VU Research Portal

A simple performability estimate for Jackson networks with an unreliable output channel

van Dijk, N.M.

1989

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van Dijk, N. M. (1989). *A simple performability estimate for Jackson networks with an unreliable output channel*. (Serie Research Memoranda; No. 1989-32). Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

SERIE RESEARCH MEMORANDA

A SIMPLE PERFORMABILITY ESTIMATE FOR JACKSON NETWORKS WITH AN UNRELIABLE OUTPUT CHANNEL

N.M. van Dijk

Research Memorandum 1989-32

July 1989

**VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
EN ECONOMETRIE
AMSTERDAM**

A simple performability estimate for Jackson networks
with an unreliable output channel

Nico. M. van Dijk
Free University, the Netherlands

Abstract This paper provides a simple performance estimate with an explicit error bound for open Jackson networks with departure blocking such as due to an output channel with breakdowns.

Keywords Jackson network * performability * breakdown * throughput * estimate *

1 Introduction

At present reliability aspects in computer performance evaluation receive considerable attention under the name of performability (cf. [2], [4], [8], [9]). Here one may typically think of breakdowns, error detection and fault-tolerancy (cf. [6]).

As such features are most complicated to analyze analytically, both numerical and approximate, performance evaluation techniques are extensively studied (cf. [8], [9]). Usually, these are computationally most expensive. Particularly, reduction methods have therefore been developed. Most notably, in an elegant paper [9], results on decomposability as developed in [2] and [3] were applied and extended to approximate system availability in repairable multi-computer systems. These results however still require a fair amount of computational effort while in engineering situations one may just be interested in a quick rough estimate. In [16] a simple bound has been provided for a single station with breakdowns. For networks, however, simple estimates or bounds do not seem to be available.

This paper studies open Jackson networks, such as representing an interconnected computer network or a circuit- or packet switching communication system, for which system departures can be blocked, for instance due to a common output device which is subject to breakdowns, an error detection or a saturated next stage in an integrated network.

The prime motivation is to study the increase of station workloads due to output blocking. To this end, simple estimates for station workloads or throughputs will be provided along with an explicit error bound of their accuracy. These estimates will not be accurate but may typically prove to be useful in an engineering environment to give one:

- (i) A first indication of order of magnitude
- (ii) Quantitative or qualitative insight.

The proof technique, which is based on Markov reward theory, is of interest in itself and seems promising for extension to more complex performability structures.

2.1 Model

Consider an open Jackson queueing network of N service stations, with jobs routing from one station to another to receive certain amounts of service and jobs arriving from outside according to a Poisson process with parameter λ . A job requires an exponential amount of service at station i with parameter μ_i and station i provides an amount $f_i(n_i)$ of service per unit of time when n_i jobs are present, where $f_i(n_i)$ is assumed to be non-decreasing in n_i as natural and to be bounded for technical convenience (see remark 4 in section 2.3).

An arriving job is assigned station j with probability p_{0j} and upon service completion at station i routes to station j with probability p_{ij} or request to leave the system with probability $p_{i0} = 1 - [p_{i1} + \dots + p_{iN}]$. This request, however, can be blocked depending upon the output status being "up" or "down" as will be described below. When "up" it is granted and the job departs the system. When "down" it is blocked and the job has to stay at station i where it receives a new service.

The "up" and "down" output status is determined by exponential times with parameters v_1 and v_0 , respectively. This may for example reflect a common output device or link that once in a while breaks down and requires repair or, as in packet switching communication networks, an error that arises when information on the job completion is to be stored. Typically, the time fraction $v_1/(v_0 + v_1)$ that the system is down should hereby be thought of as being small.

The above system does no longer exhibit a simple closed form expression such as Jackson's celebrated product form when blocking wouldn't occur ($v_1=0$). Only for the special case of a single service facility subject to this type of breakdown (independent case), a closed form expression for the generating fraction of the queue length has been reported (cf [7], p. 103).

2.2 Simple performance estimate

Performance measure of interest We aim to evaluate the effect of output breakdowns on stations with $p_{i0} > 0$ and denote by L_i the mean number of service completions per unit of time (workload or throughput) at station i .

Modification As the quantity doesn't seem to have a simple closed form expression we propose an estimate $\hat{L}_i$ by modifying the system such that departures are never blocked. This is naturally attained by assuming that the output states can never become down, that is $v_1=0$, but for a technical purpose in the next section we model this as follows.

Special interpretation Consider an open Jackson queueing network as described of which the output status can also alternatively be "up" and "down" for exponential times with parameters v_1 and v_0 respectively, but where the output status has no impact on the queueing network. More precisely that is, when a job completes a service station i it always leaves the system with probability $p_{i0} = 1 - \sum_{j \neq 0} p_{ij}$, regardless of the "up" or "down" status.

Estimates Clearly, under the above interpretation the behaviour of the population vector $\vec{n} = (n_1, \dots, n_N)$ describing the number of jobs n_i at all stations i , is stochastically equivalent to that of the standard Jackson network when $v_1=0$. As a result, from Jackson's celebrated product form we thus conclude

$$\bar{L}_l = \frac{\sum_{k=0}^{\infty} \mu_l f_l(k) (\lambda_l / \mu_l)^k \left[\prod_{j=1}^k f_l(j) \right]^{-1}}{\sum_{k=0}^{\infty} (\lambda_l / \mu_l)^k \left[\prod_{j=1}^k f_l(j) \right]^{-1}}$$

where λ_l is uniquely determined, up to normalization, by

$$\lambda_j = \lambda p_{0j} + \sum_{i=1}^N \lambda p_{ij} \quad (j=1, \dots, N).$$

A comparison of the value L_l of actual interest and this approximate value $\bar{L}_l$ is this to be made. Most notably, we aim to find an error bound for

$$|\bar{L}_l - L_l|$$

Remark Clearly, also other performance measures can be of interest, for which similar approaches as following herein can be employed. Some direct consequences of the one studied above are the following

- (i) When station l has an infinite number of servers, then by Little's law we immediately obtain an estimate for the mean station length.
- (ii) When the stations can be ordered such that feedback among stations j with $p_{j0} = 0$ is not possible, the throughput can be estimated at these stations also by setting

$$\bar{L}_j = \lambda p_{0j} + \sum_i \bar{L}_i p_{ij}$$

where $\bar{L}_i$ for $p_{i0} > 0$ is the estimate given by (i) for $l = i$.

- (iii) If all station lengths can be estimated (e.g. as under (i) and either $p_{j0} > 0$ for all j or (ii)), the total system size and thus also (by Little's law again) response time can be estimated. $\square$

2.3 Error bound

For both the original and modified model, let the state of the system be represented by $[\bar{n}, \theta]$ where $\bar{n} = (n_1, \dots, n_N)$ denotes the number of jobs at station i and where $\theta = 1$ or 0 depending upon whether the output status is "up" or "down" respectively. Further, by $\bar{n}$ up to one job more (+) or less (-) at station i and we read $1_{\{A\}}$ for an indicator of an event A , i.e. $1_{\{A\}} = 1$ if A is satisfied and $1_{\{A\}} = 0$ otherwise.

Throughout, we use an upper bar symbol "-" for expressions corresponding to the modified model 1 and the symbol "(-)" to indicate that the expression is to be read both for the original and modified model.

By virtue of the above interpretation of the modified model both the original and modified model then constitute a continuous time irreducible Markov chain at the same set of reachable states $R = \bar{R} = \{[\bar{n}, \theta] \mid n_i \geq 0, i=1, \dots, N, \theta = 0, 1\}$. Let $(\bar{q})([\bar{n}, \theta], [\bar{n}', \theta'])$ denote the corresponding transition rates for a transition from state $[\bar{n}', \theta']$ in $[\bar{n}, \theta]$. Then we have

$$(\bar{q})([\bar{n}, \theta], [\bar{n}', \theta']) =$$

$$\left\{ \begin{array}{ll} \lambda p_{0j} & \bar{n}' = \bar{n} + e_j \\ 1_{\{\theta=1\}} v_1 & \bar{n}' = \bar{n}, \theta=0, \theta'=1 \\ 1_{\{\theta=0\}} v_0 & \bar{n}' = \bar{n}, \theta=1, \theta'=0 \\ \mu_i f_i(n_i) p_{ij} & \bar{n}' = \bar{n} - e_i + e_j \quad (j \neq 0) \end{array} \right.$$

but

$$\begin{aligned} q([\bar{n}, \theta], [\bar{n}-e_j, \theta]) &= 1_{\{\theta=1\}} \mu_i f_i(n_i) p_{i0} \\ q([\bar{n}, \theta], [\bar{n}-e_j, \theta]) &= \mu_i f_i(n_i) p_{i0} \quad (\theta=0, 1) \end{aligned} \quad (2)$$

Now, let Q be such that for all $[\bar{n}, \theta] \in \bar{R}$

$$\lambda + v_0 + v_1 + \sum_i \mu_i f_i(n_i) \leq Q \quad (3)$$

Then we can apply the standard uniformization technique (cf. [15], p. 110) to transform the continuous-time Markov chains in discrete-time Markov chains and employ inductive Markov reward theory. More precisely, first define the transition probabilities

$$\begin{aligned} & - \quad \bar{p}^{(-)}([\bar{n}, \theta], [\bar{n}', \theta']) = \\ & = \begin{cases} \bar{q}^{(-)}([\bar{n}, \theta], [\bar{n}', \theta'])/Q & [\bar{n}', \theta'] \neq [\bar{n}, \theta] \\ 1 - \sum_{[\bar{n}', \theta'] \neq [\bar{n}, \theta]} \bar{q}^{(-)}([\bar{n}, \theta], [\bar{n}', \theta'])/Q & [\bar{n}', \theta'] = [\bar{n}, \theta] \end{cases} \end{aligned} \quad (4)$$

and next, let the functions $\bar{V}_t^{(-)}(\cdot)$ at $\bar{R}$ be recursively determined by

$$\begin{aligned} & \bar{V}_{t+1}^{(-)}([\bar{n}, \theta]) = \\ & \mu_k f_k(n_k) p_k 0/Q + \sum_{[\bar{n}', \theta']} \bar{p}^{(-)}([\bar{n}, \theta], [\bar{n}', \theta']) V_t([\bar{n}', \theta']) \end{aligned} \quad (5)$$

Then by standard Markov reward arguments (cf. [10]) and the uniformization technique (cf. [15], p.110) we conclude that for any $[\bar{n}, \theta] \in \bar{R}$:

$$\bar{L}_1^{(-)} = \lim_{T \rightarrow \infty} \frac{Q}{T} \bar{V}_T^{(-)}([\bar{n}, \theta]) \quad (6)$$

The next lemma will be essential in the key-theorem.

Lemma 2.1 For $i=1, \dots, N$ all $[\bar{n}+e_i, \theta] \in R - \bar{R}$ and all t , we have

$$0 \leq \Delta_i V_t([\bar{n}, \theta]) = V_t([\bar{n}+e_i, \theta]) - V_t([\bar{n}, \theta]) \leq 1. \quad (7)$$

Proof We will apply induction to t . Clearly, (7) holds for $t=0$ as $V_0(\cdot) = 0$. Suppose that (7) holds for $t \leq m$. Then by (2) and (5) and substituting $h = 1/Q$ we obtain for $t = m+1$:

$$\Delta_i \tilde{V}_{m+1}([\tilde{n}, \theta]) =$$

$$\left\{ h \mu_l f_l(n_l+1) p_{l0} + \right.$$

$$h \lambda \sum_j p_{0j} \tilde{V}_m([\tilde{n}+e_i+e_j, \theta]) +$$

$$h v_0 1_{\{\theta=0\}} \tilde{V}_m([\tilde{n}+e_i, 1]) + h v_1 1_{\{\theta=1\}} \tilde{V}_m([\tilde{n}+e_i, 0]) +$$

$$h \sum_j \mu_j f_j(n_j) \left[\sum_{s=1}^N p_{js} \tilde{V}_m([\tilde{n}+e_i-e_j+e_s, \theta]) + 1_{\{\theta=1\}} p_{j0} \tilde{V}([\tilde{n}+e_i-e_j, 1]) \right] +$$

$$h \mu_i [f_i(n+1) - f_i(n)] \left[\sum_{s=1}^N p_{is} \tilde{V}_m([\tilde{n}+e_s, \theta]) + 1_{\{\theta=1\}} p_{i0} \tilde{V}([\tilde{n}, \theta]) \right] +$$

$$\left[1 - h \left\{ \lambda + v_\theta + \sum_j \mu_j f_j(n_j) \left[\sum_{s=1}^N p_{js} + 1_{\{\theta=1\}} p_{j0} \right] + \right. \right.$$

$$\left. \mu_i [f_i(n+1) - f_i(n)] \left[\sum_{s=1}^N p_{is} + 1_{\{\theta=1\}} p_{i0} \right] \right\} \right] \tilde{V}([\tilde{n}+e_i, \theta]) \Big\}$$

—

$$\left\{ h \mu_l f_l(n_l) p_{l0} + \right.$$

$$h \lambda \sum_j p_{0j} \tilde{V}_m([\tilde{n}+e_j, \theta]) +$$

$$h v_0 1_{\{\theta=0\}} \tilde{V}_m([\tilde{n}, 1]) + h v_1 1_{\{\theta=1\}} \tilde{V}([\tilde{n}, 0]) +$$

$$h \sum_j \mu_j f_j(n_j) \left[\sum_{s=1}^N p_{js} \tilde{V}_m([\tilde{n}-e_j+e_s, \theta]) + 1_{\{\theta=1\}} p_{j0} \tilde{V}_m([\tilde{n}-e_j, 1]) \right] +$$

$$\left[1 - h \left\{ \lambda + v_\theta + \sum_j \mu_j f_j(n_j) \left[\sum_{s=1}^N p_{js} + 1_{\{\theta=1\}} p_{j0} \right] \right\} \right] \tilde{V}_m([\tilde{n}, \theta]) \Big\} =$$

$$\begin{aligned}
 & \left[h\mu_i [f_i(n_i+1) - f_i(n_i)] p_{i0} \right] + \\
 & h\lambda \sum_j p_{0j} \Delta_i \tilde{V}_m([\tilde{n}+e_j, \theta]) + \\
 & h v_0 1_{\{\theta=0\}} \Delta_i \tilde{V}_m([\tilde{n}, 1]) + h v_1 1_{\{\theta=1\}} \Delta_i \tilde{V}_m([\tilde{n}, 0]) + \\
 & h \sum_j \mu_j f_j(n_j) \left[\sum_{s=1}^N p_{js} \Delta_i \tilde{V}_m([\tilde{n}-e_j+e_s, \theta]) + 1_{\{\theta=1\}} p_{j0} \Delta_i \tilde{V}_m([\tilde{n}-e_j, 1]) \right] + \\
 & h\mu_i [f_i(n_i+1) - f_i(n_i)] \left\{ \sum_{s=1}^N p_{is} \Delta_s \tilde{V}_m([\tilde{n}, \theta]) + 1_{\{\theta=1\}} p_{i0} \left[\tilde{V}_m([\tilde{n}, 1]) - \tilde{V}_m([\tilde{n}, 1]) \right] \right\} + \\
 & \left[1 - h \left\{ \lambda + v_0 + \sum_j \mu_j f_j(n_j) \left[\sum_{s=1}^N p_{js} + 1_{\{\theta=1\}} p_{j0} \right] + h\mu_i [f_i(n_i+1) - f_i(n_i)] \right. \right. \\
 & \left. \left. \left[\sum_{s=1}^N p_{is} + 1_{\{\theta=1\}} p_{i0} \right] \right\} \right] \Delta_i \tilde{V}_m([\tilde{n}, \theta]) \quad (8)
 \end{aligned}$$

For $i \neq \ell$ the induction hypothesis (7) for $t = m$ and the fact that $f_i(n_i)$ is nondecreasing now directly proves (7) also for $t = m + 1$. For $i = \ell$, the lower estimate 0 in (7) for $t = m + 1$ also follows directly by substituting the induction hypothesis: $\Delta_j V_m(\cdot) \geq 0$. To conclude the upper estimate, first note that the one but last term is equal to 0. Then, by letting the first additional positive term: $h\mu_i [f_i(n_i + 1) - f_i(n_i)] p_{i0}$ replace this one but last term and using (3), the upper estimate 1 in (7) for $t = m + 1$ and $i = \ell$ is also concluded. $\square$

We are now able to prove the main result. It shows that station throughputs are enlarged due to repeated services upon output blocking, as intuitively obvious (also see remark below), with an error bound of linear order in the probability that the system is down.

Theorem

$$0 \leq \bar{L}_1 - L_1 \leq [v_1/(v_0+v_1)] \sup_{\tilde{n}} [\sum_i \mu_i f_i(n_i) p_{i0}] p_{i0}^{-1}. \quad (9)$$

Proof Consider an arbitrary t and $[\bar{n}, \theta] \in \bar{R}$. Then from (5) we derive

$$\begin{aligned} \bar{V}_t([\bar{n}, \theta]) - V_t([\bar{n}, \theta]) = \\ \sum_{[\bar{n}', \theta']} \left[\bar{p}([\bar{n}, \theta], [\bar{n}', \theta']) - p([\bar{n}, \theta], [\bar{n}', \theta']) \right] \bar{V}_{t-1}([\bar{n}', \theta']) + \\ \sum_{[\bar{n}', \theta']} p([\bar{n}, \theta], [\bar{n}', \theta']) \left[\bar{V}_{t-1}([\bar{n}', \theta']) - V_{t-1}([\bar{n}', \theta']) \right] \end{aligned}$$

Further, from (3) and (4) we conclude

$$\begin{aligned} \sum_{[\bar{n}', \theta']} \left[\bar{p}([\bar{n}, \theta], [\bar{n}', \theta']) - p([\bar{n}, \theta], [\bar{n}', \theta']) \right] V_{t-1}([\bar{n}', \theta']) = \\ 1_{\{\theta=0\}} \sum_i \mu_i f_i(n_i) p_{i0} \left[V_t([\bar{n}-e_i, 0]) - V_t([\bar{n}, 0]) \right] Q^{-1} \quad (11) \end{aligned}$$

By iteratively substituting (11) in (10) and repeating (10) for $t = T, T-1, \dots, 1$, and using that $V_0(\cdot) = 0$, we obtain

$$\begin{aligned} \bar{V}_T([\bar{n}, \theta]) - V_T([\bar{n}, \theta]) = \\ \sum_{k=0}^{T-1} P^k([\bar{n}, \theta], [\bar{n}', \theta']) 1_{\{\theta'=0\}} \sum_i \mu_i f_i(n_i') p_{i0} \\ \left[V_{T-k-1}([\bar{n}'-e_i, 0]) - V_{T-k-1}([\bar{n}', 0]) \right] Q^{-1} \quad (12) \end{aligned}$$

where $P^k(\dots)$ denotes the k -th power of the one-step transition probabilities (or matrix) $P(\dots)$. The lower estimate 0 from (7) together with (6) directly gives the lower estimate 0 in (9) if in (12) we let $T \rightarrow \infty$ and divide by T . The upper estimate in (9) is concluded similarly by using the upper estimate 1 from (7), recalling (3) and noting (e.g. by simple renewal theory) that the fraction of time that the system is down tends to $v_1 / (v_0 + v_1)$ if $T \rightarrow \infty$.

Example (Single server tandem line) Consider a tandem line of N -single server station and let $\ell=N$ so that $p_{i0}=0$ for $i \neq \ell$ and $p_{\ell 0}=1$. Then by (9)

$$0 \leq \bar{L}_N - L_N \leq [v_1 / (v_0 + v_1)] \mu_N$$

Remarks

1. (Upper bound) The lower estimate 0 in (9) or equivalently the fact that the station throughput in the modified model dominates that of the original model may seem trivial. However, as in [1], [11] or [13] one can sometimes give counterintuitive examples in which a throughput is increased by rejecting specific arrivals. Relatedly, a throughput can sometimes be decreased by blocking specific departures.

2. (Error bound) The absolute error bound in (9) will be quite large when the service capacities can become large. In that case, however, also the values $\bar{L}_i$ are likely to be large. Roughly, (9) may then lead to relative error bounds of order $v_1 / (v_0 + v_1)$ which is to be thought of as being small, say in the order of 1 à 2%.

3. (Prooftechnique) Monotonicity prooftechniques have received considerable attention over the last decade (e.g. [11], [14], [18], [19]). The prooftechnique employed herein can be seen as an extension as it also provides error bounds. This technique has already been succesful in simple one- or two-station models (cf. [16], [18]). The present application, in contrast, is multidimensional. Moreover, some other technical details such as the use and limiting argument of equation (12) are new. The prooftechnique thus seems of interest in itself as well as for further application.

4. (Bounded rates) The proof seems to essentially require the restrictive boundedness assumption (3) in order to apply the uniformization technique (4). However, by using an approximate uniformization as recently developed in [17], the results can be relaxed to allow for unbounded rates. The details are rather technical and therefore omitted.

References

- [1] Adan, I.J.B.F. and van der Wal, J. (1987) "Monotonicity of the throughput of a closed queueing network in the number of jobs", To appear: *Opns. Res.*

- [2] Courtois, P.J., "Decomposability: Queueing and Computer System Application", New York: Academic (1977).
- [3] Courtois, P.J., and Semal, P., "Bounds on Conditional Steady-State Distributions in Large Markovian and Queueing Models", *Teletraffic Anal. and Computer Perf. Eval.*, O.J. Boxma, J.W. Cohen and H.C. Tijms editors, North Holland (1986).
- [4] Goyal, A., Lavenberg, S.S. and Trivedi, K.S. "Probabilistic Modeling of Computer System Availability", *Ann. of Oper. Res.*, vol. 8, pp. 285-306 (1986).
- [5] Goyal, A. and Tantawi, A.N., "Evaluation of performability for degradable computer systems", *IEEE Transactions on Computers*, C-36(6):738-744 (June, 1987).
- [6] Haverkort, B.R. and Niemegeers, J.G., "A theory of describing performability models", Research report, Twente University (1988).
- [7] Jaiswal, N.K., "Priority Queues, Academic Press, New York (1968).
- [8] Meyer, J.F., "Closed form solutions of performability", *IEEE Transactions on Computers* 31, pp. 648-657 (1982).
- [9] Muntz, R.R., de Souza et Silva, E. and Goyal, A., "Bounding availability of repairable computer systems", *Performance Evaluation Review*, vol. 17, 29-38 (1989).
- [10] Ross, S.M., "Applied probability models with optimization application", Holden day, San Francisco (1970)
- [11] Shanthikumar, J.G. and Yao, D., "Stochastic monotonicity of the queue lengths in closed queueing networks", *Operations Research*, 35, 583-588 (1987).
- [12] Smith, R.M., Trivedi, K.S. and Ramesh, A.V., "Performability analysis: measures, an algorithm and a case study", *IEEE Transactions on Computers*, C-37 4, 406-417 (1988).
- [13] Souza e Silva, E. de, and Gail, H.R., "Calculating availability and performability measures of repairable computer systems using randomization", *JACM*, vol. 36 (1989).
- [14] Stoyan, D., "Comparison methods for queues and other stochastic models", Wiley, New York (1983).
- [15] Tijms, H.C., "Stochastic Modelling and Analysis: A computational approach", Wiley, New York (1986).
- [16] Van Dijk, N.M., "Simple bounds for queueing systems with break-downs", *Performance Evaluation*, 8, 117-128 (1988).

- [17] Van Dijk, N.M., "Approximate uniformization for continuous time Markov chains with an application to performability analysis", Research report, Free University (1988).
- [18] Van Dijk, N.M., and Van der Wal, J., "Simple bounds and monotonicity results for multi-server exponential tandem queues", *Queueing Systems* (1987).
- [19] Whitt, W., "Comparing counting processes and queues", *Adv. Appl. Prob.*, 13, 207-220, (1981).