Collective delusions behind how capacity gets shared

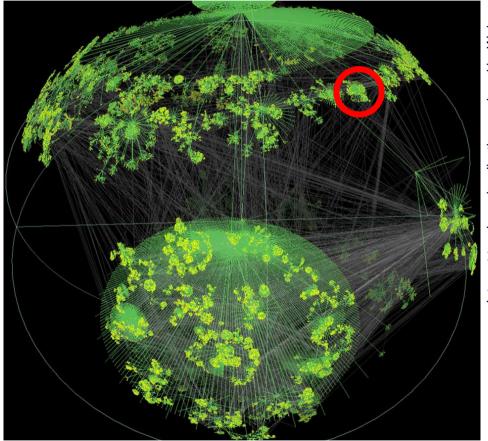


Bob Briscoe with Toby Moncaster & Lou Burness presented by Dirk Trossen Jan 2008



freedom to limit the freedom of others?

- tremendous idea
 - anyone can use any link anywhere on the Internet without asking
- when any link is overused
 - who decides how big a share each gets?
 - 1. TCP
 - 2. Comcast
 - 3. The Oval Office



for scale: ~10M lines ringed in red



fair bottleneck bit-rate?

two incompatible partial worldviews

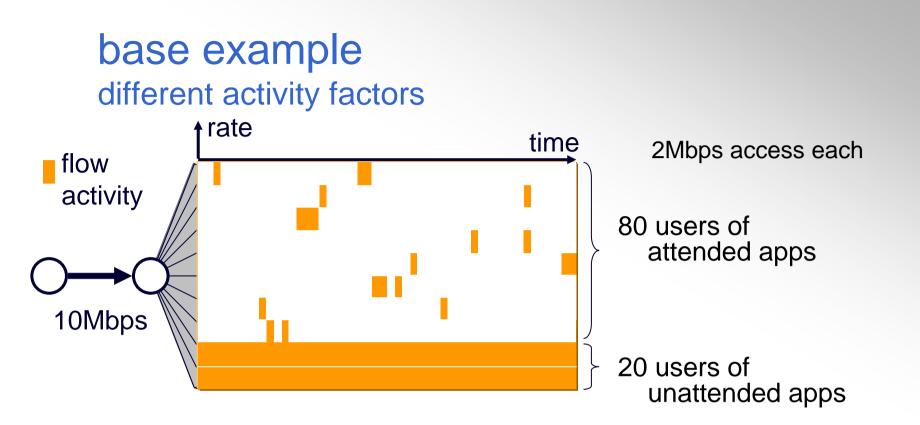
the Internet way (TCP) operators (& users)

'flow rate equality'	'volume accounting'
per data flow	per user
instantaneous	over time

this talk

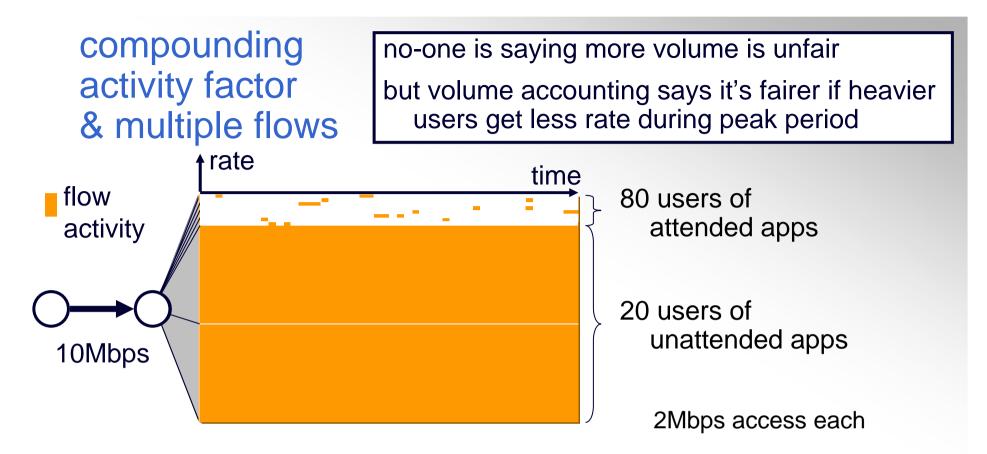
- status report on our attempts to unveil multiple delusions
- the standards and research community's double delusion
 - TCP's equal flow rates are no longer fair at all (by any definition)
 - TCP protocol increasingly doesn't determine capacity shares anyway





usage type	no. of users	activity factor	ave.simul flows /user	TCP bit rate /user	vol/day (16hr) /user	traffic intensity /user
attended	80	5%	=	417kbps	150MB	21kbps
unattended	20	100%	=	417kbps	3000MB	417kbps
				x1	x20	x20

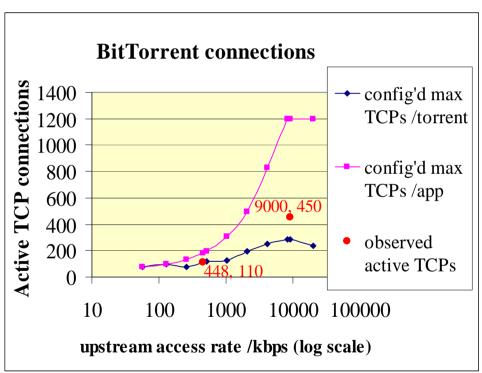




usage type	no. of users	activity factor	ave.simul flows /user	TCP bit rate /user	vol/day (16hr) /user	traffic intensity /user
attended	80	5%	2	20kbps	7.1MB	1kbps
unattended	20	100%	50	500kbps	3.6GB	500kbps
				x25	x500	x500



realistic numbers? there are elephants in the room



- number of TCP connections
 - Web1.1: **2**
 - BitTorrent: ~100 observed active
 - varies widely depending on
 - no. of torrents per user
 - maturity of swarm
 - config'd parameters

details suppressed:

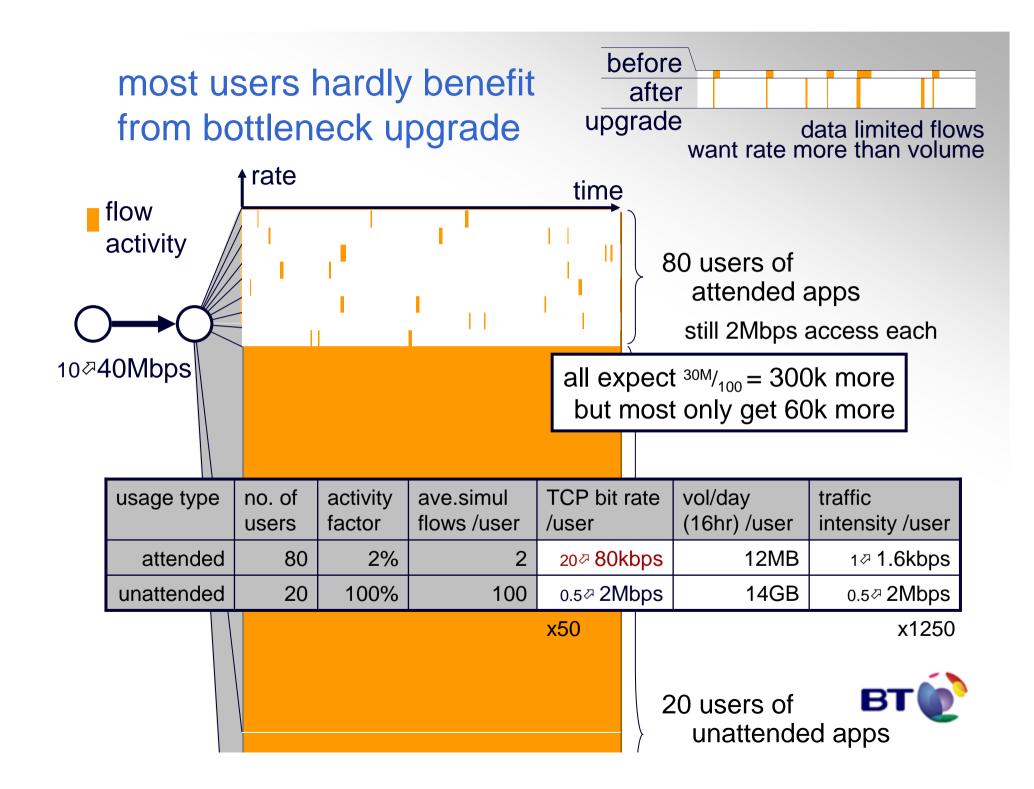
- utilisation never 100%
 - but near enough during peak period
- on DSL, upstream constrains most p2p apps
 - other access (fixed & wireless) more symmetric



typical p2p file-sharing apps

• 105-114 active TCP connections altogether

4	Azureus											8
Fil	ile Transfers To	orrent View Tools	Plugins H	Help								
	A A 6	2004	00) 🕢 🕨 🖬 🗶								
			~~~~~				_					
My	y Torrents	00.0% : Nigella Expres	s S01E0	100.0% : Atom 67.1% : N	Vigella Expr	ress S01E07 🖇	3					
	General Peers	Swarm Pieces Files	s Info	Options Console Geo Map								
	IP	Client	Т	Pieces	%	D 🐨	Up Speed	State	Encryption	Down	Up	I 🗠
1.00	78.86.8.10	Azureus 3.0.2.2	L		100.0%	14.5 kB/s		Fully established	RC4-160	6.87 MB	25.8 kB	
	76.65.28.192	µTorrent 1.7.5	R		100.0%	11.1 kB/s	S.S. 2.S.	Fully established	None	10.52 MB	14.6 kB	
1 of 3 torrents shown	.199	Azureus 3.0.3.4	- L -		100.0%	10.7 kB/s	10.00	Fully established	RC4-160	7,24 MB	26.6 kB	
I UI S IUITEITIS SHOW		Azureus 3.0.2.2	- L -		100.0%	18.8 kB/s	0.000	Fully established	RC4-160	18.91 MB	59.8 kB	
	.114	Mainline 6.0.0	R		100.0%	11.8 kB/s	12.02.0	Fully established	None	8.12 MB	12.1 kB	
15 TCPs por torro	nt 17	µTorrent 1.7.5			100.0%	13.5 kB/s	10000	Fully established	RC4-160	7.16 MB	11.2 kB	
<ul> <li>~45 TCPs per torre</li> </ul>	16	µTorrent 1.7.5	L		100.0%	6.8 kB/s		Fully established	RC4-160	5.58 MB	9.4 kB	
		µTorrent 1.7.5	R L		100.0%	9.0 kB/s 9.6 kB/s	100.000	Fully established	RC4-160 RC4-160	4.85 MB 8.43 MB	8.6 kB 12.4 kB	
<ul> <li>but ~40/torrent activity</li> </ul>	Ve 99	µTorrent 1.7.5 µTorrent 1.7.5	R			9.6 KB/S 12.1 kB/s	22.02.0	Fully established Fully established	RC4-160 RC4-160	5.30 MB	12.4 KB 8.3 kB	
	22	Contraction of the state of the	L		100.0%		1.000	0.04.04.04.04.04.04.04.04.04.04.04.04.04		2000 C C C C C C C C C C C C C C C C C C		
	1.3.20000	µTorrent 1.7.5 µTorrent 1.7.5	R		100.0%	7.4 kB/s 6.5 kB/s	1000	Fully established Fully established	RC4-160 RC4-160	6.59 MB 4.27 MB	10.5 kB 8.1 kB	
	66.214.134.174		L		100.0%	8.0 kB/s		Cost 23. Second Se	RC4-160 RC4-160	4.27 MB 4.91 MB	о.1 кв 8.9 kB	
	24.108.88.117	µTorrent 1.6.0 µTorrent 1.7.2	R		100.0%	12.0 kB/s		Fully established Fully established	None	4.91 MB 8.91 MB	0.9 KB 12.9 kB	
	87.194.119.77	uTorrent 1.7.3	L		100.0%	7.7 kB/s		Fully established	RC4-160	5.43 MB	9.3 kB	
	121.45.133.231	µTorrent 1.7.5	R		100.0%	7.7 kB/s	2012	Fully established	None	2.54 MB	5.1 kB	
	220.245.217.58	KTorrent 2.2	L		100.0%	5.8 kB/s		Fully established	RC4-160	5.15 MB	9.5 kB	
11.53	124.102.103.7	µTorrent 1.7.5	R		100.0%	6.0 kB/s	13 B/s		RC4-160	6.17 MB	10.0 kB	
	121.45.153.84	µTorrent 1.7.5	L		100.0%	4.8 kB/s		Fully established	RC4-160	5.29 MB	9,2 kB	
	121110.100.01	int 1.7.5	R		100.0%	4.9 kB/s	12 B/s	Fully established	RC4-160	2.08 MB	5.9 kB	
		nt 1.6.1	L		100.0%	4.4 kB/s	0.000	Fully established	RC4-160	5.01 MB	8.9 kB	
environment		JS 3.0.2.2	R		100.0%	4.3 kB/s	26 B/s	Fully established	None	1.28 MB	6.1 kB	
		nt 1.7.5	L		100.0%	4.8 kB/s	669000 m	Fully established	RC4-160	3.79 MB	7.6 kB	
Azureus BitTorren	t app	nt 1.7.5	i —		100.0%	4.7 kB/s	15 B/s	Fully established	RC4-160	3.13 MB	6.8 kB	
	app	et 0.93	i i		100.0%	3.8 kB/s	1010000	Fully established	RC4-160	2.85 MB	6.5 kB	
ADSL+ 448kb ups	troom	e 6.0.0	R		100.0%	4.6 kB/s	10 B/s		None	2,54 MB	5.3 kB	
ADOL+ 440KD ups	ueam	nt 1.6.1			100.0%	3.2 kB/s	1.5.5.5.5.	Fully established	RC4-160	5.89 MB	9.7 kB	
OS: Windows XP			i —		100.0%	4.7 kB/s	12 B/s	Fully established	RC4-160	3.00 MB	6.7 kB	
	P10 3F	L nt 1.7.5	L		100.0%	3.4 kB/s	10 B/s	Fully established	RC4-160	2.02 MB	5.8 kB	
	07.434.447.144	HZUTEUS 3.0.2.2	L		100.0%	3.8 kB/s	30 B/s	Fully established	RC4-160	2.05 MB	10.7 kB	6
	000 171 64 174				100.002	L.D	0.02-		DC4 100	E OC MD	0.710	0.000
8	<u> </u>											2
P	Piece Map Consol	e										-
	🔲 Both have 🦳 Peer has; You don't 💼 You have; Peer doesn't 🔄 Neither has 📁 Transferring 📰 Next Request 📰 Availability Count											
Azi	Azureus 3.0.2.2					* 2.2 kE						
	💾 start	Dinbox - Micr.	2	Azureus 🛛 🙆 2 Firefox	- A	Aicrosoft O	1 📥 2 W					-
	Stant	Mindox-Strift.		AZUICOS	1 P	ner osor e os n			Province of the left		203	- aprica



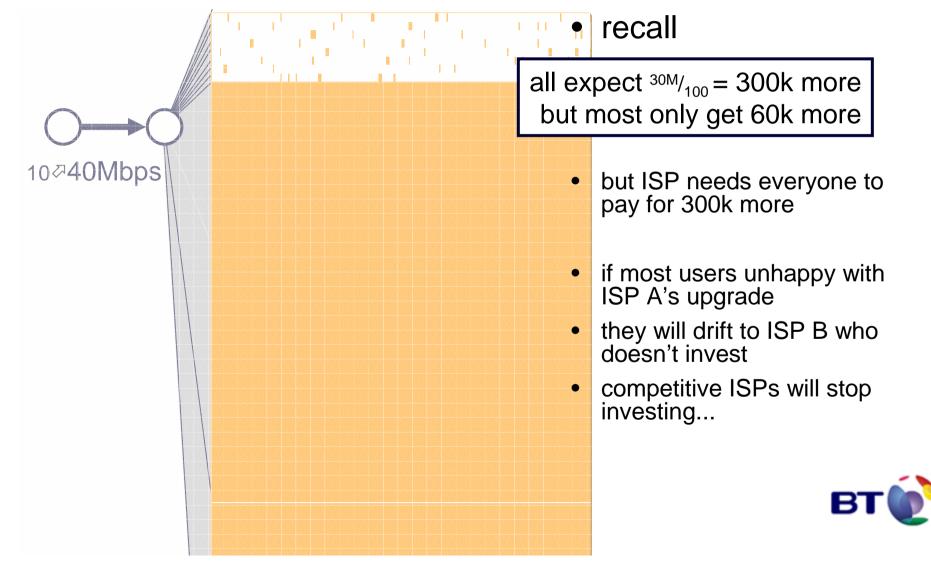
# so what?

- fairness can't be such a problem, the Internet works
  - we all have enough most of the time, even if A has more than B

- Internet technical community likes to think this is due to its protocols
- next few slides cast doubt on this complacency



#### concrete consequence of unfairness #1 higher investment risk



## ...but we still see enough investment

- main reasons
  - subsidies (e.g. Far East)
    - light users get 'enough' if more investment than they pay for
  - weak competition (e.g. US)
    - operators still investing because customers will cover the costs
  - throttling heavy users at peak times (e.g. Europe)
    - overriding TCP's rate allocation



#### concrete consequence of unfairness #2 trend towards bulk enforcement

- as access rates increase
  - attended apps leave access unused more of the time
  - anyone might as well fill the rest of their own access capacity
- operator choices:
  - a) either continue to provision sufficiently excessive shared capacity
  - b) or enforce tiered volume limits

see CFP white paper "Broadband Incentives"



# so the Internet way was wrong and the operators were right?

• no, both were part right, part wrong

	<b>)</b> ( )	
degree of freedom	'flow rate equality'	'volume accounting'
multiple flows	×	$\checkmark$
activity factor	×	$\checkmark$
application control	$\checkmark$	×
congestion variation*	$\checkmark$	×

the Internet way (TCP) operators (& users)

*another story

• both sides are failing to understand the strengths of the other



#### concrete consequence of unfairness #3 networks making choices for users

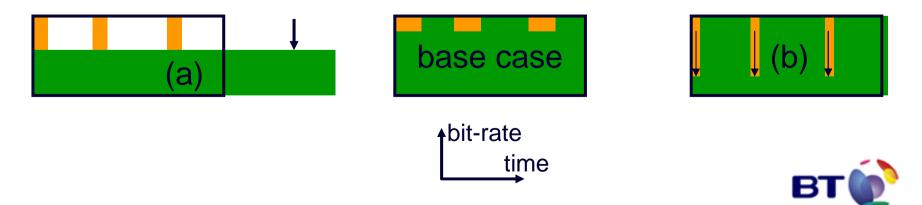
- characterisation as two user communities over-simplistic
  - heavy *users* mix heavy and light *usage*
- two enforcement choices
  - a) bulk: network throttles all a heavy user's traffic indiscriminately
    - encourages the user to self-throttle least valued traffic
    - but many users have neither the software nor the expertise
  - b) selective: network *infers* what the user would do
    - using deep packet inspection (DPI) and/or addresses to identify apps
- even if DPI intentions honourable
  - confusable with attempts to discriminate against certain apps
  - user's priorities are task-specific, not app-specific
  - customers understandably get upset when ISP guesses wrongly



# there are better solutions than fighting think on this

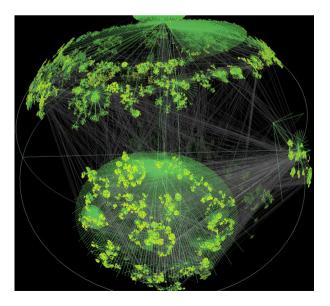
- are these marketing spin for the same thing?
  - a) slowing down heavy users
  - b) allowing light users to go faster

 light usage can go much faster without appreciably affecting completion times of heavy usage



## BT's two solutions (each yet another story)

- tactical (operational architecture)
  - "long term fair queuing"
- strategic (future Internet arch)
  - bulk edge congestion policing using "re-feedback"
  - encourages evolution of weighted TCP



anyone will (still) be able to use any link on the Internet ...without asking

whether NGN, cellular, ad hoc wireless, public Internet, satellite, cable...



#### Further reading

Problem Statement: We [the IETF] don't have to do fairness ourselves <<u>www.cs.ucl.ac.uk/staff/B.Briscoe/</u> projects/refb/#relax-fairness>





# freedom to limit the freedom of others

