックが非対称な分布を「クライアント型利用者」、右側にある少数でIN/OUTのトラフィックが対称なヘビーユーザの分布を「ピア型利用者」と呼ぶことにします。

表-1に、トラフィック量の平均値と、分布の頂点である最頻出値の推移を示します。平均値はグラフ右側に存在するヘビーユーザの使用量の影響を受けるので、2011年の平均値はIN(アップロード)が432MB、OUT(ダウンロード)が1,001MBでした。2010年では、それぞれ469MBと910MBでしたので、INが減少しOUTが増えています。

クライアント型利用者での分布の最頻出値を2010年と2011年で比較すると、IN(アップロード)で7MBから8.5MB、OUT(ダウンロード)で145MBから223MBにそれぞれ増え、特にダウンロード量が増えていることが分かります。

2005年と2011年を比較している図-3の右側に注目すると、一般利用者の使用量が着実に増えているのに対して、量的に大勢を占めるヘビーユーザの使用量が横ばいであり、その割合が減ってきていることが分かります。

図-4では、ランダムに抽出した利用者5,000人のIN/OUT使用量をプロットしています。X軸にOUT(ダウンロード)、Y軸にIN(アップロード)を採り、両対数グラフで表しています。IN/OUTのトラフィックが同量である利用者は、グラフの対角線上にプロットされます。

年	IN (MB/day)		OUT (MB/day)	
	平均値	最頻出値	平均値	最頻出値
2005	430	3.5	447	32
2007	433	4	712	66
2008	483	5	797	94
2009	556	6	971	114
2010	469	7	910	145
2011	432	8.5	1,001	223

表-1 利用者の1日のトラフィック量の平均値と最頻出値の推移

ここでは、2つのクラスタが見られます。対角線の下側にあり、対角線に沿って広がるクラスタは、OUT(ダウンロード)量がIN(アップロード)量よりも1桁多いクライアント型利用者です。一方、右上の対角線上あたりを中心に薄く広がるクラスタは、ピア型利用者です。しかし、この2つのクラスタの境界はあいまいです。これは、実際には、クライアント型利用者もSkype等のピア型アプリケーションを利用し、ピア型利用者もWeb等のダウンロード型アプリケーションを利用しているためです。つまり、多くの利用者は両タイプのアプリケーションを異なる割合で使用しています。また、各利用者の使用量やIN/OUT比率のばらつきも大きく、多様な利用形態が存在することがうかがえます。この傾向は、2010年と比較しても、ほとんど違いがありません。

3.5 ポート別使用量

次に、トラフィックの内訳をポートごとの使用量から見ていきます。最近では、ポート番号からアプリケーションを特定することが困難です。P2P系アプリケーションでは、双方で動的ポートを使うものが多い一方で、クライアント・サーバ型アプリケーションの多くでは、ファイアウォールを回避するためにHTTPで使われる80番ポートが利用されています。大雑把に分けると、双方が1024番以上の動的ポートを使っていればP2P系アプリケーションの可能性が高く、片方が1024番未満のいわゆるウェルノウンポートを使っていればクライアント・サーバ型アプリケーションの可能性が高いと言えます。そこで、ここでは、TCPとUDPでソースとデスティネーションのポート番号の小さいほうを採り、ポート番号ごとの使用量を見ることにします。

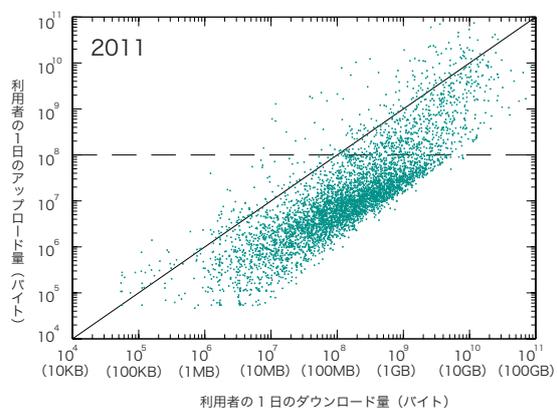


図-4 利用者ごとのIN/OUT使用量

また、トラフィックの全量では、利用量の多いピア型利用者のトラフィックの影響が大きくなります。このため、クライアント型利用者の動向を見るために、少々乱暴なやり方ですが、1日のアップロード量が100MB未満のユーザを抜き出し、それらをクライアント型利用者としてしました。これは、図-3でIN(アップロード)の2つの分布の中間にあたり、図-4でIN=10⁸(100MB)である水平線の下側の利用者になります。

図-5は、ポート別使用状況を、全体とクライアント型利用者に分け、2010年と2011年で比較したものです。また、表-2に、その詳細を数値で示します。

2011年では、トラフィック量の86%がTCPです。しかし、全体のトラフィックでは、2010年に総量の64%だった1024番以上のTCPの動的ポートの割合が、2011年は50%まで減少しています。動的ポートでの各ポート番号が占める個々の割合はわずかで、最大でも総量の1%にしか過ぎません。一方、80番ポートの割合は、2010年の23%から32%に増加しています。

クライアント型利用者に注目してみると、2010年に75%を占めていた80番ポートの割合が、2011年は67%に減少しています。2番目に大きな割合を占める554番ポートは、2009年以前に7%程度ありましたが、2010年に2%まで減った後7%に回復しています。このポートは、RTSP(Real-Time Streaming Protocol)で使われ、ビデオコンテンツの増加に関連しています。また、動的ポートの割合は15%から11%に減少しています。

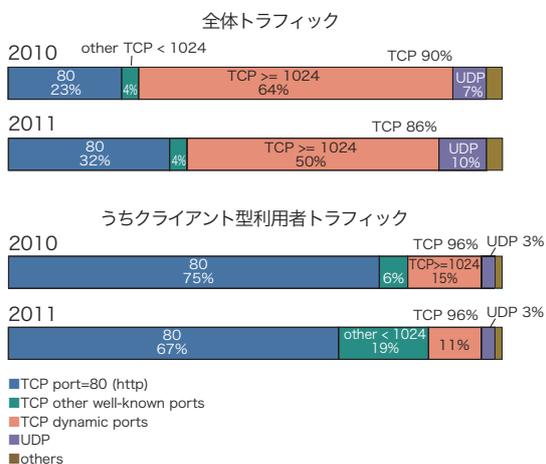


図-5 ポート別使用状況

これらのことから、昨年報告したTCP80番ポートのトラフィックの増加が一般利用者に加えてヘビーユーザにも広がっているという傾向が、いまだに続いていることが確認できます。80番ポートにはビデオコンテンツやソフトウェアアップデート等も含まれるため、コンテンツタイプは特定できませんが、クライアント・サーバ型の通信量が増えていることは確かです。

図-6は、全トラフィックでのTCPポート利用状況の週間推移を示しています。ここでは、TCPでのポート利用を、80番ポート、その他のウェルノウンポート、動的ポートの3つに分けています。また、トラフィック量は、ピーク時の総トラフィック量を1として正規化して表しています。全トラフィックでは、依然として動的ポートの割合が大きく、そのピークが23:00~1:00にあります。土日には昼間のトラフィックも増加し、家庭での利用時間が反映されています。しかし、2010年と2011年のデータを比較してみると、2011年では80番ポートの割合が増え、動的ポートの割合に肩を並べるほどになっています。

protocol	port	2010		2011	
		total (%)	client type	total (%)	client type
TCP	*	90.09	95.82	85.95	96.28
	(<1024)	26.46	80.87	36.24	85.69
	80 (http)	23.00	75.12	32.10	67.30
	443 (https)	0.98	2.28	1.33	1.91
	554 (rtsp)	1.15	2.45	1.33	6.89
	22 (ssh)	0.14	0.10	0.27	0.17
	(>=1024)	63.63	14.95	49.71	10.59
	1935 (rtmp)	1.04	2.91	1.58	1.51
	6346 (gnutella)	0.86	0.33	0.68	0.60
	6699 (winmx)	0.65	0.17	0.40	0.24
	7144 (peercast)	0.34	0.04	0.38	0.00
UDP		6.79	2.76	10.01	2.61
ESP		2.91	1.30	3.56	1.02
GRE		0.14	0.06	0.15	0.05
L2TP		0.00	0.00	0.13	0.00
ICMP		0.02	0.04	0.10	0.04

表-2 ポート別使用量詳細

図-7は、図-6と同様にクライアント型利用者によるTCPポート利用状況の週間推移を示したものです。2011年では、80番ポートの割合が2010年よりも増えています。また、ピーク時間は21:00～23:00で図-6に比べて少し早くなっていますし、土日は朝から利用が増えています。

3.6 おわりに

前回のレポートでは、それまで一般利用者に顕著だったP2Pファイル共有からWebサービスへの移行が、ヘビーユーザにも広がってきたことを報告しました。また、一般利用者の使用量が、ビデオコンテンツや他のWeb 2.0系のリッチコンテンツによって着実に増加してきていることも示しました。今回の調査でも、これらの傾向に変化はなく、Webサービスへの移行が一層進んだことが確認できます。

また、今回は、東日本大震災の影響について、マクロレベルの視点で見たときには限定的であったことも報告しました。社会が大きく影響を受けた災害でも、インターネットのトラフィックがあまり影響を受けなかったことは、インターネットが生活に欠かせないインフラになっていることを示しています。

ところで、私自身は、今回の震災でもっとトラフィックが減ると予想していました。計画停電や節電のために、P2Pファイル共有等に使われている自宅サーバーを停止する利用者がもっといると思ったからです。

2010年1月の改正著作権法施行の時に比べて、トラフィックへの影響が少なかった理由にはふたつの見方があります。ひとつは、一部のヘビーユーザの行動はダウンロード違法化のような強制措置を取らないと変わらないというものです。別な見方は、前回の報告で述べたように、ダウンロード違法化はすでにあった流れを加速するトリガーになったに過ぎないという視点から、最初のトリガーで移行が進んだので、2回目のトリガーの効果が少なかったというものです。もし、法改正と震災が逆順で起こっていたら、やはり最初のトリガーで移行が進んだ可能性もあります。

もし近い将来に別の強制措置がP2Pファイル共有の利用に関して取られるようなことがあれば、その時にトラフィックが減れば前者の見方が、減らなければ後者の見方が有力だと考えることができそうです。

IJでは、今回の震災のような大きな出来事や、ユーザの利用形態の変化にも素早く対応できるよう、継続的にトラフィックの観測を行っています。今後も、定期的にレポートを提供していく予定です。

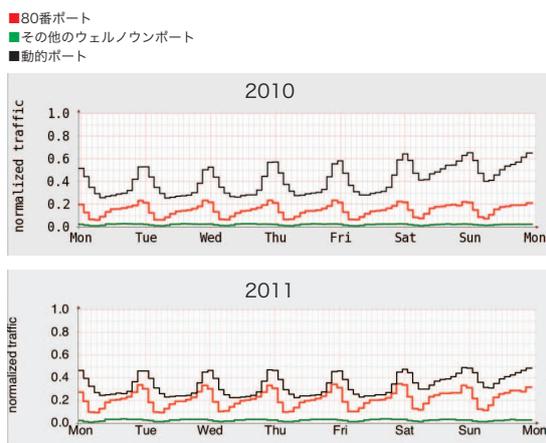


図-6 TCPポート利用の週間推移

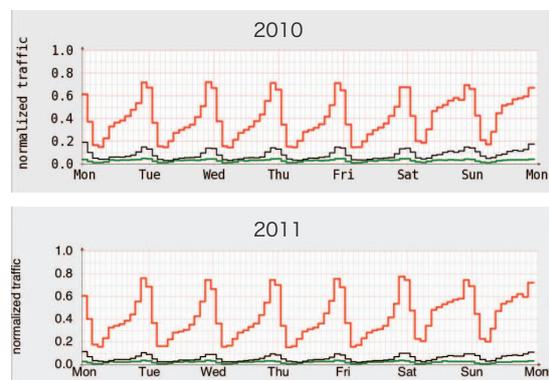


図-7 クライアント型利用者のTCPポート利用の週間推移

執筆者：
長 健二朗(ちよう けんじろう)
株式会社IJイノベーションインスティテュート
技術研究所 所長