InternetWeek 2006

T25 フローベースのトラフィック計測と解析

Part II: フローデータを使ったトラフィック解析の実際

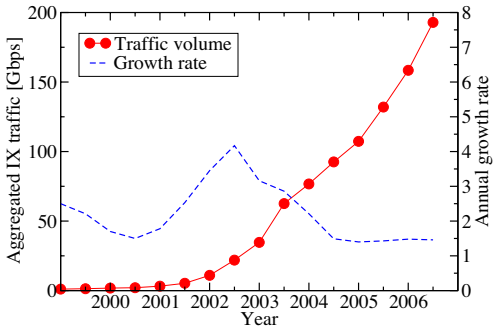
長 健二朗
株式会社インターネットイニシアティブ

はじめに

- 本チュートリアル前半でフローベースの計測技術を解説
- 後半では、フローデータを使ったトラフィック解析について
  - 具体的な応用例をデータ処理手法や解析結果を交えて解説

背景(1) バックボーントラフィックの急増

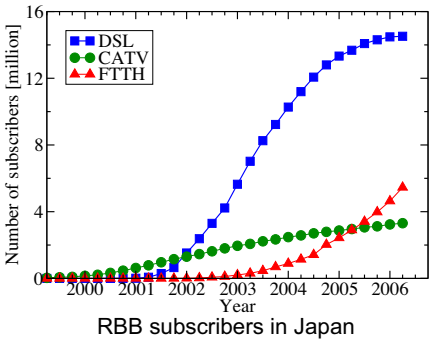
◦ **主要IXにおけるピークトラフィック** (JPIX/JPNAP/NSPIXP)
- **指数関数的増加**
- 2002**年には年率**4.5**倍、最近は**50%**程度の増加**



Traffic growth at major Japanese IXes

背景(2) ブロードバンドユーザの増加

◦ 2,330**万、**DSL:1,450**万** CATV:330**万** FTTH:550**万** (2006**年**3**月**)
- **ここ数年は**DSL**が鈍化、**FTTH**が急増**
- 2007**年中にも**FTTH**が**DSL**を抜く予想**



RBB subscribers in Japan

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, Gyao, YouTube, ...

フローベースの計測

- SNMPによるインターフェイスカウンタ値による計測の限界
  - 総量は分かるが、それ以上の情報取得が困難
- フローベースの計測
  - フロー (5-tuple) 毎の統計情報
  - プロトコル: NetFlow、 sFlow、 IPFIX、 ...
    - プロトコルバージョンや実装による違いも
- 今回はSampled NetFlow (version 5)を利用した例
  - 解析の考え方は共通

NetFlowの基本

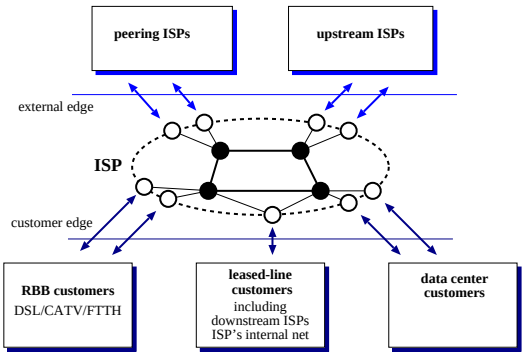
- インターフェイス毎のキャッシュ情報をUDPでコレクタに送信
- パケットがインターフェイスに到着すると
  - 新規エントリを作成
  - または、既存のエントリをアップデート
    - バイトカウント、パケットカウント、エンドタイム、TCPフラグ (ORed)
  - エクスパイア条件 (４種類)：
    - キャッシュがフル、TCP RST or FIN
    - 非アクティブフロー15秒、アクティブフロー30分
    - エクスパイアしたフローエントリはコレクタに送信される
- フロー情報
  - saddr, daddr, sport, dport, proto, ToS, input ifIndex
  - byte count, packet count, start time, end time, output ifIndex
  - TCP flags, next hop, src AS, dst AS

フロー計測の準備

- どこで測るか
  - 合算が可能か？ (ダブルカウント問題)
- 計測の精度
  - サンプリング値
    - サンプリングの影響
  - 時間粒度

どこで測るか

- (ISP)ネットワーク全体のトラフィックを測るとすると
- アップストリームが1本なら、そこで測るのが簡単
- 通常は、カスタマーエッジ または 外部エッジ(他ISPへの接続)
  - 仮定： ネットワークに入ったパケットは全て出ていく
  - リンク情報を合算して合計値を求める



結果の合算

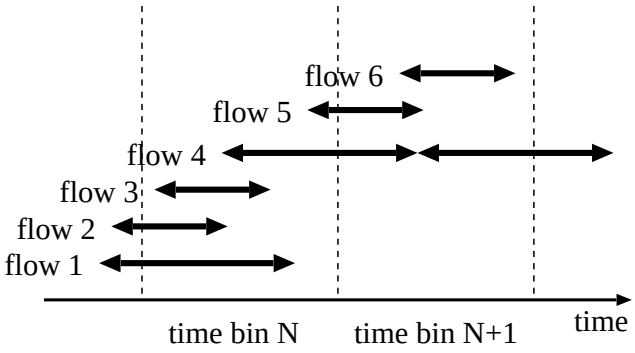
- 加算可能性
  - 平均値は加算可能
  - ピーク値等は加算できない
- ダブルカウントの問題
  - 同じパケットを複数の箇所でカウント
    - ネットワーク内部(バックボーン等)で測った値は合算できない
  - フローの両エンドがカスタマーの場合
    - 総トラフィック量の視点からはダブルカウントになる
    - カスタマートラフィック量の視点からはダブルカウントではない
  - エッジでも(構成によって)折り返しトラフィックが存在
    - 今回は量的にもわずかなので考慮していない

サンプリング値

- 考慮すべき点
  - ルータの負荷
  - データ量
  - コレクタの処理能力
- サンプリングの影響
  - 測定結果は、測定値にサンプリング値の逆数を乗じて補正
    - 使用量が大きいフローはいいが、小さいフローは精度がでない
    - 例： サンプリング値:1/100, 100ユーザがそれぞれ1KBパケットを1個送った
      - 測定結果: 100KBを送ったユーザが1人いると誤認
  - 必要な精度に応じたサンプリング値の設定が必要
    - 実際には、サンプリング値による精度の限界を理解して解析

時間粒度

- アクティブなフロー情報は30分に1度しかエクスポートされない
  - 単位時間(ビンサイズ)は小さく出来ない
- 簡単のためエンドタイムでカウント
  - より正確にはスタートタイムも使い比例割り当て



ブロードバンドトラフィック解析の実際

- 調査対象はブロードバンドユーザ
  - ファイバーとDSL接続
- 測定場所
  - ブロードバンドユーザ収容エッジルータ群
    - キャリアからは集約した回線で提供される
      - 回線種別(DSL/ファイバ)
      - 都道府県別
- サンプリング値: 2048-8192 (逆数を乗算して補正)
- データ収集: 2005年2月と7月の1週間分 (曜日による違いを見るため)
- 時間粒度: 1時間 (時間帯による違いを見るため)

解析の下準備

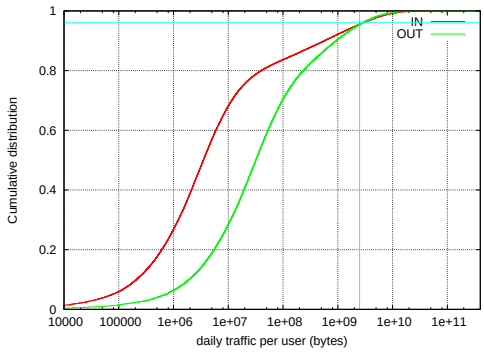
- 利用者別のトラフィック量を1時間毎に集計
  - スクリプト処理用にバイナリをテキスト形式に変換
  - 利用者毎にフロー情報を集約
    - アドレス割り当て記録を使用しユーザID(のハッシュ値)にマップ
    - IPアドレスで処理する場合
      - 管理用トラフィック等測定対象外のアドレスを持つデータをフィルタリングする必要

データ解析

- まずはユーザ毎の使用量分布を調べる
  - ヒストグラムは適切なサンプル数、ビンサイズが必要なため使い難い
    - ビンが小さ過ぎると空のビンができて分かり難い
    - ビンが大き過ぎると情報量小
- 累積分布 (Cumulative Distribution Function, CDF)
  - 分布密度関数: 事象 x を観測する確率
    f(x) = P[X = x]
  - 累積分布関数: x またはそれ以下の事象を観測する確率
    F(x) = P[X <= x]
  - サンプル数が少ない場合や極端な値がある場合にも有効
- 相補累積分布 (Complementary Cumulative Distribution Function, CCDF)
  - CDFの補数
    F(x) = 1 - P[X <= x]
    - 分布のテイルやスケール特性を見るのに有効

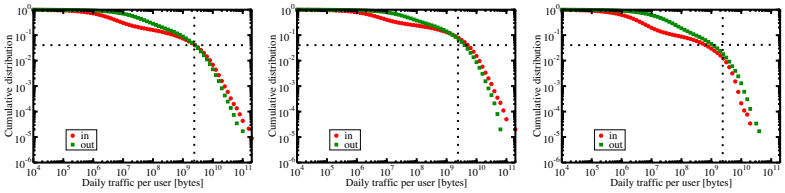
トラヒック使用量に対するユーザの累積分布

- ユーザ毎のトラフィック: 各ユーザの1週間分のトラフィックから1日平均の使用量を得る
  - ユーザID、1日平均INトラフィック量、1日平均OUTトラフィック量のテーブルを作り、使用量順にソート



トラヒック使用量に対するユーザの相補累積分布

- ヘビーユーザに注目
- ヘビーユーザは広範囲にわたり統計的に分布
  - べき分布 (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%

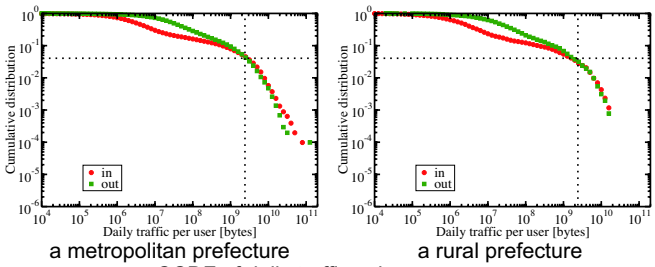


total　　fiber　　DSL

CCDF of daily traffic volume per user

地域別のトラヒック使用量

- 回線の都道府県属性で分類して、同様に処理
- トラヒ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

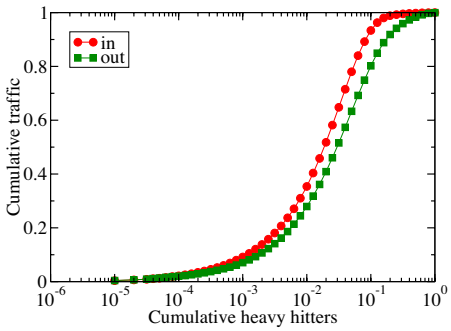


a metropolitan prefecture　　a rural prefecture

CCDF of daily traffic volume per user

ヘビーユーザへのトラヒックの偏り度合

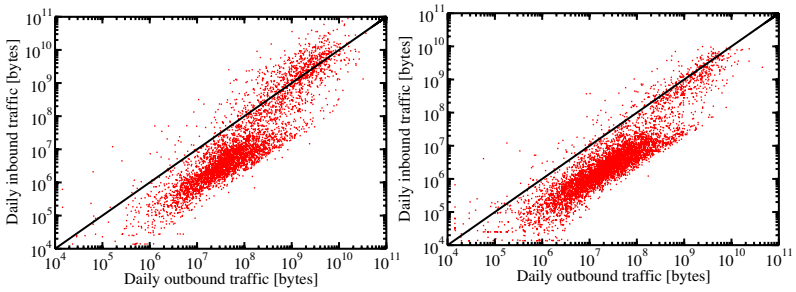
- 使用量上位何%のユーザが全体トラヒックの何%を占めるか
  - 使用量の多い順にソートしてCDFを作る
- トラフィック使用量に大きな偏り
  - IN側：上位4%が全体の75%を占める
  - OUT側：上位4%が全体の60%を占める



Cumulative Traffic of heavyhitters

ユーザのIN/OUTトラヒックの相関

- 各ユーザの1日平均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:一般ユーザ層 (ダウンロード中心)
  - 対角線上右上のクラスタ: ヘビーユーザ (IN/OUT対称)
- ファイバとDSLは同様の傾向
  - 質的な違いはない、単にヘビーユーザ比率が違う
- ここでもヘビーユーザと一般ユーザの境界はあいまい



fiber DSL

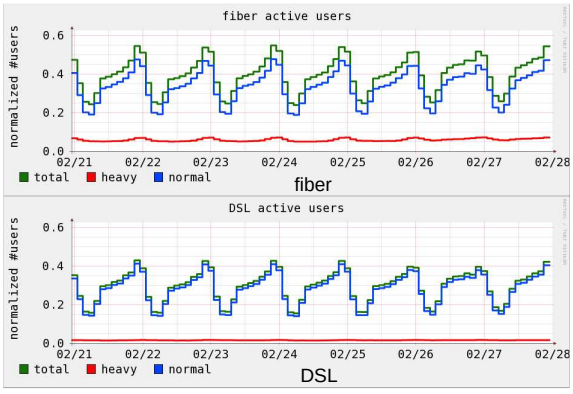
daily per-user inbound/outbound volumes

ヘビーユーザとその他のユーザの区別

- 1日平均INトラフィック2.5GB以上のユーザ
  - まずユーザを分けて、それを元に時系列データを再分類
- トラフィック挙動の違いを見る
  - 1週間のアクティブなユーザ数
  - ファイバ/DSL週間トラフィック

1週間のアクティブなユーザ数

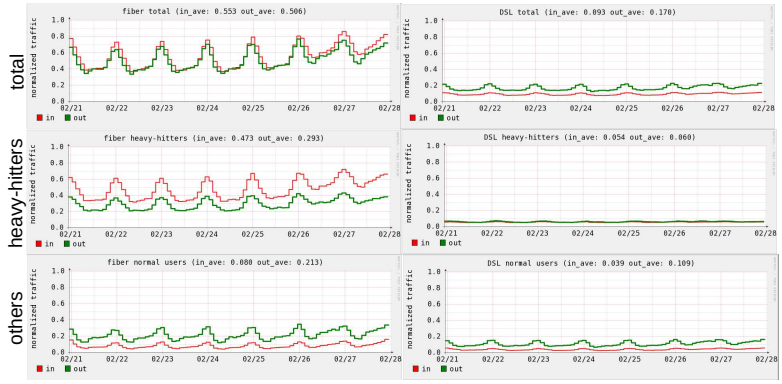
- ファイバ/DSL合計のピーク値に正規化(絶対値を出さないように)
  - RRDtoolで時系列グラフ化
- アクティブなユーザ数はファイバよりDSLが少し多い
- ヘビーユーザ数はほぼ一定



Normalized number of active users

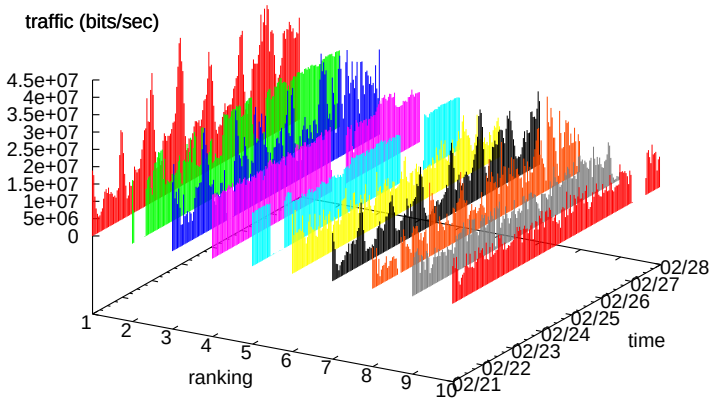
ファイバ/DSL週間トラフィック

- これも全体ピーク値に正規化
- INをみると86%がファイバから、DSLはわずか14%
- 全体はファイバのヘビーユーザの影響大



トップ10ヘビーユーザのアップロード挙動

- 1時間粒度、1週間のトラフィック
  - 挙動の違い、異なるアプリケーションの利用



プロトコル/ポート

- 元のNetFlowデータから直接集計
  - sport/dportの小さい方をカウント
- 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**			

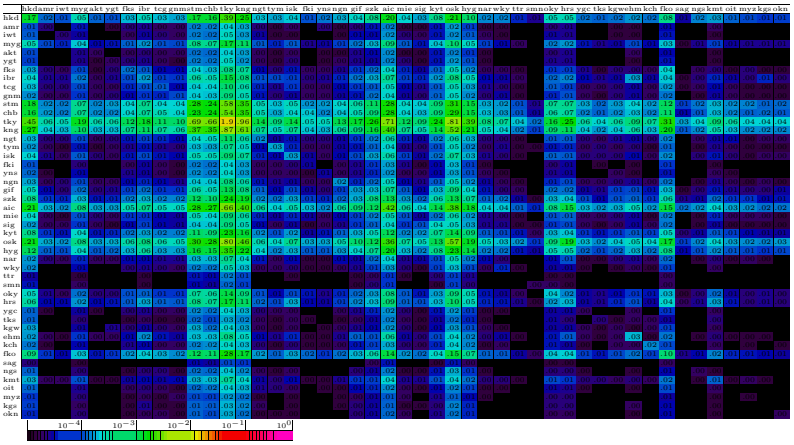
地域別トラフィックマトリクス

- 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
    - サイバーエリアリサーチ SUTFPOINT: ISPの一般ユーザ用アドレスを都道府県にマップ
    - Digital Envoy Netacuity: それ以外を国内国外に分類
- 62%は一般ユーザ同士の通信
- 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?

$src \backslash dst$	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

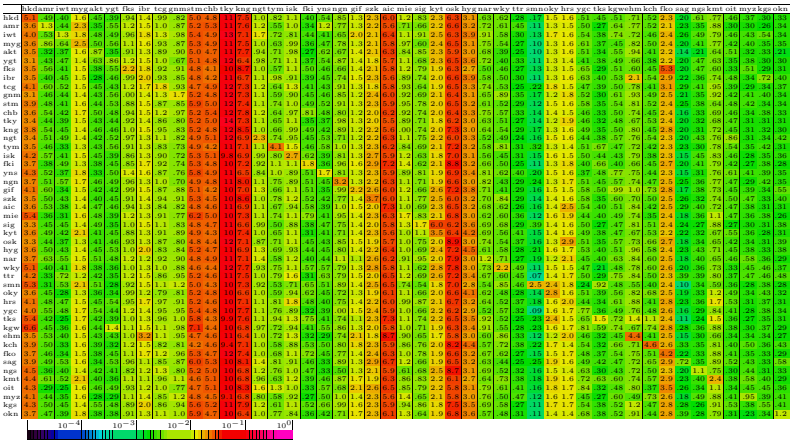
都道府県別トラフィックマトリクス

- 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)
    - 図はTeXのcolortblを使用
  - 人口に応じたトラフィック量分布



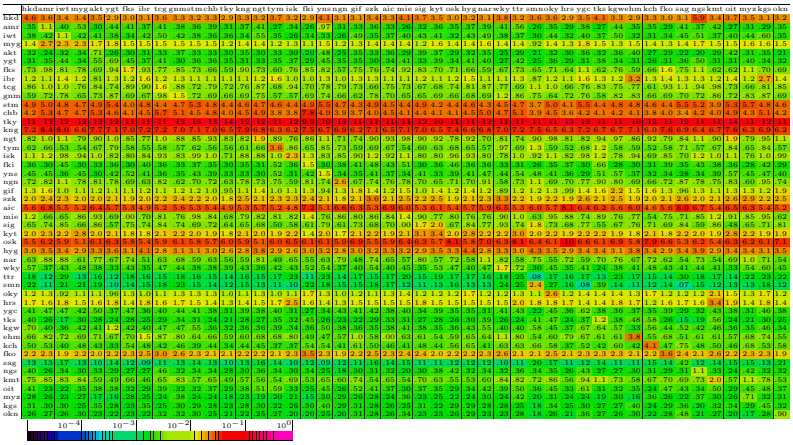
ソースに正規化したマトリクス

- 各行の合計が100%: 各県からの送出分布を見る
- 各県共通の分布、地域性がない、県内トラフィックは1-3%程度



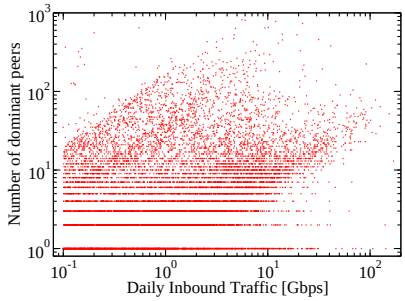
デスティネーションに正規化したマトリクス

- 同様の結果



ユーザ毎のドミナントピア数

- 使用アプリケーションタイプを見るため通信相手数を調査
  - ドミナントピア数: 50%トラフィックを占める上位ピア数をカウント
    - web広告等の影響を低減
- 調査前には2種類のピア数分布を予測
  - ストリーミングやダウンロード中心のユーザは小数のピア数
  - P2Pファイル共有を利用するユーザは多くのピア数
- ピア数分布結果からはユーザ層の区別ができない



まとめ

- フローベースのトラフィック解析の実際
  - 某ISPのブロードバンドトラフィック解析を紹介
- 解析はやり方が分かれば簡単に実行できる
  - 一方で運用部隊だけではなかなか解析まで手がまわらない
  - ある程度公開できる統計情報に加工する工夫
- 以下、時間があればブロードバンドトラフィックに関して少し議論

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

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

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

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
  - コモンズの悲劇
    - 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
- 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

ブロードバンド第二幕への課題

- ブロードン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