Observing Slow Crustal Movement in Residential User Traffic

Kenjiro Cho (IIJ), Kensuke Fukuda (NII),
Hiroshi Esaki (Univ. of Tokyo), Akira Kato (Keio Univ.)

ACM CoNEXT2008, December 11 2008

WIDE

explosive traffic growth by video content?

many media reports on explosive traffic growth by video content

OPINION | JANUARY 20, 2007

The Coming Exaflood

By BRET SWANSON

Article

Email Printer Friendly Share: Yahoo Buzz

Today there is much praise for YouTube, MySpace, blogs and all the
technologies that are allowing you and me to transform media and com
Internet applications are at risk, thanks to the regulatory implications of
Proponents of this concept -- including Democratic Reps. John Dingell

Forbes.com

Search: Forbes.com Quotes Video

U.S. EUROPE ASIA HOME PAGE FOR THE WORLD'S BUSINESS LE

HOME BUSINESS TECH MARKETS ENTREPRENEURS LEADERSHIP

E-Mail | Print | Comments | Request Reprints | E-Mail Newsletters | RSS

Commentary

Information Super Traffic Jam

Phil Kerpen 01.31.07, 6:00 AM ET

Phil Kerpen

WASHINGTON, D.C. -A new assessment from Deloitte & Touche predicts that global traffic will exceed the Internet's capacity as soon as this year. Why? The rapid growth in the number of global Internet users, combined with the rise of online video services and the lack of investment in new infrastructure. If Deloitte's predictions are accurate, the traffic on many Internet backbones could slow to a crawl this year absent substantial new infrastructure investments and deployment.

Uncertainty over potential network neutrality requirements is one of the major factors delaying necessary network upgrades. The proponents of such regulations are back on the offensive, heartened by sympathetic new Democratic majorities and the concession made by **AT&T** (nyse: T - news - people) in its **BellSouth** (nyse: BLS - news - people) acquisition. The Google/MoveOn.org coalition fighting for network neutrality mandates calls itself "Save the Internet." But the Internet doesn't need to be saved--it needs to be improved, expanded and bulked up. An attempt to "save" the Internet in its current state would be something akin to saving the telegraph from the telephone.

USA TODAY

Home News Travel Money Sports Life Tech

Money Markets Economy Companies & Execs Media Cars Personal Finance Real E

GET A QUOTE: Enter symbol(s) or Keywords GO DJIA 8,149.09 -679.95 NASDAQ 1,3

Video, interactivity could nab Web users by '10

Updated 11/20/2007 12:29 AM | Comments 52 | Recommend 14 E-mail | Save | Print | Reprints & Permissions | RSS

By David Lieberman, USA TODAY

Mixx it
Other ways to sha
Digg
del.icio.us
Newsvine
Reddit
Facebook
What's this?

NEW YORK — Enjoy your speedy broadband Web access while you can.

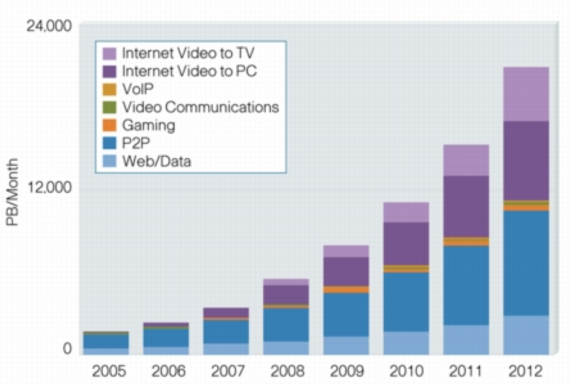
The Web will start to seem pokey as early as 2010, as use of interactive and video-intensive services overwhelms local cable, phone and wireless Internet providers, a study by business technology analysts Nemertes Research has found.

"Users will experience a slow, subtle degradation, so it's back to the bad old days of dial-up," says Nemertes President Johna Till Johnson. "The cool stuff that you'll want to do will be such a pain in the rear that you won't do it."

modest traffic growth?

but technical sources report only modest traffic growth worldwide

- MINTS: 50-60% in U.S. and worldwide
- Cisco visual networking index: worldwide growth of 50% per year over last few years
- TeleGeography: network capacity also grows by 50% per year



source: Approaching the Zettabyte Era (Cisco 2008/6)

motivation

why is traffic growth important?

- one of the key factors driving research, development and investiment in technologies and infrastructures
  - with annual growth of 100%, it grows 1000-fold in 10 years
  - with annual growth of 50%, it grows 58-fold in 10 years
- crucial is the balance between demand and supply
  - balanced growth makes both users and ISPs happy
  - traffic surged in 2003-2004 by p2p file sharing
  - might need to worry about oversupply in the future?

key question: what is the macro level impact of video and other rich media content on traffic growth at the moment?

- measurements: 2 data sets
  - aggregated SNMP data from 6 ISPs covering 42% of Japanese traffic
  - Sampled NetFlow data from 1 ISP

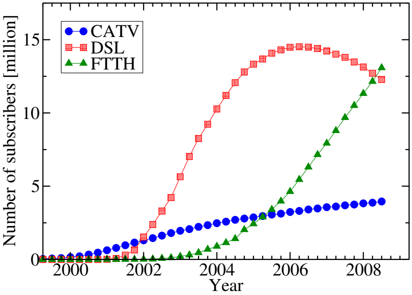
residential broadband subscribers in Japan

29.3 million broadband subscribers as of June 2008

- reached 56% of households, increased by only 5% in 2007
- FTTH:13.1 million, DSL:12.3 million, CATV:3.9 million

shift from DSL to FTTH: FTTH has exceeded DSL

- 100Mbps bi-directional fiber access costs 40USD/month
  - effects of sales promotion for VoIP and IPTV?
- significant impact to backbones



residential broadband subscribers in Japan

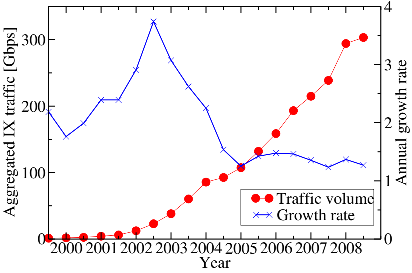
traffic growth in backbone

rapidly growing residential broadband access

- low-cost high-speed services, especially in Korea and Japan
- Japan is the highest in Fiber-To-The-Home (FTTH)

traffic growth of the peak rate at major Japanese IXes

- modest growth of about 40% per year since 2005



traffic growth of the peak rate at major Japanese IXes

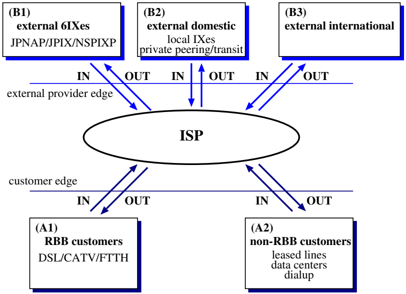
SNMP data collection from 6 ISPs

focus on traffic crossing ISP boundaries (customer and external)

- tools were developed to aggregate MRTG/RRDtool traffic logs

only aggregated results published not to disclose individual ISP share

challenges: mostly political or social, not technical



5 traffic groups at ISP cusomer and external boundaries

IN/OUT from ISPs' view

methodology for aggregated traffic analysis

month-long traffic logs for the 5 traffic groups with 2-hour resolution

- each ISP creates log lists and makes aggreagated logs by themselves without disclosing details

biggest workload for ISP

- creating lists by classifying large number of per-interface logs
  - some ISPs have more than 100,000 logs!
- maintaining the lists
  - frequent planned and unplanned configuration changes

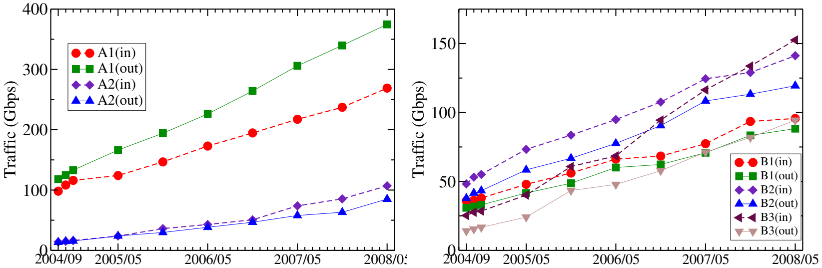
data sets

- 2-hour resolution interface counter logs
  - from Sep/Oct/Nov 2004, May/Nov 2005-2008
  - by re-aggregating logs provided by 6 ISPs
- our data consistently covers 42% of inbound traffic of the major IXes

traffic growth

22-68% increase in 2007

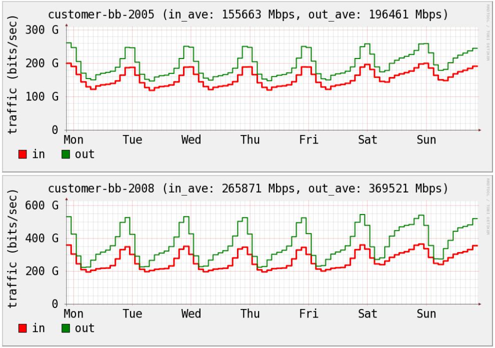
- RBB: 22% increase for inbound, 29% increase for outbound
- a sharp increase in international inbound due to popular video and other web2.0 services



measured traffic growth: customer traffic(left) external ISPs(right)

changes in RBB weekly traffic

- traffic patterns by home users (peak at 21:00-23:00)
- 2005: in/out were almost equal (dominated by p2p)
- 2008: outbound (downloading to users) became larger
  - both constatnt portion and daily fluctuations grew



weekly RBB traffic: 2005(top) 2008(bottom)

aggregated traffic summary

in 2008, we observed

- larger download volume, larger evening-hour volume in RBB
- RBB traffic decreased share in customer traffic
- larger growth of international inbound
- change in volume is comparable to p2p file sharing

implies a shift from p2p to video and other web2.0 services

analysis of per-customer traffic in one ISP

one ISP provided per-customer traffic data (RBB traffic only)

- Sampled NetFlow data
  - from edge routers accommodating fiber/DSL RBB customers
- week-long data from Apr 2004, Feb 2005, Jul 2007, Jun 2008
  - focus on Feb 2005 and Jun 2008, before and after the advent of YouTube and others

ratio of fiber/DSL active users and total traffic volumes

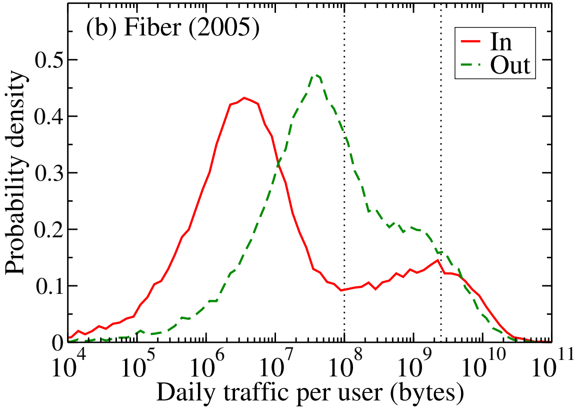
- in 2008, 80% of active users are fiber users, consuming 90% of traffic
  - active user: unique customer IDs observed in the data set

		active users (%)	total volume (%)
2005	fiber	46	79
	DSL	54	21
2008	fiber	79	87
	DSL	21	13

PDF of daily traffic per user

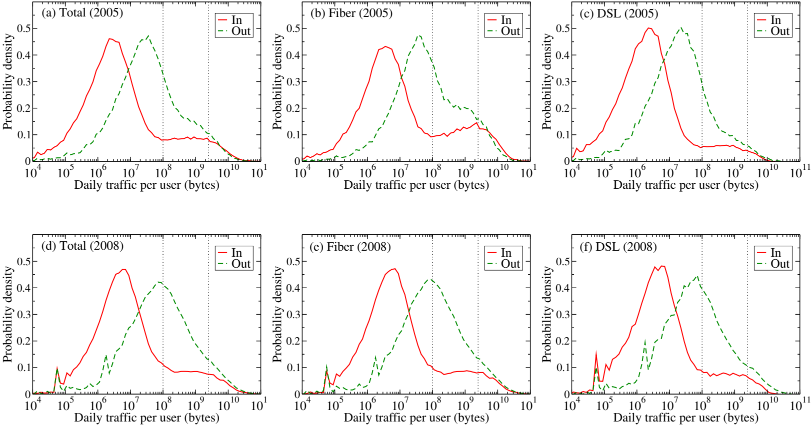
each distribution consists of 2 roughly lognormal distributions

- client-type: asymmetric (majority)
- peer-type: symmetric high-volume



PDF of daily traffic per user: 2005 and 2008

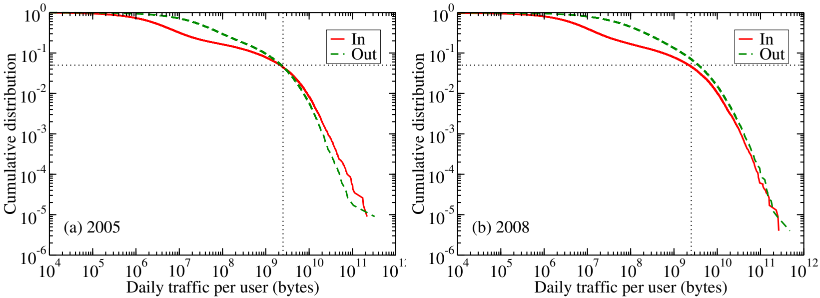
- increase in download volume of client-type users
  - mode: from 32MB/day to 94MB/day (similar in fiber/DSL)
- while peer-type dist. isn't growing much (mode:2GB/day)



daily traffic per user: total(left) fiber(middle) DSL(right): 2005(top) 2008(bottom)

CCDF of daily traffic per user

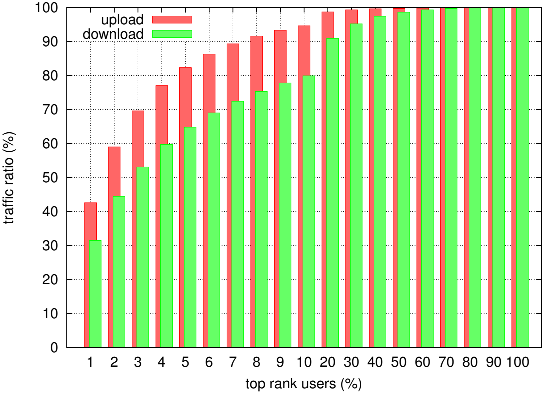
- heavy-tailed distribution
  - the tail exceeds 200GB/day
- larger increase in outbound (download for users)
- the tail becomes symmetric (no longer need to compensate upstream shortage of DSL)



CCDF of daily traffic per user: 2005(left) 2008(right)

skewed traffic usage among users

- highly skewed distribution in traffic usage
  - top 10% users consume 80% of download, 95% of upload volumes
- no noticeable change from 2005 to 2008
  - long-tailed distribution (common to other Internet data)
  - looks similar even if p2p traffic is removed

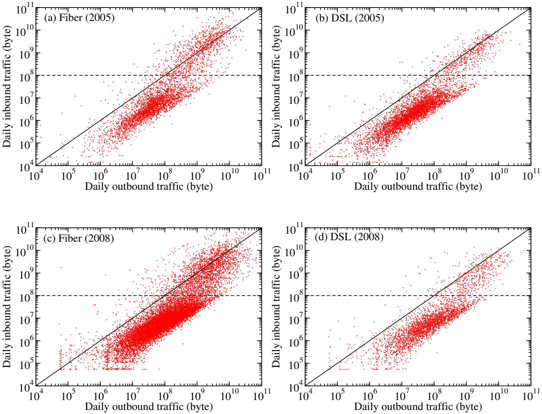


traffic share of top-ranking heavy-hitters

correlation of inbound/outbound volumes per user

2 clusters: client-type users and peer-type heavy-hitters

- difference between fiber and DSL: only heavy-hitter population
- no clear boundary: heavy-hitters/others, client-type/peer-type
- actual individual users have different traffic mix



in/out volumes per user: fiber(left) DSL(right) 2005(top) 2008(bottom)

protocols/ports ranking

classify client-type/peer-type with threshold: 100MB/day upload

- to observe differences in protocol/port usage
- port number: min(sport, dport)

observations

- dominated by TCP dynamic ports (often used by p2p)
  - 83% in 2005, 78% in 2008
  - but each port is tiny
- TCP port 80 is increasing (again)
  - 9% in 2005, 14% in 2008
  - client-type: 51% in 2005, 65% in 2008

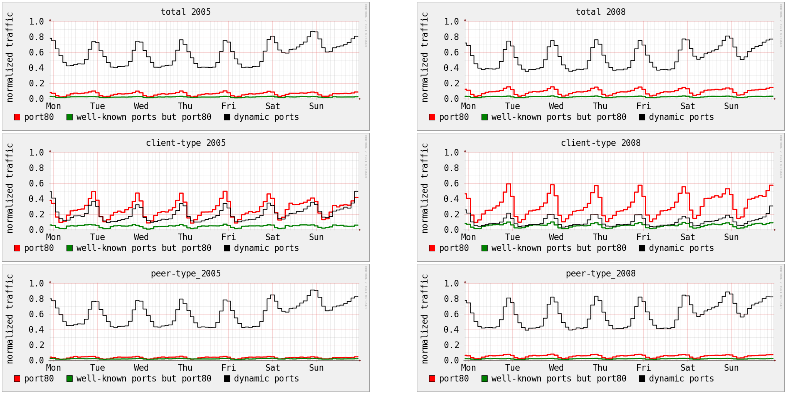
protocols/ports ranking data

protocol	port	2005 total (%)	2005 client type	2005 peer type	2008 **total** (%)	2008 client type	2008 peer type
TCP	*	**97.43**	**94.93**	**97.66**	**96.00**	**95.51**	**96.06**
	(< 1024)	*13.99*	*58.93*	*8.66*	*17.98*	*76.16*	*11.35*
	80 (http)	9.32	50.78	5.54	14.06	64.96	8.26
	554 (rtsp)	0.38	2.44	0.19	1.36	8.21	0.58
	443 (https)	0.30	1.45	0.19	0.58	1.63	0.46
	20 (ftp-data)	0.93	1.25	0.90	0.24	0.17	0.25
	(>= 1024)	*83.44*	*36.00*	*89.00*	*78.02*	*19.35*	*84.71*
	6346 (gnutella)	0.92	0.84	0.93	0.94	0.67	0.97
	6699 (winmx)	1.40	1.14	1.43	0.68	0.24	0.73
	7743 (winny)	0.48	0.15	0.51	0.30	0.04	0.33
	1935 (rtmp)	0.20	0.81	0.14	0.22	0.73	0.16
	6881 (bittorrent)	0.25	0.06	0.27	0.22	0.02	0.24
UDP	*	**1.38**	**3.41**	**1.19**	**1.94**	**2.50**	**1.88**
	53 (dns)	0.03	0.14	0.02	0.04	0.12	0.03
	others	1.35	3.27	1.17	1.90	2.38	1.85
ESP		**1.09**	**1.35**	**1.06**	**1.93**	**1.85**	**1.94**
GRE		**0.07**	**0.12**	**0.06**	**0.09**	**0.08**	**0.09**
ICMP		**0.01**	**0.05**	**0.01**	**0.02**	**0.05**	**0.02**

temporal behavior of TCP port usage

3 types: port 80, well-kown port but 80, dynamic ports

- total traffic heavily affected by peer-type traffic
- shift from dynamic ports to port 80 for client-type users
- daily fluctuations also observed in dynamic ports
    - slow decay of dynamic port traffic over night



TCP usage: total(top) client-type(middle) peer-type (bottom) 2005(left) 2008(right)

summary of per-customer traffic analysis

- overall traffic still dominated by heavy-hitters, mainly using p2p
  - but p2p traffic decreased in population share and volume share
- client-type traffic slowly moving towards high-volume
  - circumstantial evidence: driven by video and web2.0 services
- current slow growth is due to stalled growth of dominant aggressive p2p traffic
- meanwhile, network capacity also grows 50% per year (by various sources)
  - seems faster than the traffic growth

growth model based on lognormal distributions

fitting client-type outbound volumes to lognormal distribution

$$p(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp(\frac{-(\ln x - \mu)^2}{2\sigma^2})$$

$$E(x) = \exp(\mu + \sigma^2/2)$$

- by definition, mean grows much faster than mode
- simplistic growth projections by exponential model for outbound traffic per user (MB/day) for client-type users
  - mean is less predictable (easily affected by various constraints)

	mode	mean
2004 Apr	26.2MB	110.6MB
2005 Feb	32.0MB	162.7MB
2007 Jul	65.7MB	483.2MB
2008 Jun	94.1MB	862.6MB
growth/yr	1.38	1.72
2009 Jun	130MB	1480MB
2010 Jun	179MB	2540MB
2011 Jun	248MB	4359MB

conclusion

apparent slow growth attributed to decline of p2p traffic

- but p2p willl not go away anytime soon
- p2p could evolve for large scale distribution

crustal is slowly swelling with video and other web2.0 content

- similar to how web traffic was perceived in late 90es
- still, will take a while to catch up with p2p

network capacity is growing faster than traffic at the moment

- no need to worry too much about video traffic

our observations seem to be common to other countries

- though exact ratio of traffic mix and growth are different

it is difficult to predict future traffic (lognormal!)

many challenges ahead

- technical factors: content caching, CDN, QoS
- economic factors: access cost, capacity/equipment costs
- political/social factors: net-neutrality, content management

acknowledgments

- IIJ, SoftBank Telecom, K-Opticom, KDDI, NTT Communications, SoftBank BB for data collection support
- ministry of internal affairs and communications for coordination

references

[CFEK2008] K. Cho, K. Fukuda, H. Esaki, and A. Kato.
Observing Slow Crustal Movement in Residential User Traffic.
To appear in *ACM CoNEXT2008*, Madrid, Spain, Dec. 2008.

[CFEK2006] K. Cho, K. Fukuda, H. Esaki, and A. Kato.
The impact and implications of the growth in residential user-to-user traffic.
In *SIGCOMM2006*, Pisa, Italy, Aug. 2006.

[Cisco2008a] Cisco.
visual networking index – forecast and methodology, 2007-2012.
June 2008.

[Cisco2008b] Cisco.
Approaching the zettabyte era.
June 2008.

[Odlyzko2008] A. M. Odlyzko.
Minnesota Internet traffic studies.
`http://www.dtc.umn.edu/mints/home.html`.

[TeleGeography2007] TeleGeography Research.
Globel Internet Geography.
2007.