Internet Engineering Task Force
INTERNET-DRAFT
Intended status: Informational
Expires: April 2008

Sally Floyd
Editor

5 October 2007

Metrics for the Evaluation of Congestion Control Mechanisms
draft-irtf-tmrg-metrics-11.ps

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. These include metrics for the evaluation of new transport protocols, of proposed modifications to TCP, of application-level congestion control, and of Active Queue Management (AQM) mechanisms in the router. This document is the first in a series of documents aimed at improving the models that we use in the evaluation of transport protocols.

This document is a product of the Transport Modeling Research Group (TMRG), and has received detailed feedback from many members of the Research Group (RG). As the document tries to make clear, there is not necessarily a consensus within the research community (or the IETF community, the vendor community, the operations community, or any other community) about the metrics that congestion control mechanisms should be designed to optimize, in terms of tradeoffs between throughput and delay, fairness between competing flows, and the like. However, we believe that there is a clear consensus that congestion control mechanisms should be evaluated in terms of tradeoffs between a range of metrics, rather than in terms of optimizing for a single metric.

Table of Contents

TO BE DELETED BY THE RFC EDITOR UPON PUBLICATION:

Changes from draft-irtf-tmrg-metrics-10.txt:

* Added tunnels (e.g., IPsec) and non-router queues
  (e.g. Ethernet switches) to the discussion on deployability.
  Feedback from Aaron Falk in the IRSG poll.

* Minor editing in response to feedback from Lachlan Andrew.

* Added a new citation to [DM06].

Changes from draft-irtf-tmrg-metrics-09.txt:

* Copied a paragraph from the Abstract to the Introduction, and
  updated references. From feedback from Tony Li in the IRSG
  review.

Changes from draft-irtf-tmrg-metrics-08.txt:

* General feedback from Michael Welzl, adding explanations
  to some of the fairness discussions.

Changes from draft-irtf-tmrg-metrics-07.txt:

* General feedback from Mark Allman.

* Feedback and contributions from Lachlan Andrew about fairness.

Changes from draft-irtf-tmrg-metrics-06.txt:

* Added a paragraph about tradeoffs between
  fairness and throughput.

* Minor editing changes.

Changes from draft-irtf-tmrg-metrics-05.txt:

* Added a sentence about drop synchronization.
  Feedback from David Ros.

* Added a citation to fairness draft by Briscoe.

Changes from draft-irtf-tmrg-metrics-04.txt:

* Added text to the abstract.

Changes from draft-irtf-tmrg-metrics-03.txt:

* Added a paragraph about sudden changes due to mobility

and the heterogeneity of wireless access types.
Suggestion from Andras Veres.

* Add covariance as one of the metrics for oscillations.
Suggestion from Saverio Mascolo, original text
contribution from Injong Rhee.

Changes from draft-irtf-tmrg-metrics-02.txt:

* Added a few sentences to the Abstract about the
status of the document.

Changes from draft-irtf-tmrg-metrics-01.txt:

* Added a discussion about the metrics in IPPM.

Changes from draft-irtf-tmrg-metrics-01c.txt:

* Added to the discussion of network-based, flow-based,
and user-based metrics, based on email from Dado Colussi,
Sean Moore, Damon Wischik, Dah Ming Chiu, and others.

* Changed "packet drop rate" to "packet loss rate".
Suggestion from Nelson Fonseca.

* Added a discussion of the Colussi et al. paper on a new
definition of fairness.

* Added a discussion of the Chiu and Tan paper on redefining
fairness for inelastic traffic.

Changes from draft-irtf-tmrg-metrics-01b.txt:

* Added a discussion of goodput vs. throughput.
Suggestion from Nelson Fonseca.

Changes from draft-irtf-tmrg-metrics-01a.txt:

* Added to the discussion of packet drop rate metrics.
Suggestions from Janardhan Iyengar, Sean Moore,
Armando Caro, and Nelson Fonseca.

* Added a sentence about throughput used as a metric for
transfer times for very short flows.
Response to email from Sean Moore.

Changes from draft-irtf-tmrg-metrics-00.txt:

* Added a list of relevant congestion control mechanisms to
the abstract. Suggestion from Sean Moore.

* Added to the Introduction. Suggestion from Dado Colussi.

* Added a sentence about jitter to the discussion of minimizing oscillations. Suggestion from Wesley Eddy.

* Added a note about convergence between existing flows after a change in bandwidth. Suggestion from Wesley Eddy.

* Added to the section on deployability. Suggestion from Wesley Eddy.

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.

1. 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. This document doesn't attempt to specify every metric that a study might consider; for example, there are domain-specific metrics that researchers might consider that are over and above the metrics laid out in this document.

We consider metrics for aggregate traffic (taking into account the effect of flows on competing traffic in the network) as well as the heterogeneous goals of different applications or transport protocols (e.g., of high throughput for bulk data transfer, and of low delay for interactive voice or video). Different transport protocols or AQM mechanisms might have goals of optimizing different sets of metrics, with one transport protocol optimized for per-flow throughput and another optimized for robustness over wireless links, and with different degrees of attention to fairness with competing traffic. We hope this document will be used as a step in evaluating proposed congestion control mechanisms for a wide range of metrics, for example, noting that Mechanism X is good at optimizing Metric A, but pays the price with poor performance for Metric B. The goal would be to have a broad view of both the strengths and weaknesses of newly-proposed congestion control mechanisms.

Subsequent documents are planned to present sets of simulation and testbed scenarios for the evaluation of transport protocols and of congestion control mechanisms, based on the best current practice of the research community. These are not intended to be complete or final benchmark test suites, but simply to be one step of many to be used by researchers in evaluating congestion control mechanisms. Subsequent documents are also planned on the methodologies in using these sets of scenarios.

This document is a product of the Transport Modeling Research Group (TMRG), and has received detailed feedback from many members of the Research Group (RG). As the document tries to make clear, there is not necessarily a consensus within the research community (or the IETF community, the vendor community, the operations community, or any other community) about the metrics that congestion control mechanisms should be designed to optimize, in terms of tradeoffs between throughput and delay, fairness between competing flows, and the like. However, we believe that there is a clear consensus that congestion control mechanisms should be evaluated in terms of tradeoffs between a range of metrics, rather than in terms of optimizing for a single metric.

2. Metrics

The metrics that we discuss are the following:

- Throughput;
- Delay;
- Packet loss 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;
- User-based metrics.

We consider each of these below. Many of the metrics have network-based, flow-based, and user-based interpretations. For example, network-based metrics can consider aggregate bandwidth and aggregate drop rates, flow-based metrics can consider end-to-end transfer times for file transfers or end-to-end delay and packet drop rates for interactive traffic, and user-based metrics can consider user wait time or user satisfaction with the multimedia experience. Our main goal in this document is to explain the set of metrics that can be relevant, and not to legislate on the most appropriate methodology for using each general metric.

For some of the metrics, such as fairness, there is not a clear agreement in the network community about the desired goals. In these cases, the document attempts to present the range of approaches.

2.1. Throughput, Delay, and Loss Rates

Because of the clear tradeoffs between throughput, delay, and loss rates, it can be useful to consider these three metrics together. The tradeoffs are most clear in terms of queue management at the router; is the queue configured to maximize aggregate throughput, to minimize delay and loss rates, or some compromise between the two? 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.

Using flow-based metrics instead of router-based metrics, the relationship between per-flow throughput, delay, and loss rates can often be given by the transport protocol itself. For example, in TCP, the sending rate s in packets per second is given as

s = 1.2/(RTT*sqrt(p)),

for the round-trip time RTT and loss rate p [FF99].

2.1.1. Throughput

Throughput can be measured as a router-based metric of aggregate link utilization, as a flow-based metric of per-connection transfer times, and as user-based metrics of utility functions or user wait 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.

Throughput is sometimes distinguished from goodput, where throughput is the link utilization or flow rate in bytes per second, and goodput, also measured in bytes per second, is the subset of throughput consisting of useful traffic. That is, ‘goodput’ excludes duplicate packets, packets that will be dropped downstream, packet fragments or ATM cells that are dropped at the receiver because they can’t be re-assembled into complete packets, and the like. In general, this document doesn’t distinguish between throughput and goodput, and uses the general term "throughput".

We note that maximizing throughput is of concern in a wide range of environments, from highly-congested networks to under-utilized ones, and from long-lived flows to very short ones. As an example, throughput has been used as one of the metrics for evaluating Quick-Start, a proposal to allow flows to start-up faster than slow-start, where throughput has been evaluated in terms of the transfer times for connections with a range of transfer sizes [RFC4782] [SAF06].

In some contexts, it might be sufficient to consider the aggregate throughput or the mean per-flow throughput [DM06], 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.

2.1.2. Delay

Like throughput, delay can be measured as a router-based metric of queueing delay over time, or as a flow-based metric in terms of per-packet transfer times. Per-packet delay can also include delay at the sender waiting for the transport protocol to send the packet. For reliable transfer, the per-packet transfer time seen by the application includes the possible delay of retransmitting a lost 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.

Note that queueing delay at a router is shared by all flows at a FIFO (First-In First-Out) queue. Thus, the router-based queueing delay induced by bulk data transfers can be important even if the bulk data transfers themselves are not concerned with per-packet transfer times.

2.1.3. Packet Loss Rates

Packet loss rates can be measured as a network-based or as a flow-based metric.

When evaluating the effect of packet losses or ECN marks (Explicit Congestion Notification) [RFC3168] on the performance of a congestion control mechanism for an individual flow, researchers often use both the packet loss/mark rate for that connection, and the congestion event rate (also called the loss event rate), where a congestion event or loss event consists of one or more lost or marked packets in one round-trip time [RFC3448].

Some users might care about the packet loss rate only in so far as it affects per-connection transfer times, while other users might care about packet loss rates directly. RFC 3611, RTP Control Protocol Extended Reports, describes a VoIP performance-reporting standard called RTCP XR, which includes a set of burst metrics. In RFC 3611, a burst is defined as the maximal sequence starting and ending with a lost packet, and not including a sequence of Gmin or more packets that are not lost [RFC3611]. The burst metrics in RFC 3611 consist of the burst density (the fraction of packets in bursts), gap density (the fraction of packets in the gaps between bursts), burst duration (the mean duration of bursts in seconds), and gap duration (the mean duration of gaps in seconds). RFC 3357 derives metrics for "loss distance" and "loss period", along with statistics that capture loss patterns experienced by packet streams on the Internet [RFC3357].

In some cases it is useful to distinguish between packets dropped at routers due to congestion, and packets lost in the network due to corruption.

One network-related reason to avoid high steady-state packet loss rates is to avoid congestion collapse in environments containing paths with multiple congested links. In such environments, high packet loss rates could result in congested links wasting scarce bandwidth by carrying packets that will only be dropped downstream, before being delivered to the receiver [RFC2914]. We also note that in some cases the retransmit rate can be high, and the goodput correspondingly low, even with a low packet drop rate [AEO03].

2.2. 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.

2.2.1. 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.

Congestion control mechanisms also have to contend with sudden changes in the bandwidth-delay product due to mobility. Such bandwith-delay product changes are expected to become more frequent and to have greater impact than path changes today. As a result of both mobility and of the heterogeneity of wireless access types (802.11b,a,g, WIMAX, WCDMA, HS-WCDMA, E-GPRS, Bluetooth, etc.), both the bandwidth and the round-trip delay can change suddenly, sometimes by several orders of magnitude.

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 [RFC3448], and for AIMD (Additive Increase Multiplicative Decrease) or related mechanisms using parameters that make them more slowly-responding than 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] [YKL01]. 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 congestion 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. This aggressiveness can be a function of the mode of the transport protocol; for TCP, the aggressiveness of slow-start is quite different from the aggressiveness of congestion avoidance mode.

2.2.2. Minimizing Oscillations

One goal is that of stability, in terms of minimizing oscillations of queueing delay or of throughput. In practice, stability is frequently associated with rate fluctuations or variance. Rate variations can result in fluctuations in router queue size and therefore of queue overflows. These queue overflows can cause loss synchronizations across co-existing flows and periodic under-utilization of link capacity, both of which are considered to be general signs of network instability. Thus, measuring the rate variations of flows is often used to measure the stability of transport protocols. To measure rate variations, [JWL04], [RX05], and [FHPW00] use the coefficient of variation (CoV) of per-flow transmission rates and [WCL05] suggests the use of standard deviations of per-flow rates. Since rate variations are a function of time scales, it makes sense to measure these rate variations over various time scales.

Measuring per-flow rate variations, however, is only one aspect of transport protocol stability. A realistic experiment setting always involves multiple flows of the transport protocol being observed, along with a significant amount of cross traffic, with rates varying over time, on both the forward and reverse paths. As a congestion control protocol must adapt its rate to the varying rates of competing traffic, just measuring the per-flow statistics of a subset of the traffic could be misleading because it measures the rate fluctuations due in part to the adaptation to competing traffic on the path. Thus, per-flow statistics are most meaningful if they are accompanied by the statistics measured at the network level. As a complementary metric to the per-flow statistics, [HKLRX06] uses measurements of the rate variations of the aggregate flows observed in bottleneck routers over various time scales.

Minimizing oscillations in queueing delay or throughput has related per-flow metrics of minimizing jitter in round-trip times and loss rates.

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 in Section 2.2.1 above.

As discussed in [FK07], the synchronization of loss events can also affect convergence times, throughput, and delay.

2.3. Fairness and Convergence

Another set of metrics is that of fairness and convergence times. Fairness can be considered between flows of the same protocol, and between flows using different protocols (e.g., TCP-friendliness, referring to fairness between TCP and a new transport protocol). Fairness can also be considered between sessions, between users, or between other entities.

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].

2.3.1. Metrics for fairness between flows

This section discusses fairness metrics that consider the fairness between flows, but that don't take into account different characteristics of flows (e.g., the number of links in the path, or the round-trip time). For the discussion of fairness metrics, let x_i be the throughput for the i-th connection.

Jain's fairness index: The fairness index in [JCH84] is

$$((\text{sum}_i\ x_i)^2) / (n * \text{sum}_i ((x_i)^2)),$$

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

product_i x_i ,

the product of the throughput of the individual connections, is also used as a measure of fairness. (In some 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 [MS91] 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.

Epsilon-fairness: A rate allocation is defined as epsilon-fair if

(min_i x_i) / (max_i x_i) >= 1 - epsilon.

Epsilon-fairness measures the worst-case ratio between any two throughput rates [ZKL04]. Epsilon-fairness is related to max-min fairness, defined later in this document.

2.3.2. Metrics for fairness between flows with different resource requirements

This section discusses metrics for fairness between flows with different resource requirements, that is, with different utility functions, round-trip times, or numbers of links on the path. Many of these metrics can be described as solutions to utility maximization problems [K01]; the total utility quantifies both the fairness and the throughput.

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. (Max-min fairness can be explained by the progressive filling algorithm, where all flow rates start at zero, and the rates all grow at the same pace. Each flow rate stops growing only when one or more links on the path reaches link capacity.)

Proportional fairness: In contrast, a feasible 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]. (Using the language of utility functions, proportional fairness can be achieved by using logarithmic utility functions, and maximizing the sum of the per-flow utility functions; see [KMT98] for a fuller explanation.)

Minimum potential delay fairness: Minimum potential delay fairness has been shown to model TCP [KS03], and is a compromise between max-min fairness and proportional fairness. An allocation x is defined as having minimum potential delay fairness if

sum_i (1/x_i)

is smaller than for any other feasible allocation. That is, it would minimize the average download time if each flow was an equal-sized file.

In [CRM05], Colussi et al. propose a new definition of TCP fairness, that "a set of TCP fair flows do not cause more congestion than a set of TCP flows would cause", where congestion is defined in terms of queueing delay, queueing delay variation, the congestion event rate [e.g., loss event rate], and the packet loss rate.

Chiu and Tan in [CT06] argue for redefining the notion of fairness when studying traffic controls for inelastic traffic, proposing that inelastic flows adopt other traffic controls such as admission control.

Briscoe in [B07] argues that flow-rate fairness should not be a desired goal, and that instead "a realistic fairness mechanism must share out the ‘cost’ of each users actions on others".

2.3.3. Comments on fairness

Tradeoffs between fairness and throughput: The fairness measures in the section above generally measure both fairness and throughput, giving different weights to each. Potential tradeoffs between fairness and throughput are also discussed by Tang et al. in [TWL06], for a framework where max-min fairness is defined as the most fair. In particular, [TWL06] shows that in some topologies throughput is proportional to fairness, while in other topologies throughput is inversely proportional to fairness.

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]. Utility maximization provides one framework for describing this tradeoff in fairness.

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). One common argument against the goal of fairness between flows with different round-trip times has been that flows with long round-trip times consume more resources; this aspect is covered by the previous paragraph. Researchers have also noted the difference between the RTT-unfairness of standard TCP, and the greater RTT-unfairness of some proposed modifications to TCP [LLS05].

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 [RFC3714], with [RFC4828] 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 long-delay flows. Convergence times also concern the time for convergence to fairness after a sudden change such as a change in the network path, the competing cross-traffic, or the characteristics of a wireless link. 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].

2.4. 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, reordering, 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 [GF04]. In general, the 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].

The goodput ratio: For wireless networks, the goodput ratio can be a useful metric, where the goodput ratio can be defined as the useful data delivered to users as a fraction of the total amount of data transmitted on the network. A high goodput ratio indicates an efficient use of the radio spectrum and lower interference with other users.

2.5. 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.

2.6. 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.

One key aspect of deployability concerns the range of deployment needed for a new congestion control mechanism. Consider the following possible deployment requirements:

* Only at the sender (e.g., NewReno in TCP [RFC3782]);
* Only at the receiver (e.g., delayed acknowledgements in TCP);
* Both the sender and receiver (e.g., SACK TCP [RFC2018]);
* At a single router (e.g., RED [FJ93]);
* All of the routers along the end-to-end path;
* Both end nodes and all routers along the path (e.g., XCP [KHR02]).

Some congestion control mechanisms (e.g., XCP [KHR02], Quick-Start [RFC4782]) may also have deployment issues with IPsec, IP in IP, MPLS, or other tunnels, or with non-router queues such as those in Ethernet switches.

Another deployability issue concerns the complexity of the code. How complex is the code required to implement the mechanism in software? Is floating point math required? How much new state must be kept to implement the new mechanism, and who holds that state, the routers or the end-nodes? We note that we don't suggest these questions as ways to reduce the deployability metric to a single number; we suggest them as issues that could be considered in evaluating the deployability of a proposed congestion control mechanism.

2.7. Metrics for Specific Types of Transport

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

2.8. User-Based Metrics

An alternate approach that has been proposed for the evaluation of congestion control mechanisms would be to evaluate in terms of user metrics such as user satisfaction, or in terms of application-specific utility functions. Such an approach would require the definition of a range of user metrics or of application-specific utility functions for the range of applications under consideration (e.g., FTP, HTTP, VoIP).

3. Metrics in the IP Performance Metrics (IPPM) Working Group

The IPPM Working Group [IPPM] was established to define performance metrics to be used by network operators, end users, or independent testing groups. The metrics include metrics for connectivity [RFC2678], delay and loss [RFC2679] [RFC2680] [RFC2681], delay variation [RFC3393], loss patterns [RFC3357], packet reordering [RFC4737], bulk transfer capacity [RFC3148], and link capacity [CI07]. The IPPM documents give concrete, well-defined metrics, along with a methodology for measuring the metric. The metrics discussed in this document have a different purpose from the IPPM metrics; in this document we are discussing metrics as used in analysis, simulations, and experiments for the evaluation of congestion control mechanisms. Further, we are discussing these metrics in a general sense, rather than looking for specific concrete definitions for each metric. However, there are many cases where the metric definitions from IPPM could be useful, for specific issues of how to measure these metrics in simulations or in testbeds, and for providing common definitions for talking about these metrics.

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.

As an example from the literature on evaluating transport protocols, Li et al. in "Experimental Evaluation of TCP Protocols for High-Speed Networks" [LLS05] focus on the performance of TCP in high-speed networks, and consider metrics for aggregate throughput, loss rates, fairness (including fairness between flows with different round-trip times), response times (including convergence times), and incremental deployment. More general references on methodology include [J91]. Papers that discuss the range of metrics for evaluating congestion control include [MTZ04].

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 Lachlan Andrew, Mark Allman, Armando Caro, Dah Ming Chiu, Eric Coe, Dado Colussi, Wesley Eddy, Nelson Fonseca, Janardhan Iyengar, Doug Leith, Tony Li, Saverio Mascolo, Sean Moore, Injong Rhee, David Ros, Juergen Schoenwaelder, Andras Veres, Michael Welzl, and Damon Wischik, and members of the Transport Modeling Research Group for feedback and contributions.

Informative References

[AEO03] M. Allman, W. Eddy, and S. Ostermann, Estimating Loss Rates With TCP, ACM Performance Evaluation Review, 31(3), December 2003.

[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.29 N.4, April 1981.

[B07] B. Briscoe, Flow Rate Fairness: Dismantling a Religion, internet-draft draft-briscoe-tsvarea-fair-02.txt, work-in-progress, July 2007.

[CI07] P. Chimento and J. Ishac, Defining Network Capacity, internet-draft draft-ietf-ippm-bw-capacity-05, work-in-progress, May 2007.

[CRM05] D. Colussi, A New Approach to TCP-Fairness, Report C-2005-49, University of Helsinki, Finland, 2005.

[CT06] D. Chiu and A. Tam, Redefining Fairness in the Study of TCP-friendly Traffic Controls, Technical Report, 2006.

[DM06] N. Dukkipati and N. McKeown, Why Flow-Completion Time is the Right Metric for Congestion Control, ACM SIGCOMM, January 2006.

[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.

[FF99] Floyd, S., and Fall, K., "Promoting the Use of End-to-End Congestion Control in the Internet", IEEE/ACM Transactions on Networking, August 1999.

[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/".

[FHPW00] S. Floyd, M. Handley, J. Padhye, and J. Widmer, Equation-Based Congestion Control for Unicast Applications, SIGCOMM 2000, August 2000.

[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.

[FJ93] S. Floyd and V. Jacobson, Random Early Detection gateways for Congestion Avoidance, IEEE/ACM Transactions on Networking, V.1 N.4, August 1993,

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

[FK07] S. Floyd and E. Kohler, Tools for the Evaluation of Simulation and Testbed Scenarios, internet-draft draft-irtf-tmrg-tools-04.txt, July 2007.

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

[HKLRX06] S. Ha, Y. Kim, L. Le, I. Rhee, and L. Xu, A Step Toward Realistic Evaluation of High-speed TCP Protocols, technical report, North Carolina State University, January 2006.

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

[IPPM] IP Performance Metrics (IPPM) Working Group, URL "http://www.ietf.org/html.charters/ippm-charter.html".

[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.

[JWL04] C. Jin, D. Wei, and S. Low, FAST TCP: Motivation, Architecture, Algorithms, Performance, IEEE INFOCOM, March 2004.

[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.

[KS03] S. Kunniyur and R. Srikant, End-to-end Congestion Control Schemes: Utility Functions, Random Losses and ECN Marks, IEEE/ACM Transactions on Networking, 11(5):689-702, October 2003.

[LLS05] Y-T. Li, D. Leith, and R. Shorten, Experimental Evaluation of TCP Protocols for High-Speed Networks, Hamilton Institute, 2005. URL "http://www.hamilton.ie/net/eval/results-HI2005.pdf".

[MS91] D. Mitra and J. Seery, Dynamic Adaptive Windows for High Speed Data Networks with Multiple Paths and Propagation Delays, INFOCOM '91, pp 39-48.

[MTZ04] L. Mamatas, V. Tsaoussidis, and C. Zhang, Approaches to Congestion Control in Packet Networks, 2004.

[RFC2018] TCP Selective Acknowledgement Options. M. Mathis, J. Mahdavi, D. Floyd, and A. Romanow. RFC 2018, Proposed Standard, April 1996.

[RFC2680] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Packet Loss Metric for IPPM", RFC 2680, September 1999.

[RFC2678] Mahdavi, J. and V. Paxson, "IPPM Metrics for Measuring Connectivity", RFC 2678, September 1999.

[RFC2679] Almes, G., Kalidindi, S. and M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, September 1999.

[RFC2681] Almes, G., Kalidindi, S., Zekauskas, M., "A Round-Trip Delay Metric for IPPM". RFC 2681, STD 1, September 1999.

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

[RFC3148] Mathis, M. and M. Allman, "A Framework for Defining Empirical Bulk Transfer Capacity Metrics", RFC 3148, July 2001.

[RFC3168] K.K. Ramakrishnan, S. Floyd, and D. Black, The Addition of Explicit Congestion Notification (ECN) to IP, RFC 3168, September 2001.

[RFC3357] R. Koodli and R. Ravikanth, One-way Loss Pattern Sample Metrics, RFC 3357, Informational, August 2002.

[RFC3393] Demichelis, C. and P. Chimento, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)", RFC 3393, November 2002.

[RFC3448] M. Handley, S. Floyd, J. Padhye, and J. Widmer, TCP Friendly Rate Control (TFRC): Protocol Specification, RFC 3448, Proposed Standard, January 2003.

[RFC3611] T. Friedman, R. Caceres, and A. Clark, RTP Control Protocol Extended Reports (RTCP XR), RFC 3611, November 2003.

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

[RFC3782] S. Floyd, T. Henderson, and A. Gurtov, The NewReno Modification to TCP's Fast Recovery Algorithm, RFC 3782, Proposed Standard, April 2004.

[RFC4737] A. Morton, L. Ciavattone, G. Ramachandran, S. Shalunov, J. Perser, Packet Reordering Metrics, RFC 4737, November 2006.

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

[RFC4828] S. Floyd and E. Kohler, TCP Friendly Rate Control (TFRC): the Small-Packet (SP) Variant, RFC 4828, April 2007.

[RX05] I. Rhee and L. Xu, CUBIC: A New TCP-Friendly High-Speed TCP Variant, PFLDnet 2005.

[SAF06] P. Sarolahti, M. Allman, and S. Floyd, Determining an Appropriate Sending Rate Over an Underutilized Network Path, Computer Networks, September 2006.

[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] S. Savage, N. Cardwell, D. Wetherall, and T. Anderson, 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.

[TWL06] A. Tang, J. Wang and S. H. Low. Counter-Intuitive Throughput Behaviors in Networks Under End-to-End Control, IEEE/ACM Transactions on Networking, 14(2):355-368, April 2006.

[WCL05] D. X. Wei, P. Cao and S. H. Low, Time for a TCP Benchmark Suite?, Technical Report, Caltech CS, Stanford EAS, Caltech, 2005.

[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.

[YKL01] Y. Yang, M. Kim, and S. Lam, Transient Behaviors of TCP-friendly Congestion Control Protocols, Infocom 2001.

[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 at icir.org>
ICSI Center for Internet Research
1947 Center Street, Suite 600
Berkeley, CA 94704
USA

Full Copyright Statement

Intellectual Property