

Measuring Internet Evolution

or ...

**If we don't Measure, we don't
Know What's Happening!**

Measurement WG, APAN 26, Queenstown, 2008

Nevil Brownlee

Introduction: Why we Measure

- The Internet is an un-natural system, in widespread use. Its providers and users need to understand it!
 - ISPs need to manage traffic loads, detect unusual behaviour
 - Researchers need to observe how traffic behaves, so as to understand changes in traffic patterns
- The Internet is *always changing*
 - we need *lots* of data, to ensure our measurement consistency
 - we need to *share* our data, so others can verify our work
- Measurement capability was not designed into the Internet
 - we have developed many different measurement techniques
 - some of them are 'direct,' e.g:
 - send a test packet to measure transit time
 - others are much less direct, e.g:
 - Packet dispersion to measure available bandwidth
 - Passive methods, e.g. observing links to measure properties (sizes, lifetimes) of traffic *flows*

Overview

- In this talk I'll review five papers and one report on:
 - 1990: Self-Similarity on Ethernet
 - 1995: NSFNET Traffic Mix (rise of the web)
 - 2002: Stream Lifetimes (dragonflies and tortoises)
 - 2005: Residential Broadband traffic in Japan
 - 2006: Skype – how does it's traffic behave?
 - 2008: Content Providers and their Global WANs
- I've selected these papers because
 - they provide a history of Internet development
 - each one demonstrates some interesting and effective measurement techniques

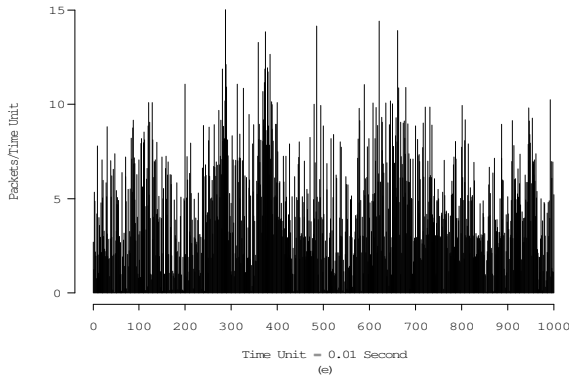
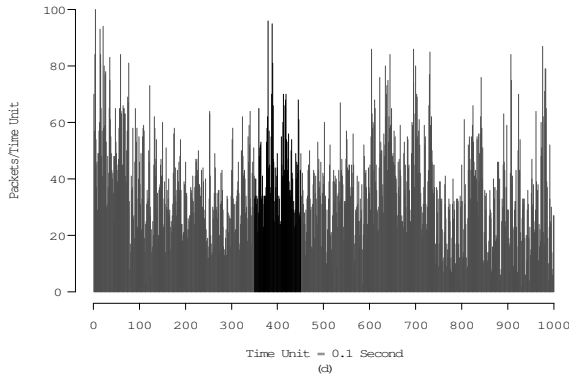
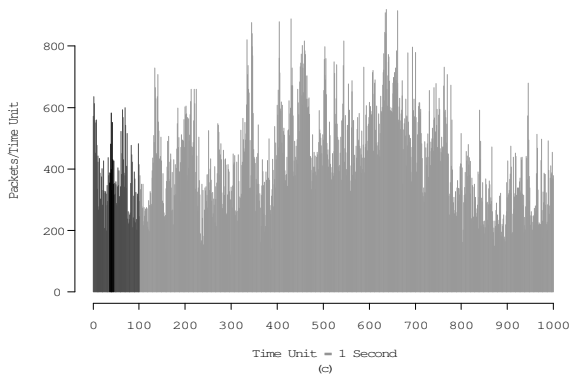
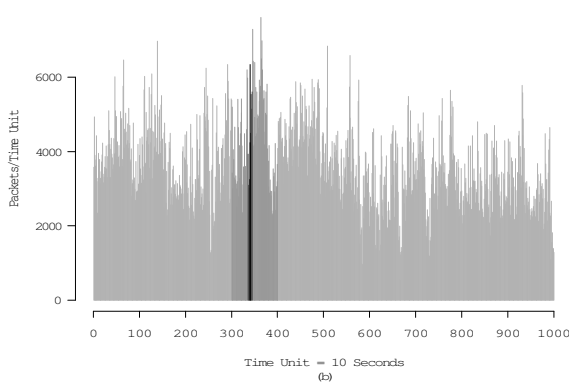
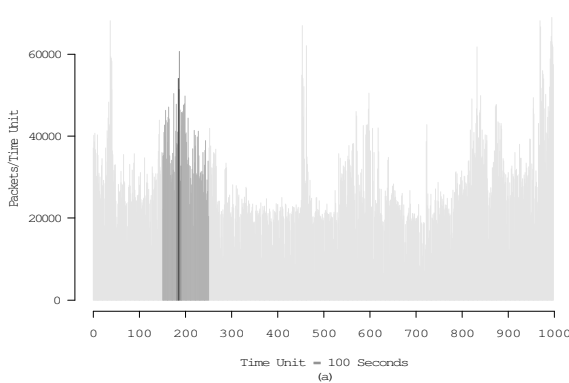
1990: Packets on Ethernet, self-similarity

- We assumed that telephone traffic modelling would work for network traffic. That assumed:
 - Poisson arrival process
 - Erlang call duration
 - Queueing models like $M/E_k/n$
- A Poisson process
 - When observed on a fine time scale will appear bursty
 - When aggregated on a coarse time scale will flatten (smooth) to white noise
- Very successful for telephone engineering
 - *but it didn't seem to hold for LAN traffic . . .*
- *On the Self-Similar Nature of Ethernet Traffic.* Will Leland, Murad Taqqu, Walter Willinger, Daniel Wilson. SIGCOMM, 1993

Measurements on Ethernet

- In 1989, Leland and Wilson begin taking high resolution traffic traces at Bellcore
 - Ethernet traffic from a large research lab
 - $20\mu s$ time stamp resolution
 - Packet length + status + 60 bytes of data
 - Mostly IP traffic (a little NFS)
 - Four data sets over three year period
 - Over 100 million packets in traces
 - Traces considered representative of normal use

Ethernet traffic rate plots



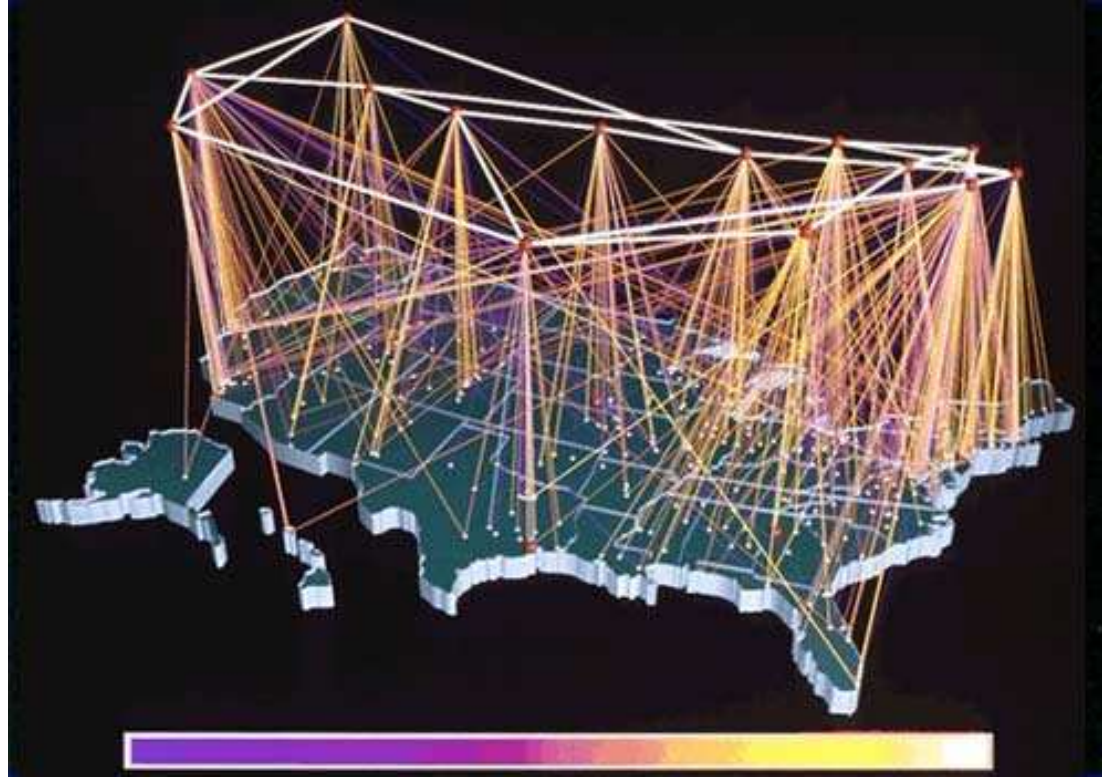
Different grey levels identify the same segment of traffic on the different time scales

*Figure 3.1 “Pictorial proof of self-similarity: Ethernet traffic (packets per time unit) on five different time scales”
Leland, Taqqu, Willinger and Wilson, 1994*

Self-similar traffic: Summary

- A Self-Similar (fractal) process
 - When aggregated over wide range of time scales will maintain its bursty characteristic
 - There is no natural length of a burst
 - At any time, we can't predict when a burst will appear, or how long it will last

Early days: NSFNET Traffic Mix



- <http://www.merit.edu/networkresearch/projecthistory/nsfnet>
 - Started in 1987, 56kb/s backbone linking 5 supercomputer centres
 - Upgraded to T1 in 1988, T3 in 1992, becoming the backbone linking regional networks (illustrated above)
 - Links to 93 countries by 1995. Phased out 30 Apr 95

Traffic Mix on NSFNET T3

● November 1992

Service Name	Port	Rank	Packets	% Pkt	Rank	Bytes	% Byte
ftp-data	20	1	5346086600	23.4	1	1879761013600	44.4
telnet	23	2	3294996450	14.4	4	224789605450	5.3
nntp	119	3	1846051150	8.1	2	387621790550	9.2
smtp	25	4	1493729000	6.5	3	248933009450	5.9
domain	53	5	1035987000	4.5	5	99056612200	2.3
icmp	-1	6	543199650	2.4	6	50648808200	1.2

● Mosaic Browser

- Feb 1993: NCSA released first version
- Aug 1994: Rights assigned to Spyglass, Inc.

● November 1994

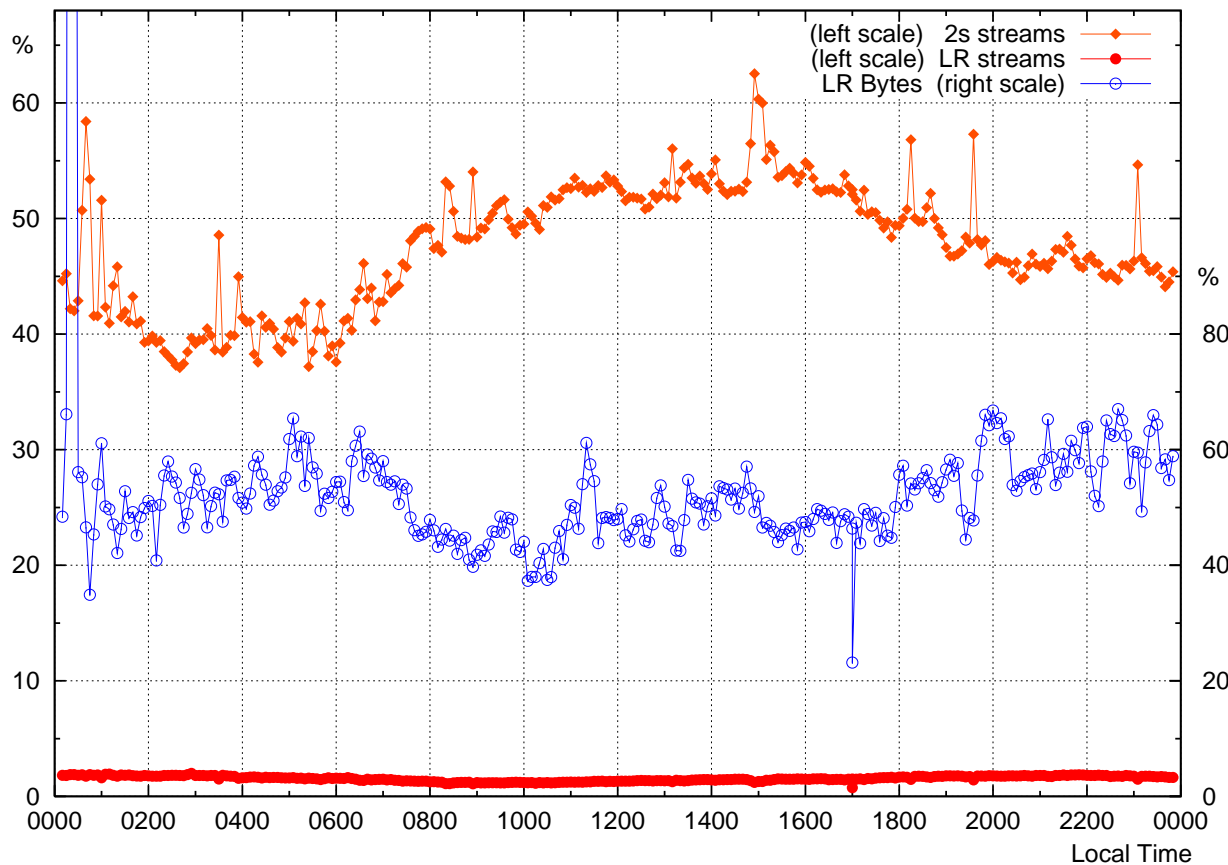
Service Name	Port	Rank	Packets	% Pkt	Rank	Bytes	% Byte
ftp-data	20	1	20460016300	19.2	1	7130364445000	31.7
telnet	23	2	12416140350	11.6	6	973506738300	4.3
www	80	3	10402528150	9.8	2	3126195255050	13.9
nntp	119	4	10110453700	9.5	3	2324060845300	10.3
smtp	25	5	8348142300	7.8	4	1380684299300	6.2
domain	53	6	5601761900	5.3	8	581492142200	2.6

data from <http://www.ifla.org.sg/documents/internet/Nsf9411.txt>

2002: Internet Traffic Stream Behaviour

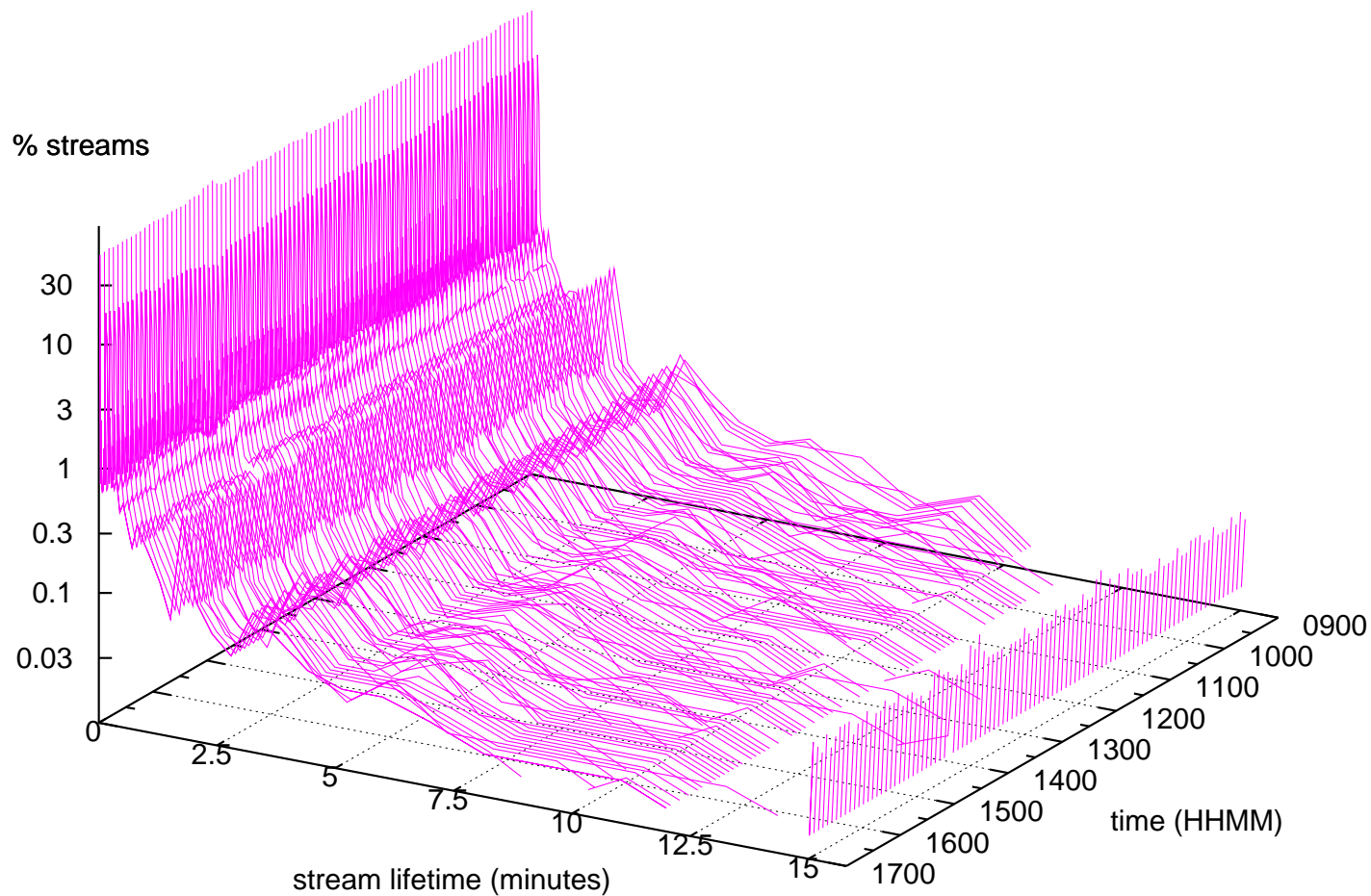
- A *Stream* is a *bi-directional 5-tuple flow*
- Accepted Wisdom:
 - “most bytes on a link are carried by large streams (*elephants*)”
 - “the rest of the bytes are carried by small streams (*mice*)”
 - “one can ignore the mice – they will not influence the elephants”
- New approach: stream lifetimes
 - how long do streams last?
 - how are stream lifetimes distributed?
 - what proportion of a link’s Bytes are carried by short-lived streams?
- Nevil Brownlee, *Understanding Internet Traffic Streams: Dragonflies and Tortoises*, IEEE Communications, Oct 2002
- Used NeTraMet (streams extension to RTFM) to analyse traffic at U Auckland and UC San Diego in June 2002

Stream Lifetimes at UCSD



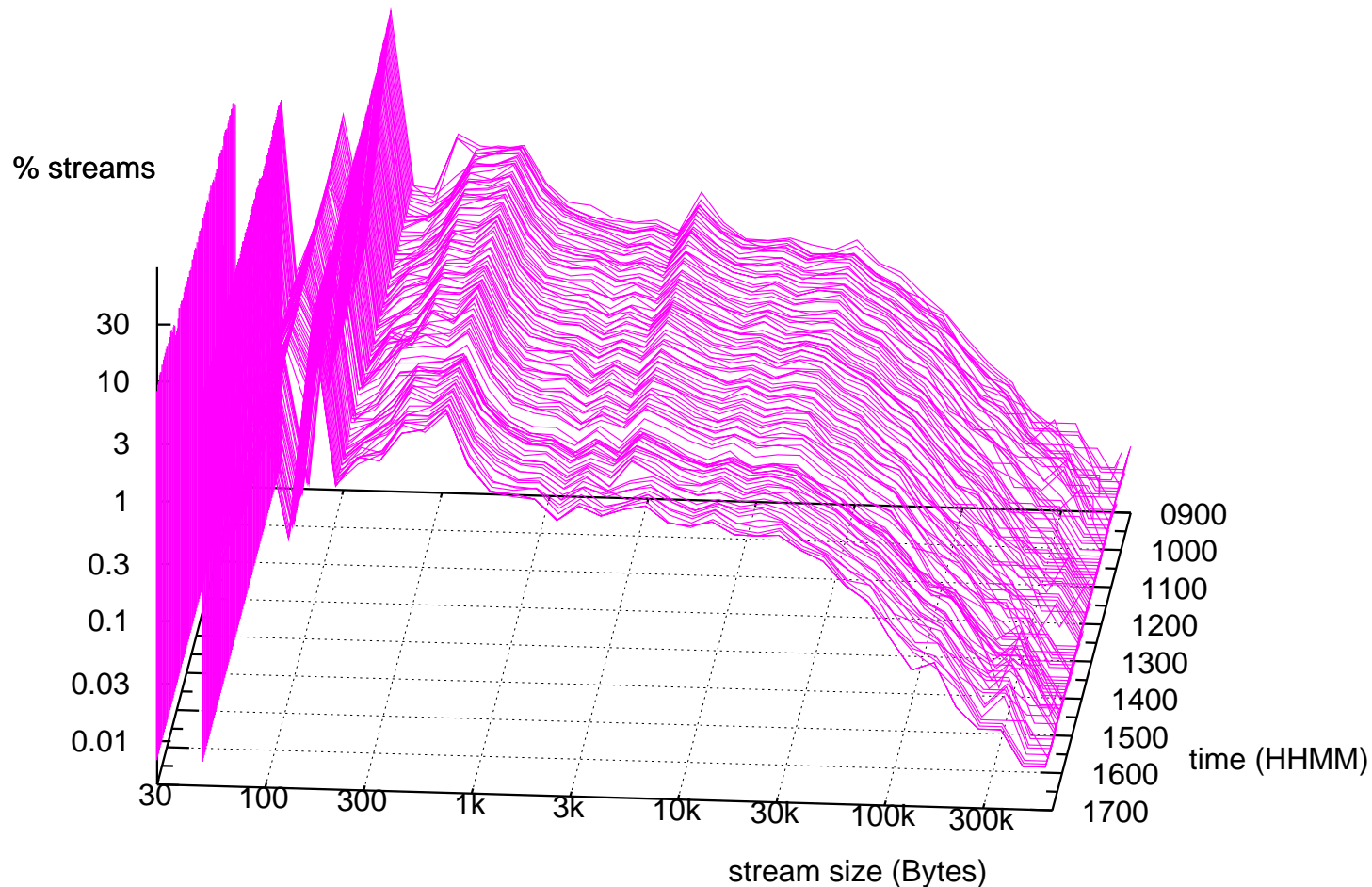
- 40% to 70% of streams are dragonflies, (orange diamonds) lasting 2 seconds or less
- About 1.5% of streams (red dots) are tortoises, i.e. LR streams lasting > 15 minutes. They contribute 5% to 50% of a link's bytes (blue circles)

Stream Lifetime Changes over Time



- UCSD lifetimes; OC12 link, 5m intervals, 0900–1700 on 12 Jun 2002
- 45% of streams < 2 s, in first bin
- Spike at 15m is all streams at end of interval
- Lifetime distribution is surprisingly *stable*

Stream Size Changes over Time

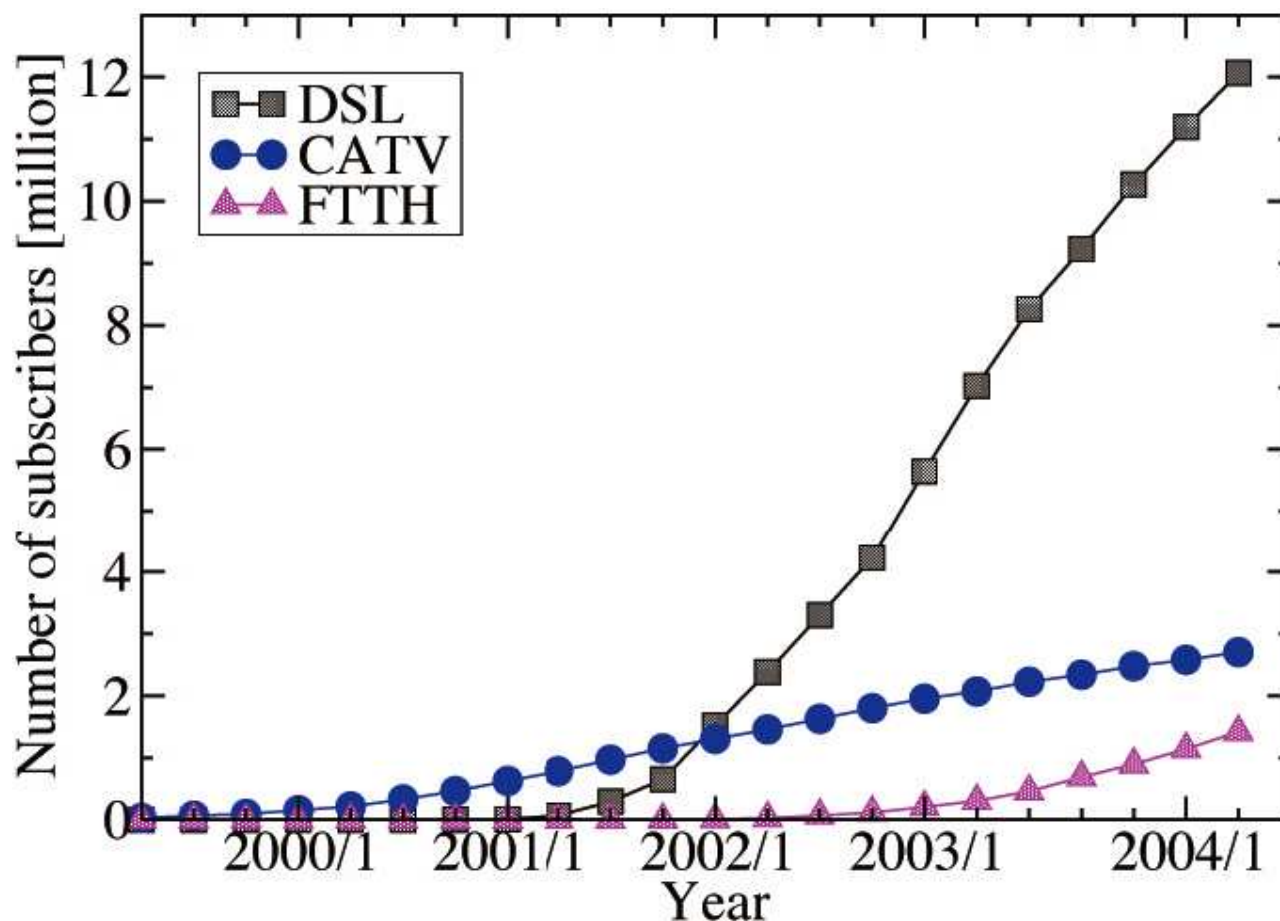


- UCSD web stream size distributions on Wed, 12 Jun 02
 - local maximum 300~800B for small web objects
 - plateau up to 40kB, power law (approx linear) fall for large files
- Size distribution is (again) surprisingly *stable*

Stream Lifetimes: Summary

- At least 45% of streams have lifetimes of 2 seconds or less, (dragonflies), and about 98% of them last less than 15 minutes
- The remaining one or two percent, which we call Long-Running (LR) streams (tortoises), have lifetimes of hours to days and can carry a high proportion (50% to 60%) of the total bytes on a link
- Implications
 - hardware and software that work with flows must allow for the large proportion of Very Short flows
 - tortoise flows can be a significant fraction of packets and bytes, reducing available bandwidth

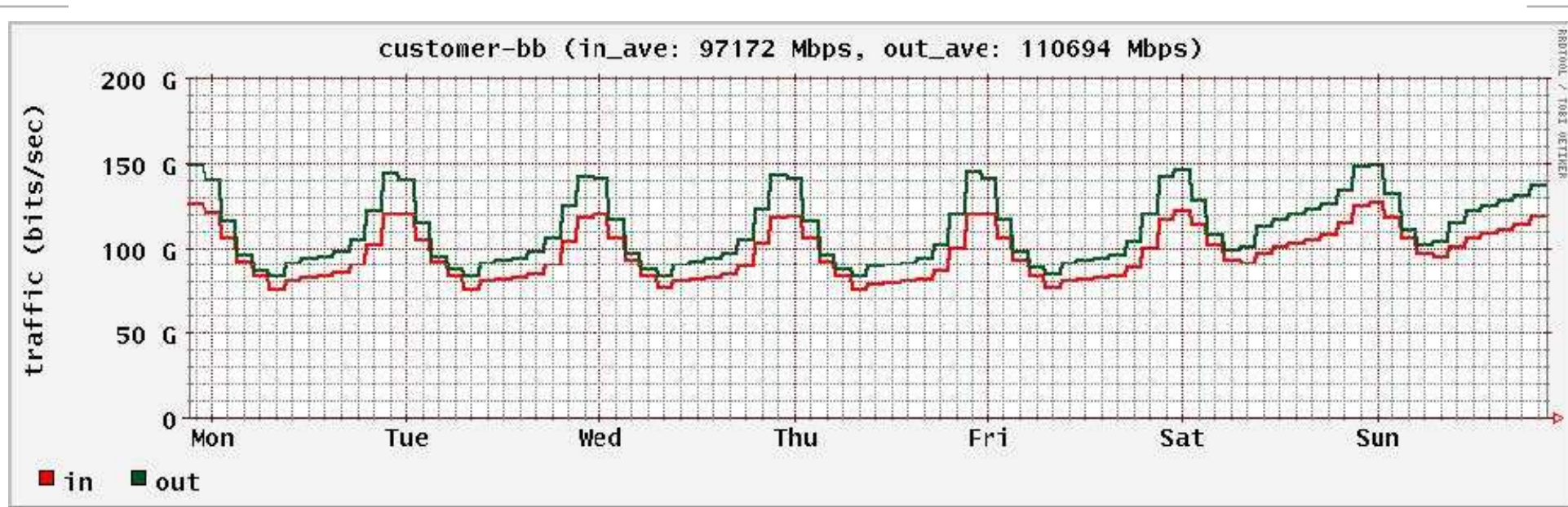
2005: Residential Broadband (RBB) Traffic in Japan



ISP Subscribers in Japan, 1999–2005

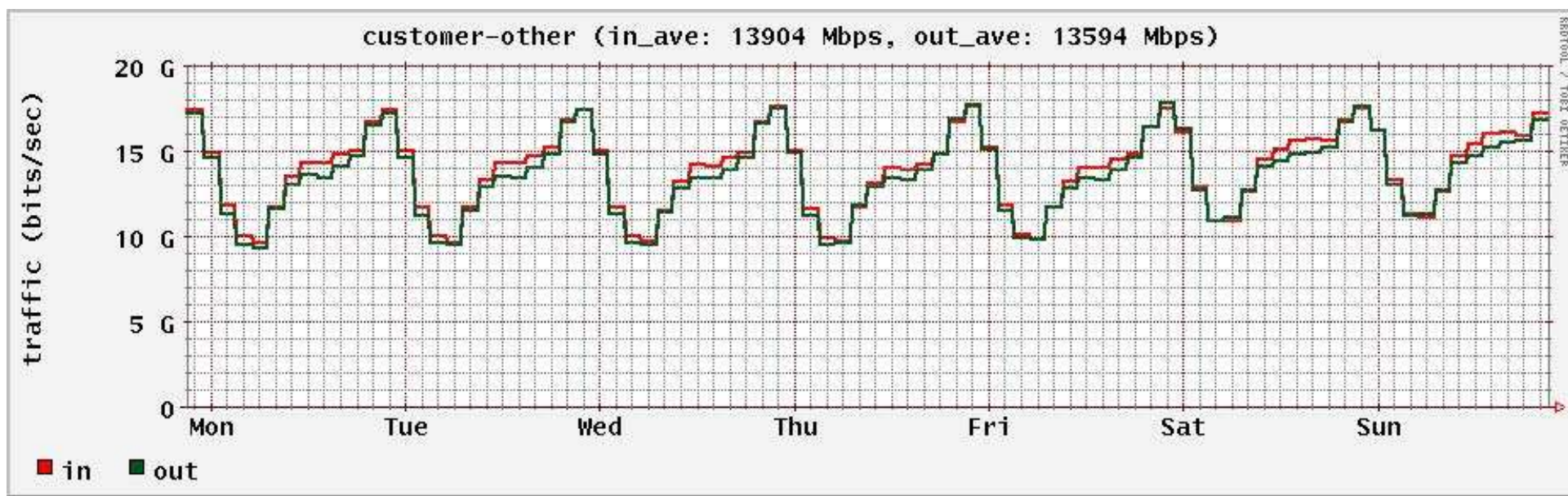
- *The impact of residential broadband traffic on Japanese ISP backbones.* Kensuke Fukuda, Kenjiro Cho, Hiroshi Esaki. CCR Jan 2005

Residential Broadband traffic



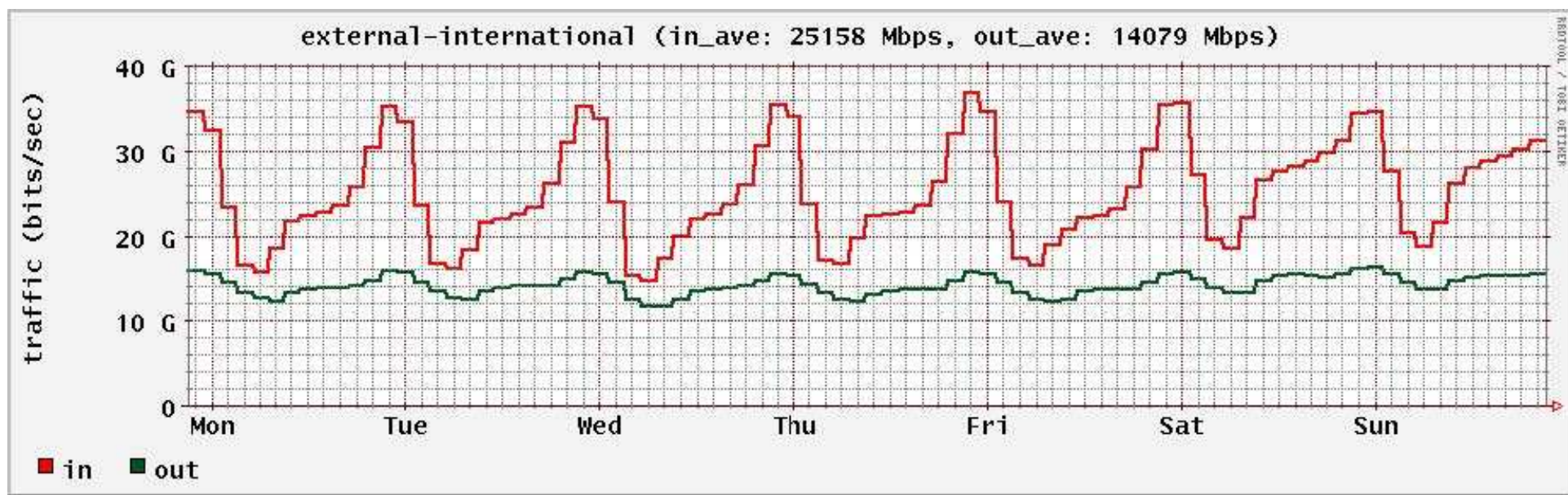
- WIDE project; study group looked at traffic from 7 large ISPs
 - data collected via SNMP at each ISP, and at IXPs
 - separate counts for RBB, non-RBB, non-residential, international and regional traffic
 - data combined to produce overall plots
- RBB plot:
 - about 100Gb/s, 70% of it steady. (Note: 'in' means 'from user')
 - peak hours 2100-2300, different behaviour at weekend

Non-RBB traffic



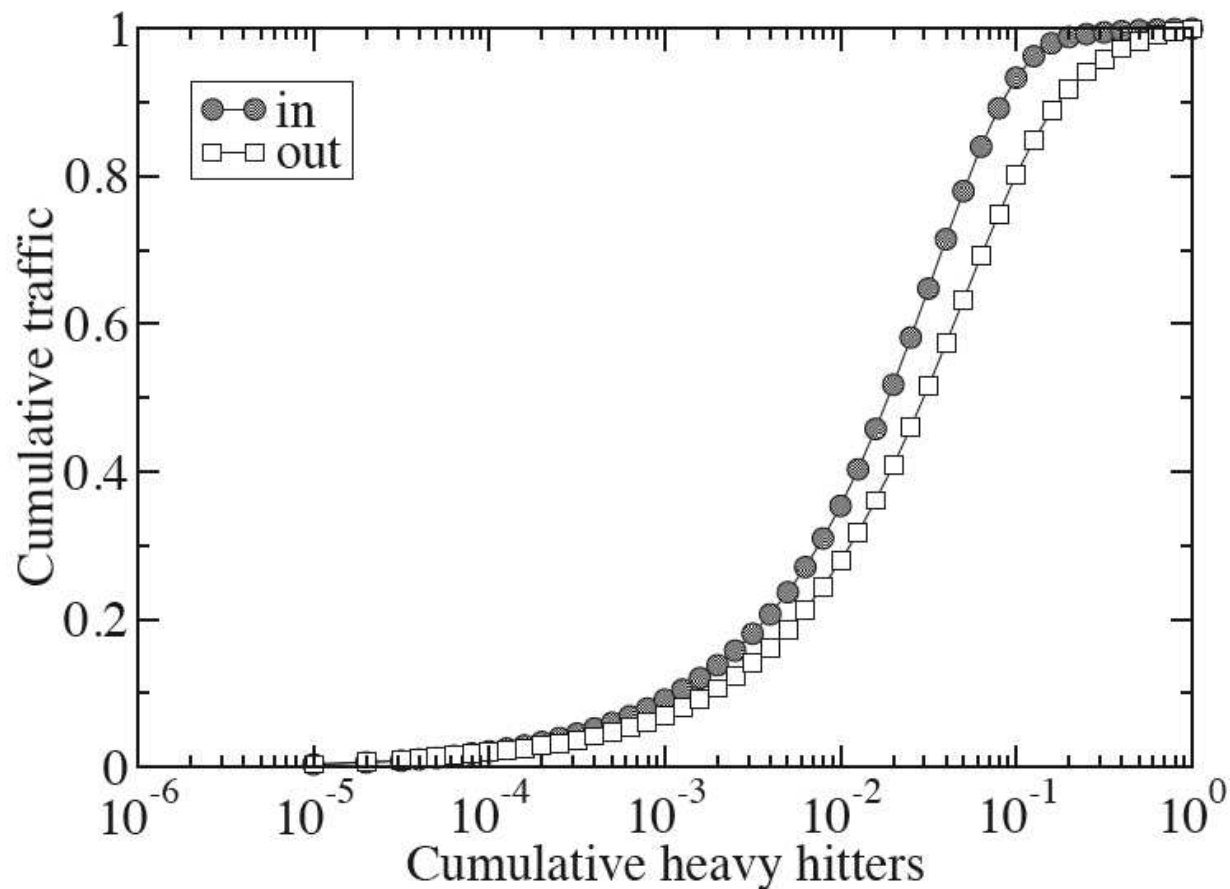
- Non-RBB plot:
 - data from only four (of the seven) ISPs
 - includes some leased lines used to provide RBB, hence dominated by RBB usage pattern
 - about 12Gb/s, roughly 10% of RBB
 - 'office hours' 0800-1800
 - in \approx out volumes

International traffic



- International plot:
 - about 20Gb/s, little difference at weekend
 - in > out volumes – ‘traditional’ pattern (here, ‘in’ means ‘from overseas’)
 - more like RBB (*residential*) than non-RBB (*office*) traffic

Heavy Hitters: Traffic per Host



Cumulative traffic volume, heavy hitters in decreasing order of volume

- Top 4% of customers use 75% of total inbound traffic, and 60% of the outbound
- A few hosts can – and do – use high upstream bandwidth

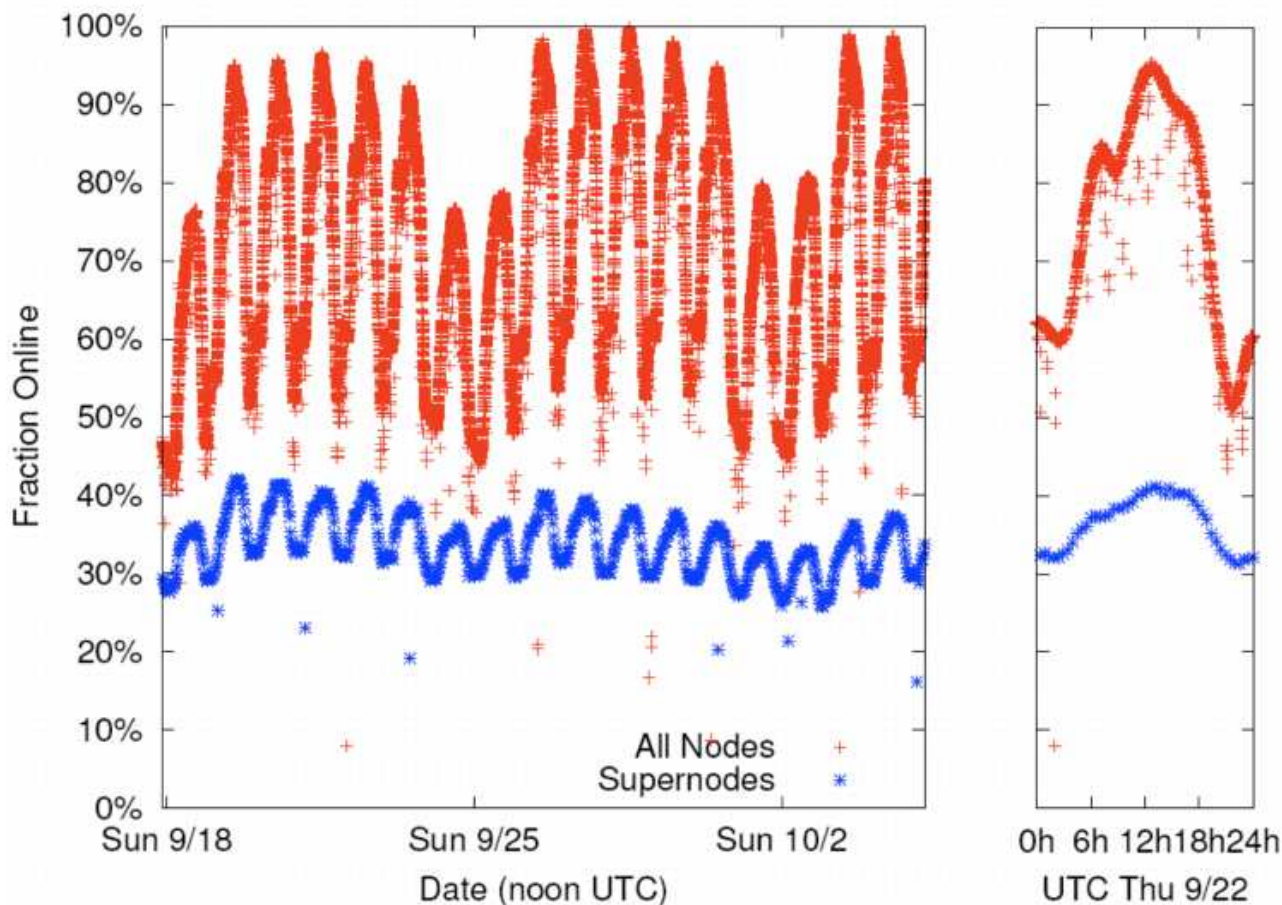
Peer to Peer (P2P) Applications

- Napster (1999):
 - central server, clients asked “who has a copy of this file?”
- Gnutella (2000):
 - no central server, clients search for file by flooding enquiry messages to known peers. Slow searches
- Fasttrack network (2001):
 - clients Kazaa, Grokster, Morpheus, ...
 - all clients access same peers, hence can find same files
- BitTorrent (2002): Focussed on efficient fetching
- Now used for:
 - software distribution (coping with ‘flash crowds’), multi-player gaming, telephony, etc.
- No change to Internet protocols. New distributed applications, where all peers can interact with each other
- Big changes to Internet traffic patterns

2006: Skype – P2P IP Telephony

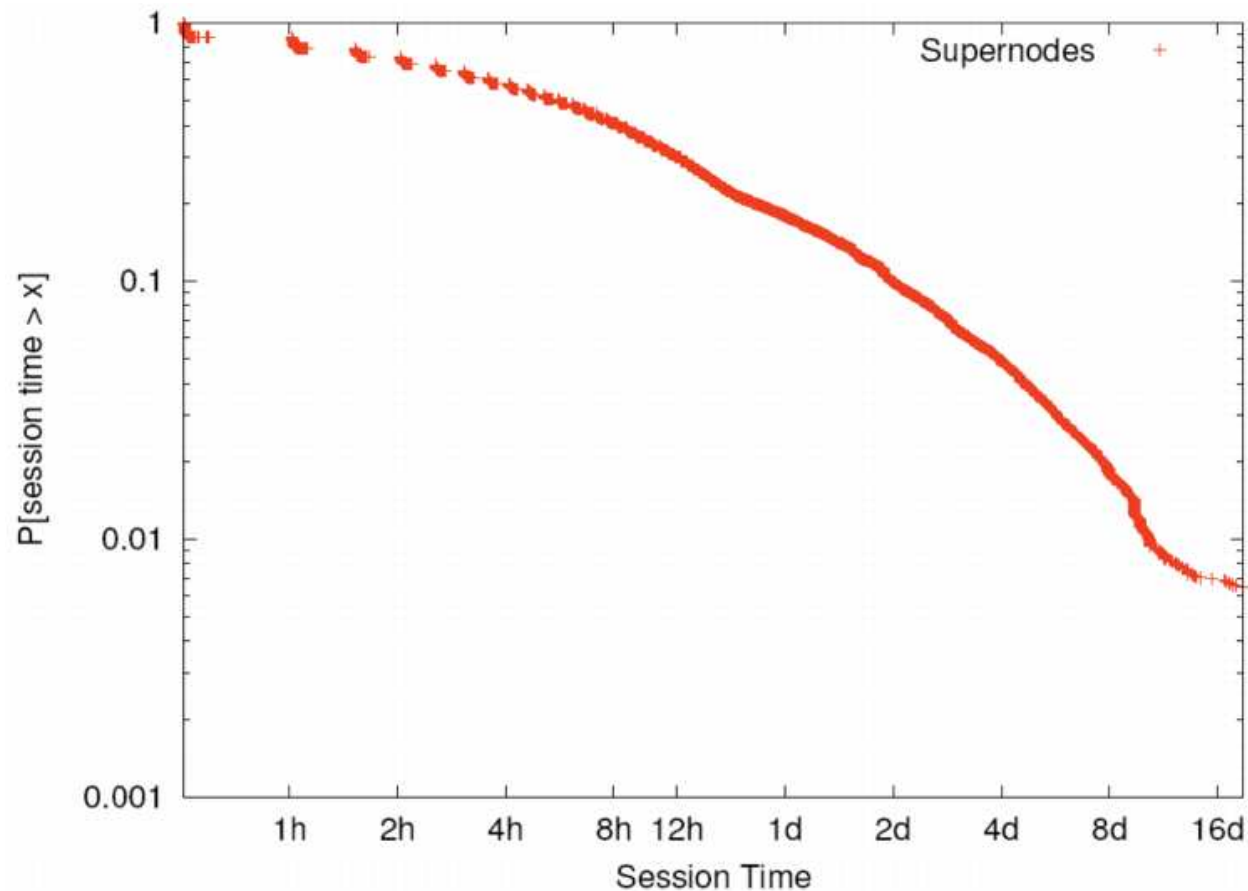
- Provides voice and video over IP, gateways in/out of global telephone network, instant messaging and file transfer
 - 2003: initial beta version released
 - 2006: 100 million registered users
 - 2007: video support introduced
- Proprietary system, most Skype traffic is encrypted
 - easy to use (good user interface)
 - uses its own codecs – good voice and video quality
 - works through (or relays around) firewalls
 - what effect does Skype have on network security?
 - how much traffic does Skype need per user?
- *An Experimental Study of the Skype Peer-to-Peer VoIP System.* Saikat Guha, Neil Dswani, Ravi Jain. IPTPS '06 (5th International Workshop on Peer-to-Peer Systems)

Skype Users (Nodes and Supernodes) Online



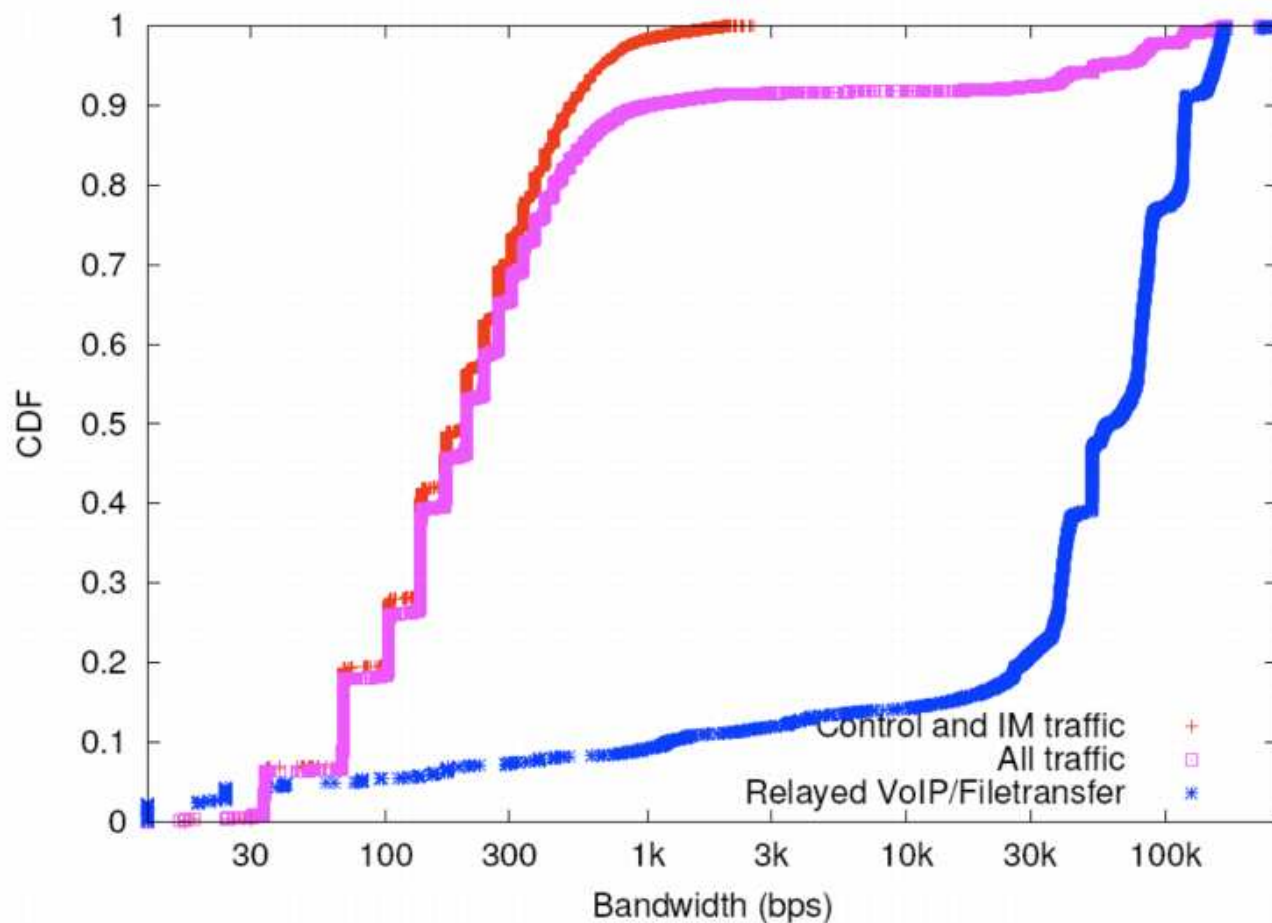
- Supernodes (blue) and All nodes (red), Sep 05–Jan 06
 - supernodes observed by monitoring client's supernode cache
 - clear diurnal variations, peak during working hours
 - supernodes more stable than ordinary nodes

Node Session Times



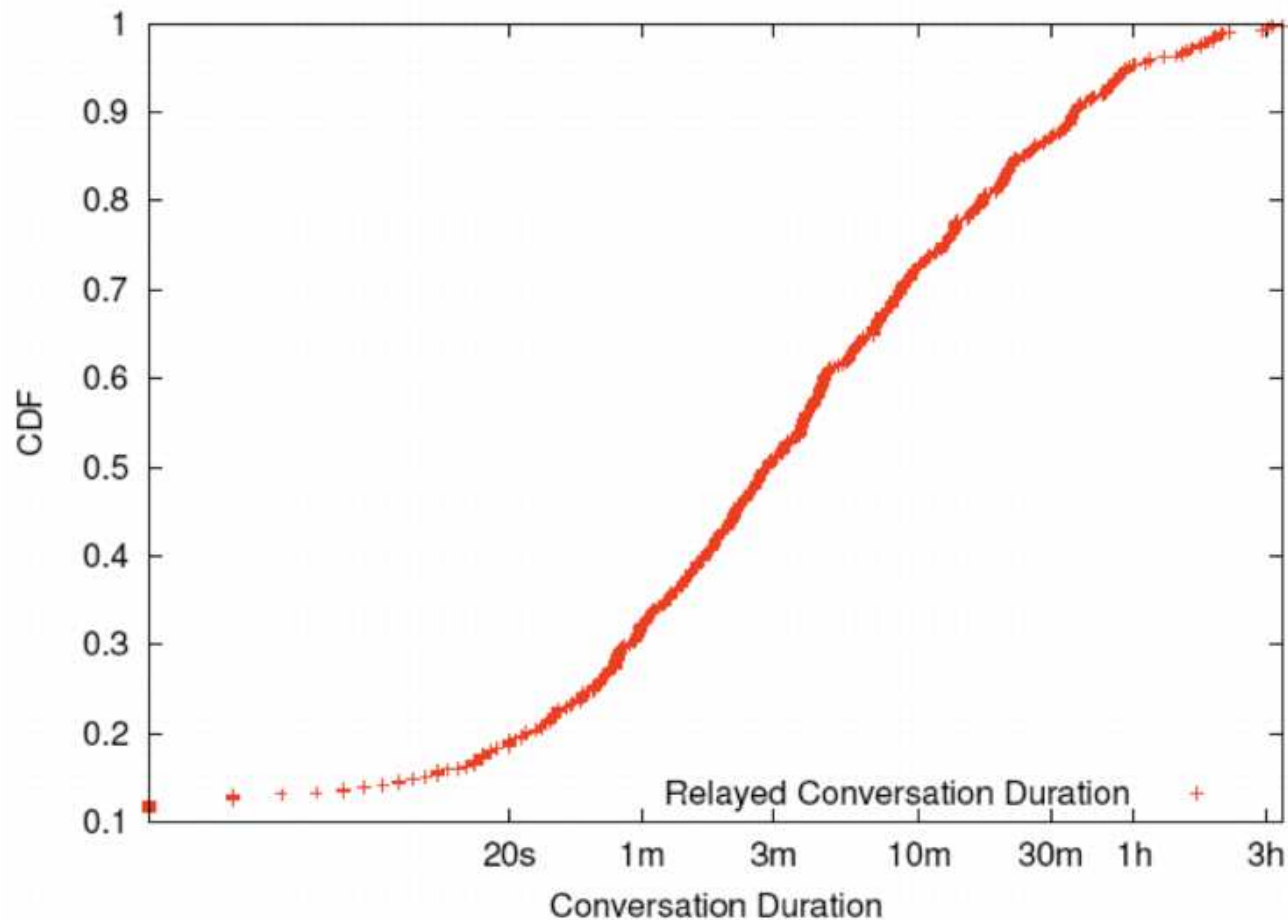
- Node session times tend to follow working hours
 - file-sharing users “do batch downloads over days and weeks”
- Plot: log-log Complementary CDF, *Supernode* session times
 - median supernode session time is about 5.5 hours

Supernode Bandwidth



- Semi-log (linear vs log) CDF of supernode bandwidth usage
 - 50% of the time, a supernode uses less than about 205b/s
 - 90% of the time, it uses less than about 100 kb/s
 - relayed traffic (blue trace) suggests relaying not often needed

Skype VoIP Call Duration



- Semi-log CDF of VoIP call duration
 - could only measure duration of relayed calls
 - median 2m 50s, average 12m 50s, maximum 3h 26m
 - “much higher than traditional telephone calls 3m average”

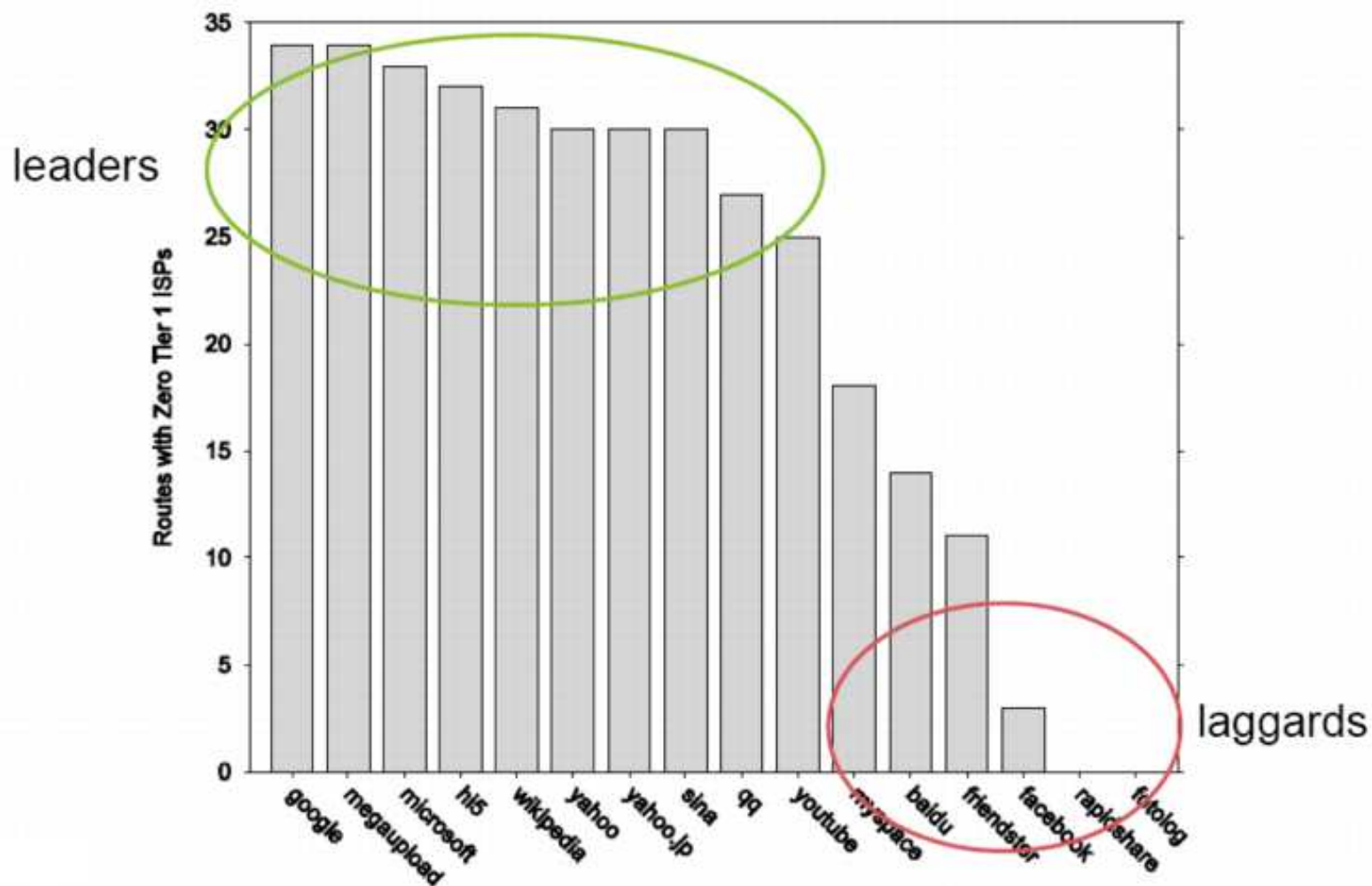
2008: Content Providers and Global Infrastructure

- ‘Cloud’ providers are deploying private WANs. For example:
 - since mid-2006 Google has been building 10 large data centres
 - they have also built a global WAN to provide transport to them
- Gill, Arlitt, Li & Mahanti looked at ‘top 20’ content providers
 - traced routes to each of them from 50 public traceroute servers
 - looked up AS numbers for each IP address on the traceroutes
 - example traceroute on next slide . . .
- *The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse?* Phillipa Gill, Martin Arlitt, Zongpeng Li, Anirban Mahanti. PAM 2008

traceroute from Othello server (UK) to Google server (US)

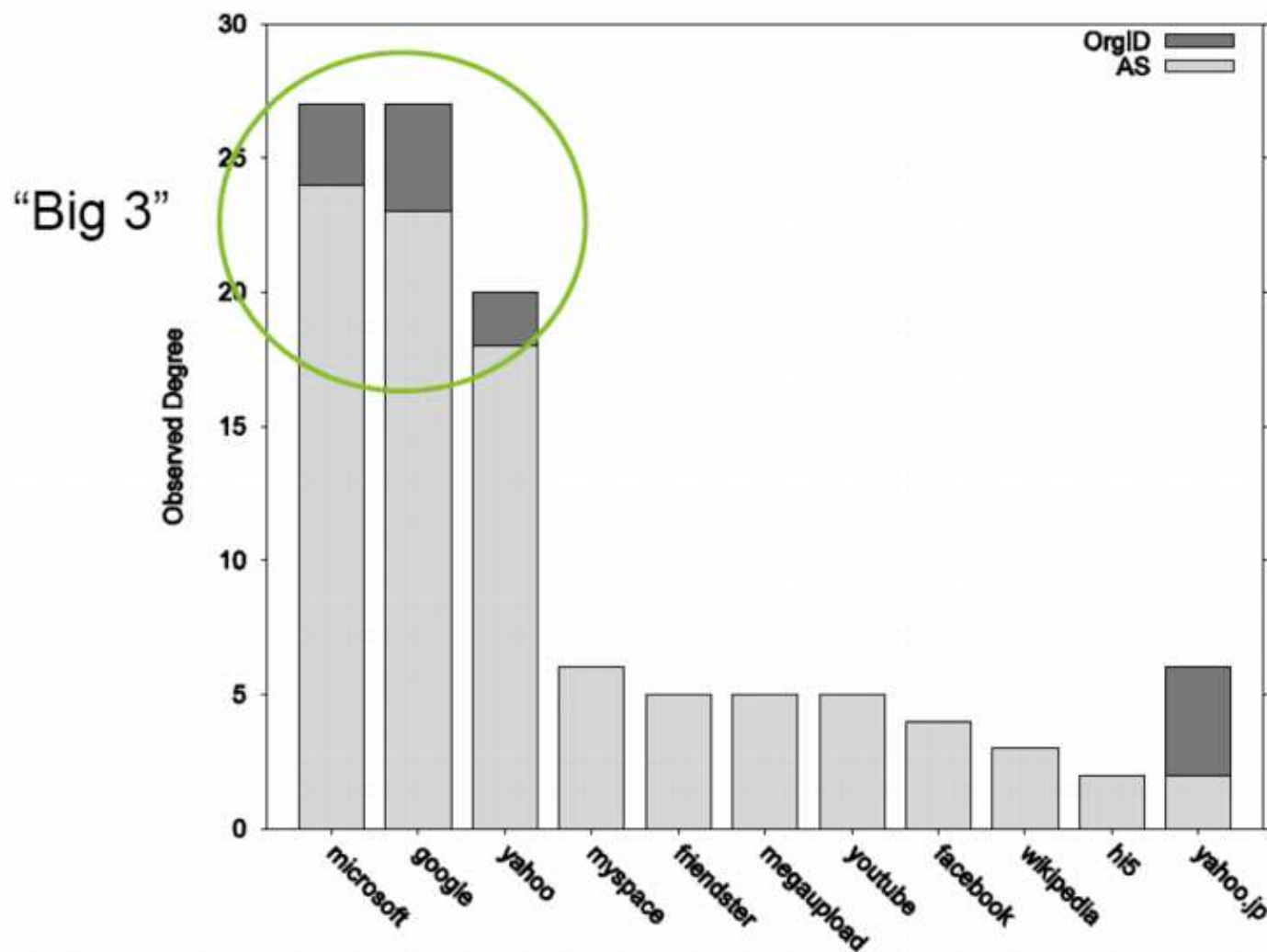
Hop	Hostname	IP address	AS	Latency Measurements		
1	*.uk.othellotech.net	80.82.140.227	29527	0.239 ms	0.201 ms	0.187ms
2	transit2.as29527.net	80.82.140.42	29527	0.412 ms	0.401 ms	0.389 ms
3	peering1.as29527.net	80.82.140.43	29527	0.484 ms	0.476 ms	0.463 ms
4	unknown-LIPEX NA	193.109.219.50	N/A	0.941 ms	1.048 ms	1.163 ms
5	unknown-LINX-PEER-1	195.66.224.125	702	0.890 ms	1.102 ms	1.089 ms
6	unknown-GOOGLE	209.85.252.40	15169	1.323 ms	1.238 ms	1.224 ms
7	unknown-GOOGLE	72.14.236.216	15169	71.532 ms	71.396 ms	68.750 ms
8	unknown-GOOGLE	216.239.46.227	15169	90.721 ms	90.755 ms	90.744 ms
9	unknown-GOOGLE	72.14.233.115	15169	82.481 ms	136.491 ms	136.489 ms
10	unknown-GOOGLE	72.14.233.56	15169	136.401 ms	136.359 ms	136.349 ms
11	unknown-GOOGLE	72.14.233.119	15169	139.580 ms	140.230 ms	140.225 ms
12	unknown-GOOGLE	72.14.233.54	15169	186.172 ms	143.874 ms	146.345 ms
13	unknown-GOOGLE	72.14.233.27	15169	141.707 ms	144.377 ms	147.607 ms
14	unknown-GOOGLE	216.239.47.34	15169	145.104 ms	157.081 ms	147.812 ms
15	po-in-f104.google.com	72.14.253.104	15169	145.284 ms	145.394 ms	145.386 ms

'No Tier-1 Hops' paths to Content Providers



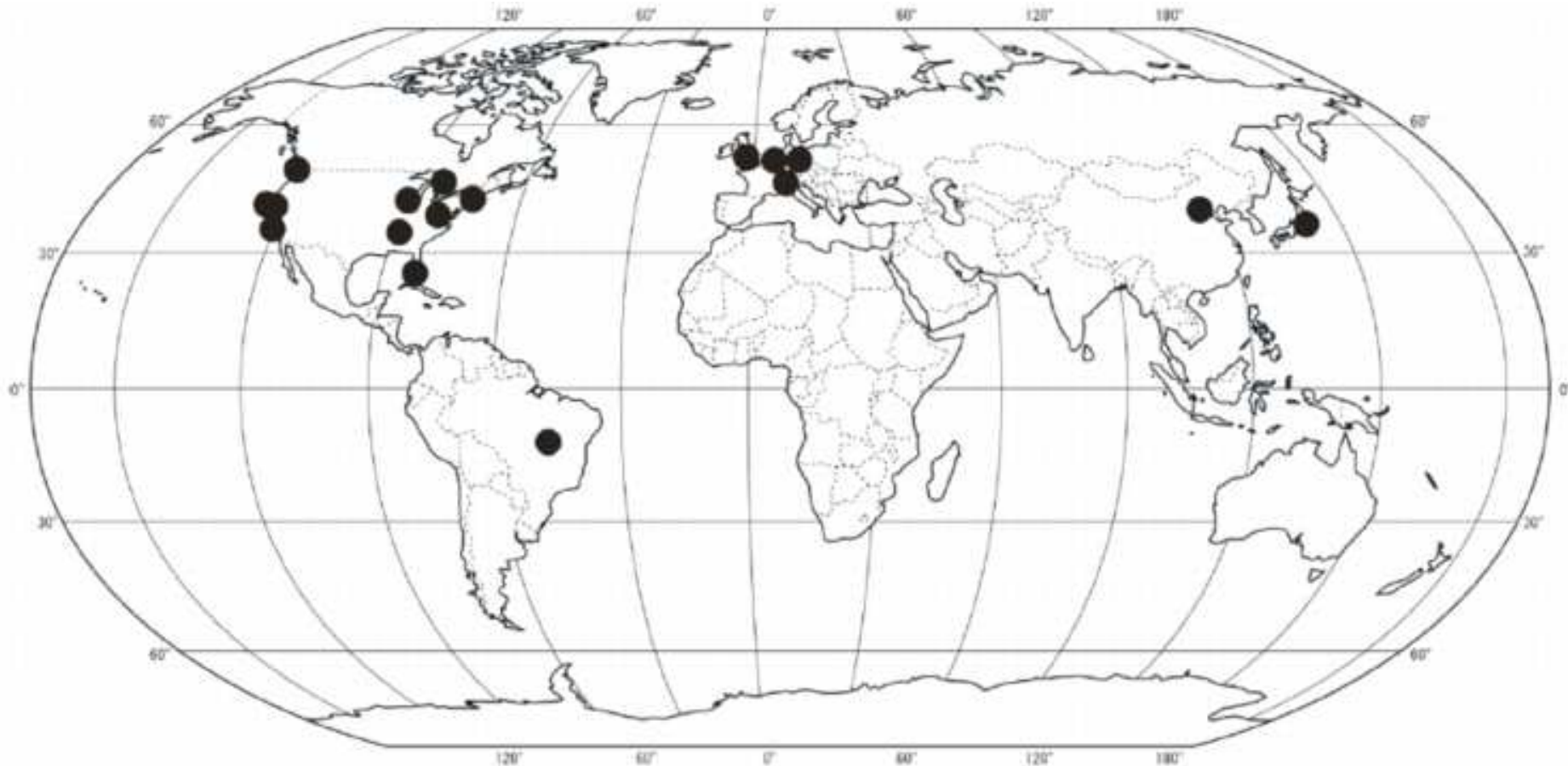
- 'Leaders' had 25~34 (of 50) paths with no tier-1 hops
- 'Laggards' had significantly fewer 'no tier 1' paths

Connectivity of Large Content Providers



- ‘Big 3’ content providers connect to (27, 27, 20) ISPs
- Others have connections to fewer ISPs

Google's Global WAN



- Gill, Arlitt, Li & Mahanti's traceroutes show that Google's WAN has 10 North American, 4 European, 2 Asian and 1 South American cities
- This is clearly flattening the Internet global topology

Conclusion

- I've presented my own view of Internet evolution, as revealed by measurement since the late 80's
- The Internet will continue to develop and change
- If it is to continue to grow, we face many development challenges, for example:
 - scaling issues in the global routing system
 - security issues, particularly in vital areas such as routing and the Domain Name System
- I hope you're now convinced that as well as being important, *Internet measurement is fun!*

Bibliography

- *On the Self-Similar Nature of Ethernet Traffic.* Will Leland, Murad Taqqu, Walter Willinger, Daniel Wilson. SIGCOMM, 1993
- *NSFNET: Final Report, 1987-1995.* Available from:
http://www.merit.edu/documents/pdf/nsfnet/nsfnet_report.pdf
- *Understanding Internet Traffic Streams: Dragonflies and Tortoises.* Nevil Brownlee. IEEE Communications, Oct 2002
- *The impact of residential broadband traffic on Japanese ISP backbones.* Kensuke Fukuda, Kenjiro Cho, Hiroshi Esaki. CCR Jan 2005
- *An Experimental Study of the Skype Peer-to-Peer VoIP System.* Saikat Guha, Neil Dswani, Ravi Jain. IPTPS '06 (5th International Workshop on Peer-to-Peer Systems)
- *The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse?* Phillipa Gill, Martin Arlitt, Zongpeng Li, Anirban Mahanti. PAM 2008