

The Internet is Broken, and How to Fix It

Jim Gettys
Bell Labs
July 27, 2012

Bufferbloat Status

- Many problems all over the network: the edge is likely the most severe, though it is endemic in hosts, home routers, broadband gear, 3g, some switches, overloaded routers.... Be paranoid!
- Reminder: *two* bottlenecks are in play in the home
 - Broadband hop (single queue!)
 - Wireless hop (potentially four HW queues in 802.11)
- Most computer nerds have at least heard the word “bufferbloat”, but don't understand the severity
- Articles are out in CACM
- DOCSIS (cable) change is deploying

Isolated Changes Collude to Produce Systemic Disaster

- Only a single, bloated queue is available today for customers' Internet service
- Applications maximize their benchmarks oblivious to apps sharing the link: e.g. web browsers using many TCP connections.; web site sharding; IW10 change multiplies by all this by 2.5x; we're ten times worse than RFC 2616 era
 - You can easily have transient HOL blocking of > 100 packets
- Ack clocking often MIA: broadband may merge acks
- “Smart” hardware on servers transmits large line-rate bursts to the same client w. no pacing nor muxing between clients: these arrive **splat** at the broadband head end
- Bandwidth in the network edge is highly variable: static buffer sizing is a non-starter

There is No Single Bullet

- Multiple queues are essential in broadband equipment
- AQM needed to avoid elephant flows: TCP's responsiveness is quadratic in the delay
- “Fair” Queuing is also needed
 - “Fair” depends on where you are: I don't mean simply TCP fair queuing but “fair” among TCP flows, among devices, among customers, among policies
 - TCP fair queuing helps RTT fairness, ack compression, interactive versus non-interactive bulk transfers, etc.
- Port based Classification & diffserv with multiple queues are still essential: one 1500 byte packet @ 1Mbps == >13ms

CoDel Status

- Ethernet easy (with BQL): wireless not so easy, due to driver buffering for aggregation
- Linux 3.5 has codel and fq_codel queue disciplines
- We really, really like fq_codel
 - Fair queuing is only 2% of CPU on 10GigE
- A problem was recently discovered in CoDel under high load with many flows; possible solutions are being simulated and implemented

Really Big Headaches

- Current broadband has a single bloated queue
 - The technologies admit to additional queues, but these are today only available to the ISP's telephony services
- Broadband splits the diffserv domain between the customer & the broadband head end
 - How to communicate the customer's classification preferences? (At least two possibilities...)
 - A explicit protocol
 - Andrew McGregor's idea to infer incoming classification & marking from outgoing marking on flows

Home Router Disaster

- Home routers are busted in 4 ways
 - Horribly antique and insecure; ships 5 year old software on new hardware
 - IPv6 deployment is now gated by the home routers
 - Bufferbloat
 - Funding model of that market is broken
- Come help debug CeroWrt, a home router you might like

CeroWrt

- An advanced build of OpenWrt
- Running Linux 3.3.8 release with CoDel, BQL
- Current Bind & DNSsec in chroot jail
- Routes, not bridges
- Running CoDel, which is currently only partially effective due to buffering in the drivers
- Real web server, proxy, IPv6 support, etc.....
- Come help test, develop, and improve