Operational Guidance for Deployment of L4S in the Internet

draft-ietf-tsvwg-l4sops-02

Greg White, Editor

TSVWG @ IETF112

Nov. 8, 2021

Contributors include: Bob Briscoe, Jake Holland, Koen De Schepper, Olivier Tilmans, Tom Henderson, Asad Ahmed, Gorry Fairhurst, Sebastian Moeller, Pete Heist

Scope & Status

- Addresses the concerns raised about possible rate-imbalance in shared-queue RFC3168 bottlenecks
  - Guidance for Operators of End-hosts, Operators of Networks, Researchers

- WG Draft Adopted on March 26
  - Draft-00 (May 2021)
  - Draft-01 (July 2021)
  - Draft-02 (October 2021)
  - Not seeking WGLC currently

Outline

1. Introduction
2. Per-Flow Fairness
3. Flow Queuing Systems
4. Detection of Classic ECN Bottlenecks
    4.1. Recent Studies
    4.2. Future Experiments
5. Operator of an L4S host
    5.1 Server Type
    5.2 Server deployment environment
6. Operator of a Network Employing RFC3168 FIFO Bottlenecks
    6.1 Preferred Options
    6.2 Non-Preferred Options
    6.3 Last Resort Options
7. Operator of a Network Employing RFC3168 FQ Bottlenecks
8. Conclusion of the L4S experiment
    8.1. Termination of a successful L4S experiment
    8.2. Termination of an unsuccessful L4S experiment

Summary of Deltas in Draft-02

1. Added text that discusses the risk of incorrectly classifying a path as containing/not-containing a RFC3168 bottleneck
2. Changed the text describing short transactional transfers to address https://mailarchive.ietf.org/arch/msg/tsvwg/VHtB6QpSyMo33fgE7vT0slNR0Q4/
3. Added guidance that L4S experiments start with edge servers as opposed to hosts that send traffic to a wide variety of networks.
4. For hosts that serve a wide variety of networks, added a comment that the Akamai study found a small number of networks with heavier deployment of RFC3168, and that these networks could be avoided for use of L4S.
5. For networks that contain RFC3168 equipment, updated the description of "6.1 Preferred Options" to make it clear why they are preferred.
6. Added a ref to section 4.2 of L4S-Arch for L4S-aware FQ.
7. Added a paragraph describing the use of alternative (local-use) identifiers to facilitate updating RFC3168 nodes
8. Changed "Less Preferred Options" to "Non-Preferred Options" and added a description as to why they are non-preferred.

1. Risk of incorrectly classifying a path

New text in Section 5:

“Some of the recommendations in this section involve the sender determining (through various means) the likelihood of a particular path having a bottleneck that implements single queue RFC3168 AQM. Since this determination can be imprecise, there exists some risk that a path is incorrectly classified. In the case of false-positives (where a path is erroneously believed to contain RFC3168), discontinuing the use of L4S on that path would result in a lost opportunity for low-latency low-loss service, and thus likely an unnecessary degradation in the quality of experience for the user. In the case of false-negatives, the use of L4S has the potential to result in a reduction in the throughput of non-L4S flows while the L4S flow is active. In environments where the risk of false-negatives is significant, it is recommended that hosts limit the use of L4S congestion control to application-limited flows that are especially sensitive to latency, latency variation and loss.”

2. Short transactional transfers

Revised bullet in 5.1.1 to read:

“• Depending on the details of the L4S congestion control implementation, taking action based on the detection of RFC3168 FIFO bottlenecks may not be needed for short transactional transfers that are unlikely to achieve the steady-state conditions where unfairness is likely to occur.”

3. New guidance that L4S experiments start with edge servers

New statement in 5.2.1 (Edge Servers):

"It is recommended that L4S experimental deployments begin with such servers."

New statement in 5.2.2 (Other hosts):

"It is recommended that operators of such hosts consider carefully whether these hosts are appropriate for early experimentation with L4S."

4. Avoid using L4S on networks with high 3168 deployment.

New text in 5.2.2 (Other hosts):

“Additionally, the most recent large scale study [Holland] indicated that there were a small number of networks in which RFC3168 bottlenecks are more prevalent than the global average. Therefore, it may be possible for a host to maintain a list of networks where L4S should not be enabled, and, for other networks, to cache a list of end host ip addresses where a RFC3168 bottleneck has been detected.”

5. Update "Preferred Options" to make it clear why they are preferred.

New text in 6.1:

“The options in this section preserve the ability of the bottleneck to CE-mark ECT(1) packets as well as ECT(0) packets. The result of these options is that hosts utilizing classic (RFC3168) ECN and hosts utilizing L4S ECN receive the benefit of ECN. Further with these options, the hosts that choose to use L4S ECN see the benefit of reduced latency and latency-variation compared to hosts that choose instead to use classic ECN.”

6. Added a ref to section 4.2 of L4S-Arch for L4S-aware FQ.

New text in 6.1.1 (Upgrade AQMs to an L4S-aware AQM)

"Section 4.2 of [I-D.ietf-tsvwg-l4s-arch] contains a description of the options available, including a discussion about L4S-aware FQ implementations."

7. Local-use identifiers

New text in 6.1.2 (non-coupled dual queue), and referenced in related options (6.2.1 & 6.2.2):

“If classification based on the ECN field isn't possible in the bottleneck, this option may still be useful if an external system can be configured to reflect the ECN codepoint to another field that could then be used as an alternative identifier to classify traffic into Queue #1. For example, if at network ingress an edge router can apply a local-use DSCP to ECT(1) & CE packets, the bottleneck can then utilize a DSCP classifier. Similarly, in MPLS networks, ECT(1) & CE packets could use a different EXP value [RFC5129] than classic packets. More generally, any tunnelling protocol can be used to proxy the ECN value of the encapsulated packet to its outer header, enabling bottlenecks to classify packets based on their input virtual interface.”

8. Changed "Less Preferred Options" to "Non-Preferred Options"

“ 6.2. Non-Preferred Options

The options in this section come with a downside that they treat ECT(1) packets as NotECT, and thus don't provide the latency/loss benefit to flows marked ECT(1) (i.e. L4S flows). In the case that there is a strong concern about per-flow fairness between L4S flows and Classic flows in an RFC3168 FIFO bottleneck, and none of the remedies in the previous section can be implemented, the options listed in this section could be considered. These options are non-preferred because bottlenecks that implement them create a dilemma for operators of hosts, in that the application could see better performance if it uses classic (RFC3168) ECN rather than L4S ECN.”

Next Steps

- Please review latest draft and send comments to the list