

ISPから見たブロードバンドトラフィックの傾向

長 健二郎

株式会社インターネットイニシアティブ

JPNIC総会 2007年3月9日

1

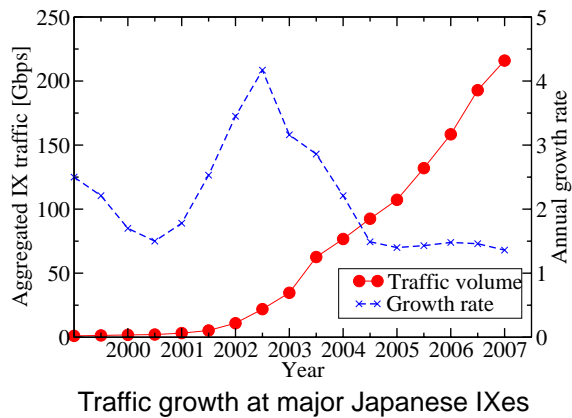
はじめに

- ISPから見たブロードバンドトラフィックの傾向
 - 国内ISP6社によるSNMPベースのトラフィック調査 (3年分)
 - うち1社のブロードバンドトラフィックをSampled NetFlowで解析
- ブロードバンドトラフィックのISPバックボーンへの影響を把握
 - インターネットに関し各方面からさまざまな議論
 - 憶測や偏ったデータをもとにした議論が多い
- インターネットのデータの重要性
 - インターネット屋が公正なデータを公開する努力が必要

2

背景(1) バックボーントラフィックの急増

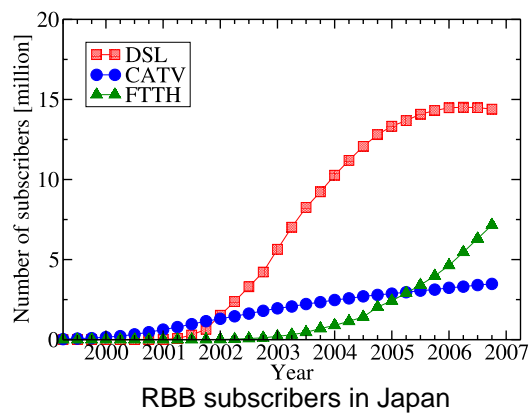
- 主要IXにおけるピークトラフィック (JPIX/JP NAP/NSPIX)
 - ブロードバンド普及による指数関数的増加
 - 2002年には年率4.5倍、最近は50%程度の増加



3

背景(2) ブロードバンドユーザの増加

- 2,500万、DSL:1,440万 CATV:350万 FTTH:720万 (2006年9月)
 - ここ数年はDSLが鈍化し減少へ転換、FTTHが急増
 - 2008年にはFTTHがDSLを抜く予想



4

背景(3) ブロードバンドトラフィック傾向調査の必要

- 急激なブロードバンドトラフィックの増加への懸念
 - バックボーン技術がトラフィック増加に追い付かない
 - ISPはブロードバンドでは儲からないので投資が困難
- ブロードバンド利用者のトラフィック動向の把握
 - 新しいアプリケーションの利用拡大
 - Winny, Gyaο, YouTube, ...

5

ISP6社による傾向調査

6

ブロードバンドトラフィック量傾向調査

- 急激なブロードバンドトラフィックの増加への懸念
 - バックボーン技術がトラフィック増加に追い付かない
 - ISPはブロードバンドでは儲からないので投資が困難
- トラフィック量の把握のため関係者が集まり勉強会をスタート
 - 国内ISP6社、総務省データ通信課(次世代IPインフラ研究会の流れ)
 - IJ, K-Opticom, KDDI, NTT Communications, SoftBank BB, SoftBank Telecom, (POWEREDCOM)
 - 研究者: 江崎(東大) 加藤(東大) 長(IJ) 福田(NII)
 - 2004年9,10,11月 2005年5,11月、2006年5,11月に測定を実施
- プレス発表: 総務省 2005/01/25, 2005/7/27, 2006/3/10, 2006/7/31
- 論文: ACM SIGCOMM CCR, Jan 2005, ACM SIGCOMM, Sep 2006
- 外部発表: SRCCS workshop, Seoul 2005/01. IEPG meeting, Minneapolis 2005/03. Interop Tokyo 2005/06. インターネットコンファレンス 2005/10. NANOG36, Dallas 2006/02. Asia Broadband Summit, Bali 2007/02.

7

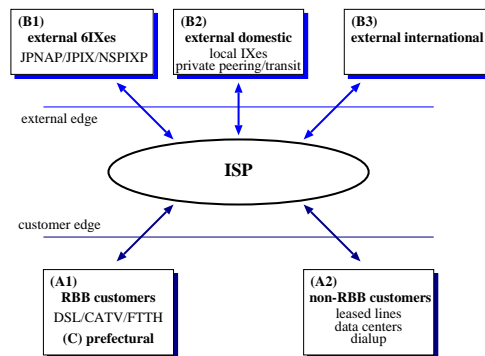
目的

- 今後のインターネットインフラ構築のための基礎データ収集
 - 多くのISPは内部的にデータを持っているが公開されない
 - 参照可能な基礎データを得る
- ブロードバンド普及のバックボーンへの影響をマクロに把握
 - 複数ISPにまたがるトラフィック計測
 - ブロードバンドトラフィック量にフォーカス
 - ブロードバンドとその他のトラフィック量の割合
 - トラフィックパターンの変化
 - トラフィックの地域差

8

主要ISPをまたいだデータの収集

- ISPの顧客境界と外部境界のトラフィックを収集
 - MRTG/RRDtoolのログを集計するツールを開発
- 個別ISPのシェア等が分からないように合算した結果のみ利用



5 traffic groups at ISP customer and external boundaries

9

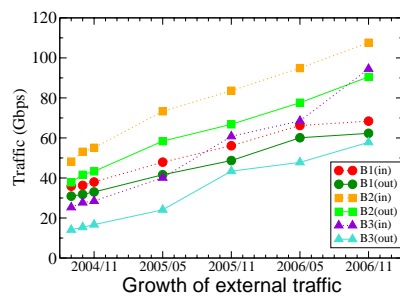
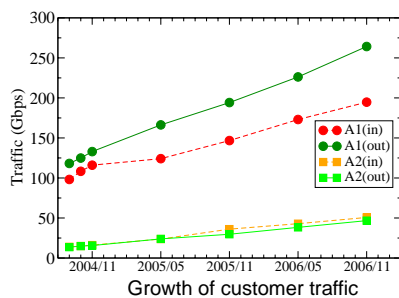
データの収集方法

- トラフィックグループ毎にSNMPインターフェイスカウンタ値を集計
 - 2時間粒度、1ヶ月分
- ほとんどのISPはMRTGまたはRRDtoolでSNMPデータを保存
 - MRTG/RRDtoolのログを集計するスクリプトを開発
 - 各ISPは自社で集計を行ない結果のみを提出
- 参加ISPの負荷
 - インターフェイス毎のログリストの作成
 - ISPによっては10万以上のインターフェイスログが存在
 - ログリストの維持管理
 - 頻繁な構成変更に対応する必要
- データ
 - 6社7ネットワーク分のデータの合算値
- IN/OUTはISPからの視点

10

トラフィックの増加傾向

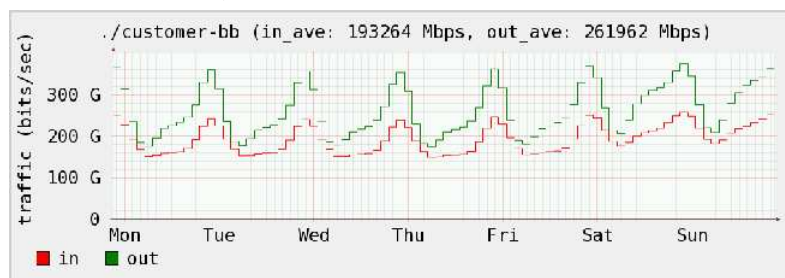
- 2006年には 26-66% の増加
 - ブロードバンド: IN:33% OUT:36%
- 増加率は2002年頃の 100% に比べ鈍化
 - (香港や米国でも同様の報告)



11

ブロードバンドカスタマー (2006/11)

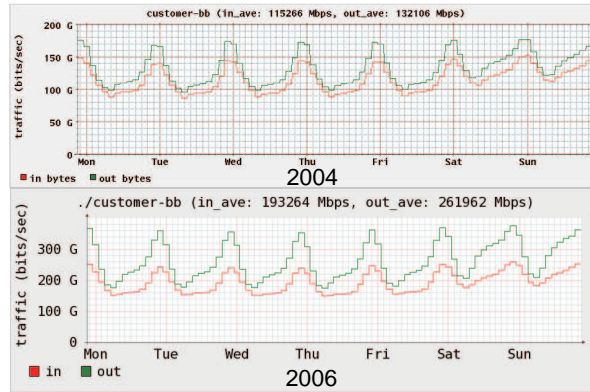
- 6社のDSL/CATV/FTTHカスタマー
 - IN/OUTがほぼ同量 (IN/OUTはISPから見た方向)
 - 平均で200Gbpsの流量、150Gbpsは定常的
 - 一日のピークは、21:00から23:00



12

ブロードバンドトラフィックの変化

- 2004年にはINとOUTがほぼ同量
- 2006年にはOUT(利用者のダウンロード)が拡大



13

ブロードバンド INトラフィック (2004/2005/2006)

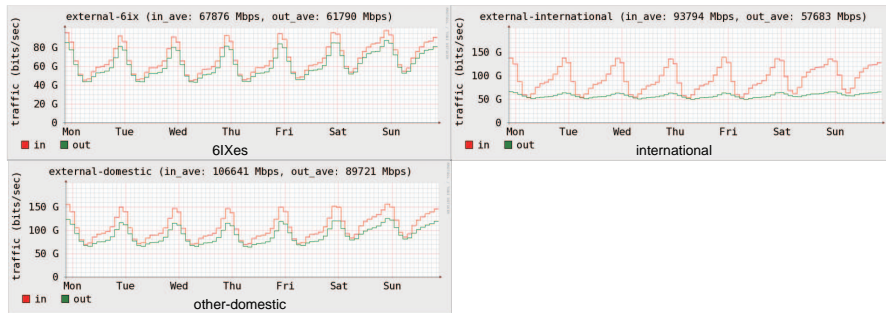
- 2005年は定常部分が増加
- 2006年には定常部分、変動部分共に増加



14

外部トラフィック

- 外部トラフィックもブロードバンドの影響を大きく受けている
 - 主要IXトラフィック：
 - その他国内：プライベートピアリング、トランジット、地域IXなど
 - 主要IXトラフィックを超えてきた
 - 国際：流入(従来型ダウンロード)が支配的

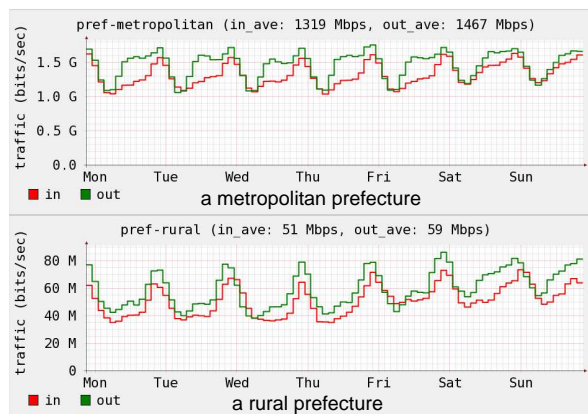


External weekly traffic in November 2006

15

県別トラフィック

- 全ての県に共通して同様のトラフィックパターン
 - ピーク時間、7割が定常的
- 都市部は昼間のトラフィックが大きい

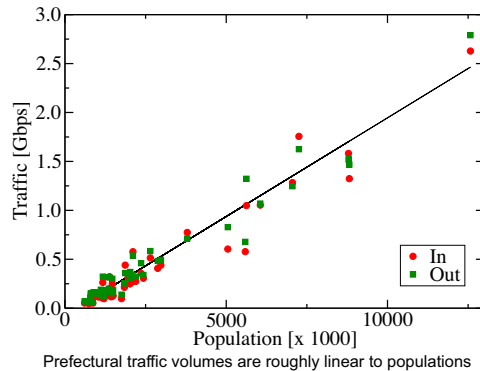


Example prefectural traffic

16

県別人口とトラフィック

- 県別人口とトラフィック量の相関
 - トラフィック量はほぼ人口に比例している
 - インターネットユーザ数で見ても同様の傾向
- (国内のユニバーサルなブロードバンドサービス展開の成果)



17

主な結果

- 協力ISPは主要IXトラフィックの42%のシェア
 - 国内BBトラフィック総量は637Gbpsと計算できる (2006/11)
- BBトラフィックは21:00-23:00にピーク、70%は定常的(自動生成?)
 - 2004年にはファイル交換が支配的、2006年にはビデオダウンロードが急増
- バックボーンにおいてもBBトラフィックが支配的
- 主要IX以上の量がプライベートピアリングで交換されている
- 県別BBトラフィック量はほぼ人口に比例

- 今後は年2回の計測を継続
- 本活動は国際的にも注目

18

ISPのブロードバンドユーザ傾向調査

19

目的

- 6社調査のカウンタ値だけでは分からなかった詳細の調査
- 利用量からみたユーザ分布とトラフィックパターン
 - ユーザ間のトラフィックの偏り度合の検証
 - ヘビーユーザと一般ユーザの違い
 - ファイバユーザとDSLユーザの違い
- トラフィックの地域性調査
 - トラフィックマトリクス
 - どの県からどの県にどれくらい通信があるか
 - 将来のバックボーン設計へのインプット
 - インターネット通信には地域性が少ない仮定の検証

20

調査方法

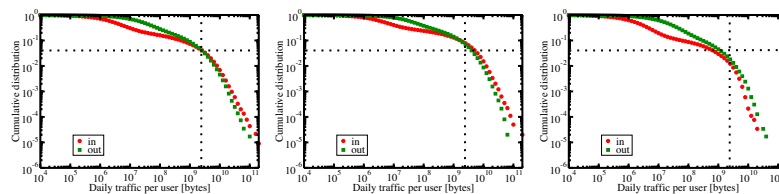
- ブロードバンドユーザ収容ルータでSampled NetFlowを利用
 - 2005/2、2005/7の3回実施
- 調査対象はブロードバンドユーザ
 - 全体: ファイバー/DSL接続

- 調査結果は6社による調査結果と整合、一般性を持つ

21

トラフィック使用量に対するユーザの相補累積度分布

- ヘビーユーザは広範囲にわたり統計的に分布
 - ベキ分布 (フラクタル): 右端は200GB/day (19Mbps)!
 - ヘビーユーザと一般ユーザの境界はあいまい
- 変曲点 2.5GB/day (230kbps)、上位4%ヘビーユーザ (全体)
 - ここでは上り2.5GB/day以上のユーザをヘビーユーザと定義
- ヘビーユーザの割合: 全体の4%、ファイバの10%、DSLの2%

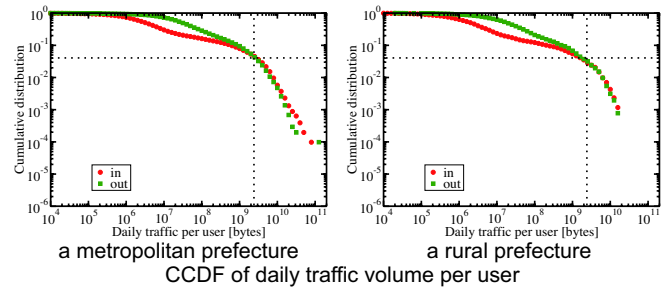


total fiber DSL
CCDF of daily traffic volume per user

22

地域別のトラフィック使用量

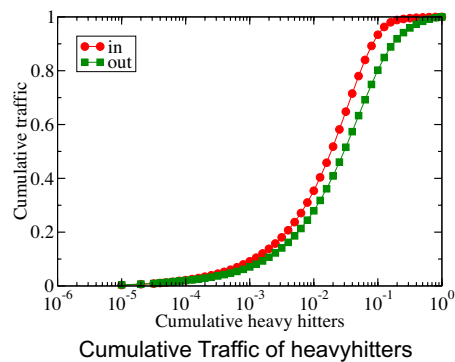
- トラフィック使用量分布の形は全国に共通
 - 母数による分布右端の長さの違い
 - ファイバとDSLの割合を反映



23

ヘビーユーザへのトラフィックの偏り度合

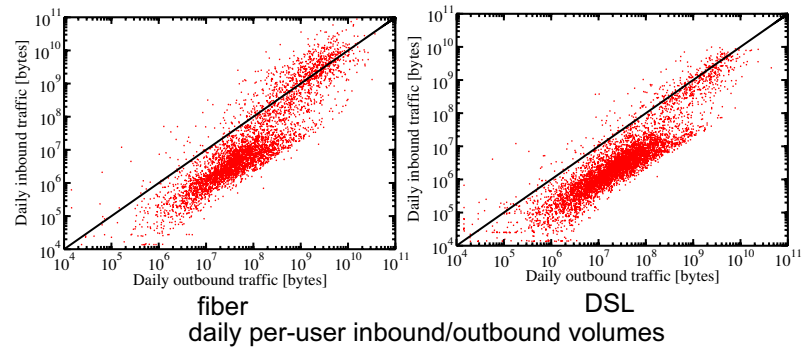
- 使用量上位何%のユーザが全体トラフィックの何%を占めるか
- トラフィック使用量に大きな偏り
 - IN側：上位4%が全体の75%を占める
 - OUT側：上位4%が全体の60%を占める



24

ユーザのIN/OUTトラフィックの相関

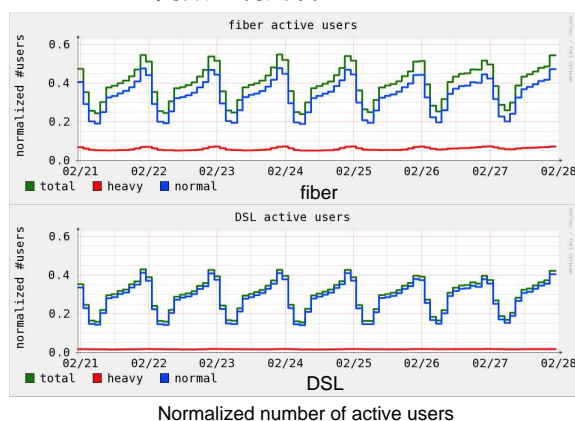
- 各ユーザのIN/OUT量をログ・ログスケールでプロット
 - 対角線下方のクラスター: 一般ユーザ層 (ダウンロード中心)
 - 対角線右上のクラスター: ヘビユーザ (IN/OUT対称)
- ファイバとDSLは同様の傾向
 - 質的な違いはない、単にヘビユーザ比率が違う
- ここでもヘビユーザと一般ユーザの境界はあいまい



25

アクティブな利用者数

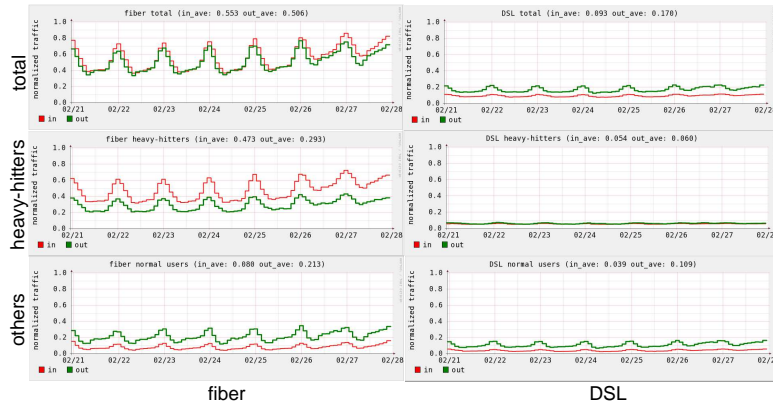
- 観測されたファイバ/DSLの利用者数を全体ピーク値に正規化
- ファイバ、DSLでほぼ同数の利用者



26

ファイバ/DSL トラフィック

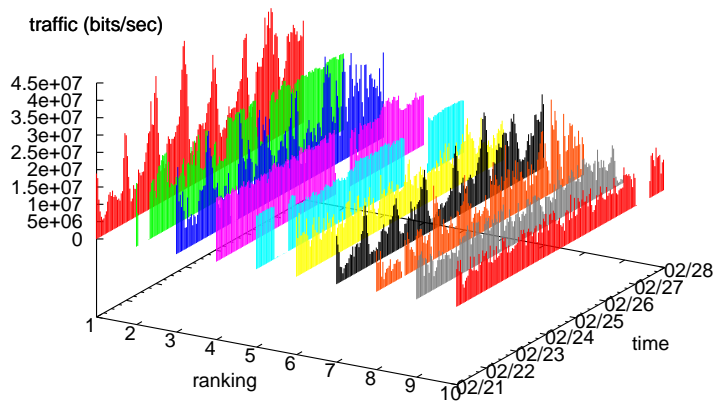
- ファイバ/DSLのトラフィック量を全体ピーク値に正規化
- IN: 86%はファイバ利用者トラフィック, DSLは14%にすぎない
- 全体はファイバ利用のヘビーユーザの影響大



27

上位10ヘビーユーザの挙動

- 1週間分の1時間平均トラフィック
- 利用者による挙動の違い、異なる使い方を示唆



28

プロトコル/ポート

- TCP 80番 (http) は高々9%
- 83%はTCPのダイナミックポート
 - ポート番号ではアプリケーションの識別不能

protocol	port	name	(%)	port	name	(%)
TCP	*		97.43			
	(< 1024		13.99)	81	-	0.15
	80	http	9.32	25	smtp	0.14
	20	ftp-data	0.93	119	nntp	0.13
	554	rtsp	0.38	21	ftp	0.11
	443	https	0.30	22	ssh	0.09
	110	pop3	0.17		others	2.27
	(>= 1024		83.44)	1935	macromedia-fsc	0.20
	6699	winmx	1.40	1755	ms-streaming	0.20
	6346	gnutella	0.92	2265	-	0.13
	7743	winny	0.48	1234	-	0.12
	6881	bittorrent	0.25	4662	edonkey	0.12
	6348	gnutella	0.21		others	79.41
UDP	*		1.38	6257	winmx-	0.06
	6346	gnutella	0.39		others	0.93
ESP			1.09			
GRE			0.07			
ICMP			0.01			
others			0.02			

29

地域別トラフィックマトリクス

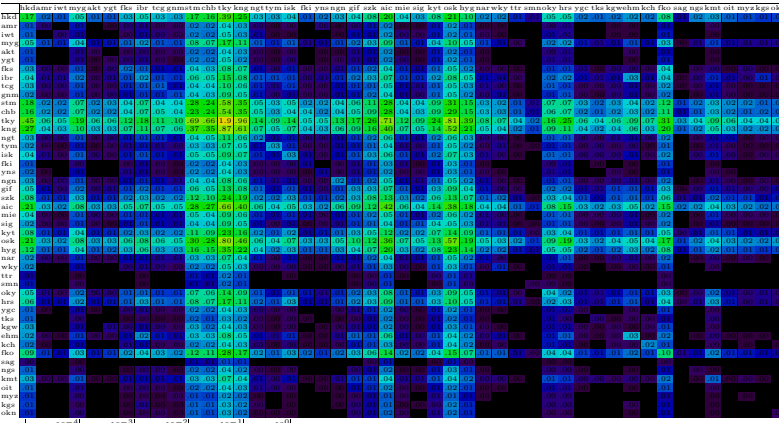
- RBB(residential broadband), DOM (other domestic), INTL (international)
 - 両エンドのIPアドレスを商用Geo-IPデータベースで識別
- 62%はユーザ・ユーザの通信
- 90%は国内に閉じたトラフィック(RBBまたはDOM)
 - 言語、文化の壁
 - 国内ファイバユーザがP2Pのスーパーノード網を構成か?

<i>src\dst</i>	ALL	RBB	DOM	INTL
ALL	100.0	84.8	11.1	4.1
RBB	77.0	62.2	9.8	3.9
DOM	18.0	16.7	1.1	0.2
INTL	5.0	4.8	0.2	0.0

30

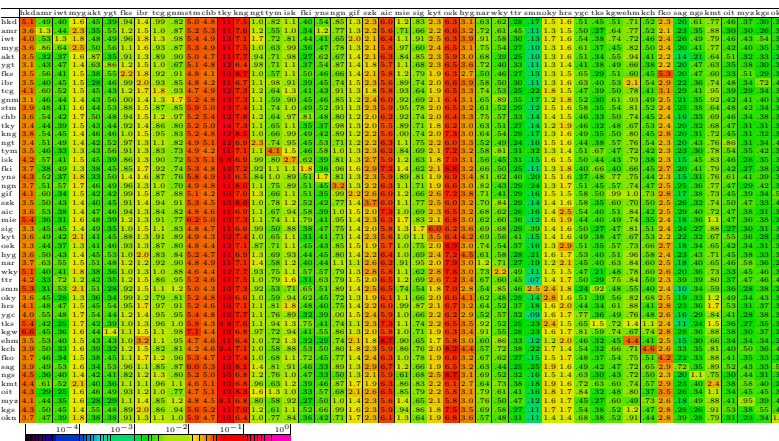
都道府県別トラフィックマトリクス

- 全流量に占める割合
 - ソース(Y軸) デスティネーション(X軸)
 - 人口に応じたトラフィック量分布



ソースに正規化したマトリクス

- 各行の合計が100%
- 各県共通の分布、地域性がほとんどない、県内トラフィックは1-3%程度
 - デスティネーションに正規化しても同様の結果



議論(1) ブロードバンドトラフィック増加

- 日本は世界のモデル、直面する問題をどう解決していくか
 - 突出したファイバの普及率の影響
 - ブロードバンドがバックボーントラフィックの2/3
 - 全国的なユニバーサルな展開には成功
- トラフィック増加率
 - 一時より増加が鈍っている
 - ブロードバンド一巡
 - 度重なる情報漏洩事件の影響
 - このままなら既存技術の延長で対応も可能?
 - 年率100%増加なら10年で1000倍、50%なら58倍

33

議論(2) 二極化問題

- 一見二極化に見える
 - 4%のヘビーユーザが75%の上りトラフィックを使用
 - ファイバが全体の上りトラフィックの85%を占める
- 実は多様かつべき分布、境界ははっきりしない
 - 一般ユーザのヘビーユーザ化
 - 数が多過ぎる(ファイバの10%)、統計的に分布
 - 一般ユーザがヘビーユーザ化しファイバへ移行
 - 一般ユーザがファイバ契約し帯域の使い方を求める
- いままでなかった現象、ブロードバンド第二幕

34

議論(3) アプリケーションの多様性

- ユーザ間トラフィック量の増加
 - ISPやコンテンツプロバイダはあまり考えていなかった
- ストリーミング、ビデオチャットも増えてきている
- ファイル共有はブロードバンド初期の常時接続化の産物
 - 広帯域があれば必ずしも必要ない
 - 現状はいつまでも続かないのでは？
- 新しいアプリケーションの出現で事態が一変する可能性
 - 一般ユーザはあまり意識せずに使っている
 - 実際にファイル共有からビデオダウンロードへのシフトが起きている

35

議論(4) コスト負担の不公平性

- 一部のヘビーユーザトラフィックをみんなで負担している現状
 - 技術的な解決：帯域制限、通信品質劣化
 - 課金による解決
- いずれにせよタイトなシステムは構築、運用とも高コスト
 - 不公平を許容するコストとの比較
- インターネットは統計多重による低コストアーキテクチャ
 - コモنزの悲劇
 - アーキテクチャが破綻、単にISPビジネスモデルではない
 - ユーザの意識改革が必要(エコと同様)、共有を認識する仕組み
- インフラ投資ができない問題
 - 囲い込みモデルが機能しなくなっている
 - Net Neutrality議論
 - ISP、キャリアのコスト構造の再検討が必要
 - インフラコストはインフラで回収すべき

36

ブロードバンド第二幕への課題

- ブロードバンドの恩恵を受けたエンドユーザによる革新
 - 予想していなかった展開が始まっている
 - ISP、キャリアは革新を受け止める準備が必要
- 将来のインターネットの発展を考える
 - 一部の極端な使用を低コストで防止する
 - 技術、課金、ユーザ啓蒙のルーズな組合せ
 - 普通の使い方なら固定料金的に使える必要
 - 新しい使い方が出てくるための十分なマージン確保
- インターネットのデータの重要性
 - 憶測や偏ったデータをもとにした議論が多い
 - インターネット屋が公正なデータを公開する努力が必要