ITC19/ Performance Challenges for Efficient Next Generation Networks 1275-1284
LIANG X.J. and XIN Z.H.(Editors)
V.B. IVERSEN and KUO G.S.(Editors)
Beijing University of Posts and Telecommunications Press

Resource Sharing Models for Quality-of-Service with Multi-Rate Traffic

Torgny Holmberg and Villy B. Iversen

Research Center COM
Technical University of Denmark, DTU
Building 345V
DK-2800 Kgs. Lyngby
E-mail: {th, vbi}@com.dtu.dk.

Abstract. Dimensioning of data packet systems such as the Internet is a notoriously difficult problem where the previously reliable classical telecommunication theories based on statistical methods often have failed. Alternative methods such as simulations, min-max bandwidth allocation, elastic traffic descriptions and network calculus, to mention some, have thus become the alternative approach thought they are not as descriptive as the stochastic methods. However, recent studies have shown that bandwidth sharing transport protocols may allow that the given link can be well modelled as a Processor Sharing node, and furthermore, the sources and link may be viewed upon as a closed-queueing network. Luckily, such systems are well covered and understood and can be described in a statistical sense. At least for single-rate traffic.
In this paper, we analyse multi-rate traffic sources in the described scenario. We will prove that the insensitivity property of single-rate traffic do not apply for the multi-rate traffic. Fortunately, although we cannot achieve insensitivity, we will show by means of simulations that the assumption of exponential idle and service times is a surprisingly good approximation of a system with arbitrary distributions of the ON and OFF times.
Keywords: multi-rate traffic, processor sharing, closed network, insensitivity, feedback.

1 Introduction

The Internet is constantly evolving, introducing new applications almost on a daily basis, introducing new protocols and modifications of existing protocols, and introducing new network elements which change the characteristics of the network traffic. Complex services such as the World Wide Web (WWW), providing the possibility of experiencing a large set of media objects, and new services with a diverse set of Quality of Service (QoS) requirements put great demands on the existing network elements as well as service providers. Thus, the dimensioning methods need also be kept up to date in order to provide the increasing demand for QoS.

In this paper, we provide means for modelling and dimensioning of Internet access networks carrying heterogenous elastic multi-rate traffic with throughput requirements. The motivation for this study are the verifications that, see below, Transmission Control Protocol (TCP) and other feedback protocols can be well modelled by a Processor Sharing (PS) server due to their capacity sharing property and ability to adapt to the transmission rate of the access trunk line. The assumption is though that the feedback procedure of such transport protocols is close to ideal, i.e., instantaneous feedback. These positive verifications are explained by the lesser need of having large buffers in the network nodes imposing a relatively small buffer delay thus enabling bufferless processor sharing models [2]. Any delay imposed by the PS node can thus be interpreted as delay imposed by the multiplexer.

The modelling and analysis presented in this paper is based on classical telecommunication theory, more notably closed-queueing networks in the form of the Palm machine repair model with several PS servers. For such model, we show that the probability state distribution is *almost* insensitive towards the active (ON-period) and idle phase (OFF-period) distributions of the ON/OFF sources. The almost insensitivity property enables far less complex analysis of the targeted system and we may derive simple expressions on performance.

1.1 Related Work

In [4], the authors present methods for analysing feedback based protocols in the context of Web-traffic. By modelling the access network as a closed-queueing network, they can apply Engset model and they further show that this system model is insensitive, i.e., the distribution of the sources' active and idle phases only influence the performance by their respective means. In the saturated scenario when the number of active flows exceed the link capacity, the TCP bandwidth sharing features on the target link, as well as the use of fair queueing, are modelled as a Processor Sharing (PS) node. Furthermore, the authors also consider the loss of capacity due to packet loss in the saturated link buffer where the available link capacity is lowered by a factor which depends on the given queue management scheme.

In [8], Nabe et al. make a fairly elaborate characterization of Internet traffic extracted from various traces and adopt the observed service rate to the $M/G/1-PS$ queueing model for defining a dimensioning methodology. Their assumptions are that the service rate is state dependent and that efficient sharing of the link is well performed by the underlying TCP protocol. It is observed that the transmission time of web documents generally is heavy tailed and well modelled by the log-normal distribution. The expected conditional waiting time is analytically tractable and the authors show that the trace-driven simulation coincides well with analysis.

In [10], Roberts presents a survey on statistical bandwidth sharing, mainly based on his and his frequent co-authors results. Roberts model the network as a Whittle network (a Jackson network with PS nodes and service rate restrictions) in which the nodes implement the PS discipline. Unlike in above references, Roberts also model the thinking phase as a PS node with rate equal to the number of flows in the thinking phase. However, in practice this makes a little difference as both presentations result in product form solutions of the state probabilities.

In [3], the authors investigate the single server processor-sharing system with Poissonian arrivals and with a general service time distribution serving M heterogeneous job classes. Each customer class receives service in proportion to its weight and the authors calculate the average response time for each class. The derived expressions are non-trivial, even for exponentially distributed service times. Furthermore, the authors show that the average response time depends on the service time distribution. The conclusion is thus that multi-rate traffic impose great complexity in the analysis.

See also [1] and [2] for further information and references.

1.2 Outline of Paper

In Section 2, we recall the concept of closed-queueing networks and put it in a historical context and also describe the network model used in this paper. In Section 3, the proof of the sensitivity property of the considered network is presented, followed by simulation verifications in Section 4. Finally, in Section 5 we conclude our results and observations.

2 Closed-Queueing Network Model

The closed-queue network with limited number of customers, shown in Figure 2.1, was first used for modelling and analysis of machine-repair problems. The first paper published on this topic was by B. W. Gnedenko in 1934 in which the reliability of the machines used in textile manufacturing was considered. This closed-queueing network was later revived in 1947 by C. Palm [9] and has thereafter often been referred to as the Palm machine repair model.

Later, with the introduction of computing and communicating devices, the network has been used to model a finite pool of clients (users) which need to execute (transmit) data on a shared resource. In a computing time-sharing system, the users alternate between their thinking phase and their waiting phase. The thinking phase is modelled by an Infinite Server (IS) ($M/G/\infty$) node which represents the time when the user reflects over or formulates the problem to be executed. Traditionally, the thinking time has often been assumed to be exponentially distributed. In the waiting phase, the execution unit serves the user requests. A new request which finds the execution unit busy may, depending on service discipline, need to be buffered in a queue awaiting service. This problem is well covered in [7].

In a communication environment, the queueing network may be used to model a communication node where the thinking phase may represent the idle time between two consecutive session activation times (e.g., time between two file transfers) and the delay in the remaining network transmission components. The transmission phase is modelled by a wait system with an appropriate service discipline.

The success of the Palm machine-repair model relies on the fact that it can be generalised to cover non-exponential thinking and service time phases (with certain restrictions) yet providing a product form state probability.

In this paper, we generalise above network by allowing several PS servers and multi-rate traffic sources, i.e., sources may request an integral number of units of system capacity, and the aim of this paper is to show that this more general system shows almost insensitivity.

2.1 System Parameters

The assumptions enabling the analysis in the following sections are based on the following. The IS node represents the ON/OFF sources which are fixed in number. Each source is independent of any other source and the duration of the ON period is drawn from a given distribution of arbitrary type, as is the OFF period. Thus, we study an alternating renewal process. The PS node represents the targeted network node's output port and buffer.

This closed-queueing network model is very general as it allows ON and OFF periods to model the elastic flow activity on several time scales. At the lowest level, the ON period can represent the transmission of a data packet and the OFF period as the idle time between two consecutive packets. On a higher time scale, the ON period could be the download of a single file and the OFF period the idle time between file transmissions. At the next higher level, a burst of files with an inter arrival time shorter than a certain limit can be be interpreted to be logically connected and may be handled accordingly, hence, the ON period may represent the length of the flow duration. The OFF period thus models the logical separation of the packet bursts. Finally, at the highest level, the ON period could model the duration of a session, e.g., a user activity with a duration of a connection life time. The OFF period would then correspond to the user changing activity or application. Hereafter, we will view the system on the second lowest time scale and view the ON periods as file transmissions.

The shared resource, e.g., the access trunk line or network bottleneck link has S servers each of capacity C yielding a total capacity $S \cdot C$. Alternatively, we may model the shared resource as a PS node with a single server of capacity $S \cdot C$ which often is seen in the literature.

Furthermore, it will be convenient to define a class of sources as the set of sources sharing the same mean idle time, mean intended holding time and multi-rate demand. Such partitioning is of course easier in theory, but may in practice correspond to, e.g., different applications showing the same or sufficiently similar behaviour. The number of classes will be denoted as M. The number of, e.g., sources of a class i will be indexed as N_i and where $\sum_{i=1}^{M} N_i = N$.

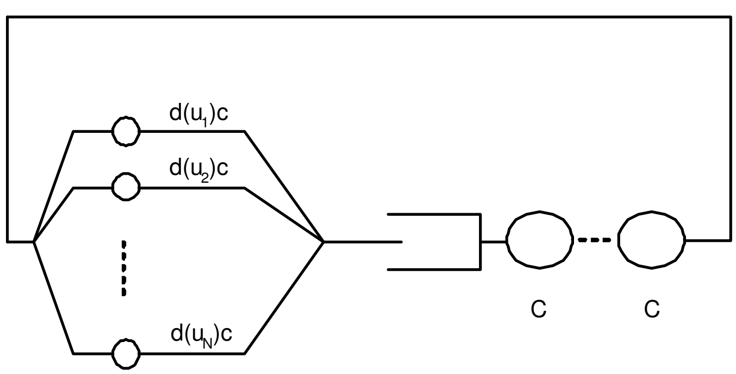

Fig. 1. The Palm machine repair model.

3 Almost Insensitivity

In this section we will derive the system state probability[1]. By defining the state as the set of active sources, we can easily obtain important performance measures as presented in Section 4. We start by calculating the system service rate in case of no overload and overload respectively which will better motivate the derivation of probability state distribution below. In case of no overload, the service rate of source u, $\mu(u)$, with rate demand $d(u) \cdot c$ when downloading a file of size $h(u)$ is

$$m(u) = \frac{1}{\mu(u)} = \frac{h(u)}{d(u)c} = \frac{1}{\frac{d(u)c}{h(u)}} \Rightarrow \mu(u) = \frac{d(u)c}{h(u)}.$$

In the case of overload, i.e., when $\sum_{v \in j} d(v)c > S \cdot C$ we get the scaled service rate

$$m(u)' = \frac{1}{\mu(u)'} = \frac{h(u)}{\frac{d(u)}{\sum_{v \in j} d(v)} S \cdot C} \Rightarrow \mu(u)' = \frac{S}{\sum_{v \in j} d(v)} \mu(u).$$

where j is the set of active sources. Hereafter, we will assume that $C = c$ and the number of servers will then be equal to the maximum number of capacity units that may be utilised by the multi-rate sources.

Proposition 31 *In the case of exponential idle and service times, the above multi-rate system has the probability state distribution*

$$\pi(j) = c\, \frac{g(j)}{S^{|j|}} \prod_{u \in j} \gamma(u), \qquad \gamma(u) = \lambda(u)/\mu(u) \tag{1}$$

where j is the set of active sources, c is the normalization constant and $g(j)$ is a non-trivial function which depends on the given state j and neighbouring states.

Proof. The proof is omitted due to lack of space. It is available in the full length paper.

Thus, we see that the given system is complex and that the state probabilities are not easily obtained. The probability state distribution is thus found by either solving the flow equations directly or by finding the function g with its constraints, which is likely to be at least as difficult though it requires fewer equations to be solved. Consequently, the larger number of PS servers and the lower multi-slot requirements, the fewer equations.

Having found the state distribution for the observed system in the case of exponentially distributed idle and service times, we will now show that we will not obtain insensitivity for the multi-rate system using the given state distribution. This is perhaps easiest shown by finding a certain arrival or service process for which the state distribution is not valid. We will do this the hard way by showing that the Coxian distribution fails to provide insensitivity.

[1] The state distribution with an alternative state representation is found in the extended paper which is available upon request.

Lemma 32 *Considering the above multi-rate system, the probability state distribution given in Proposition 31 is not valid in the case of Coxian idle and service times, where*

$$\frac{1}{l(u)} = \lambda(u) = \sum_{i=1}^{r(u)} l_i(u), \qquad l_i(u) = \frac{\alpha_i(u)}{\lambda_i(u)},\ \alpha_i(u) = a_0(u) \cdot a_1(u) \cdots a_{i-1}(u)$$
$$\frac{1}{m(u)} = \mu(u) = \sum_{i=1}^{s(u)} m_i(u), \qquad m_i(u) = \frac{\beta_i(u)}{\mu_i(u)},\ \beta_i(u) = b_0(u) \cdot b_1(u) \cdots b_{i-1}(u).$$

Proof. (Sketch of proof. See extended version for full proof.) Define the state

$$[j, X, Y] = [(v_1, v_2, ..., v_n), (x(u_1), ..., x(u_m)), (y(v_1), ..., y(v_n))], \tag{2}$$

$m + n = M$, $1 \leq x(u_i) \leq r(u_i)$, $1 \leq y(v_i) \leq s(v_i)$. We will thus show that the joint probability distribution

$$P(j, X, Y) = \prod_{u \notin j} \frac{l_{x(u)}(u)}{l(u)} \prod_{v \in j} \frac{m_{y(v)}(v)}{m(v)} \pi(j) \tag{3}$$

does not fulfill the flow equations. Further, since we allow irregular point-processes in the case of $a_0(u) < 1$ or $b_0(u) < 1$, we define $p(u)$ and $q(u)$ to be the probability of a successful call arrival and call termination, respectively [6]. Setting up the equations for the output and input flow (details in the extended paper) and dividing both sides with $P(j, X, Y)$, and using that

$$p(u) \cdot \frac{a_0(u)}{b_0(u)} = \frac{b_0(u)}{a_0(u) + b_0(u) - a_0(u)b_0(u)} \frac{a_0(u)}{b_0(u)} = \frac{a_0(u)}{a_0(u) + b_0(u) - a_0(u)b_0(u)} = q(u), \tag{4}$$

we get the following equations:
Case 1:
Considering a state j which is included in the set of states for which processor sharing is not active, we get that

$$\text{Output flow} = \sum_{u \notin j} \lambda_{x(u)}(u) + \sum_{v \in j} \mu_{y(v)}(v)$$

$$\begin{aligned}\text{Input flow} = &\sum_{u \notin j} \lambda_{x(u)}(u) + \sum_{v \in j} \mu_{y(v)}(v) - \\ &- \sum_{v \in j} \mu_{y(v)}(v) q(v) \left(1 - \tfrac{g(j \setminus \{v\}) S}{g(j)}\right) \cdot 1[y(v) = 1] - \\ &- \sum_{u \notin j} \lambda_{x(u)}(u) p(u) \Big(1 - W\Big) \cdot 1[x(u) = 1]\end{aligned}$$

where

$$W = \left(\tfrac{g(j \cup \{u\})}{g(j) S} \cdot 1[j \cup \{u\} \in G_1] + \tfrac{g(j \cup \{u\})}{g(j)} \tfrac{1}{\sum_{w \in j} d(w) + d(u)} \cdot 1[j \cup \{u\} \in G_2] \right). \tag{5}$$

In the proof, we have divided the state space into the sets G_1 and G_2. If for a state j the required resources is less or equal to the system capacity, then $j \in G_1$, otherwise $j \in G_2$.

from which we get the time average of the carried traffic of the source, $A_c(v)$, as well as of the total system, A_c,

$$A_c(v) = \sum_j A_c(v|j)\pi(j), \qquad A_c = \sum_{v \in J} A_c(v). \tag{7}$$

In case of $S \cdot C \geq \sum_{v \in j} d(v)$, the carried traffic is equal to the offered traffic and the virtual waiting time is zero. Since the observed system will carry the offered traffic as the system serves the traffic at a lower rate rather than blocking the traffic, no traffic will be lost though it will take longer time to transmit, but the offered traffic will be carried. From this fact, we can calculate the virtual waiting time, $w(v)$, and actual service time, $x(v)$, from the time average carried traffic. The virtual waiting time is easily obtained from

$$m(v)d(v) = \Big(l(v) + m(v) + w(v)\Big)\, A_c(v) \Rightarrow w(v) = m(v)\left(\tfrac{d(v)}{A_c(v)} - 1\right) - l(v).$$

Average actual service time is the sum of the virtual waiting time and the intended service time

$$x(v) = w(v) + m(v).$$

4.1 Simulation Results and Numerical Examples

In this section, we will compare carried traffic and virtual waiting time between the Engset $M/M/S-PS$, for which we have the exact solution, and the Engset system with heavy tailed idle and service time distributions. In particular, we will consider the log-normal distribution for which we can define arbitrary positive mean and standard deviation, denoted σ_i and σ_s for the arrival and service distribution respectively. In Table 1, the carried traffic for the three traffic classes, $A_{c,j}$ $j = 1, 2, 3$, and the virtual waiting time are presented. In the table, we see that the carried traffic and the virtual waiting time correspond surprisingly well with the theoretical Markov/Markov model. Even for the traffic scenario where both the arrival and service process are very heavy-tailed with respective squared coefficient of variation $C^2 \geq 10^8$. In the second simulation setup, two classes of sources

	$A_{c,1}$	$A_{c,2}$	$A_{c,3}$	w_1	w_2	w_3
M/M/S-PS	0.2878	0.6644	3.7609	0.2371	0.1700	0.5279
M/log-n/S-PS, $\sigma_s = 1$	0.2881	0.6654	3.7637	0.2354	0.1684	0.5266
M/log-n/S-PS, $\sigma_s = 100$	0.2881	0.6657	3.7649	0.2352	0.1682	0.5256
M/log-n/S-PS, $\sigma_s = 10000$	0.2706	0.6321	3.7372	0.2127	0.1511	0.4883

Table 1. Performance results of simulations with parameters $N = (2, 2, 2)$, $\lambda = (1, 2, 3)$, $\mu = (2, 3, 1)$, $d = (1, 2, 7)$, $S = 10$. (Poor significance in simulation results with $\sigma_s = 10000$.)

are simulated where each class contains 10 sources. The idle time and service time distributions are both log-normal with standard deviations taking values in $\{1, 10, 100\}$. In Table 2, the results from a number of simulations with varying combinations of standard deviations of the idle and service distributions are presented. From the results, it is seen that the carried traffic and consequently the virtual waiting time coincide surprisingly well with the theoretical results obtained for the Engset $M/M/S-PS$ system.

Processor Sharing node we can view the system as a closed-queueing network, and more precisely, as the Palm machine repair model.

For the given system, we have derived important performance measures suitable for calculating entities such as average virtual waiting time and average throughput. The derived expressions of performance are general, but we can only perform the calculations under the assumption of exponentially distributed ON and OFF times of the multi-rate sources since the investigated system does not possess the insensitivity property which also is proved in this paper.

Fortunately, we have shown by means of simulations, that the Markovian assumption is surprisingly accurate also in the general case, i.e., allowing the sources' ON and OFF times to be of arbitrary distributions. Thus, we have a system which is almost insensitive. Furthermore, from the simulations, we have seen that the broadband traffic gets, at the expense of the low-rate traffic, a very large share of the shared resource and thus gets a very high throughput and low waiting times compared to the low-rate traffic. This phenomenon is explained by the arrival theorem.

In this paper, we have not analytically evaluated the accuracy of the Markovian assumption to the general case. By finding a good measure on how well the approximation is, if such exists, we would be able to accept or reject the presented approximation. This problem we leave for further studies.

References

1. A. W Berger and Y. Kogan. Dimensioning bandwidth for elastic traffic in high-speed data networks. *IEEE/ACM Transactions on Networking*, 8(5):643–654, 2000.
2. J. Charzinski. Fun factor dimensioning for elastic traffic. *Proceedings of the ITC Specialist Seminar on Internet Traffic Measurement*, pages 37/1–37/9, Sept. 2000.
3. G. Fayolle, I. Mitrani, and R. Iasnogorodski. Sharing a processor among many job classes. *J. ACM*, 27(3):519–532, 1980.
4. D. P. Heyman, T. V. Lakshman, and A. L. Neidhardt. A new method for analyzing feedback-based protocols with applications to engineering Web traffic over the Internet. *IEEE Computer Communications*, 26(8):785–803, 2003.
5. A. Hordijk and R. Schassberger. Weak convergence of generalized semi-markov processes. Technical Report 77–8, Institute of Applied Mathematics and Computer Science, University of Leiden, The Netherlands, Feb. 1979.
6. V. B. Iversen. The A-Formula. *Teleteknik*, 23(2):64–79, 1980.
7. V. B. Iversen. *Teletraffic Engineering Handbook*. ITU-D & ITC, Available online at http://www.com.dtu.dk/education/34340, 2004.
8. M. Nabe, M. Murata, and H. Miyahara. Analysis and modeling of World Wide Web traffic for capacity dimensioning of Internet access lines. *Performance Evaluation*, 34(4):249–271, 1998.
9. C. Palm. The assignment of workers in servicing automatic machines. *Journal of Industrial Engineering*, 9:28–42, 1958. First published in Swedish in 1947.
10. J. W. Roberts. A survey on statistical bandwidth sharing. *Computer Networks*, 45(3):319–332, 2004.