Performance modeling of elastic traffic in overload

Thomas Bonald
France Telecom R&D
thomas.bonald@francetelecom.com

James Roberts
France Telecom R&D
james.roberts@francetelecom.com

ABSTRACT

While providers generally aim to avoid congestion by adequate provisioning, overload can clearly occur on certain network links. In this paper we propose some simple preliminary models for an overloaded link accounting for user impatience and reattempt behavior.

1. INTRODUCTION

We consider a bottleneck link of capacity C receiving traffic from non-persistent TCP flows. We assume flows arrive according to a given random process and that the distribution of the amount of data transferred, or flow size, has a Pareto distribution:

$$F(s) = \Pr[\text{size} > s] = (k/s)^{\beta}, \quad (1)$$

where k is a minimum size and $1 < \beta \leq 2$ [2]. This distribution has the particularity that most flows are very small (so-called "mice") while most of the traffic is contained in large flows (so-called "elephants"). Let λ denote the flow arrival rate and $\rho = \lambda \mathrm{E}[s]/C$ the link offered load.

We model traffic at flow level using a fluid model and assume TCP shares bandwidth perfectly fairly. This model has been used to evaluate throughput performance when $\rho < 1$ (see [1], for example). In the present case we consider what happens when $\rho > 1$. Note that this overload situation is not exceptional and can occur due to planning errors, outages, traffic surges or deliberate underprovisioning of costly transoceanic routes or strategic peering links. A preliminary study of demand overload and the way stability is maintained by user impatience was described in [4] under Markovian assumptions. We adopt an alternative approach here accounting for the Pareto size distribution (1) and more general assumptions regarding impatience and reattempt behavior.

2. FLOW LEVEL CONGESTION

In overload, the number of flows in progress tends to increase and their throughput tends to zero since the arrival rate λ is greater than the average rate of flow completion. This behavior is illustrated in Figure 1 which shows results from an ns-2 simulation of a 10 Mbit/s link under 20% overload. The link is empty at time zero and fed by Poisson arrivals of short-lived TCP flows with a Pareto distributed document size (for numerical results, we take $E[s] = 45\text{KB}$, $\beta = 1.5$). The dots represent the throughput realized by flows at the instant of their completion.

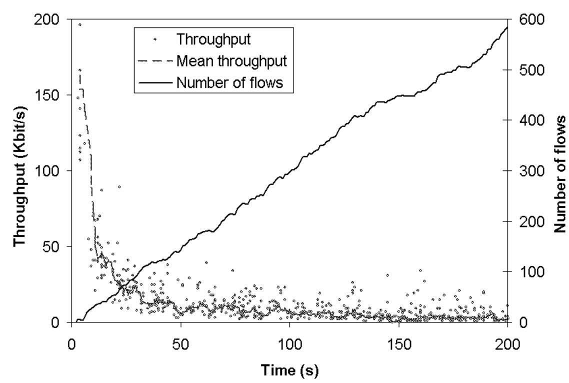

Figure 1: Performance during transient overload

It is interesting to note that the rate of increase in the number of flows actually depends significantly on the tail of the size distribution [3]. The rate of increase is smaller as the tail is heavier. The very short mice manage to complete while the elephants remain indefinitely as realized throughput continues to decrease.

Note that the degree of congestion on a network link may not be easy to detect simply by observing the packet arrival process. The packet arrival rate is controlled by TCP, ideally at a value close to the service rate, even though the number of contributing flows may be increasing rapidly. Elastic traffic congestion is manifested essentially at flow level rather than at packet level.

3. IMPATIENCE

In a real network, if demand exceeds capacity, the number of flows in progress does not increase indefinitely. As per-flow throughput decreases some flows or sessions will be interrupted, due either to user impatience or to aborts by TCP or higher layer protocols. In the following we use the term impatience for all causes of premature abandon.

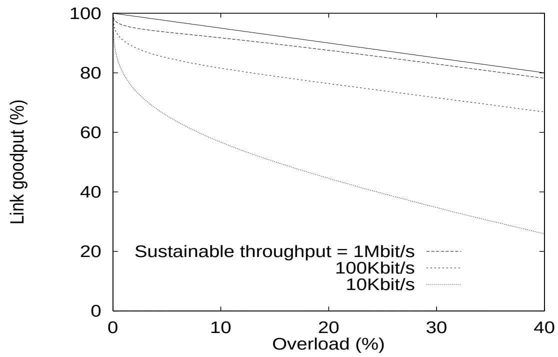

Figure 2: Impact of user behavior on link goodput

We are unaware of any published statistics on user impatience. The phenomenon is clearly very difficult to observe in practice, notably because all flow aborts are not in reaction to excessive response time and because most impatience is manifested by the interruption of a session and may not be detectable as an abnormal event. However, to gain some insight we here propose a simple hypothetical model.

We suppose a flow of size s will be interrupted if and when its response time exceeds a patience duration $\delta(s)$. It seems natural to assume that δ is an increasing but concave function of s since users have a response time expectation which increases with the flow size but need proportionally more throughput. Such impatience causes the number of flows in progress to stabilize and vary slightly about a certain mean value, resulting in an approximately equal bandwidth share for each flow. To simplify we assume here that each flow receives exactly the same bandwidth γ, typically much smaller than the link capacity. A flow of size s is then completed if and only if $s \leq \gamma\delta(s)$.

It follows from the concavity of δ that there exists a critical flow size $s^\star$ satisfying $s^\star = \gamma\delta(s^\star)$ such that any flow of size smaller than $s^\star$ is completed while all the others are aborted. We also note that γ is necessarily smaller than the limit of $s/\delta(s)$ when s tends to infinity. This limit is referred to as the *sustainable throughput*, i.e., the minimum throughput required to transfer documents of arbitrarily large size. The precise value of γ depends strongly on δ and the flow size distribution. It can be determined by arguing that, since the link is always saturated, we must have:

$$C = \lambda \int_0^\infty \min(s, \gamma\delta(s))dF(s). \tag{2}$$

The link goodput, i.e., the fraction of link capacity used by flows that are effectively completed, is then given by:

$$U = \frac{\lambda}{C} \int_0^{s^\star} s.dF(s). \tag{3}$$

This model provides some useful insights into the impact of congestion. Both realized throughput and link goodput deteriorate with longer patience and increasing overload, but are otherwise independent of link capacity. This is illustrated by Figure 2, which plots the link goodput against overload for a linear patience duration $\delta(s) = \tau + s/\alpha$, with $\tau = 20$s and different values of the sustainable throughput α. The top line corresponds to the limiting case $\alpha = \infty$ (i.e., patience duration is independent of size), for which we have the simple expression:

$$U = 1 - (\beta - 1)(\rho - 1). \tag{4}$$

On the other hand, confirming the positive impact of a heavy-tailed size distribution noted in Section 2, both throughput and goodput improve as β decreases from 2 to 1. This is explained by the fact that impatience discriminates against elephants and interrupts these large flows after only a small fraction of their data has been tranferred.

4. REATTEMPTS

Aborted flows are not generally abandoned at the first attempt as users will frequently make a repeat attempt. The impact of this behavior is to exacerbate the loss of goodput due to impatience as it is likely that the reattempts will also be interrupted. Consider the following simple model.

User behavior is modeled by a size dependent patience duration as introduced in the previous section. If a user aborts it reattempts with fixed probability p. Reasoning as above, in place of equation (2), the maximum completed flow size $s^\star$ now satisfies:

$$C = \lambda \int_0^{s^\star} s.dF(s) + \frac{\lambda}{1-p} \int_{s^\star}^\infty \gamma\delta(s)dF(s). \tag{5}$$

The loss of goodput due to reattempts can be considerable. In the particular case of a 20% overloaded link with a fixed patience duration, goodput decreases from 90% to 60% as p increases from 0 to 1/2. While the model is overly simple, it does illustrate the negative impact user behavior can have in case of overload. Network efficiency would gain from a more proactive reaction to congestion.

5. ADMISSION CONTROL

An alternative to allowing an overloaded link to stabilize through impatience is to perform admission control. If flows arriving when the bandwidth they would achieve is less than an acceptable threshold were rejected immediately, there would be no cause for impatience. The advantage is that goodput is maintained close to 100% and is unaffected by reattempts. On the other hand, admission control would tend to increase the proportion of uncompleted flows since it applies equally to both mice and elephants. Since elephants may well be more important than mice, this is not necessarily a disadvantage. Indeed, one advantage of admission control is that it can be applied selectively with different admission thresholds applying to different classes of traffic.

6. REFERENCES

[1] T. Bonald and L. Massoulié, "Impact of fairness on Internet performance", *ACM SIGMETRICS'01*, 2001.

[2] M. Crovella and A. Bestravos, "Self-similarity in World Wide Web traffic: Evidence and possible cause", *ACM SIGMETRICS'96*, 1996.

[3] A. Jean-Marie and P. Robert, "On the transient behavior of the processor sharing queue", Queueing Systems Theory and Applications 17 (1994) 129–136.

[4] L. Massoulié and J.W. Roberts, "Arguments in favour of admission control for TCP flows", *ITC 16th*, 1999.