Socioeconomic issues in the Internet of the future:
the case of congestion control

Prof. Costas Courcoubetis
Athens University of Economics and Business
courcou@aueb.gr

Joint work with A. Dimakis, B. Briscoe in the context of Trilogy FP7 Project

http://www.bobbriscoe.net/present.html

trilogy
re-architecting the Internet

the neck of the hourglass, for control

www.trilogy-project.eu

This work is partly funded by Trilogy, a research project (ICT-216372) supported by the European Community under its Seventh Framework Programme. The views expressed here are those of the author(s) only. The European Commission is not liable for any use that may be made of the information in this document.

Some questions

- What are the challenges in building the future Internet?
- Just build “good” technology?
- **Basic question**: is good engineering enough?
  - What is “good” engineering?
  - Why bother with socio-economics?

Good engineering (D. Clark):

- Development of models, techniques, tools that deliver *predictable desired behaviour*
- Criteria: robustness, scalability, stability…

But Internet is more a mirror of the societies in which it operates!

- Criteria: dynamic management of evolving and conflicting interests, *no prediction*
- *Technical architecture must accommodate tussles in society AND traditional eng. goals*

The Internet landscape

- Many different stakeholders: users, ISPs, private sector networks, governments, content providers, …
- Unsurprisingly different stakeholders exhibit different conflicting interests, continuously contending in the Internet ecosystem (tussle)
- The Internet architecture and its protocols only rarely allow the involved parties to resolve in-band conflicts of interest.
- Instead they violate many times the architecture to achieve their ends, compromising the intended benefits
- Can we correct that in the future development of the Internet?

CNET News
February 22, 2008 12:00 PM PST

Comcast vs. BitTorrent to be focus of FCC hearing

By Anne Broache
Staff Writer, CNET News

The high-profile squabble over Comcast's slowdown of BitTorrent file-sharing traffic--and broader questions of network handling by Internet service providers--is set for public scrutiny Monday at a federal hearing.

This time, the Federal Communications Commission will depart its headquarters just off the National Mall in Washington and head north to a courtroom on Harvard Law School's campus in Cambridge, Mass. (The FCC wouldn't comment on why the site was selected, but Boston is the home turf of Democratic Rep. Ed Markey, who chairs a House Internet subcommittee.)

The hearing, which will be open to the public on a first-come, first-served basis and be otherwise accessible via an "audio-only" Webcast on the FCC site. It's an outgrowth of the agency's recently launched inquiry into what constitutes "reasonable" network management practices by Internet service providers.

Bittorrent declares war on VoIP, gamers

The next internet meltdown

By **Richard Bennett** • **Get more from this author**

Posted in Networks, 1st December 2008 12:29 GMT

Gamers, VoIP and video conference users beware. The leading BitTorrent software authors have declared war on you - and any users wanting to wring high performance out of their networks. A key design change in the P2P application promises to make the headaches faced by ISPs so far look like a party game. So what's happened, and why does it matter?

Upset about Bell Canada's system for allocating bandwidth fairly among internet users, the developers of the uTorrent P2P application have decided to make the UDP protocol the default transport protocol for file transfers. BitTorrent implementations have long used UDP to exchange tracker information – the addresses of the computers where files could be found – but the new release uses it in preference to TCP for the actual transfer of files. The implications of this change are enormous.

More in this article:

- Is bypassing TCP congestion control a good thing for the users of the network?
- Why should one persons [sic] non-interactive file sharing generating a dozen to a hundred streams be more important than my interactive VoIP call or gaming experience?
- Using it as a feature, maybe, but enabling this behavior by default is just wrong and will lead to continuing counter, counter measures and more justification for caps.

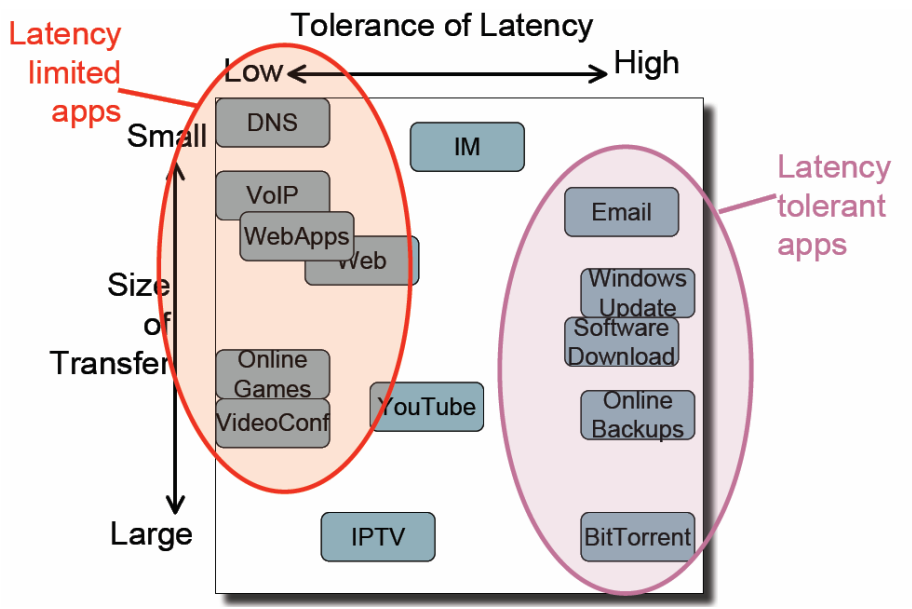
- Economics is needed to understand the wider system interactions and conflicts
  - Eng.+ Economics: Internet + stakeholders
- Engineering solutions can not be applied by force but through incentives
  - Must allow stakeholders to express their economic preferences and interact
  - Should not preclude the outcome of the tussle
- Game theory instead of optimization theory

Game model of tussle

- Architecture defines the rules of the game
- Stakeholders use these rules in the most appropriate way for them
- **Operating point** = equilibrium of the game
- Examples: Congestion control, Google, …

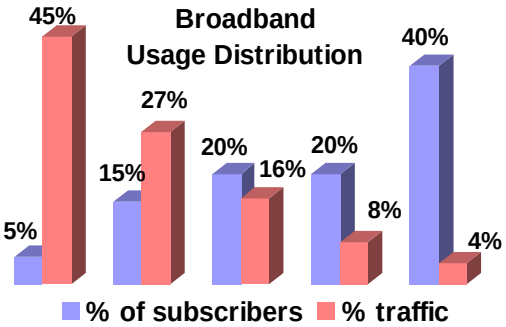
An example: the Congestion Control tussle in the Internet

- Internet bandwidth sharing: TCP, UDP,...
- What congestion?
  - Packet loss (<1%)
  - Latency
    - Packet delay
    - **Transfer completion**

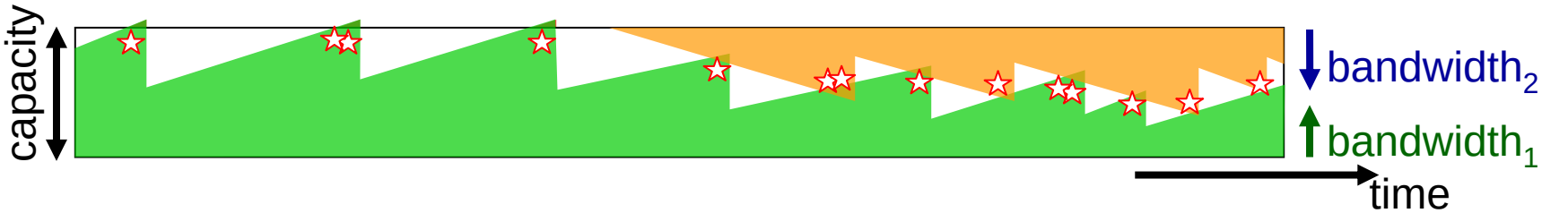


What is the problem?

Control BW sharing at run time, not design time

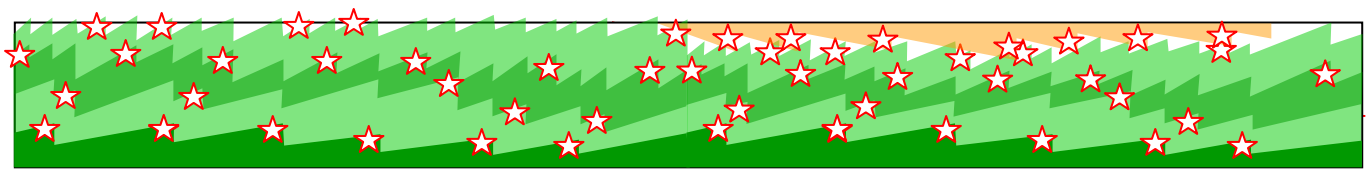


Voluntarily polite algorithm in endpoints, 'TCP-friendliness'



(VoIP, VoD Joost 700kbps)

Taking all and holding your ground pays!



Start more 'TCP-friendly' flows than anyone else (Web: x2, p2p: x5-100)
or for much longer than anyone else (file transfer x200)
net effect of both (p2p: x1,000-20,000 higher traffic intensity)

The ISP's problem

- How to maintain good service for web browsing while providing acceptable service to p2p
- Adding capacity to the network does not solve the problem!
- Use DPI technology: try to limit (throttle) traffic that causes the problem
  - But how can you see which traffic causes the problem?
  - Even users that use "nice" software are victimized regardless if they have exceeded a volume threshold
  - How can ISPs choose on behalf of their users which traffic is important?
  - Strong incentives to favour own applications

And the tussle begins…

- ISPs charge flat rate (+ volume caps)
- P2p open N x TCP, rarely inactive
- Web, VoIP and gaming loose against p2p
- ISPs use DPI to throttle p2p traffic
- P2p reacts by changing ports, encrypting
- DPI detects p2p by detecting specific traffic patterns, prioritize traffic that is “friendly”
- P2p traffic adapts…
- FCC responds with “Open Internet” rules
- Bad side-effects:
  - New applications with non-friendly traffic patterns may be killed by DPI boxes: lock in to today’s apps
  - Why not use DPI boxes to de-prioritize Skype?
- What is wrong in all this?

FCC “Open Internet” Notice of Proposed Rulemaking

“Subject to reasonable network management, a provider of broadband Internet access service **may not**:

- prevent any of its users from sending or receiving the lawful content of the user’s choice over the Internet;
- prevent any of its users from running the lawful applications or using the lawful services of the user’s choice;
- prevent any of its users from connecting to and using on its network the user’s choice of lawful devices that do not harm the network; or
- deprive any of its users of the user’s entitlement to competition among network providers, application providers, service providers, and content providers.”

“The draft non-discrimination principle would require that, subject to reasonable network management, a provider of broadband Internet access service must treat lawful content, applications, and services in a non-discriminatory manner.

The draft transparency principle would require that, subject to reasonable network management, a provider of broadband Internet access service must disclose such information concerning network management and other practices as is reasonably required for users and content, application, and service providers to enjoy the protections specified in this rulemaking.”

Some questions:

- Is the concept of TCP fairness sensible?
- Is flat rate pricing or volume pricing sensible?
- Are DPI boxes sensible?
- Is e2e sensible?
- Is BitTorrent a bad algorithm?
- What will happen with new Internet apps (Internet TV,…)?
- Is there something fundamentally wrong with the current Internet architecture?
  - provides the appropriate economic info to agents?
  - incentives are correctly set?
- All these are economic questions!

Need for new criteria besides technical correctness

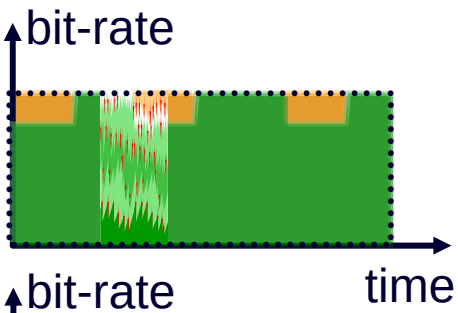
- Incentives of players to comply to the solutions
- Promote competition, low barriers of entry
- Do not force specific outcome in the market
- Incentives for investment, innovation
- Allow for conflicts to be resolved “in” the architecture, not “outside”
- Isolate tussles, avoid spillover
- More socio-economic criteria…

Back to the congestion tussle: a solution justified by economics: charge according to congestion volume! (Kelly 1998)

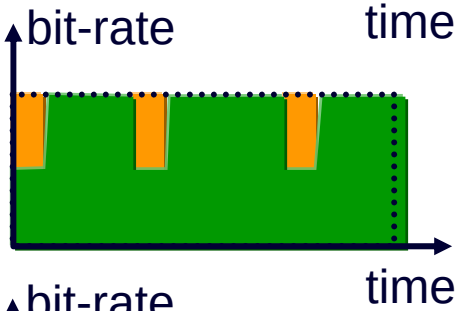
- Simple model:
  - Charge each user for the “pollution” he creates
  - Users choose a weighted TCP version
  - Packet flows are accounted for the amount of congestion they collect (congestion marks)
- Provides the right incentives
  - Interactive users **choose** to act more aggressively since latency is important
  - Volume users **choose** to act less aggressively
- Accounting problem: solved (B. Briscoe: re-ECN)

A summary of the approaches

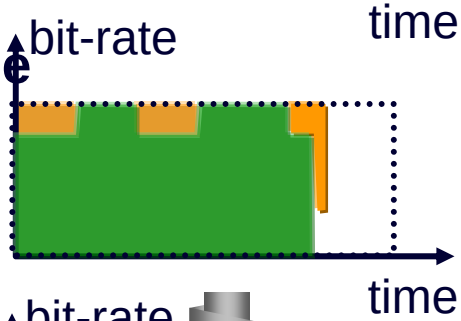
1. TCP



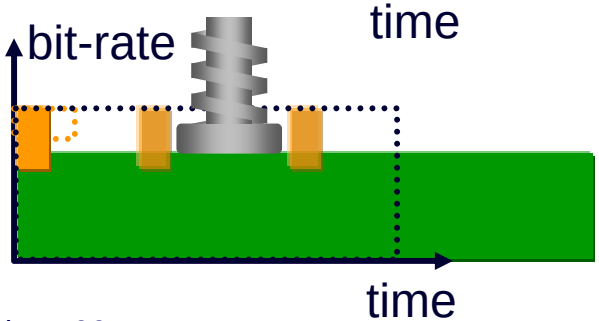
2. WFQ



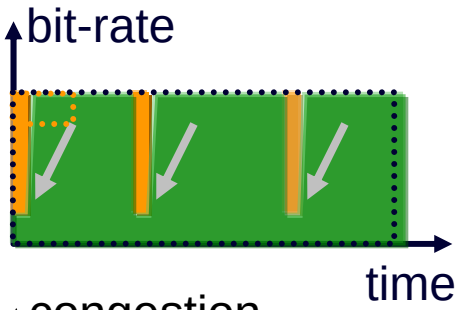
3. volume cap



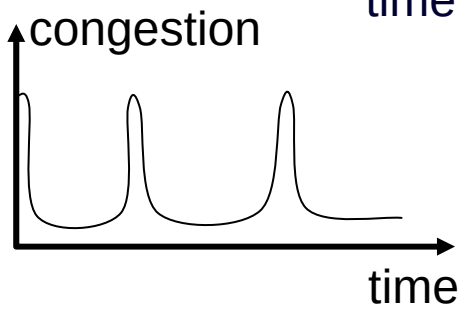
4. DPI



weighted sharing



congestion volume accounting



- light usage can go much faster
- hardly affects completion time of heavy usage
- Needs congestion volume accounting per flow to be incentive compatible

NOTE: weighted sharing doesn't imply differentiated network service, just weighted aggressiveness of end-system's rate response to congestion cf. LEDBAT

- Why is this architecture better?
  - Does not make important economic decisions on behalf of its stakeholders
  - Allows customers to control their own network experience
  - Enables continued Internet evolution
    - The ISPs respect the network neutrality rules, offer transparency of network operation to current and future applications
    - No “hide and seek” game between ISPs and specific protocols
  - Helps ISPs control better their costs and increase effectiveness of adding more resources
  - DPI boxes: police congestion allowances of users

More questions we like to answer

- Is charging by congestion marks efficient?
- What is the impact of different ISP tariff structures to user-perceived performance and ISP profits?
- How users choose congestion control parameters as a result of the ISP tariffs?
- Which tariffs are more competitive for ISPs?
- Is there a need for volume charging?

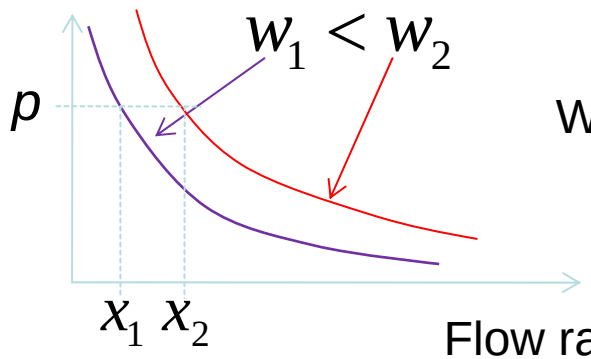
The model of the Internet:

- Involves (from fast to slow time scales):
  - Congestion control mechanisms (fast time scales)
  - Flow initiations & terminations (duration: affected by cc)
  - User demand for downloads (how often)
  - Tariff choice by ISP (charge for congestion marks, volume)

- Not yet considered, but easily modeled:
  - Network dimensioning
  - Competition between ISPs

A Congestion Control model

- Abstraction: Hide packet scale effects, consider average rate allocations after initial transient, for fixed # of active flows.
- Kelly: congestion control="demand" function, price = rate of congestion signals
  - E.g., for TCP: $x^* = \frac{c}{T\sqrt{p}} \Leftrightarrow U(x) = (c/T)^2(\alpha - 1/x)$



Weighted TCP: $x_{TCP}^{w_i}(p) = \frac{w_i c}{T\sqrt{p}}$

Sharing capacity C:

$$x_{TCP}^{w_1}(p) + x_{TCP}^{w_2}(p) = C$$

- Congestion volume charging: charge q/congestion signal generated

The time scales

- Time scale separation

$$x^w(p) = wx(p)$$

In equilibrium *p:*

$$\sum_{i=1,\ldots,n} x^{w_i}(p) = C$$

Users measure **average net value/bit** and choose non-cooperatively

w_j, λ_j

ISP fixes price/mark **q** and price/bit **v**

Fast time scale
(packet RTT)
State: # of flows

Medium time scale
(many file arrivals and completions)

Slow time scale
ISPs choose tariffs
Customers switch ISPs

Other example of congestion controllers

- Linear congestion controller
  - Corresponds to $U(x) = wx$
  - Algorithm: $\dot{x} = kx(w - p(\sum x))$

  - Flow with higher weight gets all the bandwidth
  - Price per unit of flow=w
  - Being more aggressive is more expensive
- Proportionally fair congestion controllers

Which congestion control weight w_i ?

- Users act non-cooperatively
- Learn from experience: average download delay, average charge incurred
- Maximize net benefit per bit =

$$r_i - c_i d_i - m_i$$

Delay sensitivity constant

average delay

average congestion charge

Affected by choice of congestion control

How much demand?

- Demand = frequency of file downloads
- Assumption: users would want to download more as long as they currently derive positive benefit, maximum demand exceeds capacity
- As demand increases, delay increases unboundedly (net benefit becomes negative).
- (Noncooperative) equilibrium demand
- At equilibrium (Wardrop):

$$r_i - c_i d_i - m_i = 0 \,, \qquad \forall i$$

Equilibrium definition

- Demand vector: $\lambda = (\lambda_i)$
- Congestion control vector: $\pi = (\pi_i), \pi_i \in \mathcal{P}(\mathcal{U})$

Probability distributions over set $\mathcal{U}$ of utility functions.

Equilibrium: (λ, π) *for which*

1. $r_i - c_i d_i(\lambda, \pi) - m_i(\lambda, \pi) \leq 0$

   If strict inequality, $\lambda_i = 0$.

2. π_i *maximizes* $r_i - c_i d_i(\pi'; \lambda, \pi)) - m_i(\pi'; \lambda, \pi)$ *over* $\pi' \in \mathcal{P}(\mathcal{U})$

Goodness of equilibrium?

- Social Welfare: “sum” of benefits over all flows

$$\sum_i \lambda_i (r_i - c_i d_i)$$

Demand of type *i* downloads

- Efficient equilibrium = attains maximum SW

- All this in the context of 2 user types
  - Type 1: fixed delay sensitivity $c_1 = const$
  - Type 2: delay sensitivity c_2 tends to 0

Single user type

- Assume q=1
- Inefficient equilibrium when all use proportionally fair (or TCP) congestion controllers.
- Unique efficient equilibrium when all use linear congestion controllers.

Changing q?

- Can we induce efficiency by changing the price per mark?
- No! the price/mark has no economic significance!
  - Equilibrium load and charge/bit are invariant in q
  - Affects only equilibrium choice of aggressivity parameters w
- Offers no extra degree of freedom in controlling efficiency
- Need a volume charge

Two user types

- Web-browsing: use linear c.c.
- Off-line: use w-prop. fair c.c.

$$c_1 = const, c_2 -> 0$$

- Efficient allocation: web have full priority over off-line.
- Must have m_2 =0 (If m_2 >0 off-line interfere with web)

$$r_2 - c_2 d_2 - m_2 = 0 \Rightarrow r_2 - c_2 d_2 = 0$$

(tragedy of the commons)

Off-line delays become soon too large!

Additional volume bit charge makes system efficient

- Price per bit = $v = r_2 - \epsilon$
- Net benefit = $r_2 - c_2 d_2 - m_2 - v = \epsilon - c_2 d_2 - m_2 = 0$
- Off-line mark rate is negligible.
  - No interference with web flows
- And net value per bit is positive:

$$r_2 - c_2 d_2 \geq r_2 - \varepsilon >> 0$$

- Does not contradict Kelly's congestion mark charging world; it extends it.

Conclusions

- Economic analysis is necessary when choosing new technologies for the Internet
- Good engineering must be redefined, solutions must be incentive-compatible, interest-neutral
- Examined the case of congestion control and discussed a tussle-neutral technology (charging for congestion)
- Proposed a model with time scale separation for studying the non-cooperative behaviour of Internet users using the above technology
- Analyzed the role of pricing
  - Role of congestion charge: incentivize users to choose the appropriate congestion control
  - Role of volume charge: incentivize off-line users to congest the system less

Thank you!!!