

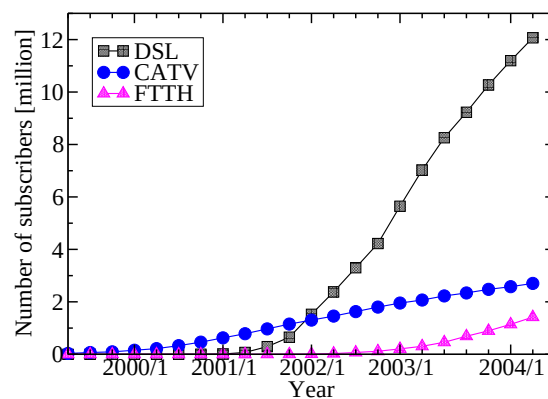
ブロードバンドユーザトラフィックの現状

長 健二郎
IIJ技術研究所

1

はじめに

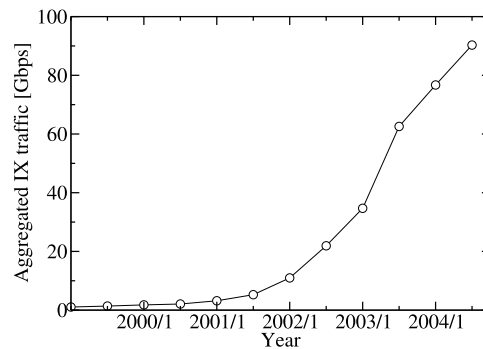
- ブロードバンドユーザの増加
 - DSL:1,200万 CATV:280万 FTTH:200万 (2004年9月)



2

バックボーントラフィックの急増

- 主要IXにおけるトラフィック (JPIX/JPNAP/NSPIX)
- このうちブロードバンドユーザが占める割合は？



3

ブロードバンドトラフィック量傾向調査

- ISP7社による傾向調査
 - 各社はSNMPでルータのインターフェイスカウンタ値を記録
 - カウンタをグループ分けしカウンタ値総和を収集解析
- IJのブロードバンドユーザ傾向調査
 - Sampled NetFlowによるブロードバンドデータの計測
 - より詳細な解析

4

ISP7社による傾向調査

5

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ISP7社、総務省データ通信課 (次世代IPインフラ研究会の流れ)
 - 研究者: 江崎(東大) 加藤(東大) 長(IIJ) 福田(NTT)
 - 2004年9、10、11月、2005年5月に測定を実施
- プレス発表: 総務省 2005/01/25, 2005/7/27
- 論文: ACM SIGCOMM CCR, January 2005
- 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

6

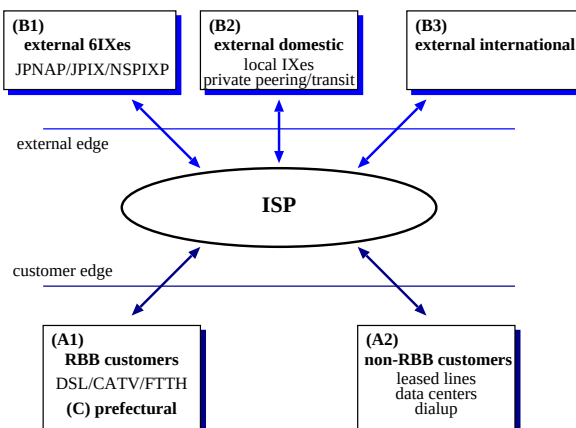
目的

- 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

7

主要ISPをまたいだデータの収集

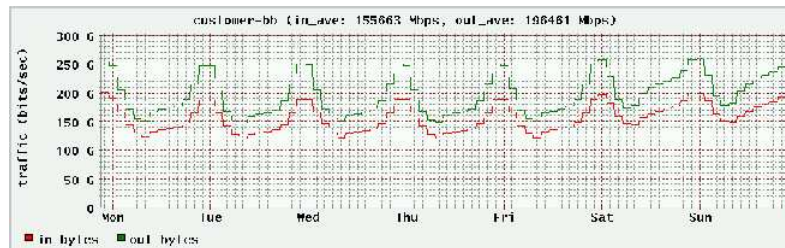
- ISPの顧客境界と外部境界のトラフィックを収集
 - MRTG/RRDtoolのログを集計するツールを開発
- 個別ISPのシェア等が分からないように合算した結果のみ利用



8

BB(ブロードバンド)カスタマー (2005/5)

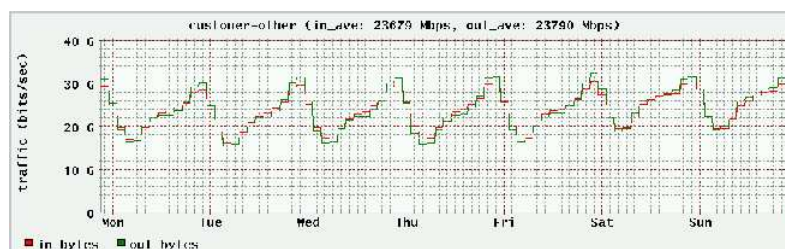
- 7社のDSL/CATV/FTTHカスタマー
 - IN/OUTがほぼ同量 (IN/OUTはISPから見た方向)
 - 平均で200Gbpsの流量、150Gbpsは定常的
 - 一日のピークは、21:00から23:00



9

BB以外のカスタマー (2005/5)

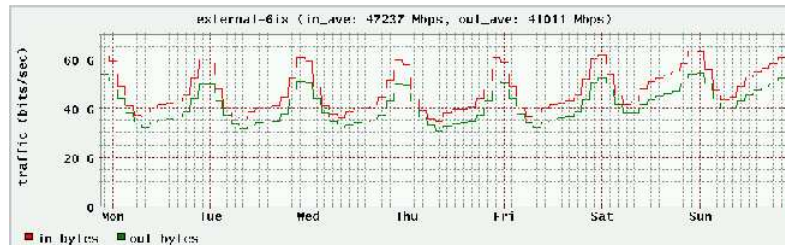
- 専用線、データセンター、その他 (4社分のみ)
 - ホームユーザの利用傾向が支配的 (ピーク時間等)
 - 専用線の先に2次プロバイダがいるため
 - 昼間のトラフィックは、BBカスタマーより多い



10

主要IX 外部トラフィック (2005/5)

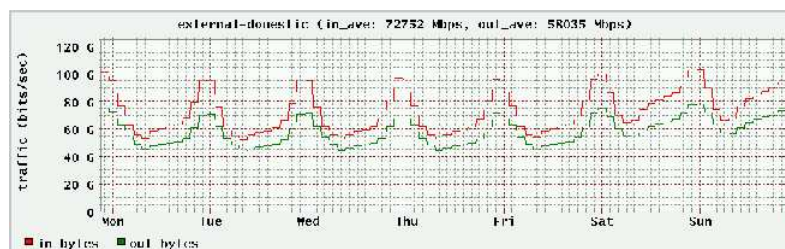
- JPIX/JPNAP/NSPIXの東京/大阪
 - BBトラフィックの強い影響が見られる



11

その他 国内外部トラフィック (2005/5)

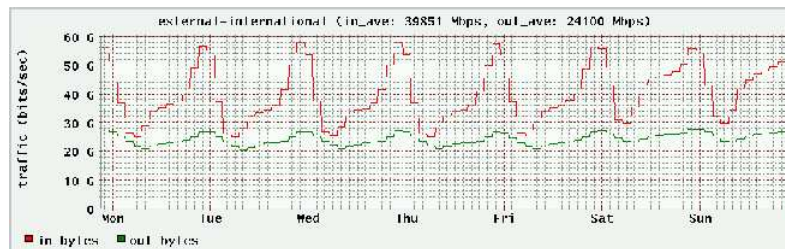
- プライベートピアリング、トランジット、地域IXなど
 - 主要IXと同様の傾向



12

国際 外部トラフィック (2005/5)

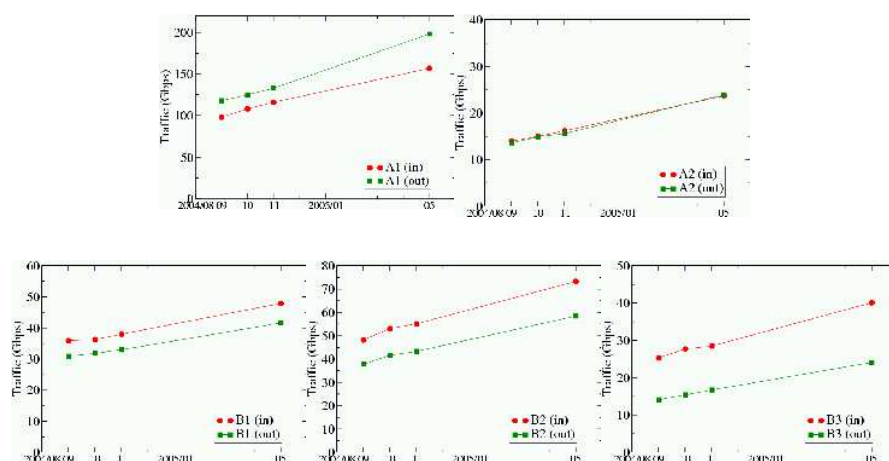
- 国際トラフィックはまだ流入が支配的
 - 従来型ダウンロードが中心
 - 日本国内のトラフィックの多くは国内に閉じている



13

トラフィックの増加傾向

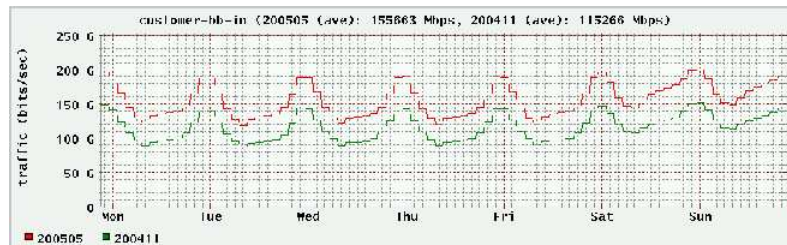
- 6ヶ月で25-50%の増加！



14

BBトラフィック 2004/11と2005/5の比較

- 定常部分が増加している！

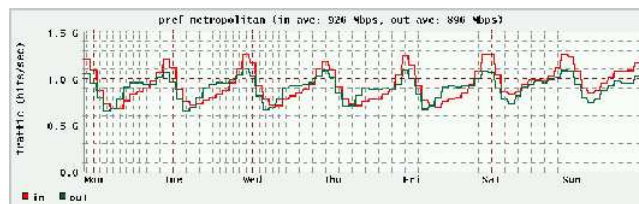


15

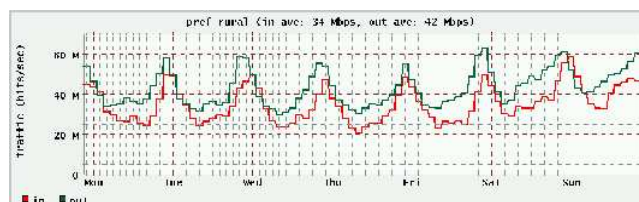
県別トラフィック (2004/09)

- 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

□ 都市部の県



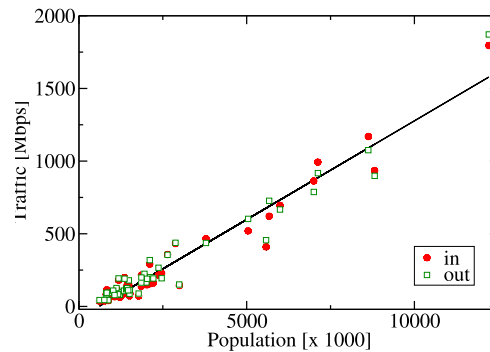
□ 地方の県



16

県別人口とトラフィック (2004/09)

- 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



17

主な結果

- 協力ISPは主要IXトラフィックの41%のシェア
 - 国内BBトラフィック総量は470Gbpsと計算できる (2005/5)
- 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?)
- 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

IIJのブロードバンドユーザ傾向調査

19

目的

- 7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ぐらい通信があるか
 - 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

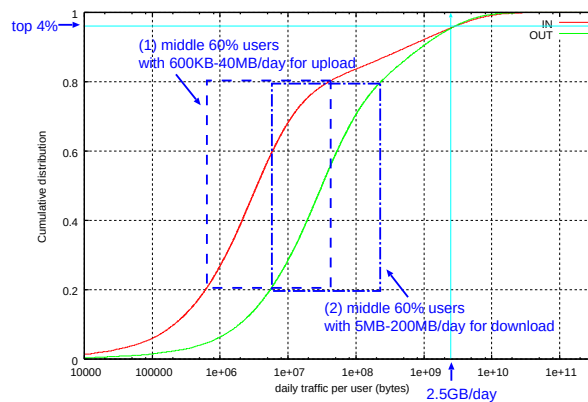
調査方法

- ブロードバンドユーザ収容ルータでSampled NetFlowを利用
 - 2004/10、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: 個人向け/企業向けサービス、ファイバー/DSL接続
 - ファイバユーザの調査: 個人向けファイバー接続
 - DSLユーザの調査: 個人向けDSL接続
- 調査結果は7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累積度分布

- 各ユーザの1日平均トラフィック量をIN/OUT独立に累積度分布
 - 1日平均使用量何バイト以下のユーザの割合を示す
 - X軸はユーザの1日平均トラフィック使用量 (ログスケール)
 - (1) IN(upload) 20%:600KB以下 80%:40MB以下
 - (2) OUT(download) 20%:5MB以下、80%:200MB以下

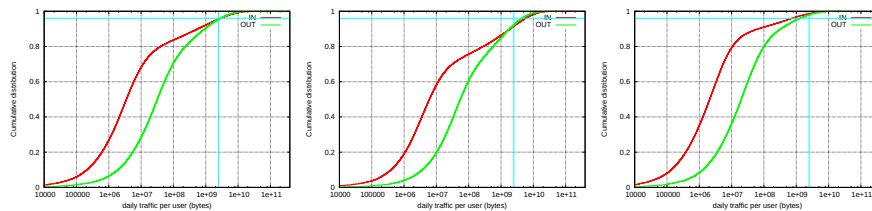


22

ユーザの累積度分布(CDF) (2005/02)

◦使用量順で20%-80%のユーザの一日のトラフィック

- ダウンロードがアップロードの約10倍
- ダウンロード全体:5-200MB ファイバ:10-600M DSL:4-100MB
- アップロード全体:0.6-40MB ファイバ:1-300M DSL:0.4-10MB

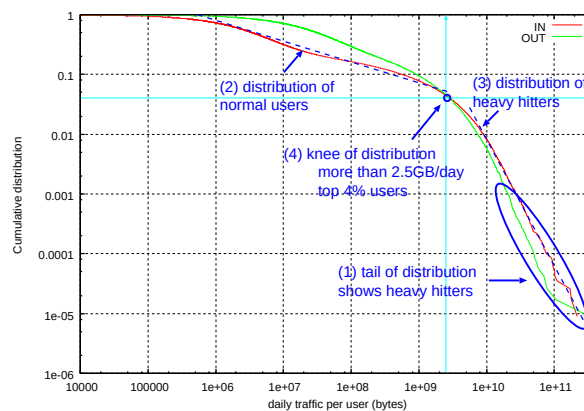


全体(左)、ファイバ(中)、DSL(右)

23

トラフィック使用量に対するユーザの相補累積度分布

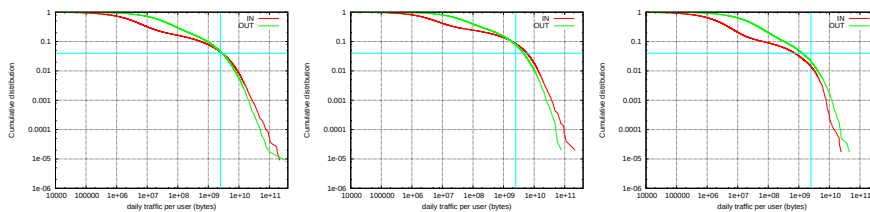
- 相補累積度分布:CCDF=(1-CDF)、ヘビーユーザ(分布右端)に注目
- 1日平均使用量何バイト以上のユーザの割合を示す(ログログ)
 - 全ての県で同様な分布(テイルの長さが違うだけ)
 - (4)分布の変曲点:4%:1日2.5GB(230kb/s)以上
 - 一般ユーザはダウンロード中心、ヘビーユーザでは逆転



24

ユーザの相補累積度分布(CCDF) (2005/02)

- 変曲点: 全体では1日平均2.5GB、上位4%のユーザ
 - アップロードが1日2.5GBを越えるユーザ
 - ファイバの10%、DSLの2%: 割合が大きく違う
- ヘビーユーザもスケールに依存しない (べき分布)
 - ヘビーユーザも一定の分布に従う: もはや特殊な存在でない
 - いくら回線容量を増やしてもそれを食い潰すような超ヘビーユーザが存在する可能性 (現実にはアクセス回線の制約)

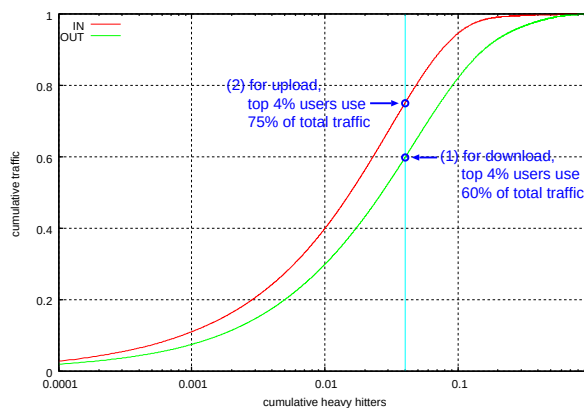


全体(左)、ファイバ(中)、DSL(右)

25

ヘビーユーザへのトラフィックの偏り度合

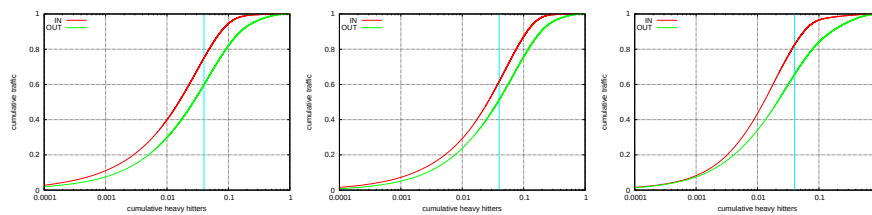
- 使用量上位何%のユーザが全体トラフィックの何%を占めるか
 - ▷ (1) OUT 上位4%: 全体の約60%を使用
 - ▷ (2) IN 上位4%: 全体の約75%を使用



26

ヘビーユーザへのトラフィックの偏り度合 (2005/02)

- 一部のユーザがほとんどのトラフィックを使っている現状
- 使用量上位4%のユーザが使用しているトラフィック量の割合
 - ダウンロード全体:60% ファイバ:50% DSL:65%
 - アップロード全体:75% ファイバ:62% DSL:93%
- DSLの方がヘビーユーザへの偏りが大きい
 - ファイバの方がヘビーユーザの割合が多いため

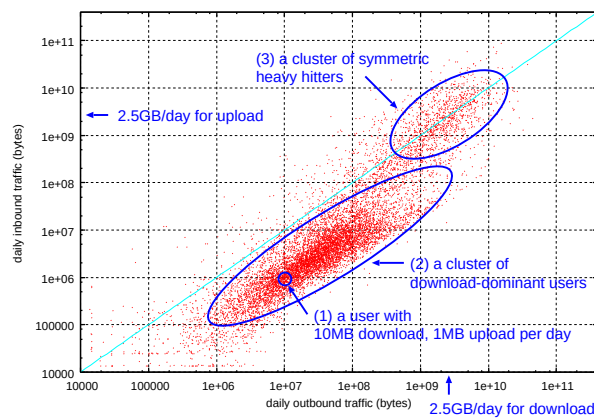


全体(左)、ファイバ(中)、DSL(右)

27

ユーザのIN/OUTトラフィックの相関

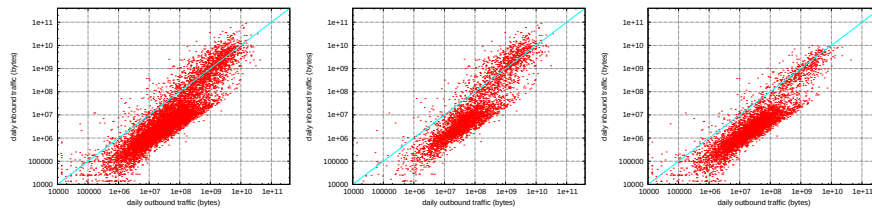
- 各ユーザのIN/OUT量をログ・ログスケールでプロット
- 対角線下方のクラスター: 典型的なユーザ層
 - ダウンロードがアップロードの10倍
- 対角線上右上のクラスター: ヘビーユーザ



28

ユーザのIN/OUTトラヒックの相関 (2005/02)

- 大きな個人差: 4桁ぐらいの幅
- ファイバ、DSLも同様の傾向
 - DSL: 一般ユーザの密度が高く、左下のライトユーザが多い
 - ファイバ: アップロードの方が多いヘビーユーザの存在



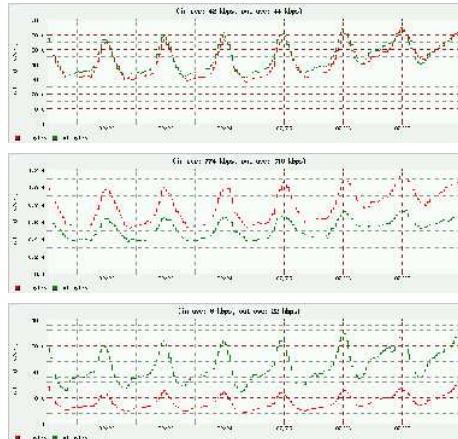
全体(左)、ファイバ(中)、DSL(右)

1時間粒度週間データの解析

- ユーザあたりの1時間粒度の平均IN/OUTトラヒック量
 - 平均値は利用量がほとんどないユーザの影響で小さい
- 1日平均2.5GB以上と未満のユーザに分けて傾向の違いを調べる
 - 全体データは一部のヘビーユーザの影響が大きい
 - 2.5GB未満のデータはヘビーユーザ分の影響を除いたもの

1時間粒度週間トラフィック量 (2005/2)

- 全体ではIN/OUTが同じぐらいだが、ユーザ層で大きな違い
- 2.5GB以上のユーザも手動と思われる変動部分が多い

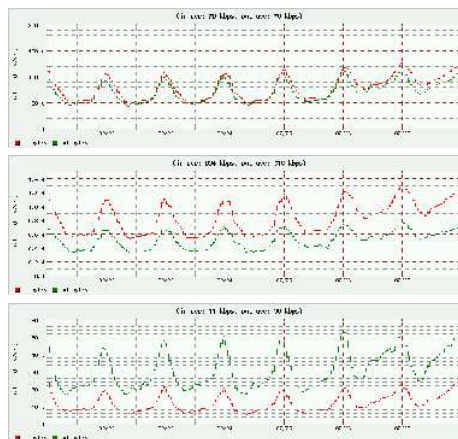


全体(上)、1日平均2.5GB以上(中)、1日平均2.5GB未満(下)

31

ファイバユーザの週間トラフィック量 (2005/2)

- 全体と同様の傾向

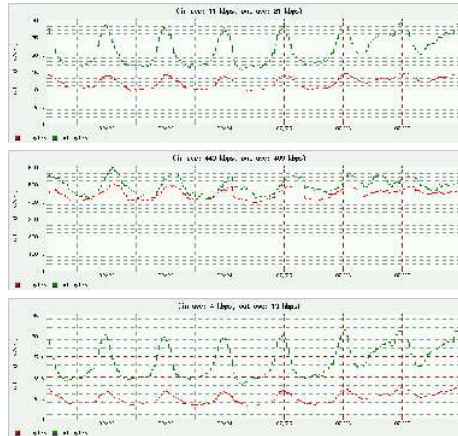


全体(上)、1日平均2.5GB以上(中)、1日平均2.5GB未満(下)

32

DSLユーザの週間トラフィック量 (2005/2)

- 2.5GB以上のユーザでDSLの上り帯域の制約の影響

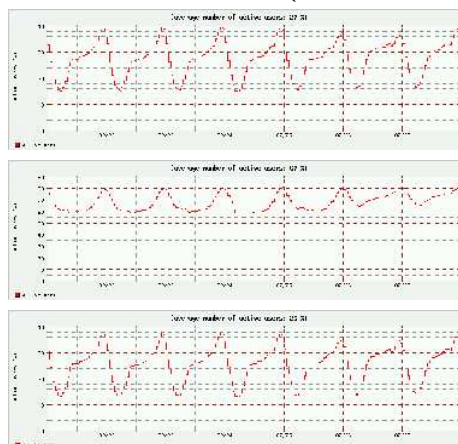


全体(上)、1日平均2.5GB以上(中)、1日平均2.5GB未満(下)

33

アクティブなユーザ数の週間変動 (2005/2)

- 1時間のアクティブユーザ数(1週間に観測した総数に対する割合)
- 2.5GB以上のユーザ数は変化が少ない(定常的に使っている)

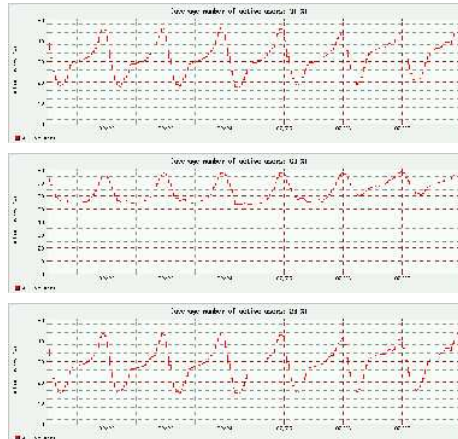


全体(上)、1日平均2.5GB以上(中)、1日平均2.5GB未満(下)

34

ファイバユーザのアクティブなユーザ数 (2005/2)

- 全体と同様の傾向
 - 朝と夕方に2段階に増えて夜中過ぎに急速に減る

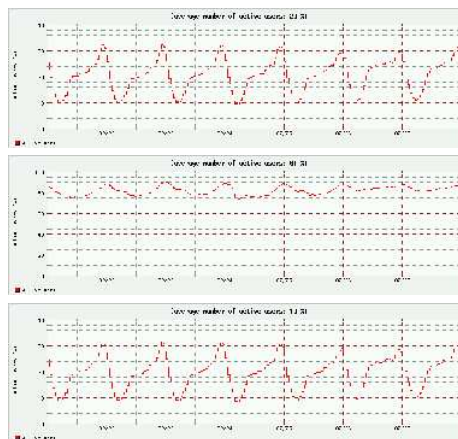


全体(上)、1日平均2.5GB以上(中)、1日平均2.5GB未満(下)

35

DSLユーザのアクティブなユーザ数 (2005/2)

- 2.5GB以上のアクティブユーザ数はほとんど変化がない

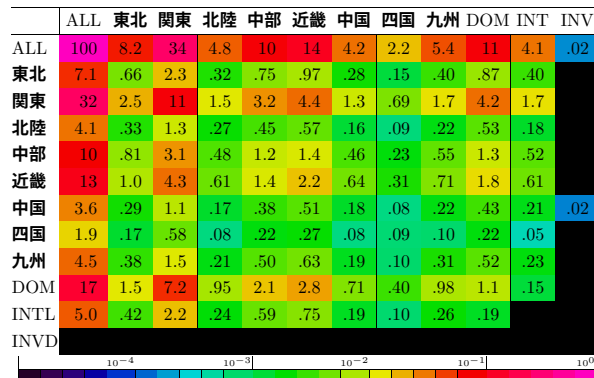


全体(上)、1日平均2.5GB以上(中)、1日平均2.5GB未満(下)

36

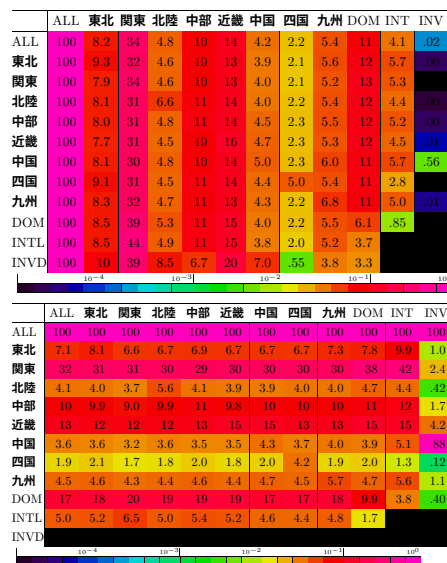
トラフィックの地域性調査 (2005/07)

- サイバーエリサーチ社IPアドレス県別データベースを利用
 - ISPの一般ユーザのみ(データセンターや専用線は含まず)
- src:Y軸 dst:X軸
 - ALL, 8地域, other DOMestic, INTernational, INValid
- user-to-userが63%、90%は国内に閉じている(言語・文化の壁)



37

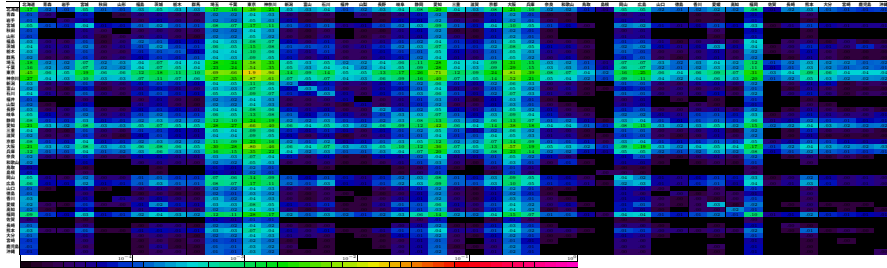
srcおよびdstに正規化した地域トラフィックの分布



38

都道府県別トラフィック分布 (2005/07)

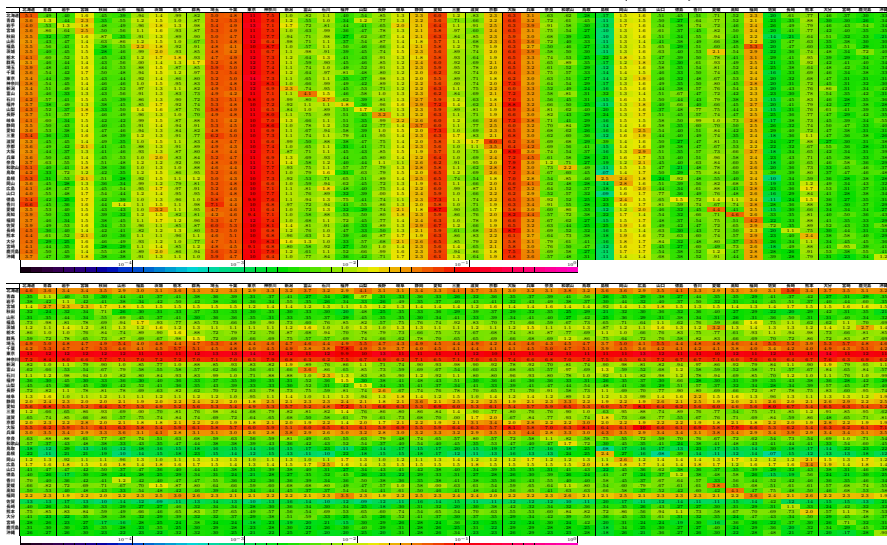
- 47都道府県間で交換されるトラフィック (地域別表の詳細)
 - 人口の多い県に多くのトラフィックが観測される



39

srcおよびdstに正規化した都道府県別トラフィック分布

- すべての県で同様な分布、わずかな県内地域性(1-3%)を確認



40

プロトコル/ポート別ランキング (2005/07)

- 80%はTCPの動的ポートが使われている
 - 従来のポートによる分類はできなくなっている

protocol	ratio(%)	port #	name	ratio(%)
TCP	97.43			
(port < 1024	13.99)	80	http	9.32
		20	ftp-data	0.93
		554	rtsp	0.38
		443	https	0.30
		110	pop3	0.17
		81	-	0.15
		25	smtp	0.14
		119	nntp	0.13
		21	ftp	0.11
		22	ssh	0.09
		-	other	2.27
(port >= 1024	83.44)	6699	winux	1.40
		6346	gnutella	0.92
		7743	winny	0.48
		6881	bittorrent	0.25
		6348	gnutella	0.21
		1935	macromedia-fsc	0.20
		1755	ms-streaming	0.20
		2265	-	0.13
		1234	-	0.12
		4662	edonkey	0.12
		8080	http-proxy	0.11
		-	other	79.30
UDP	1.38	6346	gnutella	0.39
		6257	winux-	0.06
		-	other	0.93
ESP	1.09			
GRE	0.07			
ICMP	0.01			
OTHER	0.02			

41

主な結果

- 利用量からみたユーザ分布
 - ユーザ間で大きな偏り
 - 10%のユーザが90%以上のトラフィック利用
 - ヘビーユーザも一定の分布に従う: もはや特殊な存在でない
 - 一般ユーザのヘビーユーザ化: p2p利用、ファイバへの移行
- ファイバユーザとDSLユーザの違い
 - 同様の傾向だが、ファイバーにヘビーユーザが多い
- トラフィックパターン
 - 機械生成されていると思われる定常的トラフィックの増加
- トラフィックの地域性調査
 - 90%のトラフィックが国内に閉じている日本の特殊性
 - 電話網と違い、都道府県別で見るとローカリティはほとんどない

42

まとめ

- ISP7社による傾向調査
 - 国内ブロードバンドトラフィックの現状
- IJJのブロードバンドユーザ傾向調査
 - より詳細なユーザの利用傾向解析
- ブロードバンド普及の成功の結果
 - 国内バックボーンはブロードバンドトラフィックが支配的
- 年率約2倍のトラフィック量の伸び、ユーザ間で利用量に大きな偏り
 - 機械生成されている定常的トラフィックの増加
 - 一方、一般ユーザとヘビーユーザの境界がなくなる傾向
 - 今後もDSLからファイバへの移行が進む
- トラフィックの地域性調査
 - 都道府県別で見ると地域性はほとんどない

43

最後に

- p2pトラフィックがバックボーンを圧迫してきている現状
 - 問題点: ユーザ間の不公平感、違法コンテンツ
 - しかし、p2pがいけない訳ではない
- 新しいインターネットの使い方の台頭
 - 一般ユーザにも利用が広がり、面白い使い方も増加
- インターネットを享受したエンドユーザによる利用方法の革新
 - 規制するのではなく、正しい進化を導くような対応が重要

44