Internet Engineering Task Force
INTERNET-DRAFT
draft-irtf-tmrg-metrics-00.ps
Expires: February 2006

Sally Floyd
Editor

17 August 2005

Metrics for the Evaluation of Congestion Control Mechanisms

Status of this Memo

Copyright Notice

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

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].

TO BE DELETED BY THE RFC EDITOR UPON PUBLICATION:

Changes from draft-floyd-transport-metrics-00.txt:

* Added metrics for:
  - robustness in challenging environments,
  - deployability,
  - robustness to failures and to misbehaving users

* Added a discussion of fairness and packet size.

2. Introduction

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

Subsequent documents will discuss the models that are used in analysis, simulations, or experiments for the evaluation of transport protocols in general, and of congestion control mechanisms in particular. These are intended to become documents in the newly-chartered Transport Modeling Research Group (TMRG) in the IRTF (Internet Research Task Force).

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;
- Robustness for challenging environments;
- Robustness to failures and to misbehaving users;
- Deployability;
- 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 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 networks to under-utilized ones.

In some contexts, it might be sufficient to consider the aggregate throughput or the mean per-flow throughput, while in other contexts it might be necessary to consider the distribution of per-flow throughput. 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].

Individual applications can have application-specific needs in terms of throughput. For example, real-time video traffic can have highly variable bandwidth demands; VoIP traffic is sensitive to the amount of bandwidth received immediately after idle periods. Thus, user metrics for throughput can be more complex than simply the per-connection transfer time.

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; for example, some users might care about the worst-case delay, or 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 containing paths with multiple congested links. In such environments, high packet drop rates could result in congested links wasting scarce bandwidth by carrying packets that will only be dropped downstream, before being delivered to the receiver.

3.2. Response Times and Minimizing Oscillations

In this section we consider response times and oscillations together, as there are well-known 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 using 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. These include max-min fairness [HG86], proportional fairness [KMT98, K01], the fairness index proposed in [JCH84], and the product measure, a variant of network power [BJ81].

Max-min fairness: 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.

Epsilon-fairness: A metric related to max-min fairness is epsilon-fairness, where a rate allocation is defined as epsilon-fair 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].

Proportional fairness: 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].

Jain's fairness index: 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: The product measure

$$\prod_i x_i ,$$

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] 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.

Fairness and the number of congested links: Some of 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].

Fairness and round-trip times: 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 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).

Fairness and packet size: One fairness issue is that of the relative fairness for flows with different packet sizes. Many file transfer applications will use the maximum packet size possible; in contrast, low-bandwidth VoIP flows are likely to send small packets, sending a new packet every 10 to 40 ms., to limit delay. Should a small-packet VoIP connection receive the same sending rate in bytes per second as a large-packet TCP connection in the same environment, or should it receive the same sending rate in *packets* per second? This fairness issue has been discussed in more detail in [FK04], with [FK05] also describing the ways that packet size can effect the packet drop rate experienced by a flow.

Convergence times: 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 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 [BBFS01]. A similar metric for convergence times measures the convergence time as the number of round-trip times for two flows to reach epsilon-fairness, when starting from a maximally-unfair state [ZKL04]. '

3.4. Robustness for Challenging 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 more challenging problems (e.g., corrupted packets, variable bandwidth, mobility) as well as new metrics to be considered (e.g., energy consumption).

Robustness for challenging environments: Robustness needs to be explored for paths with reordering, corruption, variable bandwidth, asymmetric routing, router configuration changes, mobility, and the like. In general, Internet architecture has valued robustness over efficiency, e.g., when there are tradeoffs between robustness and the throughput, delay, and fairness metrics described above.

Energy consumption: In mobile environments the energy consumption for the mobile end-node can be a key metric that is affected by the transport protocol [TM02].

Goodput: For wireless networks, goodput can be a useful 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. Robustness to Failures and to Misbehaving Users

One goal is for congestion control mechanisms to be robust to misbehaving users, such as receivers that 'lie' to data senders about the congestion experienced along the path or otherwise attempt to bypass the congestion control mechanisms of the sender [SCWA99]. Another goal is for congestion control mechanisms to be as robust as possible to failures, such as failures of routers in using explicit feedback to end-nodes or failures of end-nodes to follow the prescribed protocols,

3.6. Deployability

One metric for congestion control mechanisms is their deployability in the current Internet. Metrics related to deployability include the ease of failure diagnosis, and the overhead in terms of packet header size or added complexity at end-nodes or routers.

3.7. 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].

4. Comments on Methodology

The types of scenarios that are used to test specific metrics, and the range of parameters that it is useful to consider, will be discussed in separate documents, e.g., along with specific scenarios for use in evaluating congestion control mechanisms.

We note that it can be important to evaluate metrics 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]. It is also useful to compare metrics for new or modified transport protocols with those of the current standards for TCP.

More general references on methodology include [J91].

5. Security Considerations

There are no security considerations in this document.

6. IANA Considerations

There are no IANA considerations in this document.

7. Acknowledgements

Thanks to Doug Leith for feedback.

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.

[FK05] S. Floyd and E. Kohler, TFRC for Voice: the VoIP Variant, draft-ietf-dccp-tfrc-voip-02.txt, internet draft, work in progress, July 2005.

[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.

[FK04] S. Floyd and J. Kempf, IAB Concerns Regarding Congestion Control for Voice Traffic in the Internet, RFC 3714, March 2004.

[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.

[J91] R. Jain, The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling, John Wiley & Sons, 1991.

[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, "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, 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.

[SCWA99] TCP Congestion Control with a Misbehaving Receiver, ACM Computer Communications Review, October 1999.

[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

Intellectual Property