On a Theory of Interacting Queues

Alexander Stepanenko[1], Costas C. Constantinou[1], Theodoros N. Arvanitis[1],
and Kevin Baughan[2]

[1] School of Electrical, Electronic and Computer Engineering, University of
Birmingham, Edgbaston, B15 2TT, UK
[2] Nortel Networks, Maidenhead Office Park, Westacott Way, Maidenhead, Berkshire,
SL6 3QH, UK

Abstract. We present a possible way to extend queuing theory to account for interactions between adjacent queues in a packet-switched network. The interaction between queues arises because of the influence of the routing protocol on each switching decision and the stochastic nature of packet lengths and inter-arrival times.

Both the methodology and the analysis tools are adaptations of methods of statistical mechanics and are presented in outline here. The justification for their use lies in experimental evidence given in [1,2,3] that aggregate, core-network IP traffic exhibits quasi-Markovian properties.

In this paper, we focus on the interaction between pairs of queues, either in a cascaded arrangement, or connected to the same switching fabric, in the presence of an idealised routing protocol.

1 Introduction

Next generation telecommunication networks are likely to rely on an IP core network infrastructure. As a consequence, unlike today's Internet, future networks will be subject to much more demanding requirements. In order to provide operational guarantees, network owners need tools which enable the dimensioning of the core of packet-switched networks. It is customary that network design is based mainly on simulations. Replacing theory by simulations is non-tenable due to the sheer size of such networks.

A theory of traffic in packet-switched networks must be capable of predicting a number of quantities of interest: end-to-end latency, packet loss rate, etc. Based on such metrics of network performance we would then wish to quantify the external loading point of a large system of interconnected queues at which the network changes its behaviour. For example, if packet loss rate is of interest, we wish to know the loading point at which this rate exceeds a given threshold. One of the important features of such a theory is that it must model all sources of stochasticity in the system of interconnected queues that form the core network. This implies that correlations between the states of queues, which have a knock-on effect on each other, must be modelled explicitly as interactions.

Constructing of this theory has attracted a lot of attention [4,5,6,7,8,9] but this work is far from over. In this paper we take a somewhat different approach

in attempting to formulate a phenomenological theory of traffic in interacting queues. The chosen approach can be scaled to large numbers of queues which is essential in modelling a large-scale core IP network. The methodology adopted here is an adaptation of the methods of statistical mechanics, which are ideally suited to the study of large systems in the presence of sources of stochasticity.

The "microscopic" state variables which we choose to define queue dynamics, will here be the lengths of each and every buffer in the network. The sources of stochasticity we consider are the random packet lengths (measured in bits) and packet inter-arrival times. These latter random variables are distributed according to some probability distribution, which will not be discussed here (c.f. [1,2,3]).

2   Background

The average latency along some path, or the loss rate of a single buffer, or a group of buffers can be calculated using a joint probability distribution (PDF) function of all queue lengths of all routers at time t, $P = P(\{\ell_i\}_{i=1}^{N_q}; t)$, where N_q is the number of queues in the network.

It is natural to expect that some queue lengths are highly correlated, whereas others are not. So, it is quite difficult to write a dynamical equation (e. g. a Fokker-Planck equation) for P and, never mind, solve it. For this reason we try to exploit some kind of approximation scheme and the simplest possible one is the mean-field approximation originating from Quantum Field Theory.

In the simplest case the mean-field approximation implies that the joint PDF can be represented as follows

$$P(\{\ell_i\}_{i=1}^{N_q}; t) = \prod_{i=1}^{N_q} p(\ell_i, t) \tag{1}$$

where $p(\ell_i, t)$ is the PDF of an individual queue and should be determined in a self-consistent manner to account for interactions. The dynamical equation for $p(\ell_i, t)$ of an individual queue should account for the fact that the queue interacts with an "average field" of all other queues, in this case an "average" network load. Individual PDF's in the right-hand side of (1) can actually depend on more than one variable ℓ for strongly correlated subsystems. This depends on the level of approximation we wish to use to account explicitly for certain types of correlation between queues. In this paper we consider two such subsystems: two cascaded queues and n queues attached to the same router.

It has been shown in [1,2,3] that the traffic in a core IP network acquires quasi-Markovian properties in and near its congested state. Hence, for a sub-system's PDF we can write down a Fokker-Planck equation [10]. Parameters entering the Fokker-Planck equation and, as a result, the PDF's themselves will depend on overall network load, capacity of lines and switches, average states of adjacent queues, etc., and should be determined in a self-consistent manner but this is beyond the scope of this paper.

3 Dynamical Model for Interacting Queues in Subsystem

As an example of a subsystem we first consider two cascaded queues consisting of an output buffer of a router connected to an input buffer of an adjacent router (see Fig. 1). Omitting the derivation details (which can be found in [10]), the Fokker-Planck (FP) equation for this system has the following form,
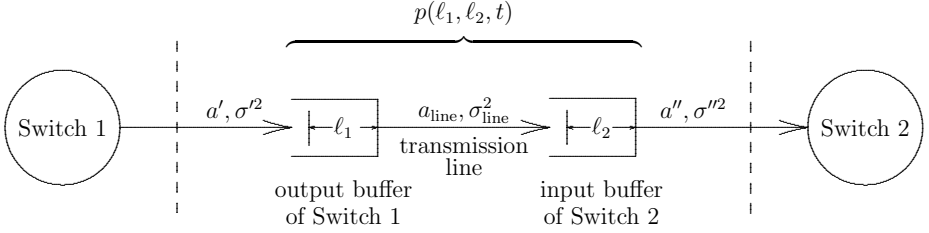

Fig. 1. Subsystem of two cascaded queues

$$
\partial_t p(\ell_1, \ell_2, t) = - \frac{\partial}{\partial \ell_1} \left[(a_1 - b_1 \ell_1 - b_2 \ell_2) p(\ell_1, \ell_2, t) \right] - a_2 \frac{\partial}{\partial \ell_2} p(\ell_1, \ell_2, t)
$$
$$
+ \frac{\sigma_1^2}{2} \frac{\partial^2}{\partial \ell_1^2} p(\ell_1, \ell_2, t) + \frac{\sigma_2^2}{2} \frac{\partial^2}{\partial \ell_2^2} p(\ell_1, \ell_2, t) - \sigma_{\text{line}}^2 \frac{\partial^2}{\partial \ell_1 \partial \ell_2} p(\ell_1, \ell_2, t)
$$
$$
\tag{2}
$$

where

$$
a_1 = \frac{a' - a_{\text{line}}}{\ell_{\max}}, \quad a_2 = \frac{a_{\text{line}} - a''}{\ell_{\max}}, \quad \sigma_1^2 = \frac{\sigma'^2 + \sigma_{\text{line}}^2}{\ell_{\max}^2}, \quad \sigma_2^2 = \frac{\sigma_{\text{line}}^2 + \sigma''^2}{\ell_{\max}^2} \tag{3}
$$

and natural boundary conditions are assumed [10]

$$
\mathbf{n} \cdot \mathbf{J}(\ell_1, \ell_2, t)|_{(\ell_1, \ell_2) \in S} = 0 \tag{4}
$$

Here we have introduced the concept of a probability current [10], which in our case is defined as follows

$$
\mathbf{J} = \begin{pmatrix} J_1 \\ J_2 \end{pmatrix} \tag{5}
$$

where

$$
J_1 = (a_1 - b_1 \ell_1 - b_2 \ell_2) p(\ell_1, \ell_2, t) - \frac{\sigma_1^2}{2} \frac{\partial}{\partial \ell_1} p(\ell_1, \ell_2, t) + \frac{\sigma_{\text{line}}^2}{2} \frac{\partial}{\partial \ell_2} p(\ell_1, \ell_2, t)
$$
$$
J_2 = a_2 p(\ell_1, \ell_2, t) - \frac{\sigma_2^2}{2} \frac{\partial}{\partial \ell_2} p(\ell_1, \ell_2, t) + \frac{\sigma_{\text{line}}^2}{2} \frac{\partial}{\partial \ell_1} p(\ell_1, \ell_2, t)
$$
$$
\tag{6}
$$

and $\mathbf{n}$ is a unit vector normal to the boundary S. The boundary S is a square with the sides at $\ell_1 = 0, \ell_1 = 1, \ell_2 = 0, \ell_2 = 1$. Note that we measure lengths of queues ℓ_1, ℓ_2 in terms of a fraction of the size of the corresponding buffer (here we set them all equal for the sake of simplicity), so that ℓ_1, ℓ_2 run from 0 to 1. Parameters a', σ'^2 characterise the mean value and the variance per unit time of the traffic coming into the first buffer [10]; $a_{\text{line}}, \sigma^2_{\text{line}}$ characterise the mean value and the variance of the traffic coming through the line; a'', σ''^2 characterise the mean value and the variance of the switching capacity available to the second buffer (traffic is measured in bits and all quantities are per unit time). We do not discuss here the nature of these parameters as they should be defined self-consistently with other parameters of a broader model for the overall network. Parameters b_1, b_2 are sensitivities of the routing protocol to the congestion of the queues. The dependence of the FP equation on the protocol sensitivities can be explained as follows. Relative queue lengths ℓ_1, ℓ_2 quantify the congestion level of the subsystem. The factor $a_1 - b_1\ell_1 - b_2\ell_2$ determines the average amount of traffic diverted to the subsystem: the more it is congested the less traffic (on average) is (or should be) diverted to it.

We seek an equilibrium (stationary) solution of the FP equation. For a equilibrium solution the detailed balance condition must be satisfied [10], and this condition demands that the parameters b_1, b_2 to be constrained as follows:

$$b_1 = b\cos\varphi \,, \quad b_2 = b\sin\varphi \,, \qquad \cos\varphi = \frac{\sigma_2^2}{\sqrt{\sigma_2^4 + \sigma_{\text{line}}^4}} \,, \quad \sin\varphi = \frac{\sigma_{\text{line}}^2}{\sqrt{\sigma_2^4 + \sigma_{\text{line}}^4}} \tag{7}$$

with b to be a single free parameter characterising the sensitivity of the protocol to the congestion of the subsystem as a whole. The stationary solution is

$$p^{(s)}(\ell_1, \ell_2) = \mathcal{N}^{-1} \exp\left[-\frac{b}{\sigma^2}\left(\lambda_1 - \frac{\bar{a}}{b}\right)^2 + \frac{2a_2\lambda_2}{\sqrt{\sigma_2^4 + \sigma_{\text{line}}^4}}\right] \tag{8}$$

where

$$\lambda_1 = \ell_1\cos\varphi + \ell_2\sin\varphi \,, \qquad \lambda_2 = \ell_2\cos\varphi - \ell_1\sin\varphi \tag{9}$$

and

$$\sigma^2 = \frac{\sigma_1^2\sigma_2^2 - \sigma_{\text{line}}^4}{\sqrt{\sigma_2^4 + \sigma_{\text{line}}^4}} \,, \qquad \bar{a} = a_1 + a_2\frac{\sigma_{\text{line}}^2[\sigma_1^2 + \sigma_2^2]}{\sigma_2^4 + \sigma_{\text{line}}^4} \tag{10}$$

The normalisation constant $\mathcal{N}$ is a lengthy linear combination of error functions of different arguments, which is omitted here for the sake of brevity.

Using the stationary PDF and the conditional PDF $w(\ell_1', \ell_2', t'|\ell_1, \ell_2, t)$ for the transition from the state ℓ_1, ℓ_2 at time t to the state ℓ_1', ℓ_2' at time t' which is a solution of the same FP equation with the initial condition $w|_{t'=t} = \delta(\ell_1' - \ell_1)\delta(\ell_2' - \ell_2)$ we define the amount of the dropped traffic per unit time by the

following expression (which should be considered as an estimation to the loss rate):

$$
R_{\text{loss}} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left(\int_1^\infty d\ell_1' \int_{-\infty}^\infty d\ell_2' \, (\ell_1' - 1) + \int_1^\infty d\ell_2' \int_{-\infty}^\infty d\ell_1' \, (\ell_2' - 1) \right)
$$
$$
\times \int_0^1 d\ell_1 \int_0^1 d\ell_2 \, w(\ell_1', \ell_2', t + \Delta t | \ell_1, \ell_2, t) p^{(s)}(\ell_1, \ell_2) \tag{11}
$$

We only present here the final expression for the packet loss rate in the case when the capacity of the transmission line between the buffers is equal to the switching capacity available to the second buffer, $a'' = a_{\text{line}}$. In this case $a_2 = 0$ and the PDF (8) takes the form:

$$
p^{(s)}_{a_2=0}(\ell_1, \ell_2) = \mathcal{N}_{a_2=0}^{-1} \exp\left[-\frac{b}{\sigma^2} \left(\lambda_1 - \frac{a_1}{b} \right)^2 \right] \tag{12}
$$

The normalisation constant $\mathcal{N}_{a_2=0}$ is now determined by the following relation

$$
\mathcal{N}_{a_2=0} = \frac{1}{\beta \sin 2\varphi} \left(e^{-\xi_{10}^2} - e^{-\xi_{20}^2} - e^{-\xi_{30}^2} + e^{-\xi_{40}^2} \right.
$$
$$
\left. + \sqrt{\pi} \left[\xi_{10} \text{erf}(\xi_{10}) - \xi_{20} \text{erf}(\xi_{20}) - \xi_{30} \text{erf}(\xi_{30}) + \xi_{40} \text{erf}(\xi_{40}) \right] \right) \tag{13}
$$

where

$$
\xi_{10} = -\frac{\alpha}{\sqrt{\beta}} \qquad\qquad \xi_{20} = \sqrt{\beta} \sin\varphi - \frac{\alpha}{\sqrt{\beta}} \quad \xi_{30} = \sqrt{\beta} \cos\varphi - \frac{\alpha}{\sqrt{\beta}}
$$
$$
\xi_{40} = \sqrt{\beta}(\sin\varphi + \cos\varphi) - \frac{\alpha}{\sqrt{\beta}} \qquad \alpha = \frac{a_1}{\sigma^2} \qquad\qquad \beta = \frac{b}{\sigma^2}
$$

For the loss rate we obtain

$$
R_{\text{loss}}^{a_2=0} = \frac{1}{8\mathcal{N}_{a_2=0}} \sqrt{\frac{\pi}{\beta}} \left\{ \frac{\sigma_1^2}{\sin\varphi} \left[\text{erf}(\xi_{40}) - \text{erf}(\xi_{30}) \right] + \frac{\sigma_2^2}{\cos\varphi} \left[\text{erf}(\xi_{40}) - \text{erf}(\xi_{20}) \right] \right\} \tag{14}
$$

The behaviour of the loss rate $R_{\text{loss}}^{a_2=0}$ is illustrated in Fig. 2.

It can be noticed that the loss rate is significantly greater than zero even below the naive capacity threshold $a' = 1$, especially in the presence of a routing protocol insensitive to congestion, or when the arriving traffic variance is relatively large.

A similar analysis for n interacting queues connected to the same switching device (see Fig. 3) has been completed. The Fokker-Planck equation for the PDF
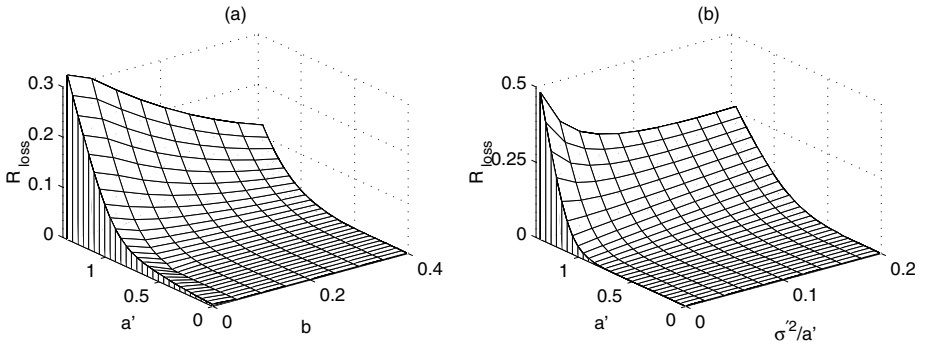
Fig. 2. Packet loss rate R_{loss} (in normalised units of buffer size $\ell_{\max}$ per unit time) plotted against traffic arrival rate a' (with a'' equal to the interconnecting line capacity $a_{\text{line}} = 1$). In subplot (a) this is also plotted against the protocol sensitivity b for a constant ratio of arriving traffic variance to the mean rate $\sigma'^2/a' = 0.1$, whereas in subplot (b) this is also plotted against the ratio σ'^2/a', for a constant protocol sensitivity $b = 0.1$. The theoretical capacity of this subsystem of queues is at $a'(= a'') = 1$.
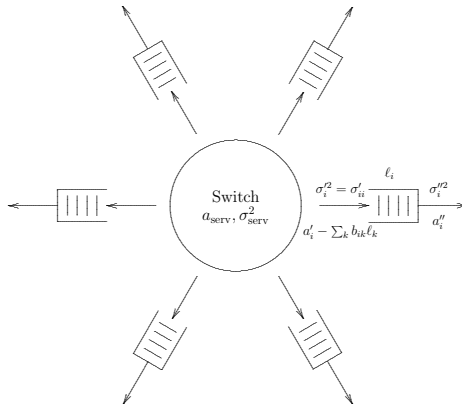

Fig. 3. Subsystem of n queues connected to the same switching device.

of this subsystem, $p = p(\ell_1, \ldots, \ell_n; t)$, is

$$\partial_t p = -\sum_i \frac{\partial}{\partial x_i}\left[a_i - \sum_k b_{ik}\ell_k\right]p + \frac{1}{2}\sum_{i,k}\frac{\partial^2}{\partial x_i \partial x_k}\sigma_{ik}p \qquad (15)$$

where

$$a_i = a_i' - a_i'' , \qquad \sigma_{ik} = \sigma_{ik}' + \delta_{ik}\sigma_i''^2 \qquad (16)$$

b_{ik} is the set of protocol sensitivities similar to the ones described earlier, a_{serv}, σ^2_{serv} is the mean value and variance (per unit time) of the overall traffic coming from the central switching device, $a'_i - \sum_k b_{ik}\ell_k$ is the set of mean values of traffic coming towards of individual queues, σ'_{ik} is the corresponding covariance matrix, a''_i, σ''^2_i the mean value and variance of switching capacities available at the other ends of individual queues. The equilibrium solution has the following form:

$$p^{(s)}(\ell_1, \dots, \ell_n) = \mathcal{N}^{-1} \exp\left[2\sum_{ik}(\sigma^{-1})_{ik}\ell_i a_k - \sum_{ikj}(\sigma^{-1})_{ij}b_{jk}\ell_i\ell_k \right] \qquad (17)$$

where $\mathcal{N}$ is the normalisation constant and the following set of relations (due to a detailed balance condition) should be imposed on the protocol sensitivities b_{ik}:

$$\sum_j b_{ij}\sigma_{jk} = \sum_j b_{kj}\sigma_{ji} \qquad (18)$$

A general solution to (18) allows a large number of free parameters. In order to arrive at a reasonable number of those we impose the following restrictions:

$$a'_i = a_{\mathrm{serv}}\frac{a''_i}{\sum_k a''_k} \ , \quad \sigma'_{ii} = \sigma'^2_i = \sigma'^2 \ \forall i \ , \qquad \rho_{ik} \equiv \frac{\sigma'_{ik}}{\sigma'_i\sigma'_k} = \rho \ i \neq k \qquad (19)$$

The first condition means that the mean value of traffic coming toward an individual queue is proportional to the switching capacity available to the corresponding queue (and vice versa), the second condition means that variances of traffic coming to all the queues are the same, and the third one is the statement that the correlation coefficients of any two pairs of incoming queue-traffic are equal as well, in order to maintain symmetry (this implicitly assumes that the system is homogeneous). In addition to this we are looking for a solution in the following class (the sensitivities of the idealised routing protocol for one particular queue to congestion on all other queues on the same switch are the same):

$$b_{ii} = \bar{b}_i \ \forall i \ , \qquad b_{ki} = \tilde{b}_i \ , \ i \neq k \qquad (20)$$

Then we have the following parameterisation for σ_{ik}, b_{ik}:

$$b_{ik} = \begin{cases} \frac{b}{\sigma^2_{\mathrm{serv}}/n + \sigma''^2_i} & i = k \\ -\frac{1}{n-1}\frac{b}{\sigma^2_{\mathrm{serv}}/n + \sigma''^2_i} & i \neq k \end{cases} \ , \quad \sigma_{ik} = \begin{cases} \sigma'^2 + \sigma''^2_i & i = k \\ \rho\sigma'^2 & i \neq k \end{cases} \qquad (21)$$

where

$$\rho = -\frac{1}{n-1} + \frac{1}{n(n-1)}\frac{\sigma^2_{\mathrm{serv}}}{\sigma'^2} \qquad (22)$$

Summarising free parameters, we have: $a_i'', \sigma_i''^2$ are the mean value and variance per unit time of the switching capacities available to each queue, $a_{\mathrm{serv}}, \sigma_{\mathrm{serv}}^2$ are the mean value and variance per unit time of the overall traffic coming from the central switching fabric, σ'^2 (which is variance of the traffic coming to each queue) and b are actually characteristics of the routing protocol. The loss rate for this idealised router subsystem is defined an analogous fashion to (10) and has been computed explicitly. For reasons of economy of space, we omit the results from this paper, but will present them at the conference.

4   Preliminary Conclusions

We have presented a theoretical framework that can be used to model the interaction between a cascade of queues in a network. The theoretical model is capable of quantitative predictions of system throughput, loss rate, end-to-end delay, etc. Here we have presented the packet loss rate for a subsystem of two strongly correlated queues.

In the presence of stochastic packet lengths and inter-arrival times, we characterise the system in terms of mean arrival bit-rates identical to those of conventional queueing theory. However, here we also employ second-order statistics for the traffic, namely the variance of the arrival bit-rates. The latter parameters can be obtained either by observation on large-scale real networks, or as part of a broader model for an entire network.

As we can see from Fig. 2, the packet loss rate is non-zero for $a' < 1$ (i. e. arrival rates less than the system capacity) due to the presence of uncertainty in the arrival rate embodied in the variance terms. The packet loss rate can also be seen to reduce in the presence of a routing protocol that is sensitive to congestion. This type of analysis is clearly more useful than the conventional mean rate analysis (leaky bucket calculations) typically used in the first-order design of networks.

Finally, this methodology can be extended to a larger number of interacting queues in a straight-forward manner using functional integral methods. Such subsystems can be incorporated directly into broader network models (which need the various PDF's as input) in order to pursue the goal of arriving at a mathematical theory for dimensioning large-scale, core, packet-switched networks. Work on the detailed methodology behind such broader network models is at hand and will be presented in a series of papers in the future.

References

1. J. Cao, W. S. Cleveland, D. Lin, D. X. Sun, *The Effect of Statistical Multiplexing on the Long Range Dependence of Internet Packet Traffic*, Bell Labs Tech. Report (2002);
2. J. Cao, W. S. Cleveland, D. Lin, D. X. Sun, *Internet Traffic Tends To Poisson and Independent as the Load Increases*, http://citeseer.nj.nec.com/426819.html;
3. A. Stepanenko, C. C. Constantinou, T. N. Arvanitis and K. Baughan, *On the Statistical Properties of Core Network Internet Traffic*, submitted to Communications Letters (2001);
4. M. Schwartz: *Telecommunication Networks, Protocols, Modeling and Analysis*, Addison-Wesley (1987);
5. F. Kelly, *Loss networks,* Ann. Appl. Probab., **1**, pp. 319–378 (1991);
6. C. Graham and S. Méléard, *Chaos hypothesis for a system interacting through shared resources,* Probability Theory and Related Fields, **100**, pp.157–173 (1994);
7. N. D. Vvedenskaya, R. L. Dobrushin, and F. I. Karpelevich, *A queueing system with a choice of the shorter of two queues — an asymptotic approach,* Problems Inform. Transmission, **32**, pp. 15–27 (1996);
8. F. Delcoigne, G. Fayolle, *Thermodynamical limit and propagation of chaos in polling systems,* Markov Processes and Related Fields, **5**, pp. 89–124 (1999);
9. Guy Fayolle, Arnaud de La Fortelle, Jean-Marc Lasgouttes, Laurent Massouli'e, James Roberts, *Best-effort networks: modeling and performance analysis via large networks asymptotics*, IEEE INFOCOM 2001;
10. H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications*, 2nd Edition, Springer Series in Synergetics, Springer-Verlag (1989);