ECN Deployment Observations

NGI POINTER

Some Congestion Experienced

HIGH FIDELITY CONGESTION CONTROL

RIPE NCC

`draft-heist-tsvwg-ecn-deployment-observations`

Quantifying L4S Deployment Risks

Jonathan Morton & Pete Heist

This project has received funding from the European Union's Horizon 2020 research and innovation programme within the framework of the NGI-POINTER Project funded under grant agreement No 871528

What is Risk?

- Every engineering discipline must manage risk.
  - **Aerospace** - risk of death, destruction, or serious injury…
  - **Internet** - risk of network, protocol, or application failure.

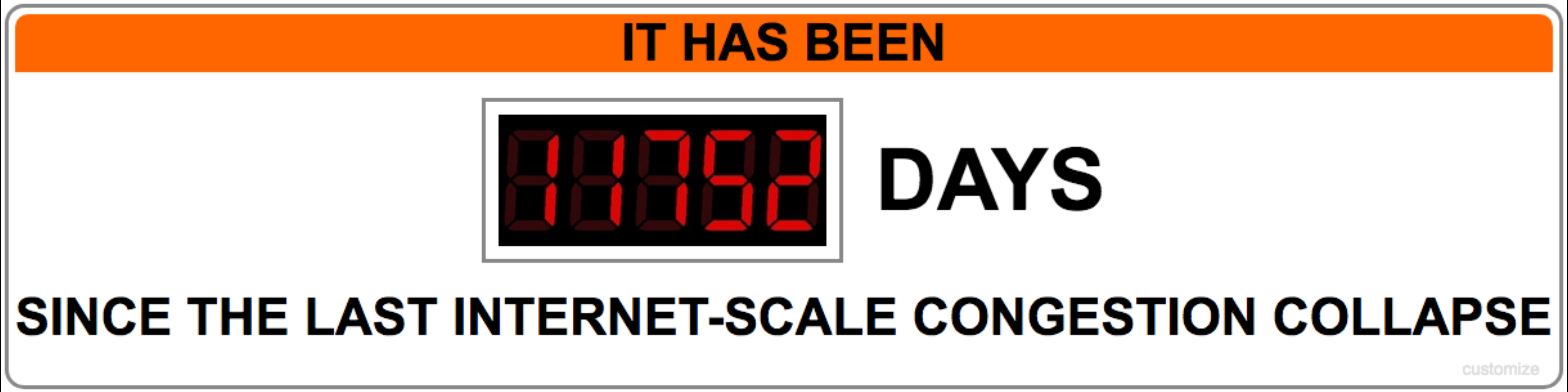
- Risk Quantity = Severity (of harm) * Likelihood (of occurrence)
  - More severe harms must be less likely, to be considered equivalent risks.
  - Threshold of acceptability depends on who the harm is inflicted upon.
    - Involved party - interested observer - innocent bystander

- Severity of harm may be established by laboratory tests.
- Likelihood of harm requires real-world data.

Severity of Harm

Aerospace		Internet
Fatality, Collision, or Hull Loss	**Catastrophic**	Whole Internet stops working
Serious Injury or Property Damage	**Serious**	AS stops working, or Class of traffic unusable
Serious Incident or Minor Injury	**Major**	Service quality seriously impaired for an AS or a class of traffic
Minor Incident	**Minor**	Moderate, sustained, localised impact on service quality
Inconsequential Nuisance	**Trivial**	Small, temporary, localised disturbance in service quality

Likelihood of Occurrence



- Catastrophic failures are unacceptable to <u>ever</u> occur.
- Serious failures tend to make the news.
- Major failures SHOULD be designed out of a system.
- Minor failures will invite a bad reputation if they occur "frequently".
- Trivial failures can be tolerated if they are predictable.

Externalities

- **Involved Party**
  - Most able to understand and control risks incurred by their experiment.

- **Interested Observer**
  - Is aware that the experiment is taking place, and of some of its characteristics.
  - Likely to be able to recognise failures and take mitigating action.

- **Innocent Bystander**
  - Does not know or care about the experiment.
  - May <u>notice</u> the effects of failures, but likely will <u>not</u> be able to recognise their source.
  - Cannot take effective mitigating action.

Principal Risk of L4S Deployment

- When L4S and "conventional" traffic share a conventional ECN AQM instance, the L4S traffic dominates, starving the "conventional" traffic almost completely.
  - Affects both throughput of long-running flows and FCT of simulated page-load traffic.

- Severity: **Major** - since the service quality of an entire class of traffic (conventional TCP) is seriously impaired.
  - Even Not-ECT traffic is equally affected.
  - Only the AQM needs to be ECN enabled to trigger this.

- Likelihood: Whenever the two classes of traffic share an RFC-3168 compliant AQM bottleneck that cannot distinguish between them.
  - Single queue ECN AQMs have not been reliably identified, but…
  - Tunnels can defeat flow-awareness mechanisms, eg. FQ, AF.
    - L4S and conventional traffic sharing a tunnel will use same queue and AQM instance in fq_codel.

L4S Domination

- Steady state of L4S flow is 2 CE marks per RTT.
- Competing AIMD flow quickly collapses to minimum cwnd at same marking rate.

- Single saturating CUBIC flow gets 1.5 Mbps on 50 Mbps, 80ms path.
- Single saturating Prague flow gets the rest, 48.5 Mbps.

- Simulated CUBIC web traffic (exponential arrival, lognormal 64K-2M 90% confidence interval flow length) in competition with single saturating flow:
- Prague background flow <u>quadruples</u> flow completion time relative to CUBIC background flow, both at median and at 95th percentile, on 10-40ms paths.

- CUBIC traffic still functions, but is seriously impaired in performance (harm coefficients 0.85+).

Measuring ECN AQM Deployment

draft-heist-tsvwg-ecn-deployment-observations

- Passive monitor attached to border gateway router of small ISP in Czech Republic.
  - Some internal paths of this ISP have known fq_codel deployment, others not.
  - Data anonymised and summarised for privacy.
  - Some tunnelled traffic was observed, but not classified in detail; probably employees working from home.

- When ECT traffic passes through a congested ECN AQM, CE marks and/or ECE echoes can be observed.
  - Not-ECT traffic can also trigger AQM, but harder to detect (didn't try).
  - Port scanners & fingerprinters added some noise; tried to filter out.
  - Uncongested AQMs are invisible; even known AQMs only detected for 60% of relevant subscribers with outbound ECT flows.

- Approximately **10%** of subscribers <u>not</u> on a known AQM path and who had some outbound ECT flows, showed some plausible ECN AQM activity.
  - This implies considerably higher ECN middlebox deployment than ECN client endpoints!

Serious Question

- Is a 10% incidence rate acceptable for a **Major** severity risk:
  - to involved parties?
  - to interested observers?
  - to innocent bystanders?

Risk Mitigation

- Existing networks & their users are "innocent bystanders".
  - Do not know about L4S & do not care. May not even understand ECN.
  - May have decades-long replacement cycle of equipment - and recently installed some that supports ECN AQM, eg. new consumer CPE.
    - Hence will be decades before RFC-3168 AQMs are expunged from the Internet, if started "tomorrow".
    - Second-hand equipment frequently transferred to economic backwaters, rather than scrapped outright.
  - Users cannot mitigate by disabling ECN at endpoint - still affected!

- The only effective mitigation is to **stop the L4S traffic**.
  - Innocent bystanders have no pre-arranged mechanism to do this.

Risk Avoidance

`Re: draft-white-tsvwg-l4sops`

- If L4S is deployed with a **Major** severity interaction with **10%** of existing networks' users:
  - MUST be on a network explicitly prepared to understand it.
  - MUST be effectively contained to that network.
  - This limits exposure of innocent bystanders to the **Major** severity risk.

- If L4S is contained to a prepared network:
  - DSCP can be used as ID, since end-to-end Diffserv handling can be part of prep.
  - No need to consume ECT(1) codepoint.
    - Leaves solution space open for potentially superior alternatives.

Potentially Superior Alternatives

- Replace ambiguous AQM signalling with unambiguous signalling.

- Jake Holland's solution:
  - Retain ECT(1) as L4S ID, but…
  - L4S AQMs change ECT(1) to ECT(0)
  - RFC-3168 AQMs change ECT(0) or ECT(1) to CE

- Some Congestion Experienced:
  - SCE AQMs change ECT(0) to ECT(1) and include FQ or AF function
  - RFC-3168 AQMs change ECT(0) or ECT(1) to CE
  - Running code, ready to try.

- Both of the above are compatible with existing deployed ECN AQMs, and fail safe.

Conclusions

- RFC-3168 middlebox deployment is significant, as seen from an Eastern European vantage point.
  - All thanks to Kathy Nichols' invention of the CoDel AQM algorithm.
- Tunnels between home offices and workplaces are a fact of life as of end 2020.

- RFC-3168 endpoint deployment is still mostly passive-only. This is largely under the control of a few large endpoint vendors' choices of defaults.
  - This makes measurements of AQM deployment more difficult, but does not lessen the potential harm of AQM+L4S interactions.

- L4S exhibits a **Major** severity failure mode in circumstances which may affect **10%** of frequent tunnel users at least several times a month, if deployed onto unprepared networks.
  - At Internet scale, that's a lot of total harm to innocent bystanders.

Questions?

Discussion on the mailing list is also welcome.

`draft-heist-tsvwg-ecn-deployment-observations`