                Specifying New Congestion Control Algorithms

Status of This Memo

   This document specifies an Internet Best Current Practices for the
   Internet Community, and requests discussion and suggestions for
   improvements. Distribution of this memo is unlimited.

Abstract

   The IETF's standard congestion control schemes have been widely shown
   to be inadequate for various environments (e.g., high-speed
   networks). Recent research has yielded many alternate congestion
   control schemes that significantly differ from the IETF's congestion
   control principles. Using these new congestion control schemes in
   the global Internet has possible ramifications to both the traffic
   using the new congestion control and to traffic using the currently
   standardized congestion control. Therefore, the IETF must proceed
   with caution when dealing with alternate congestion control
   proposals. The goal of this document is to provide guidance for
   considering alternate congestion control algorithms within the IETF.

1. Introduction

   This document provides guidelines for the IETF to use when evaluating
   suggested congestion control algorithms that significantly differ
   from the general congestion control principles outlined in [RFC2914].
   The guidance is intended to be useful to authors proposing alternate
   congestion control and for the IETF community when evaluating whether
   a proposal is appropriate for publication in the RFC series.

   The guidelines in this document are intended to be consistent with
   the congestion control principles from [RFC2914] of preventing
   congestion collapse, considering fairness, and optimizing the flow's
   own performance in terms of throughput, delay, and loss. [RFC2914]
   also discusses the goal of avoiding a congestion control "arms race"
   among competing transport protocols.

   This document does not give hard-and-fast requirements for an
   appropriate congestion control scheme. Rather, the document provides
   a set of criteria that should be considered and weighed by the IETF
   in the context of each proposal. The high-order criteria for any new
   proposal is that a serious scientific study of the pros and cons of
   the proposal needs to have been done such that the IETF has a well-
   rounded set of information to consider.

   After initial studies, we encourage authors to write a specification
   of their proposals for publication in the RFC series to allow others
   to concretely understand and investigate the wealth of proposals in
   this space.

2. Document Status

   Following the lead of HighSpeed TCP [RFC3649], alternate congestion
   control algorithms are expected to be published as "Experimental"
   RFCs until such time that the community better understands the
   solution space. Traditionally, the meaning of "Experimental" status
   has varied in its use and interpretation. As part of this document
   we define two classes of congestion control proposals that can be
   published with the "Experimental" status. The first class includes
   algorithms that are judged to be safe to deploy for best-effort
   traffic in the global Internet and further investigated in that
   environment. The second class includes algorithms that, while
   promising, are not deemed safe enough for widespread deployment as
   best-effort traffic on the Internet, but are being specified to
   facilitate investigations in simulation, testbeds, or controlled
   environments. The second class can also include algorithms where the
   IETF does not yet have sufficient understanding to decide if the
   algorithm is or is not safe for deployment on the Internet.

Each alternate congestion control algorithm published is required to
include a statement in the abstract indicating whether or not the
proposal is considered safe for use on the Internet. Each alternate
congestion control algorithm published is also required to include a
statement in the abstract describing environments where the protocol
is not recommended for deployment. There may be environments where
the protocol is deemed *safe* for use, but still is not *recommended*
for use because it does not perform well for the user.

As examples of such statements, [RFC3649] specifying HighSpeed TCP
includes a statement in the abstract stating that the proposal is
Experimental, but may be deployed in the current Internet. In
contrast, the Quick-Start document [RFC4782] includes a paragraph in
the abstract stating the mechanism is only being proposed for
controlled environments. The abstract specifies environments where
the Quick-Start request could give false positives (and therefore
would be unsafe to deploy). The abstract also specifies environments
where packets containing the Quick-Start request could be dropped in
the network; in such an environment, Quick-Start would not be unsafe
to deploy, but deployment would still not be recommended because it
could cause unnecessary delays for the connections attempting to use
Quick-Start.

For authors of alternate congestion control schemes who are not ready
to bring their congestion control mechanisms to the IETF for
standardization (either as Experimental or as Proposed Standard), one
possibility would be to submit an internet-draft that documents the
alternate congestion control mechanism for the benefit of the IETF
and IRTF communities. This is particularly encouraged in order to
get algorithm specifications widely disseminated to facilitate
further research. Such an internet-draft could be submitted to be
considered as an Informational RFC, as a first step in the process
towards standardization. Such a document would also be expected to
carry an explicit warning against using the scheme in the global
Internet.

Note: we are not changing the RFC publication process for non-IETF
produced documents (e.g., those from the IRTF or Independent
Submissions via the RFC-Editor). However, we would hope the
guidelines in this document inform the IESG as they consider whether
to add a note to such documents.

3. Guidelines

As noted above, authors are expected to do a well-rounded evaluation
of the pros and cons of proposals brought to the IETF. The following
are guidelines to help authors and the IETF community. Concerns that

fall outside the scope of these guidelines are certainly possible;
these guidelines should not be considered as an all-encompassing
check-list.

(0) Differences with Congestion Control Principles [RFC2914]

   Proposed congestion control mechanisms should include a clear
   explanation of the deviations from [RFC2914].

(1) Impact on Standard TCP, SCTP [RFC2960], and DCCP [RFC4340].

   Proposed congestion control mechanisms should be evaluated when
   competing with standard IETF congestion control [RFC2581,
   RFC2960, RFC4340]. Alternate congestion controllers that have a
   significantly negative impact on traffic using standard
   congestion control may be suspect and this aspect should be part
   of the community's decision making with regards to the
   suitability of the alternate congestion control mechanism.

   We note that this bullet is not a requirement for strict TCP-
   friendliness as a prerequisite for an alternate congestion
   control mechanism to advance to Experimental. As an example,
   HighSpeed TCP is a congestion control mechanism that is
   Experimental, but that is not TCP-friendly in all environments.
   We also note that this guideline does not constrain the fairness
   offered for non-best-effort traffic.

   As an example from an Experimental RFC, fairness with standard
   TCP is discussed in Sections 4 and 6 of [RFC3649] (HighSpeed TCP)
   and using spare capacity is discussed in Sections 6, 11.1, and 12
   of [RFC3649].

(2) Difficult Environments.

   The proposed algorithms should be assessed in difficult
   environments such as paths containing wireless links.
   Characteristics of wireless environments are discussed in
   [RFC3819] and in Section 16 of [Tools]. Other difficult
   environments can include those with multipath routing within a
   connection. We note that there is still much to be desired in
   terms of the performance of TCP in some of these difficult
   environments. For congestion control mechanisms with explicit
   feedback from routers, difficult environments can include paths
   with non-IP queues at layer-two, IP tunnels, and the like. A
   minimum goal for experimental mechanisms proposed for widespread
   deployment in the Internet should be that they do not perform
   significantly worse than TCP in these environments.

While it is impossible to enumerate all the possible "difficult
environments", we note that the IETF has previously grappled with
paths with long delays [RFC2488], high delay bandwidth products
[RFC3649], high packet corruption rates [RFC3155], packet
reordering [RFC4653], and significantly slow links [RFC3150].
Aspects of alternate congestion control that impact networks with
these characteristics should be detailed.

As an example from an Experimental RFC, performance in difficult
environments is discussed in Sections 6, 9.2, and 10.2 of
[RFC4782] (Quick-Start).

(3) Investigating a Range of Environments.

Similar to the last criteria, proposed alternate congestion
controllers should be assessed in a range of environments. For
instance, proposals should be investigated across a range of
bandwidths, round-trip times, levels of traffic on the reverse
path, and levels of statistical multiplexing at the congested
link. Similarly, proposals should be investigated for robust
performance with different queueing mechanisms in the routers,
especially Random Early Detection (RED) [FJ03] and Drop-Tail.
This evaluation is often not included in the internet-draft
itself, but in related papers cited in the draft.

A particularly important aspect of evaluating a proposal for
standardization is in understanding where the algorithm breaks
down. Therefore, particular attention should be paid to
characterizing the areas where the proposed mechanism does not
perform well.

As an example from an Experimental RFC, performance in a range of
environments is discussed in Section 12 of [RFC3649] (HighSpeed
TCP) and Section 9.7 of [RFC4782] (Quick-Start).

(4) Protection Against Congestion Collapse.

The alternate congestion control mechanism should either stop
sending when the packet drop rate exceeds some threshold
[RFC3714], or should include some notion of "full backoff". For
"full backoff", at some point the algorithm would reduce the
sending rate to one packet per round-trip time and then
exponentially backoff the time between single packet
transmissions if congestion persists. Exactly when either "full
backoff" or a pause in sending comes into play will be
algorithm-specific. However, as discussed in [RFC2914], this
requirement is crucial to protect the network in times of extreme
congestion.

If "full backoff" is used, this bullet does not require that the
full backoff mechanism must be identical to that of TCP
[RFC2988]. As an example, this bullet does not preclude full
backoff mechanisms that would give flows with different round-
trip times comparable bandwidth during backoff.

(5) Fairness within the Alternate Congestion Control Algorithm.

In environments with multiple competing flows all using the same
alternate congestion control algorithm, the proposal should
explore how bandwidth is shared among the competing flows.

(6) Performance with Misbehaving Nodes and Outside Attackers.

The proposal should explore how the alternate congestion control
mechanism performs with misbehaving senders, receivers, or
routers. In addition, the proposal should explore how the
alternate congestion control mechanism performs with outside
attackers. This can be particularly important for congestion
control mechanisms that involve explicit feedback from routers
along the path.

As an example from an Experimental RFC, performance with
misbehaving nodes and outside attackers is discussed in Sections
9.4, 9.5, and 9.6 of [RFC4782] (Quick-Start). This includes
discussion of misbehaving senders and receivers; collusion
between misbehaving routers; misbehaving middleboxes; and the
potential use of Quick-Start to attack routers or to tie up
available Quick-Start bandwidth.

(7) Responses to Sudden or Transient Events.

The proposal should consider how the alternate congestion control
mechanism would perform in the presence of transient events such
as sudden congestion, a routing change, or a mobility event.
Routing changes, link disconnections, intermittent link
connectivity, and mobility are discussed in more detail in
Section 17 of [Tools].

As an example from an Experimental RFC, response to transient
events is discussed in Section 9.2 of [RFC4782] (Quick-Start).

(8) Incremental Deployment.

The proposal should discuss whether the alternate congestion
control mechanism allows for incremental deployment in the
targeted environment. For a mechanism targeted for deployment in
the current Internet, it would be helpful for the proposal to

discuss what is known (if anything) about the correct operation
of the mechanism with some of the equipment installed in the
current Internet, e.g., routers, transparent proxies, WAN
optimizers, intrusion detection systems, home routers, and the
like.

As a similar concern, if the alternate congestion control
mechanism is intended only for specific environments (and not the
global Internet), the proposal should consider how this intention
is to be carried out. The community will have to address the
question of whether the scope can be enforced by simply stating
the restrictions or whether additional protocol mechanisms are
required to enforce the scoping. The answer will necessarily
depend on the change being proposed.

As an example from an Experimental RFC, deployment issues are
discussed in Sections 10.3 and 10.4 of [RFC4782] (Quick-Start).

4. Minimum Requirements

This section suggests minimum requirements for a document to be
approved as Experimental with approval for widespread deployment in
the global Internet.

The minimum requirements for approval for widespread deployment in
the global Internet include the following guidelines on: (1)
assessing the impact on standard congestion control, (3)
investigation of the proposed mechanism in a range of environments,
(4) protection against congestion collapse, and (8) discussing
whether the mechanism allows for incremental deployment.

For other guidelines, i.e., (2), (5), (6), and (7), the author must
perform the suggested evaluations and provide recommended analysis.
Evidence that the proposed mechanism has significantly more problems
than those of TCP should be a cause for concern in approval for
widespread deployment in the global Internet.

5. Security Considerations

This document does not represent a change to any aspect of the TCP/IP
protocol suite and therefore does not directly impact Internet
security. The implementation of various facets of the Internet's
current congestion control algorithms do have security implications
(e.g., as outlined in [RFC2581]). Alternate congestion control
schemes should be mindful of such pitfalls, as well, and should
examine any potential security issues that may arise.

6. Acknowledgments

   Discussions with Lars Eggert and Aaron Falk seeded this document.
   Thanks to Bob Briscoe, Gorry Fairhurst, Doug Leith, Jitendra Padhye,
   Colin Perkins, Pekka Savola, members of TSVWG, and participants at
   the TCP Workshop at Microsoft Research for feedback and
   contributions. This document also draws from [Metrics].

7. Normative References

   [RFC2581] Allman, M., Paxson, V., and W. Stevens, "TCP Congestion
             Control", RFC 2581, April 1999.

   [RFC2914] Floyd, S., "Congestion Control Principles", BCP 41, RFC
             2914, September 2000.

   [RFC2960] Stewart, R., Xie, Q., Morneault, K., Sharp, C.,
             Schwarzbauer, H., Taylor, T., Rytina, I., Kalla, M., Zhang,
             L., and V. Paxson, "Stream Control Transmission Protocol",
             RFC 2960, October 2000.

   [RFC4340] Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion
             Control Protocol (DCCP)", RFC 4340, March 2006.

8. Informative References

   [FJ03]   Floyd, S., and Jacobson, V., Random Early Detection
             Gateways for Congestion Avoidance, IEEE/ACM Transactions on
             Networking, V.1 N.4, August 1993.

   [Metrics] S. Floyd, Metrics for the Evaluation of Congestion Control
             Mechanisms, Work in Progress, July 2007.

   [RFC2488] Allman, M., Glover, D., and L. Sanchez, "Enhancing TCP Over
             Satellite Channels using Standard Mechanisms", BCP 28, RFC
             2488, January 1999.

   [RFC2988] Paxson, V. and M. Allman, "Computing TCP's Retransmission
             Timer", RFC 2988, November 2000.

   [RFC3150] Dawkins, S., Montenegro, G., Kojo, M., and V. Magret,
             "End-to-end Performance Implications of Slow Links", BCP
             48, RFC 3150, July 2001.

   [RFC3155] Dawkins, S., Montenegro, G., Kojo, M., Magret, V., and N.
             Vaidya, "End-to-end Performance Implications of Links with
             Errors", BCP 50, RFC 3155, August 2001.

   [RFC3649] Floyd, S., "HighSpeed TCP for Large Congestion Windows",
              RFC 3649, December 2003.

   [RFC3714] Floyd, S. and J. Kempf, "IAB Concerns Regarding Congestion
              Control for Voice Traffic in the Internet", RFC 3714, March
              2004.

   [RFC3819] Karn, P., Bormann, C., Fairhurst, G., Grossman, D., Ludwig,
              R., Mahdavi, J., Montenegro, G., Touch, J., and L. Wood,
              "Advice for Internet Subnetwork Designers", BCP 89, RFC
              3819, July 2004.

   [RFC4653] Bhandarkar, S., Reddy, A. N., Allman, M., and E. Blanton,
              "Improving the Robustness of TCP to Non-Congestion Events",
              RFC 4653, August 2006.

   [RFC4782] Floyd, S., Allman, M., Jain, A., and P. Sarolahti, "Quick-
              Start for TCP and IP", RFC 4782, January 2007.

   [Tools]   S. Floyd and E. Kohler, Tools for the Evaluation of
              Simulation and Testbed Scenarios, Work in Progress, July
              2007.

Authors' Addresses

   Sally Floyd
   ICIR (ICSI Center for Internet Research)
   1947 Center Street, Suite 600
   Berkeley, CA 94704-1198
   Phone: +1 (510) 666-2989
   EMail: floyd@icir.org
   URL: http://www.icir.org/floyd/

   Mark Allman
   ICSI Center for Internet Research
   1947 Center Street, Suite 600
   Berkeley, CA 94704-1198
   Phone: (440) 235-1792
   EMail: mallman@icir.org
   URL: http://www.icir.org/mallman/