Latency & AQM
Observations on the Internet

Jake Holland
Akamai
Outline

● Background & Goals
● Results
  ○ 2 conclusions with outline of observations
● Next Steps
  ○ Looking for a collaborator

+ Backup slides with bonus stuff
Background & Goals

Akamai enabled ECN (RFC 3168) globally in February 2020

Initial Questions:

- How prevalent is ECN usage and utilization?
- Can home router solutions fix latency variation?
  - Target: reliable 45ms End-to-End, for gaming

(PS: plz finish in ~3 weeks.)
Conclusion 1:  CE-marking low but growing

- Growth pattern suggests some ISPs have picked up marking since March
- Inspired by “Experience Enabling ECN on the Internet” @IETF 98
  - Looks a bit different from the server side
  - Looking Downstream-only (“server ever saw ECE”)
  - Looked by ASN instead of by country
  - (Note: prior 30% Argentina observation was separately discovered as a bug)
March vs. July: CE-marking prevalence, per AS#

Top 100/~800 ASNs = 1/40th of CE-marking paths (== most CE paths on internet not ISP-managed)

Very probably ISP-managed

Probably ISP-managed?
(global baseline = 0.3%)

Perhaps not ISP-managed?
Conclusion 2: Home Router AQMs not sufficient :( 

- **Note:** this study excludes TCP self-congestion (SYNACK to ACK only)
  - Home routers help a lot for self-congestion!
  - But on small packets, low throughput, in practice problems are still coming from elsewhere, even on CE-marking paths

- CE-marking paths experience ASN-correlated latency variation
  - Usually AQM helps, but not as much as a better ASN
Data Set: Latency Span over 1 day

- Latency Span within Client IP over 1 day
- Grouped by ASN
  - I chose 2 particular ASNs for illustration
  - “Good” and “Middle”, according to 91st and 98th%ile of latency for median client IP
- Inspired by “Measuring Latency on the Internet” @IETF 99’s maprg
  - Server-side passive measurements on production traffic
  - Latency Sample = Delay between SYN+ACK and ACK
Data Filters: isolate access path to end users

- Filtered for “nearby home/office consumer line”:
  - At least 50 samples/client IP from same datacenter in the same day
  - At most 300 samples/client IP in the same day (exclude CGNAT/VPN)
  - At most 20ms for the minimum sample

- Post-filter:
  - 5.3b Latency samples in 31m client IPs
  - 33k clients ever saw markings out of 11m that ever negotiated ECN

- Actually Datacenter <-> ClientIP Pair
  - Same client has different latency span to different datacenter
  - So the same client IP could appear up to 6 times after filter
“Good” ISP Latency Spans: CE-marking vs Overall

Note inversion!: most clients in these CE-marking paths had worse 98th% latency than baseline. (selection bias? Loss artifact from shorter queue with SYN?)
“Middle” ISP Latency Spans: CE-marking vs Overall

RTT Variation Spans 2020-07-19
(per ClientIP<->Datacenter pair)

- Middle ISP Baseline 101.76k pairs (8.91m RTTs)
- CE-marking 0.15k pairs (0.02m RTTs)
- 45ms=gaming target
- 50%ile RTT-minrtt Span(ms) within (ecor,client_ip) pairs
- 75%ile
- 91%ile
- 98%ile
“Good” ISP vs “Middle” ISP

RTT Variation Spans 2020-07-19
(per ClientIP<->Datacenter pair)

Percentile of (ecor,client_ip) pairs

RTT (ms)

Middle ISP 101.76k pairs (8.91m RTTs)
Good ISP 187.61k pairs (16.29m RTTs)
45ms=gaming target
50%ile RTT-minrtt Span(ms) within (ecor,client_ip) pairs
75%ile
91%ile
98%ile
Next Steps

- Find a collaborator
  - Interesting stuff, but out of my scope

Contact jholland@akamai.com if you want to give this data the analysis and write-up it deserves.
Backup Slides
and Supplemental Data
## Single day global counts, 2020-07-19

<table>
<thead>
<tr>
<th></th>
<th>Pre-filter totals</th>
<th>Post-filter totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connections</td>
<td>32.7b</td>
<td>5.3b</td>
</tr>
<tr>
<td>ECT Connections</td>
<td>1.25b</td>
<td>163m</td>
</tr>
<tr>
<td>Connections with CE marks</td>
<td>694k</td>
<td>157k</td>
</tr>
<tr>
<td>Total Client IPs</td>
<td>36.7m</td>
<td>30.7m</td>
</tr>
<tr>
<td>Client IPs that ever used ECT</td>
<td>15.1m</td>
<td>11.1m</td>
</tr>
<tr>
<td>Clients that ever saw CE</td>
<td>55k</td>
<td>33.2k</td>
</tr>
<tr>
<td>Client IPs in 100 top prevalence ASNs</td>
<td>538k</td>
<td></td>
</tr>
<tr>
<td>ECT Client IPs in top 100</td>
<td>252k</td>
<td></td>
</tr>
</tbody>
</table>
# Single day global counts, 2020-03-22

## Pre-filter totals

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Connections</td>
<td>31.2b</td>
</tr>
<tr>
<td>ECT Connections</td>
<td>1.07b</td>
</tr>
<tr>
<td>Connections with CE marks</td>
<td>245k</td>
</tr>
<tr>
<td>Total Client IPs</td>
<td>22m</td>
</tr>
<tr>
<td>Client IPs that ever used ECT</td>
<td>9.1m</td>
</tr>
<tr>
<td>Clients that ever saw CE</td>
<td>22.5k</td>
</tr>
</tbody>
</table>

## Post-filter totals

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Connections</td>
<td>3.5b</td>
</tr>
<tr>
<td>ECT Connections</td>
<td>97m</td>
</tr>
<tr>
<td>Connections with CE marks</td>
<td>54k</td>
</tr>
<tr>
<td>Total Client IPs</td>
<td>17.1m</td>
</tr>
<tr>
<td>Client IPs that ever used ECT</td>
<td>6.4m</td>
</tr>
<tr>
<td>Clients that ever saw CE</td>
<td>12k</td>
</tr>
<tr>
<td>Client IPs in 100 top prevalence ASNs</td>
<td>409k</td>
</tr>
<tr>
<td>ECT Client IPs in top 100</td>
<td>210k</td>
</tr>
</tbody>
</table>
Why cut at 300 for CGNAT/VPN?

RTT Samples per client IP
2020-07-19

Count per Client IP
aggregator exclusion threshold
(over ~300/day -> likely VPN/CGN?)
Precluded Objections

I’ve heard these raised, but I think they’re filtered:

- Variation is from geographic diversity within ISP
  - No: limited to Client IPs that saw min sample <20ms that day
- Variation is from dissimilar paths behind VPNs
  - Minor noise: excluded Client IPs with high sample count (VPN/CGNATs)
- Variation is wi-fi
  - No: major differences between ISPs, even though everyone uses wi-fi
ASN Filter Criteria

Exclude networks too small to use

- >= 450 Client IPs (post-client filters)
  - Sounds small, but by manual check includes many eyeball networks
- >= 100 ECN-using client IPs
  - May include some false positives. (Note that 3% CE-marking prevalence on 100 ECN-using clients is only 3 clients.)
  - However, low counts of CE-marking on CE-capable paths, and likely have false negatives as well.
    - CE marking observations only give lower bound on prevalence
    - Many very tiny connections won’t see congestion (there are lots)
Top 100 ASNs by CE-prevalence full view
Not All Middle ISPs--some get decent AQM benefits

RTT Variation Spans 2020-07-19
(per ClientIP<->Datacenter pair)

Percentile of (ecor,client_ip) pairs

RTT (ms)

- Middle ISP Baseline 117.55k pairs (10.12m RTTs)
- CE-marking 0.77k pairs (0.07m RTTs)
- 45ms=gaming target
- 50%ile RTT-minrtt Span(ms) within (ecor,client_ip) pairs
- 75%ile
- 91%ile
- 98%ile
Time of Day Latency vs. Sample Count (Good)

RTT Samples per 15-min time slot
ASN RTT by TOD, 2020-07-19
Time of Day Latency vs. Sample Count (Good CE-marking)
Time of Day Latency vs. Sample Count (Middle)

RTT Samples per 15-min time slot
ASN RTT by TOD, 2020-07-19

- 98% (316.19 ± 15.62)
- 91% (114.11 ± 6.43)
- 75% (59.55 ± 2.29)
- 50% (41.40 ± 1.42)
- 25% (27.56 ± 1.04)
- 9% (20.53 ± 0.58)
- 2% (0.00 ± 0.00)
Time of Day Latency vs. Sample Count (Middle CE-marking)

RTT Samples per 15-min time slot
ASN(ECT) RTT by TOD, 2020-07-19
Known potential sources of error #1

● Most flows not seeing CE marks even on CE-marking paths
  ○ Most common count of sessions with CE marks per client is 1. 2nd most common is 2, etc. thru 8. (of minimum 50 sessions)
  ○ Likely under-counting, by unknown amount
  ○ Many of our flows are tiny and may miss CE when cross-congested
  ○ Our pacing/CC strategies may often avoid causing congestion

● Downsampled input stream
  ○ Some odd and complex downsampling before I get the data
  ○ Possible source of unknown biases.

● Inter-day trends and comparisons
  ○ Different days have different traffic profiles depending on events
  ○ Not automated, no systematic inter-day comparisons.
  ○ Spot checks for robust conclusions hold up well so far, but are anecdotal
Known potential sources of error #2

- Latency variation due to wi-fi and other effects not known
  - Would be nice to know what’s achievable with ISP access AQM
  - Gamers often run non-wi-fi. Would be nice to differentiate somehow.
- Client IP remapping
  - Thought to be uncommon, but adds unknown amount of noise
- Other Weaknesses in SYNACK->ACK dataset
  - Intra-flow variation would be nice to add
  - Especially with instrumentation measuring pre- vs. post- congestion events, like Toke’s other experiment in maprg 99 talk
- Count of CE-marking paths too small for good numbers in some cases
Known potential sources of error #3

- Non-marking AQM provides same latency
  - CE-marking not a great proxy for AQM, only a lower bound
  - Especially with DOCSIS3.1.
- Unintended side-effects from filtering (trying “close”, did it pick up more?)
- “Experienced a CE” may be biased relative to “did not experience CE”
  - Could explain inversions? (where CE worse @98th %ile)
- Server-side latency (within datacenter, e.g.)
  - Believed but not proven to be negligible. If not, could corrupt results.