My Quest For Truth On The Internet

Mark Allman
International Computer Science Institute

Hacker's Society
Case Western Reserve University
March 2012

Collaborators

- Mohan Dhawan, Rutgers
- Justin Samuel, UCB
- Renata Teixeira, CNRS & UPMC
- Christian Kreibich, ICSI
- Nicholas Weaver, ICSI
- Vern Paxson, ICSI & UCB

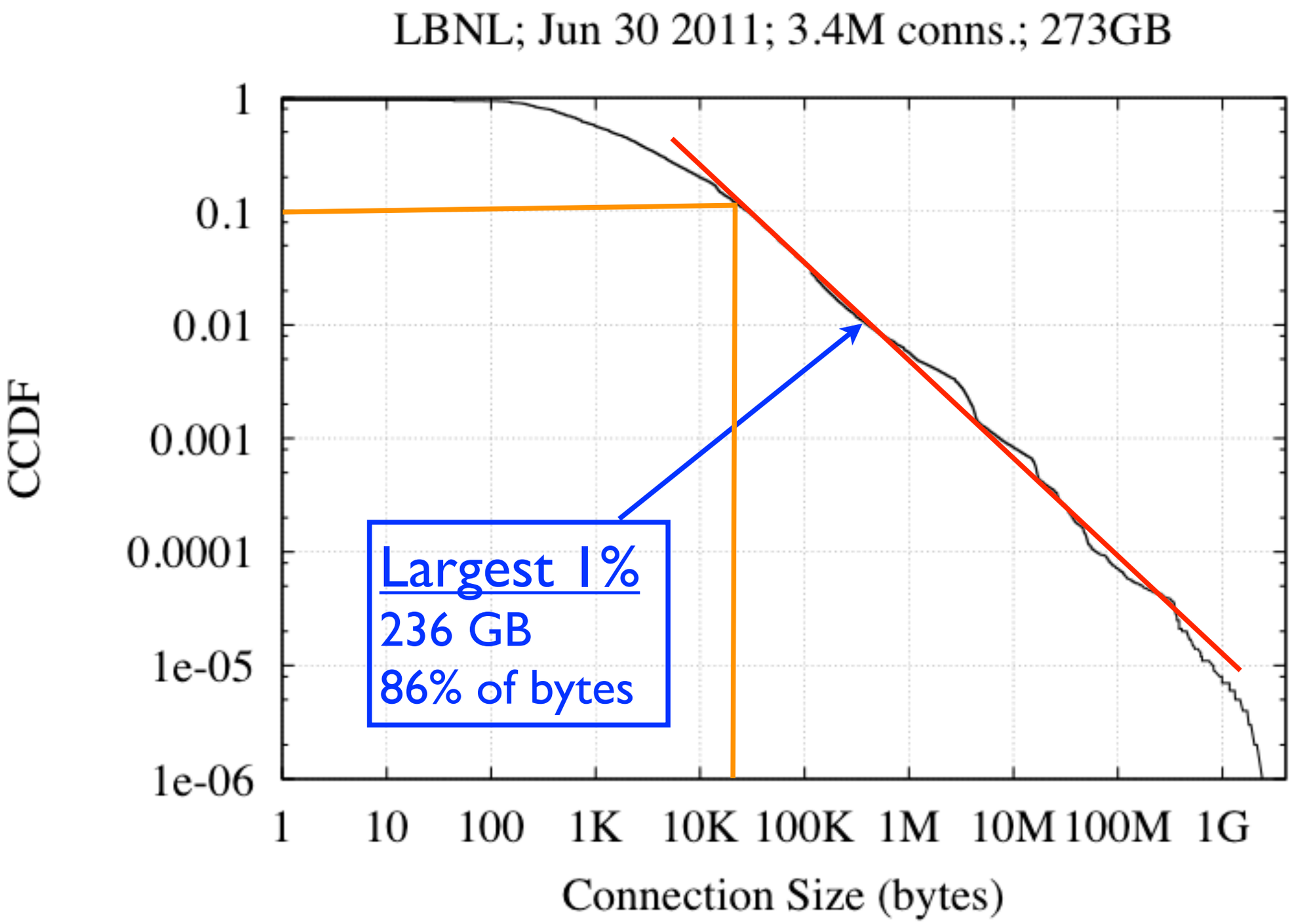
Background

- Everyone has a mental model of how the Internet works
  - .... some models informed
  - .... some not so much
- Models are very individual and come from experience
- The vast majority of such models are wrong!

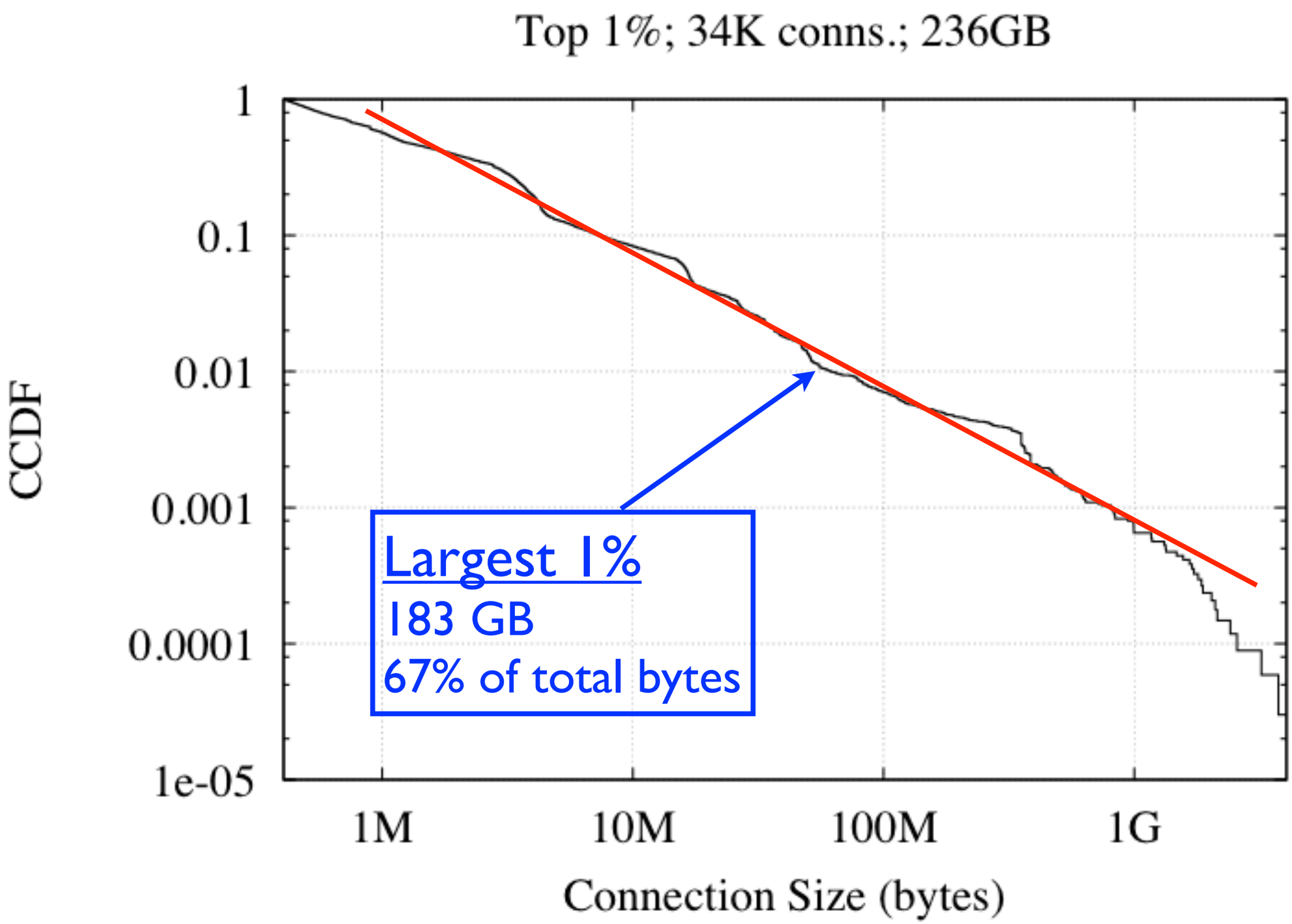
Background (cont.)

- A subset of the networking community has taken the task of understanding networks *empirically*

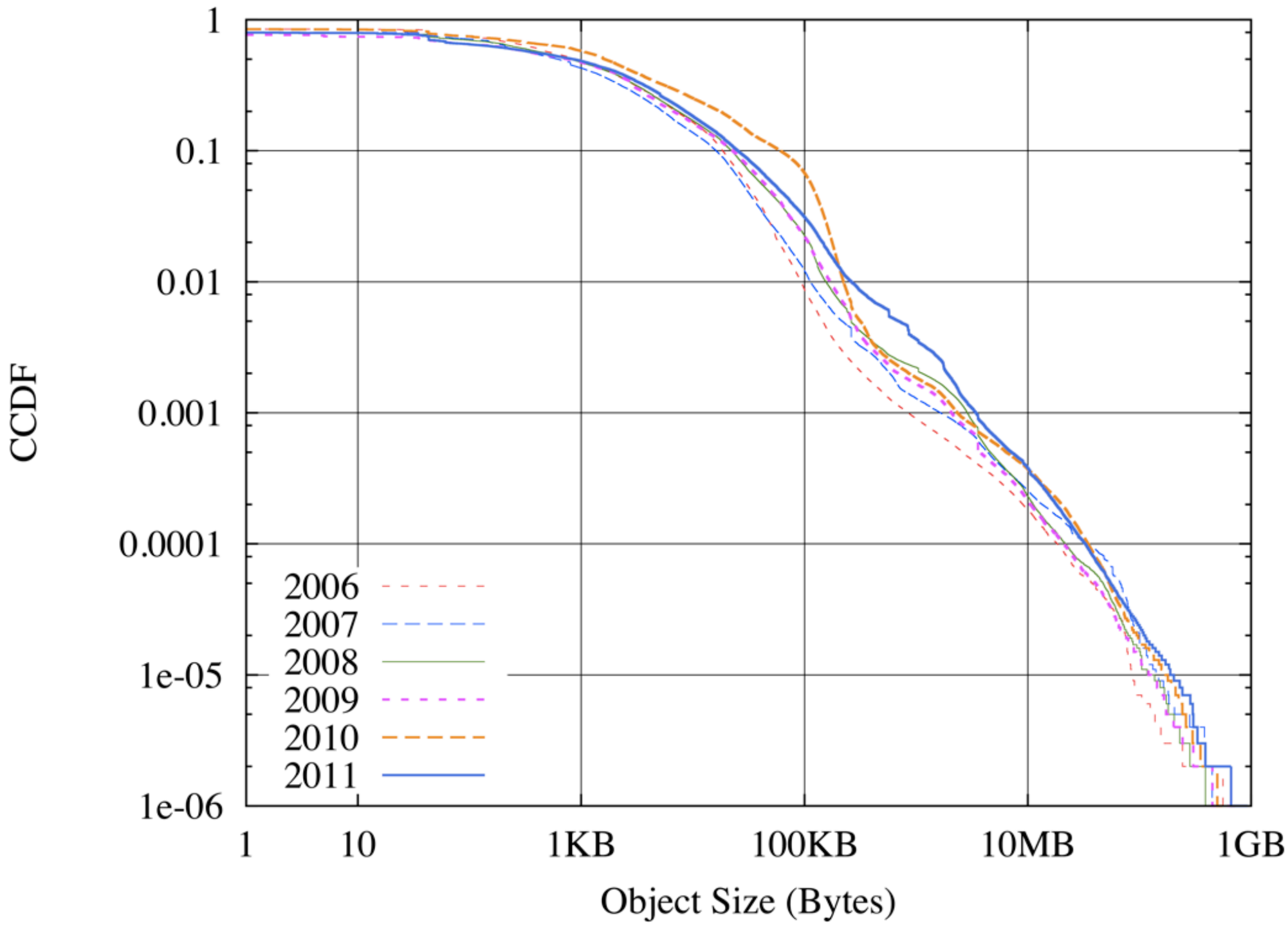
Example


LBNL; Jun 30 2011; 3.4M conns.; 273GB
CCDF
1
0.1
0.01
0.001
0.0001
1e-05
1e-06
1
10
100
1K
10K
100K
1M
10M
100M
1G
Connection Size (bytes)
Largest 1%
236 GB
86% of bytes

Example (cont.)



Example (cont.)



Courtesy Tom Callahan, Case

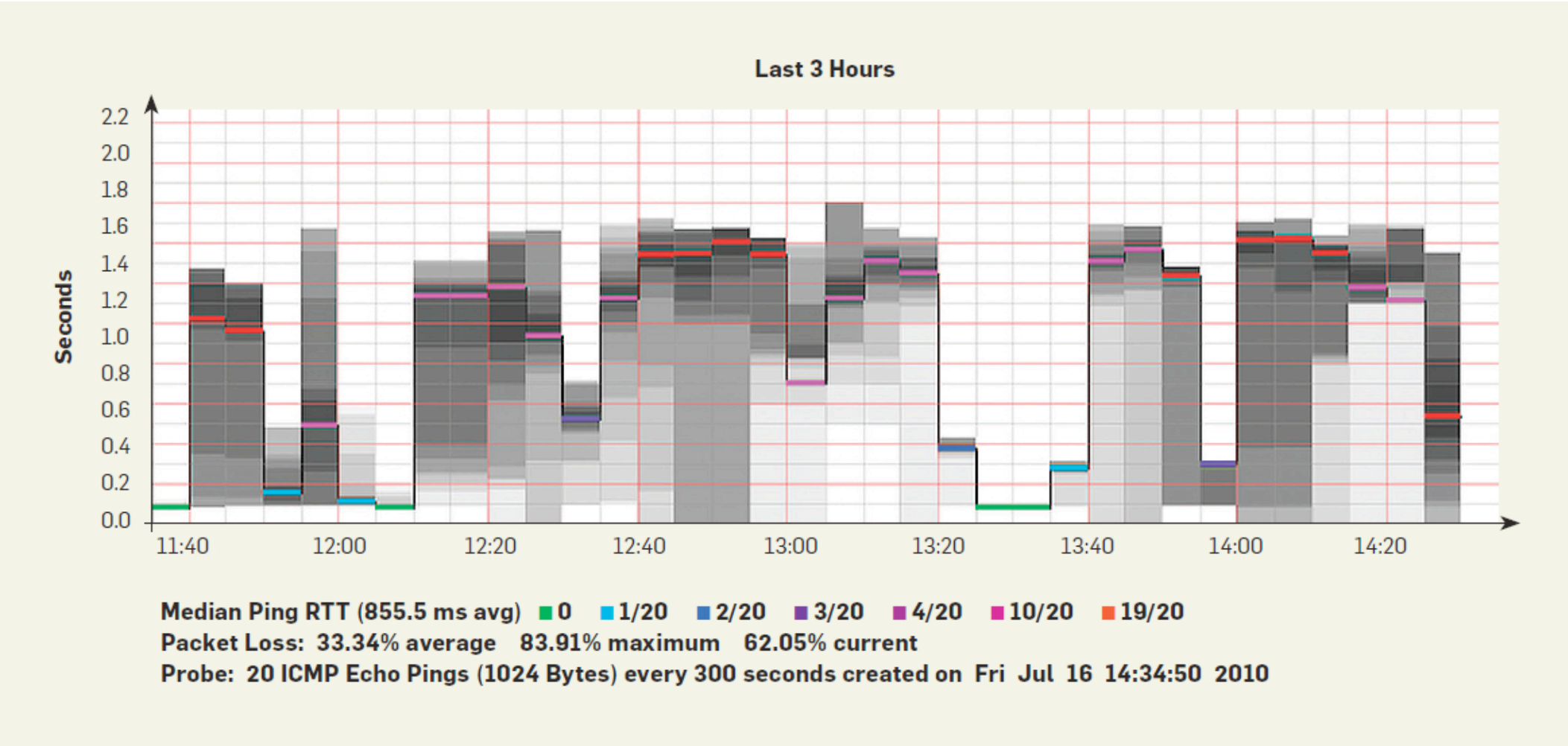
Heavy Tails

- *Most of the connections are small*
  Most of the bytes are sent on big connections

Motivating Example

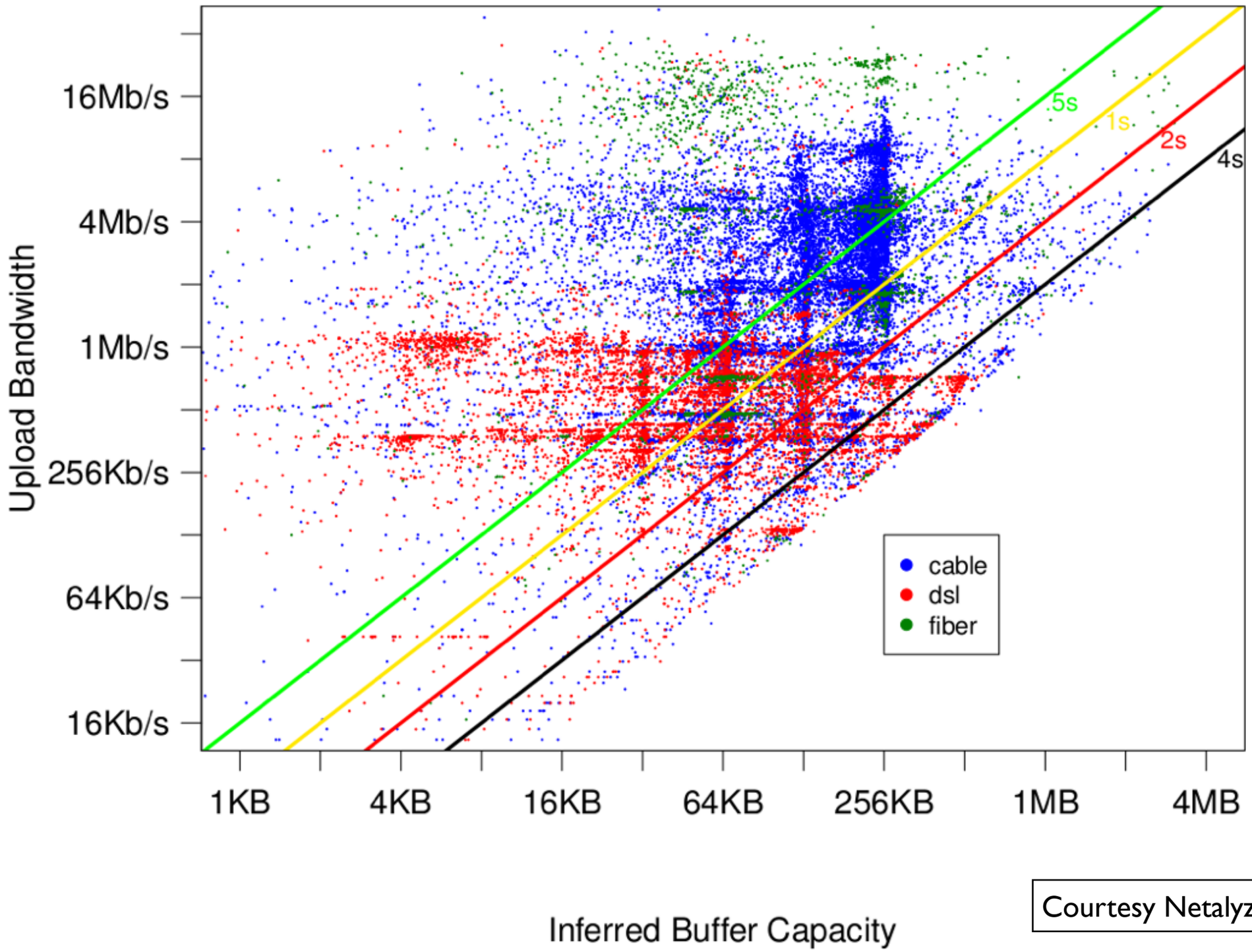
- “Bufferbloat” problem

Bufferbloat



Courtesy: Jim Gettys

Bufferbloat



Courtesy Netalyzer project

Bufferbloat

- What have I shown you?
  - that bufferbloat *can* happen

- But, what have I not shown you?
  - how often does it happen?
  - how long does it last?
  - how bad is it when it happens?

- What do we really need to reason about this "problem"?
  - *more data!*

Bufferbloat

- An experiment...
  - instrument a bajillion end hosts to do periodic pings
  - collect the data
  - assess ....
    - how often the delay increases appreciably
    - and, by how much

How To Measure?

- Two general measurement approaches:
  - active probing
  - passive observation

Active Measurement

- Send probes of various kinds into the network...
  - ... learn something from the replies

- E.g., *traceroute*, *ping*, *wget*, ..., plus lots of custom crafted tools

- Pros: easy to conduct to broad set of hosts
- Cons: synthetic workload, perturbing the system, local network bias

Passive Measurement

- Watch what is naturally occurring on the network and draw conclusions

- E.g., *NetFlow*, *tcpdump*, *wireshark*, *Bro*, etc.

- Pros: real traffic in all its "glory"
- Cons: privacy issues, limited view, local network bias

Common Problem

- Vantage point issue
  - E.g., determines what and who can be monitored
  - E.g., determines network path characteristics that color all probing

- *What we can measure is not always what we'd like to measure*

Common Problem

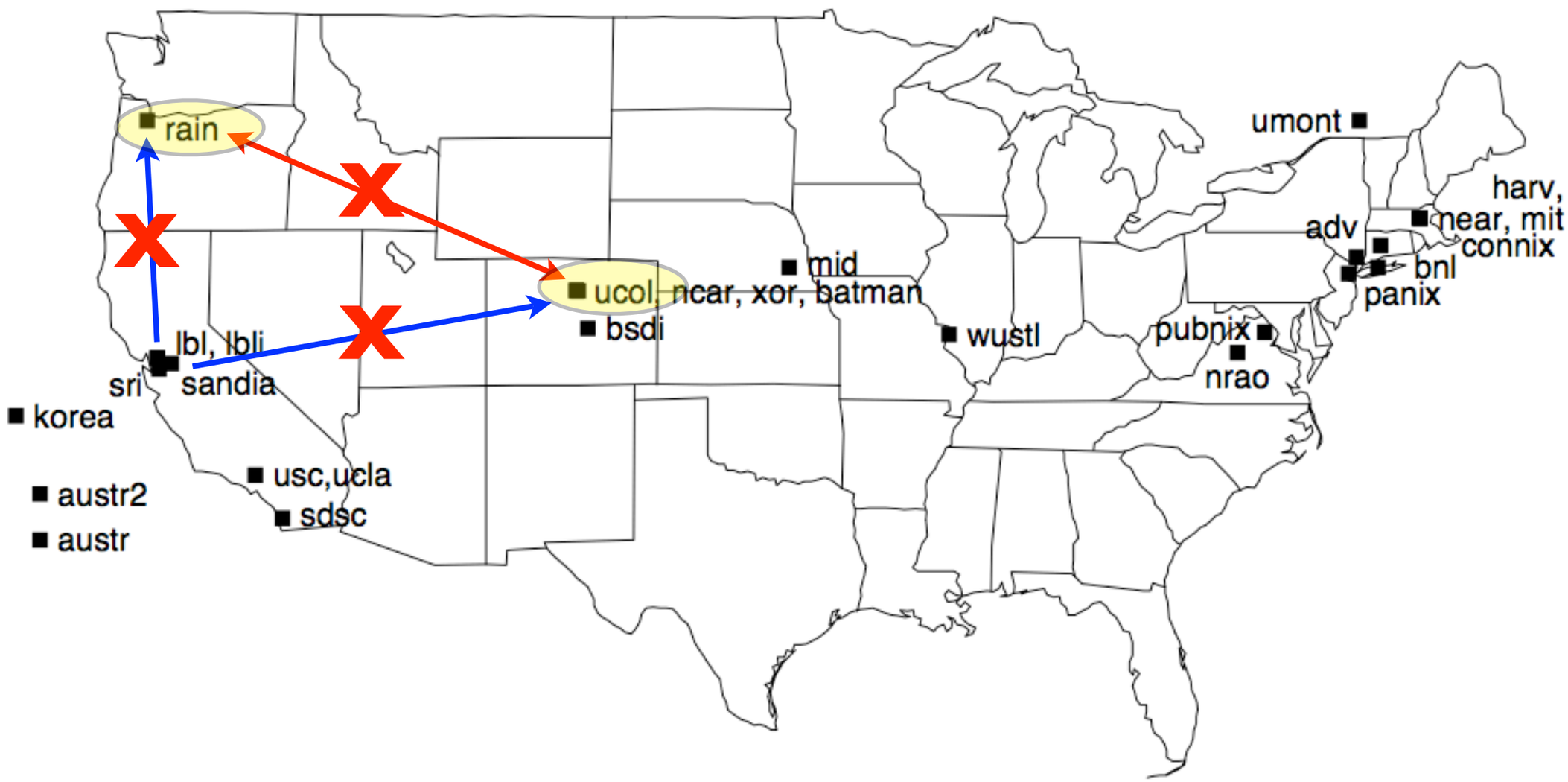
- What to do about this?
  - *more vantage points!*

- Lets solve a systems problem---understanding the Internet---with a system!

Measurement Platforms

- First attempt: *npd*, mid-90s, Vern Paxson
  - 30-some measurement hosts coordinated by a controller at LBNL
  - mostly geared towards active measurement

npd Sites



Vern Paxson, SIGCOMM 1996

Additional Platforms

- npd brought forth myriad platforms
  - NIMI, the direct followon work
  - skitter, Dimes, Neti@Home, Ark, DipZoom, PlanetLab/Scriptroute, speed tests, M-Lab, ...
  - Dasu, HostView, HomeNet

- Goal: enable widespread Internet measurement
- Result: failure
  - (for some definition of failure)

Additional Platforms

- But, even in failure ...
  - ... there are useful research results
  - ... even after taking into account:
    - narrowness
    - bias
    - etc.
- This tells us ....
  - ... the problem is not easy
  - ... but it is worthwhile
  - ... the state of the art is pathetic!

The Big Problem

- Getting software installed
  - (and updated)
- Many reasons to say "no", few incentives to say "yes"
  - security, privacy, resource consumption, who cares?, fear of the unknown, etc.

- npd: 35 sites in 1995
- PlanetLab: 539 sites in 2012

One order of magnitude in 17 years

Leveraging Apps

- One platform that has lead to some useful and wide-scale measurement research:
  - instrumenting BitTorrent clients
- Solved the incentive problem by researching how to make BitTorrent perform better
- Con: single platform one-off
- Hmm.

Academics, Stand Back!

- Researchers are not the only folk who need measurements

Google Maps Question

HarrierMan Level 1 2/11/09

What is up with Google Maps? It is either incredibly slow or pages just don't appear unless I click on the 'Still loading ...Slow? Try Basic HTML' button that has started to appear at the of of the page. This has happening for at least a week, maybe much longer. My connection runs at about 6Mb/sec and is fine for other sites so this is a google map issue. I notice many other posts over the last year where people have had problems at times when I haven't so is just my turn to be frustrated?

Google Maps "Answer"

Mike CH 3/10/09

"I'm afraid I don't have an actual answer"

I'm afraid I don't have an actual answer for you guys, as "Maps got slow" is too vague for

"When did you first notice this problem?"

an increase in latency lately, so whatever is going on is only affecting a small proportion of

"Which browsers, operating systems and (if possible) security patchlevels do you have?"

What does slow mean?

eg, "It use
loads fast

Is the CPU maxed out?

s do you have?

What other software are you running?

oolbars, chat

clients, etc installed?

Does it "seem to be waiting on the network"?

"What area of the world are you connecting from?"

A Sweet Spot?

+ Everywhere
+ All the time

- No real API
- No security model

Wouldn’t This Be Sweet?

```
<html>
<body>
  <script type="text/javascript">
    function tr_callback(results) { ... }
    measurements.traceroute("google.com",
                            tr_callback);
  </script>
</body>
</html>
```

Browser Design Space

- So, what would it take to make the browser a general measurement platform?

- Ideally, standard Javascript would do the trick
  - does suffice for some things and broadly available
  - but, it is lacking in access to sockets, host properties, etc.

Browser Design Space

- Browser extensions
  - provide a rich API to Javascript
  - run with browser privileges
  - can (with effort) access network, file system, host properties, etc.
  - compiled (via JIT) so overhead should be low and accuracy should be high
  - must be installed by the user (ugh!)

Browser Design Space

- Runtime plugins
  - Java, Flash, Silverlight, etc.
  - popular for one-off measurements
  - sandboxing keeps measurement away from web page
    - good & bad
  - runtime security policies tend to prevent things like arbitrary execution of *ping* or *traceroute*
  - portability? .... eh ....

Browser Design Space

- Custom plugins
  - if Java, Flash, etc. are too much ....
  - .... we could develop our own
  - big effort

Browser Design Space

- Browser changes
  - build a measurement API into the browser itself
  - in many ways ideal
  - major effort to implement in an open source browser
  - without buy in from a browser we'd fork ...
  - perhaps if we take another approach we could circle back to this one

Browser Design Space

- Standard in-page Javascript
- Browser extensions
- Runtime plugins
- Custom plugins
- Browser changes

Intrusiveness

Fathom

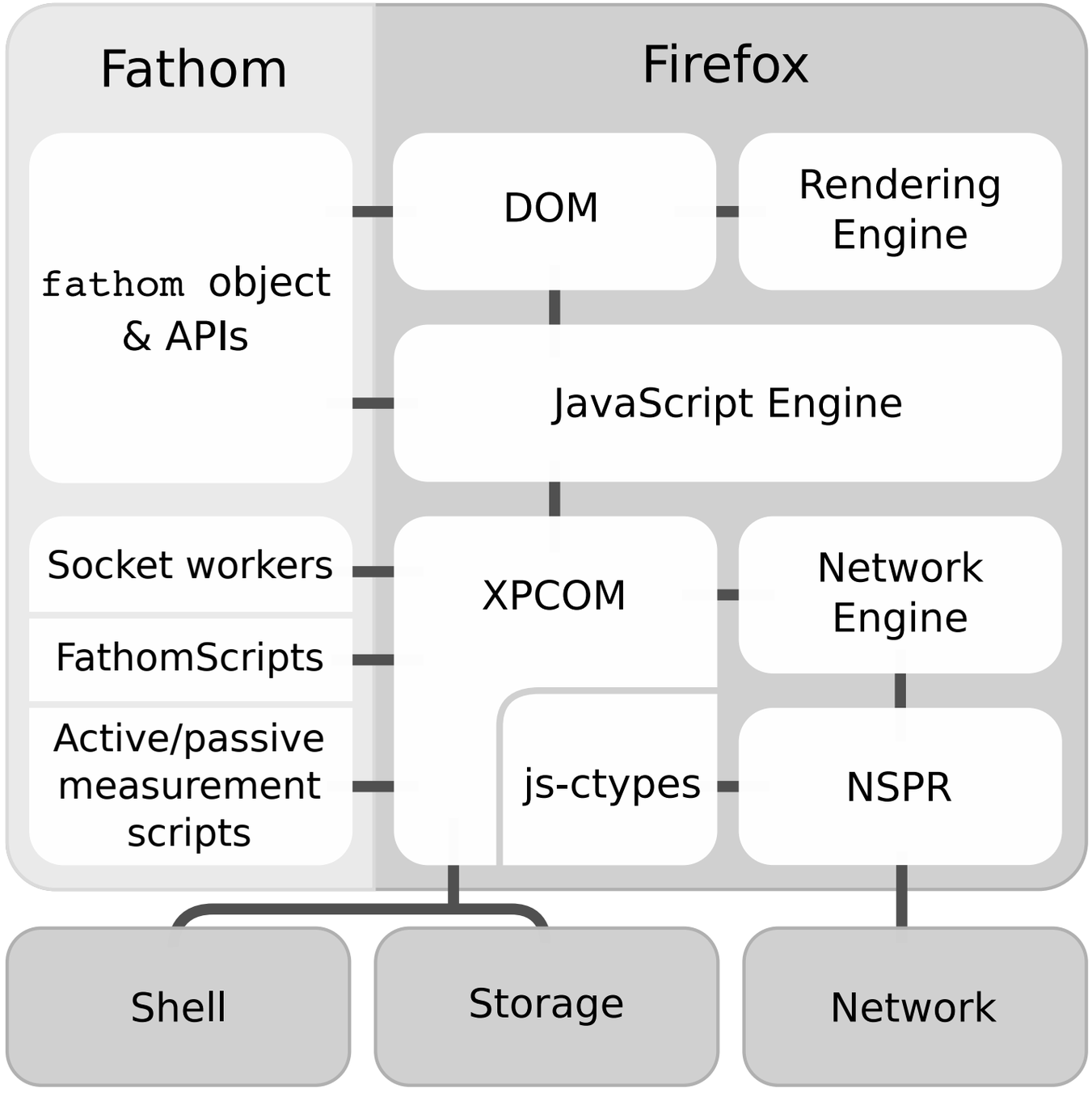
- A measurement platform within Firefox
- Just one browser, but ....
- .... a very nice sweet spot of features
  - popular browser
  - Javascript measurement code (open source)
  - portable across Firefox platforms

- Extension must be installed once, measurement code comes within web pages

Fathom API

API	Availability in JavaScript	Flash	Java Applet
`fathom.socket.tcp.*`	◐	◐	●
`fathom.socket.udp.*`	○	○	●
`fathom.socket.broadcast.*`	○	○	●
`fathom.socket.multicast.*`	○	◐	●
`fathom.proto.dns.*`	○	○	●
`fathom.proto.http.*`	◐	●	●
`fathom.proto.mdns.*`	○	○	●
`fathom.proto.upnp.*`	○	○	●
`fathom.system.getActiveInterfaces()`	○	○	●
`fathom.system.getGateway()`	○	○	◐
`fathom.system.getRoutingTable()`	○	○	◐
`fathom.system.getResourceUsage()`	○	○	◐
`fathom.system.getWifiInfo()`	○	○	◐
`fathom.system.getNetworkUsage()`	○	○	◐
`fathom.system.doTraceroute()`	○	○	◐
`fathom.system.doPing()`	○	○	◐
`fathom.utils.browser.*`	◐	◐	◐
`fathom.utils.timer.*`	●	●	●
`fathom.utils.metrics.*`	○	○	○
JavaScript & DOM access	●	◐	◐

Architecture



Overhead

Benchmark	# HTTP	# Scripts	# Images	Avg. load time (ms)
Craigslist	4	2	0	512
YouTube	28–34	2–3	17–23	869
Google Maps	39–40	7–8	26–28	1233
Yahoo	49–58	4–5	31–43	1168
Slashdot	35–60	2–12	22–30	2381
ESPN	85–89	7–8	61–63	1761
CNN	89–104	6–9	22–59	1458
NY Times	115–125	23–26	77–80	1144

								n time n (ms)	Avg. Fathom overhead(%)	Avg. Firebug overhead (%)
ESPN	85–89	7–8	61–63	1761	1805	2325		9	1.1	4
CNN	89–104	6–9	22–59	1458	150			37	3.1	10
NY Times	115–125	23–26	77–80	1144	115			60	2.9	29
								42	1.0	50
								51	1.4	71
								96	2.5	32
										55
										57

Avg. Fathom overhead(%)	Avg. Firebug overhead (%)
1.1	4
3.1	10
2.9	29
1.0	50
1.4	71
2.5	32
3.1	55
1.3	57

Timestamp Accuracy

- We timestamped sent and received packets in Fathom and via *tcpdump*
- Multiple browsing scenarios:
  - Fathom timestamps within 1msec of tcpdump
- With 4 *iperf* flows as cross-traffic on the local network:
  - a worst case of sorts
  - average difference in timestamps is 729msec

Timer Accuracy

- Send 200 byte packets every 200msec
- Under various cross traffic conditions

- Average accuracy is within 2msec
- Under heavy load we do see some outliers that are hundreds of msec

Privacy Model

- Goal: no exfiltration of sensitive information
- Fathom network I/O is independent of browser network I/O
  - e.g., script from *a.com* accesses *b.com*, but *b.com* does not get current session cookies
- Measurements open opportunities to fingerprint users (even if using proxies)
- Fathom conforms with "private browsing mode"

Security Model

- We'd like to ensure Fathom is not used as part of malicious behavior
- Ultimately, impossible while still getting work done
- But, we'd like to not exacerbate the situation
- Five inputs to policy decisions

Client Policy

- E.g., no traceroutes
- E.g., no DNS lookups
- E.g., only two simultaneous TCP connections
- Etc.

User Confirmation

- We can always fall back to asking the user if other policy components do not arrive at an answer
- Try to avoid this ...
  - users do not want to be bothered
  - its hard to ask a question most users can answer
- When we do have to ask, try to frame in terms of the high-level threat

Script Manifests

- Script must declare which elements of the API will be used
  - others will be off-limits
  - this lets us readily test whether a script is consistent with local policy

Script Manifests

```
[API subset]://destination(s):port(s)
```

```
http://*.google.com:*

udp://10.1/16:53,5353

*://{upnp}:*
```

Server Manifests

- `http://server/fathom.json`
- Akin to robots.txt
- Allows the server to green-light particular measurements

Code Signing

- Code can be signed
- Lets users reason about people, not code
- Lists of known researchers can be developed

Policy Decisions

- Default client policy
- User consent
- Script manifests
- Server manifests
- Code signing

- Stir together ....

Case Studies

- Netalyzer re-implement
  - http://netalyzer.icsi.berkeley.edu
- Local debugging button
- Google Maps debugging

- Preliminary ...

Bufferbloat ... revisited

- Fathom enables the essential aspects needed for the experiment sketched at the beginning of the talk
- I.e., periodic pings *from a zillion vantage points* to assess the delay through the network
  - and hence the buffer occupancy

Summary

- We hope Fathom hits a sweet spot for measurement platforms such that it finds broad use
- We think we have the incentives right
  - ... or, at least better
- We think we have a reasonable security model
  - ... or, at least better
- Will Fathom solve the world's problems or be join many other such efforts in the dustbin?!
  - who knows ....

Questions? Comments?

http://fathom.icsi.berkeley.edu/

Mark Allman

mallman@icir.org
http://www.icir.org/mallman/