Efficient resource allocation in bandwidth-sharing networks

Maaike Verloop*, Rudesindo Núñez-Queija*,‡

*CWI, The Netherlands
‡TNO Information and Communication Technology, The Netherlands

1. INTRODUCTION

Document transfer in the Internet is regulated by distributed packet-based congestion control mechanisms, usually relying on TCP. By dividing a document into packets, parts of one file reside at different nodes along the transmission path. The "instantaneous transfer rate" of the entire document can be thought of as being equal to the minimum transfer rate along the entire path. Bandwidth-sharing networks as considered by Massoulié & Roberts [2] provide a natural modeling framework for the dynamic flow-level interaction among document transfers. The class of α-fair policies for such networks, as introduced by Mo & Walrand [3], captures a wide range of distributed allocation mechanisms such as TCP, the proportional fair allocation and the max-min fair allocation.

Identifying optimal resource allocations in bandwidth-sharing networks is inherently complex: The distributed nature of resource allocation management prohibits global coordination for efficiency, i.e., aiming at full resource usage at all times. In addition, it is well recognized that resource efficiency may be conflicting with other critical performance measures such as flow delay. Without a notion of optimal (or "near-optimal") behavior, the performance of resource allocation schemes cannot be assessed properly.

An exact characterization of the optimal policy is in general not possible. In addition, numerically determining the optimal strategy often requires excessive computational effort. We therefore set out to study these in asymptotic regimes. In [6] this was done for a highly loaded network. In this paper, we study the asymptotically optimal strategies in the under-loaded case after scaling the state space. Armed with these, we then assess the potential gain that any sophisticated strategy can achieve over standard α-fair strategies, and confirm that α-fair strategies perform excellently. This is particularly true for the proportional fair policy ($\alpha = 1$).

Let us focus on a linear bandwidth-sharing network of two nodes, each with unit service rate, see Figure 1. There are three traffic classes, where class i requires service at node i only, $i = 1, 2$, while class 0 requires service at both nodes simultaneously. Class-i users arrive as a Poisson process of rate λ_i, and have exponentially distributed service requirements, B_i, with mean $1/\mu_i$. Let the traffic load of class i be $\rho_i := \frac{\lambda_i}{\mu_i}$, thus the load at node i is $\rho_0 + \rho_i$. The conditions $\rho_0 + \rho_i < 1$, $i = 1, 2$, are necessary for stability. For the class of α-fair policies these conditions are also sufficient [1]. This is in contrast with straight-forward extensions of size-based scheduling strategies (known to have certain optimality properties when there is a single resource [4, 5]) which may not guarantee maximum stability [7]. The full-length version of this paper is available as a report, see [8].

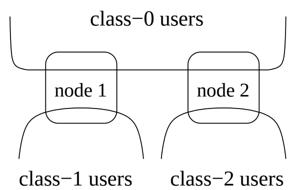

Figure 1: Linear bandwidth-sharing network

2. OPTIMALITY RESULTS

The central objective is to minimize the mean total number of users in the network among all non-anticipating policies. In the network context, besides trying to maximize the total output rate of the network, we must take into account that when serving class 1 while class 0 is present and class 2 is empty, leaves node 2 under-utilized. When $\mu_0 \geq \mu_1, \mu_2$, these two objectives are not conflicting and the optimal policy degenerates to a static priority rule.

PROPOSITION 2.1. *If $\mu_1 + \mu_2 \leq \mu_0$, then the number of users is minimized (at any time) by giving preemptive priority to class 0. Similarly, if $\mu_1, \mu_2 \leq \mu_0 \leq \mu_1 + \mu_2$, then in both nodes full service must be allocated to class 0, unless both classes 1 and 2 are present as well.*

The case that remains unsolved is when for example $\mu_0 < \mu_1$. When users of both classes 1 and 2 are present, serving them will be optimal, since $\mu_0 \leq \mu_1 + \mu_2$. When there are only users of classes 0 and 1 present (no class-2 users), no strict priority rule is optimal. It may still be better to sometimes serve class 0 even if that does not maximize the departure rate in the short run. Doing so, creates the potential to serve classes 1 and 2 simultaneously in the future and therefore offer a higher degree of parallelism. Hence as the number of users varies, the system will dynamically switch between several priority rules, characterized by a switching curve as stated in the following proposition. Denote by $N_i(t)$ the number of class-i users at time t.

PROPOSITION 2.2. *There exists a switching curve $h(\cdot)$ such that, when $N_2(t) = 0$ it is optimal to serve class 0 at full rate if $N_1(t) \leq h(N_0(t))$ and to serve class 1 otherwise.*

An exact characterization of the switching curves is in general not possible. We therefore set out to study these in asymptotic regimes.

3. FLUID SCALING

We consider a sequence of systems indexed by a superscript k. The number of class-i users in the k-th system at time t is denoted by $N_i^k(t)$. We study the fluid limits, where time is also scaled linearly: $\lim_{k\to\infty} \frac{N_i^k(kt)}{k} =: n_i(t)$. The initial queue length depends on k such that $n_i(0) = a_i$. The fluid process is described by $\frac{\mathrm{d}n_j(t)}{\mathrm{d}t} = \lambda_j - \mu_j s_j(t)$, for $j = 0, 1, 2$, with $s_0 + s_i \le 1, i = 1, 2$. Optimizing the "drain time" under a fluid scaling gives a simple linear switching strategy, which turns out to accurately approximate the numerically found optimal strategy in the stochastic model when $\rho_1 \neq \rho_2$.

Proposition 3.1. *Assume $\rho_1 \le \rho_2$, $\rho_0 + \rho_2 < 1$ and $n_2 = 0$.*
If $\mu_1, \mu_2 \ge \mu_0$, it is optimal for the fluid model to serve class 0 at rate $1 - \rho_2$ (keeping n_2 equal to zero) whenever $n_1 \le \frac{\rho_2 - \rho_1}{1 - \rho_0 - \rho_2} n_0$ and fully serve class 1 otherwise.
If $\mu_1 \ge \mu_0 \ge \mu_2$, then the corresponding condition is $n_1 \le \frac{\mu_1\mu_2/\mu_0}{\mu_1 + \mu_2 - \mu_0} \times \frac{\rho_2 - \rho_1}{1 - \rho_0 - \rho_2} n_0$.

4. CENTRAL LIMIT THEOREM SCALING

When the two nodes on the flow path are equally congested ($\rho_1 = \rho_2$), however, the fluid scaling is not appropriate, and the corresponding strategy may not even ensure stability. For example, when $\mu_1, \mu_2 \ge \mu_0$, the switching curves in the fluid model are both equal to zero, and hence the fluid policy gives preemptive priority to classes 1 and 2. However, in the stochastic model this leads to an unstable system if $\rho_0 > (1-\rho_1)(1-\rho_2)$. Let us therefore describe more carefully the behavior of the processes below a switching curve. We study the free process (denoted by the symbol $\sim$) that only serves classes 1 and 2 during (short) excursions when both of them are positive and otherwise serves class 0. From the central limit theorem we obtain that $\tilde{N}_1$ has no drift (since $\rho_1 = \rho_2$) and random fluctuations are of order $\sqrt{k}$ in a time span k. We still have a linear drift for $\tilde{N}_0$:

Proposition 4.1. *Assume $\rho_1 = \rho_2$, $\lim_{k\to\infty} \frac{\tilde{N}_1^k(0)}{\mu_1\sqrt{k}} = d_1$ and $\tilde{N}_2^k(0) \equiv 0$. Then $\tilde{n}_0(t) = \tilde{n}_0(0) - \mu_0(1 - \rho_0 - \rho_1)t$, $\tilde{n}_i(t) = \tilde{n}_i(0), i = 1, 2$ and*

$$\lim_{k\to\infty} \frac{\tilde{N}_1^k(kt)}{\mu_1\sqrt{k}} \stackrel{d}{=} \mathbf{1}_{(BM(t)+d_1\ge 0)}(BM(t) + d_1),$$

where $BM(t)$ is a Brownian motion.

Although the above provides theoretical ground for the asymptotic optimality of a square-root switching curve of the shape $N_i = c_i\sqrt{N_0}$, it is not straightforward to *analytically* determine the optimal value for c_i. It involves calculating the first passage probabilities for the switching curves.

5. PERFORMANCE EVALUATION

We investigate switching curves of shape $N_i = c_i f(N_0)$, where $f(\cdot)$ is either a square-root, linear or a threshold function. The value of c_i is varied to assess its impact. We let $\mu_0 = 2$, $\mu_1 = \mu_2 = 5$. From Figure 2 we observe that the linear (square-root) policy attains the value of the optimal policy for $\rho_1 \neq \rho_2$ ($\rho_1 = \rho_2$) given that the best coefficient c_i is chosen. Note that the proportional fair allocation (PF) is only a few percent away from the theoretical optimum and performs very well.

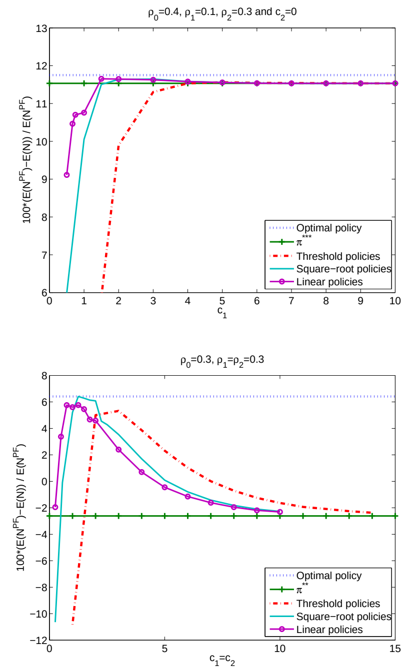

Figure 2: Performance of switching policies compared to the proportional fairness policy (PF)

6. CONCLUSIONS

Since optimal resource allocations are non-trivial in bandwidth-sharing networks, we set out to characterize their functional form under a fluid and diffusion scaling. Our aim was not to propose practically implementable strategies, but rather to provide a benchmark against which α-fair strategies can be tested. For illustration, the optimal policy depicted by a horizontal line in Figure 2 required a week's computation time, while the approximation by varying c_i can be obtained in real time.

7. REFERENCES

[1] Bonald, T., Massoulié, L. (2001). Impact of fairness on Internet performance. In: *Proc. ACM SIGMETRICS & Performance 2001 Conf.*, 82–91.

[2] Massoulié, L., Roberts, J.W. (2000). Bandwidth sharing and admission control for elastic traffic. *Telecommun. Syst.* **15**, 185–201.

[3] Mo, J., Walrand, J. (2000). Fair end-to-end window-based congestion control. *IEEE/ACM Trans. Netw.* **8**, 556–567.

[4] Righter, R., Shanthikumar, J.G. (1989). Scheduling multiclass single-server queueing systems to stochastically maximize the number of successful departures. *Prob. Eng. Inf. Sc.* **3**, 323–333.

[5] Schrage, L.E. (1968). A proof of the optimality of the shortest remaining processing time discipline. *Oper. Res.* **16**, 687–690.

[6] Verloop, I.M., Borst, S.C. (2007). Heavy-traffic delay minimization in bandwidth-sharing networks. In: *Proc. IEEE Infocom 2007.*

[7] Verloop, I.M., Borst, S.C., Núñez-Queija, R. (2005). Stability of size-based scheduling disciplines in resource-sharing networks. *Perf. Eval.* **62**, 247–262.

[8] Verloop, I.M., Núñez-Queija, R. (2007). Assessing the efficiency of resource allocations in bandwidth-sharing networks. *CWI-report PNA-E0702.*