Internet Engineering Task Force **Sally Floyd**
INTERNET-DRAFT **Editor**
draft-floyd-transport-models-00a.ps **13 May 2005**
Expires: November 2005

Metrics for the Evaluation of Congestion Control Mechanisms

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of section 3 of RFC 3667. By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she become aware will be disclosed, in accordance with RFC 3668.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on November 2005.

Copyright Notice

Copyright (C) The Internet Society (2004). All Rights Reserved.

Abstract

This document discusses the metrics to be considered in an evaluation of new or modified congestion control mechanisms for the Internet. This document is intended to be the first in a series of documents aimed at improving the models that we use in the evaluation of transport protocols.

Table of Contents

1. Conventions 3
2. Introduction 3
3. Metrics 3
3.1. Throughput, Delay, and Drop Rates 3
3.1.1. Throughput 3
3.1.2. Delay 4
3.1.3. Packet Drop Rates 4
3.2. Response Times and Minimizing Oscillations 4
3.2.1. Response to Changes 4
3.2.2. Minimizing Oscillations. 5
3.3. Fairness and Convergence 5
3.4. Metrics for Specific Environments 6
3.5. Metrics for Specific Types of Transport 7
4. Considering a Range of Environments 7
5. Security Considerations 7
6. IANA Considerations. 7
7. Acknowledgements 7
Informative References 7
Authors' Addresses 8
Full Copyright Statement 9
Intellectual Property 9

1. Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119].

2. Introduction

As a step towards improving our methodologies for evaluating congestion control mechanisms, in this document we discuss the range of metrics to be considered. We also consider the relationship between metrics, e.g., the well-known tradeoff between throughput and delay.

Subsequent documents will consider models to be used in analysis, simulations, or experiments for the evaluation of congestion control mechanisms. Our intention is to create a new IRTF (Internet Research Task Force) research group to aid collaboration in developing these models.

3. Metrics

The metrics that we discuss are the following:

- Throughput;
- Delay;
- Packet drop rates;
- Response to sudden changes or to transient events;
- Minimizing oscillations in throughput or in delay;
- Fairness and convergence times;
- Metrics for specific environments;
- Metrics for specific types of transport.

We consider each of these below. Many of the metrics have both network-based and user-based interpretations. For some of the metrics, such as fairness between flows, there is not a clear agreement in the network community about the desired goals.

3.1. Throughput, Delay, and Drop Rates

Because of the clear tradeoffs between throughput, delay, and drop rates, it can be useful to consider the three metrics together.

An alternative would be to consider a separate metric such as power, defined in this context as throughput over delay, that combines throughput and delay. However, we do not propose in this document that there should be a clear target in terms of the tradeoffs between throughput and delay; we are simply proposing that the evaluation of transport protocols include an exploration of the competing metrics.

3.1.1. Throughput

Throughput can be measured both as a router-based metric of aggregate link throughput, and as a user metric of per-connection transfer times. It is a clear goal of most congestion control mechanisms to maximize throughput, subject to application demand and to the constraints of the other metrics. We note that maximizing throughput is of concern in a

wide range of environments, from highly-congested environments to under-utilized ones.

Some researchers evaluate transport protocols in terms of maximizing the aggregate user utility, where a user's utility is generally defined as a function of the user's throughput [KMT98].

3.1.2. Delay

Like throughput, delay can be measured as a router-based metric of queueing delay over time, or in terms of per-packet transfer times. For reliable transfer, the per-packet transfer time includes the possible delay of retransmitting a dropped packet.

Users of bulk data transfer applications might care about per-packet transfer times only in so far as they affect the per-connection transfer time. On the other end of the spectrum, for users of streaming media, per-packet delay can be a significant concern. Note that in some cases the average delay might not capture the metric of interest to the users; some users might also care about the tail of the delay distribution.

3.1.3. Packet Drop Rates

Packet drop rates can be measured as a network-based or as a user-based metric.

Some users might care about packet drop rates only in so far as they affect per-connection transfer times, while other users might care about packet drop rates directly. One network-related reason to avoid high steady-state packet drop rates is to avoid congestion collapse in environments characterized by paths with multiple congested links.

3.2. Response Times and Minimizing Oscillations

In this section we consider response times and oscillations together. There are clear tradeoffs between improving response times and minimizing oscillations. In addition, the scenarios that illustrate the dangers of poor response times are often quite different from the scenarios that illustrate the dangers of unnecessary oscillations.

3.2.1. Response to Changes

One of the key concerns in the design of congestion control mechanisms has been the response times to sudden congestion in the network. On the one hand, congestion control mechanisms should respond reasonably promptly to sudden congestion from routing or bandwidth changes, or from a burst of competing traffic. At the same time, congestion control mechanisms should not respond too severely to transient changes, e.g., to a sudden increase in delay that will dissipate in less than the connection's round-trip time.

Evaluating the response to sudden or transient changes can be of particular concern for slowly-responding congestion control mechanisms such as equation-based congestion control [RFC 3448], and for AIMD (Additive Increase Multiplicative Decrease) or related mechanisms with parameters that make them more slowly-responding that TCP [BB01, BBFS01].

In addition to the responsiveness and smoothness of aggregate traffic, one can consider the tradeoffs between responsiveness, smoothness, and aggressiveness for an individual connection [FHP00]. In this case smoothness can be defined by the largest reduction in the sending rate in one round-trip time, in a deterministic environment with a packet drop

exactly every 1/p packets. The responsiveness is defined as the number of round-trip times of sustained congested required for the sender to halve the sending rate, and the aggressiveness is defined as the maximum increase in the sending rate in one round-trip time, in packets per second, in the absence of congestion.

3.2.2. Minimizing Oscillations

One goal is that of stability, in terms of minimizing oscillations of queueing delay or of throughput. Scenarios illustrating oscillations are often dominated by long-lived connections, perhaps with a small number of changes in the level of congestion.

An orthogonal goal for some congestion control mechanisms, e.g., for equation-based congestion control, is to minimize the oscillations in the sending rate for an individual connection, given an environment with a fixed, steady-state packet drop rate. (As is well known, TCP congestion control is characterized by a pronounced oscillation in the sending rate, with the sender halving the sending rate in response to congestion.) One metric for the level of oscillations is the smoothness metric given above.

3.3. Fairness and Convergence

Another set of metrics are those of fairness and of convergence times. Fairness can be considered between flows of the same protocol, and between flows using different protocols (e.g., fairness between TCP and a new transport protocol).

There are a number of different fairness measures. We discuss max-min fairness [HG86], proportional fairness [KMT98, K01], the fairness index proposed in [JCH84], and the product measure, a variant of network power [BJ81]. In order to satisfy the max-min fairness criteria, the smallest throughput rate must be as large as possible. Given this condition, the next-smallest throughput rate must be as large as possible, and so on. Thus, the max-min fairness gives absolute priority to the smallest flows.

A metric related to max-min fairness is epsilon-fairness, where a rate allocation is defined as epsilon-fairness if

  min_i x_i / max_i x_i >= 1 - epsilon.

where x_i is the resource allocation to the i-th user. Epsilon-fairness measures the worst-case ratio between any two throughput rates [ZKL04].

In contrast, an allocation x is defined as proportionally fair if for any other feasible allocation x*, the aggregate of proportional changes is zero or negative:

  sum_i (x*_i - x_i)/x_i <= 0.

"This criterion favours smaller flows, but less emphatically than max-min fairness" [K01].

The fairness index in [JCH84] is

  ((sum_i x_i)^2) / (n * sum_i (x_i)^2) ,

where there are n users. This fairness index ranges from 0 to 1, and is maximum when all

users receive the same allocation. This index is k/n when k users equally share the resource, and the other n-k users receive zero allocation.

The product measure

product_i x_i ,

the product of the throughput of the individual connections, is also used as a measure of fairness. For our purposes, let x_i be the throughput for the i-th connection. (In other contexts x_i is taken as the power of the i-th connection, and the product measure is referred to as network power.) The product measure is particularly sensitive to segregation; the product measure is zero if any connection receives zero throughput. In [MS90](p. 15) it is shown that for a network with many connections and one shared gateway, the product measure is maximized when all connections receive the same throughput. These fairness metrics are discussed in more detail in [F91].

We note that there is not a clear consensus for the fairness goals, in particular for fairness between flows that traverse different numbers of congested links [F91].

One goal cited in a number of new transport protocols has been that of fairness between flows with different round-trip times [KHR02, XHR04]. We note that there is not necessarily a consensus in the networking community about the desirability of this goal, or about the implications and interactions between this goal and other metrics [FJ92] (Section 3.3). Convergence times concern the time for convergence to fairness between an existing flow and a newly-starting one, and are a special concern for environments with high-bandwidth flows. As with fairness, convergence times can matter both between flows of the same protocol, and between flows using different protocols [SLFK03].

One metric used for convergence times is the delta-fair convergence time, defined in [BBFS01] as the time taken for two flows with the same round-trip time to go from shares of 100/101-th and 1/101-th of the link bandwidth, to having close to fair sharing with shares of (1+delta)/2 and (1-delta)/2 of the link bandwidth.

[ZKL04] defines the convergence time at the number of steps for two flows to reach epsilon-fairness, when starting from a maximally-unfair state.

3.4. Metrics for Specific Environments

While congestion control mechanisms are generally evaluated first over environments with static routing in a network of two-way point-to-point links, some environments bring up not only more challenging scenarios (e.g., corrupted links, variable bandwidth, mobility) but also new metrics to be considered.

For example, in mobile environments the energy consumption for the mobile end-node can be a key metric that is affected by the transport protocol [TM02]. For wireless networks, goodput can be a key metric, where goodput is defined as the fraction of useful data from all of the data delivered. High goodput indicates an efficient use of the radio spectrum and lower interference to other users [GF04].

3.5. Metrics for Specific Types of Transport

In some cases modified metrics are needed for evaluting transport protocols intended for QoS-enabled environments or for below-best-effort traffic [VKD02, KK03]. For example, different fairness metrics are needed for evaluating transport protocols for below-best-effort traffic.

4. Considering a Range of Environments

In this document we do not offer specific scenarios for use in evaluating congestion control mechanisms, but we do note that it can be important to evaluate over a wide range of environments, with a range of link bandwidths, congestion levels, and levels of statistical multiplexing. It is also important to evaluate congestion control mechanisms in a range of scenarios, including typical ranges of connection sizes and round-trip times [FK02].

5. Security Considerations

There are no security considerations in this document.

6. IANA Considerations

There are no IANA considerations in this document.

7. Acknowledgements

Informative References

[BB01] D. Bansal and H. Balakrishnan, Binomial Congestion Control Algorithms, IEEE Infocom, April 2001.

[BBFS01] D. Bansal, H. Balakrishnan, S. Floyd, and S. Shenker, Dynamic Behavior of Slowly-Responsive Congestion Control Algorithms, SIGCOMM 2001.

[BJ81] K. Bharath-Kumar and J. Jeffrey, A New Approach to Performance-Oriented Flow Control, IEEE Transactions on Communications, Vol.COM-29 N.4, April 1981.

[F91] S. Floyd, Connections with Multiple Congested Gateways in Packet-Switched Networks Part 1: One-way Traffic, Computer Communication Review, Vol.21, No.5, October 1991, p. 30-47.

[FHP00] S. Floyd, M. Handley, and J. Padhye, A Comparison of Equation-Based and AIMD Congestion Control, May 2000. URL "http://www.icir.org/tfrc/".

[FJ92] S. Floyd and V. Jacobson, On Traffic Phase Effects in Packet-Switched Gateways, Internetworking: Research and Experience, V.3 N.3, September 1992, p.115-156.

[FK02] S. Floyd and E. Kohler, Internet Research Needs Better Models, Hotnets-I. October 2002.

[GF04] A. Gurtov and S. Floyd, Modeling Wireless Links for Transport Protocols, ACM CCR, 34(2):85-96, April 2004.

[HG86] E. Hahne and R. Gallager, Round Robin Scheduling for Fair Flow Control in Data Communications Networks, IEEE International Conference on Communications, June 1986.

[JCH84] R. Jain, D.M. Chiu, and W. Hawe, A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Systems, DEC TR-301, Littleton, MA: Digital Equipment Corporation, 1984.

[K01] F. Kelly, Mathematical modelling of the Internet, In "Mathematics Unlimited - 2001 and Beyond" (Editors B. Engquist and W. Schmid), Springer-Verlag, Berlin, pp. 685-702, 2001.

[KHR02] D. Katabi, M. Handley, and C. Rohrs, Congestion Control for High Bandwidth-Delay Product Networks, ACM Sigcomm, 2002.

[KK03] A. Kuzmanovic and E. W. Knightly, TCP-LP: A Distributed Algorithm for Low Priority Data Transfer, Proceedings of IEEE INFOCOM 2003, April 2003.

[KMT98] F. Kelly, A. Maulloo and D. Tan, Rate Control in Communication Networks: Shadow Prices, Proportional Fairness and Stability. Journal of the Operational Research Society 49, pp. 237-252, 1998.

[MS90] D. Mitra and J. Seery, Dynamic Adaptive Windows for High Speed Data Networks: Theory and Simulations, ATT Bell Laboratories report, April 1990.

[RFC 2119] S. Bradner. Key Words For Use in RFCs to Indicate Requirement Levels. RFC 2119.

[RFC 2434] T. Narten and H. Alvestrand. Guidelines for Writing an IANA Considerations Section in RFCs. RFC 2434.

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

[RFC 3448] M. Handley, S. Floyd, J. Padhye, and J. Widmer, TCP Friendly Rate Control (TFRC): Protocol Specification, RFC 3448, Proposed Standard, January 2003.

[SLFK03] R.N. Shorten, D.J. Leith, J. Foy, and R. Kilduff, Analysis and Design of Congestion Control in Synchronised Communication Networks. Proc. 12th Yale Workshop on Adaptive & Learning Systems, May 2003.

[TM02] V. Tsaoussidis and I. Matta, Open Issues of TCP for Mobile Computing, Journal of Wireless Communications and Mobile Computing: Special Issue on Reliable Transport Protocols for Mobile Computing, February 2002.

[VKD02] A. Venkataramani, R. Kokku, and M. Dahlin, TCP Nice: A Mechanism for Background Transfers, Fifth USENIX Symposium on Operating System Design and Implementation (OSDI), 2002.

[XHR04] L. Xu, K. Harfoush, and I. Rhee, Binary Increase Congestion Control for Fast, Long Distance Networks, Infocom 2004.

[YL00] Y. R. Yang and S. S. Lam, General AIMD Congestion Control, Technical Report TR-00-09, Department of Computer Sciences, UT Austin, May 2000.

[ZKL04] Y. Zhang, S.-R. Kang, and D. Loguinov, Delayed Stability and Performance of Distributed Congestion Control, ACM SIGCOMM, August 2004.

Authors' Addresses

Sally Floyd <floyd@icir.org>
ICSI Center for Internet Research
1947 Center Street, Suite 600
Berkeley, CA 94704
USA

Full Copyright Statement

Copyright (C) The Internet Society 2004. This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.