

An Analytical Approximation for the Throughput of a Closed Fork/Join Network with Multi-Station Input Subnetworks

Erkut Sönmez

*Carroll School of Management, Boston College, 140 Commonwealth Avenue, Chestnut Hill,
MA 02467, USA
sonmeze@bc.edu*

Alan Scheller-Wolf • Nicola Secomandi

*David A. Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue,
Pittsburgh, PA 15213-3890, USA
{awolf, ns7}@andrew.cmu.edu
Tepper Working Paper 2009-E5
October 2009*

Abstract

Fork/join stations are used for modeling synchronization between entities, and fork/join queueing networks are natural models for a variety of communication and manufacturing systems: Parallel computer networks, fabrication/assembly systems, supply chains and material control strategies for manufacturing systems. Exact solutions of general fork/join networks can only be obtained by using numerical methods to analyze the underlying Markov chains. However, this method is computationally feasible only for networks with small population size and number of stations. In this paper, we present a simple approximation method to estimate the throughput of a closed queueing network that features a single fork/join station receiving inputs from multi-station subnetworks. Our technique uses aggregation to estimate the arrival process from input subnetworks. Given the estimated arrival process, we then derive closed form approximate expression for the network throughput by analyzing a simplified Markov chain. A numerical study shows that the proposed approximation is fairly effective, particularly for large network sizes.

1. Introduction

Fork/join stations are used to model synchronization between entities within a queueing network. At a fork/join station, different types of entities in the network are matched to form a single batched entity. After receiving service, this batched entity is decomposed into its component entities. Then, each entity follows a path specific to its type.

Networks containing fork/join stations are natural models for a variety of computer and manufacturing systems:

- In computer network analysis, fork/join stations have been studied in the context of parallel processing, database concurrency control, and communication protocols (Bacelli et al. 1989, Varki 1999, Prabhakar et al. 2000).
- In fabrication/assembly systems, fork/join networks represent the assembly of a product or a system that requires several parts that are processed simultaneously (Harrison 1973, Latouche 1981, Hopp and Simon 1989, Duenyas and Hopp 1993, Duenyas 1994, Rao and Suri 1994, 2000).
- Several material control strategies for manufacturing systems, such as multi-stage kanban or CONWIP systems, can be modeled as closed queueing networks (CQNs) with fork/join stations (Duenyas and Hopp 1992, 1993, Krishnamurty et al. 2004).
- In transshipment based supply chains in which holding inventory at the transshipment point is not feasible, such as shipping of liquefied natural gas, fork/join stations are used to model synchronization between arriving and departing vehicles at the transshipment location (Sönmez et al. 2009).

Unfortunately, while fork/join stations are an important modeling tool, analyzing queueing networks (open or closed) that include a fork/join station is a difficult task: These networks do not have exact product-form solutions even when all stations have exponential service times. Thus, exact solutions of these networks can only be obtained by analyzing the underlying Markov chain using numerical methods. However, such analysis is computationally feasible only for networks of small sizes (in terms of population size and number of stations).

Several researchers have developed various exact or approximation methods for analyzing the performance of the fork/join station *in isolation*. A common feature used in these analyses is a description of the inter-arrival times of the entities to the buffer queues of the fork/join station. However, in most applications this arrival process is not known a priori. For example, the arrival process to the fork/join station can be significantly influenced by the characteristics of the subnetworks from where arrivals occur, such as population limits, service characteristics, and network topology. In this case, existing approximation methods largely rely on complex iterative algorithms whose complexity and computational requirements increase with the network size.

In this paper, we present a simple approximation method for analyzing the throughput of a closed queueing network with a single fork/join station that features two buffer queues for synchronization. These buffer queues receive inputs from finite population subnetworks composed of multiple stations with a variety of queueing disciplines. We use the aggregation method (Chandy 1975) to substitute a single service center for each subnetwork, and approximate its load dependent arrival rate. Using these approximated arrival rates, we obtain a closed-form approximation for the throughput of the fork/join station by analyzing the underlying Markov chain. In computational experiments we show that the proposed approximation is fairly effective, particularly for large network sizes (when it would be most valuable).

Our approach has an important computational advantage over the existing numerical approximations, since it eliminates the computational burden associated with the iterative algorithms. Although our method is limited to networks with a single fork/join station, the closed-form expression proposed in this study could be used as an efficient tool in more general decomposition based approaches, such as those discussed in Baynat and Dallery (2000) and Krishnamurthy and Suri (2006), to analyze broader network settings.

The remainder of this paper is organized as follows. We review the related literature in §2. Section 3 presents our approximation technique for calculating the throughput of a closed queueing network with a single fork/join station and multi-station input subnetworks. We discuss a set of numerical experiments conducted to assess the accuracy of the proposed approximation in §4. We conclude in §5 by discussing the limitations of our work and further research avenues.

2. Related Literature

Previous research on fork/join stations can be classified into two main streams:

1. The arrival process into the fork/join station is known;
2. The arrival process into the fork/join station is not known a priori and is generated from multi-station subnetworks.

In the first area, Harrison (1973) and Latouche (1981) study fork/join stations with exponential inputs and provide conditions for the stability (convergence of waiting time in distribution) of fork/join stations. Later, Som et al. (1994) and Takahashi et al. (1998)

analyze the underlying Markov chain to derive closed form expressions for the throughput and mean queue lengths of the fork/join station, and characterize the departure process from the fork/join station, when the inter-arrival time distributions are exponential for each input. Takahashi et al. (2000) relax this exponential assumption and use matrix analytical methods to analyze the performance of a fork/join station in which the distribution of inter-arrival times follows a phase type distribution, assuming that the arrivals are from infinite-size populations (that is arrivals are an uninterrupted process). Krishnamurthy et al. (2004) study the performance of a fork/join station in a closed queueing network with inputs from single station servers with phase type distributions, and Krishnamurthy et al. (2003) present a two-moment approximation method for computing the throughput of the fork/join station in a closed queueing network with phase type renewal inputs. Although this literature provides valuable insights on the performance measures, it is limited to fork/join stations in isolation. For this reason, applicability of these methods in more general network settings is restricted.

In the second area, researchers have developed approximation techniques for computing the performance measures of closed fork/join queueing networks with multi-station input subnetworks. Varki (1999) proposes an iterative method based on mean value analysis for networks in which forked items must be processed by identical single server first come first serve (FCFS) stations with exponential service times. Baynat and Dallery (2000) develop an approximation method for general fork/join networks based on product-form approximations. Di Mascolo et al. (1996) apply product-form approximations to analyze the fork/join stations in a multi-stage kanban system. Krishnamurthy and Suri (2006) use a two-moment parametric decomposition technique to analyze a fork/join station in a single stage kanban system. All of the approximation methods developed in this stream of research are based on iterative algorithms.

Our work falls within the second stream of research but does not require an iterative algorithm. Moreover, existing methods approximate both the arrival and departure processes in the network and the performance of the fork/join stations. The novelty in our method is that we only approximate the arrival process into the single fork/join station which is possible because of our use of aggregation. This allows us to express the throughput estimate of the closed fork/join network with a closed form expression, which appears to be novel.

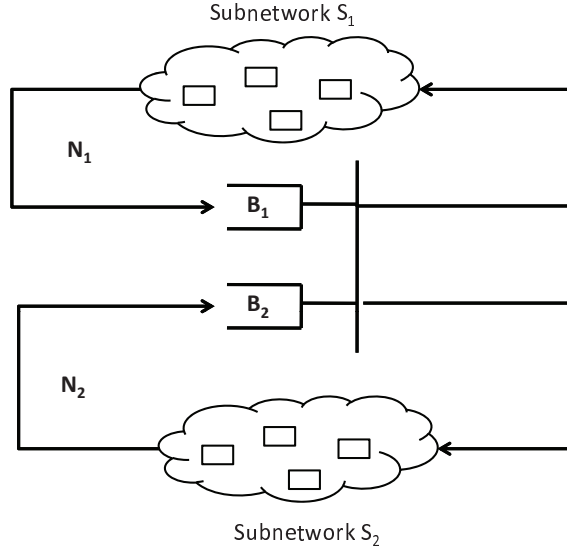


Figure 1: Closed fork/join network: Network 1.

3. Model

Figure 1 represents the closed fork/join queueing network that we study in this chapter. We refer to this network as network 1 throughout our discussion. Network 1 consists of a single fork/join station (station J) with two infinite size input buffers, B_1 and B_2 . Entities arrive to the buffer queues B_1 and B_2 from subnetworks 1 and 2 (S_1 and S_2), respectively. If an entity arriving to buffer B_1 (B_2) finds buffer B_2 (B_1) empty, it waits for the corresponding entity to arrive to buffer B_2 (B_1). As soon as this corresponding entity arrives, one entity from each buffer queue is removed; they are “matched.” The removed entities are then joined into a single entity at the fork/join station. After leaving the fork/join station, the joined entity splits back into its two component entities that are routed back to their respective subnetworks. We assume that subnetwork i ($i = 1, 2$) is a product-form network (Baskett et al. 1975) with multiple stations having FCFS, ample server (AS), last-come-first-serve (LCFS), or processor sharing (PS) service disciplines. The service time distributions at these stations can be any one having rational Laplace transform, with the exception that service times of FCFS stations must be exponential. This product form assumption is not overly restrictive, as non product-form subnetworks could be approximated with product-form networks.

Our aim is to compute the throughput of network 1, denoted by X_1 . As discussed in

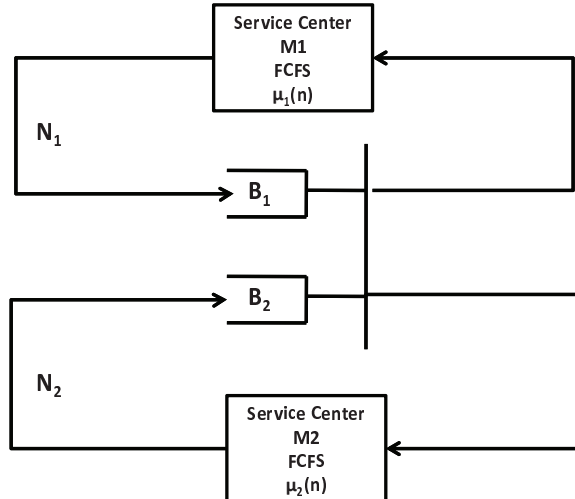


Figure 2: Aggregated closed fork/join network: Network 2.

§1, due to the presence of the fork/join station, network 1 does not have a product-form solution. Thus an exact solution can only be obtained by using numerical methods to analyze the underlying Markov chain. However, this method is computationally feasible only for networks of small sizes. For this reason we develop the following approximation technique to compute the throughput of network 1.

Our approximation technique has two main steps. In the first step, we replace subnetworks S_1 and S_2 with flow equivalent single service centers M_1 and M_2 using aggregation method. We refer to the resulting network as network 2, displayed in Figure 2. In network 2, M_1 and M_2 are defined as FCFS stations with exponential service times and load dependent service rates. We denote by $\mu_i(n_i)$ the load dependent service rate of station M_i , $i = 1, 2$, when there are n_i jobs in this station. By using load dependent service rates, we capture the throughput behavior of S_1 and S_2 in terms of the number of jobs currently within them. We set $\mu_i(n_i)$ equal to the stationary throughput, $Y_i(n_i)$, of subnetwork i analyzed *in isolation* as a closed network with n_i jobs within it. Figure 3 presents this subnetwork with n_i circulating jobs, $n_i = 1, \dots, N_i$, where N_i is the total number of type i jobs in network 1.

For a given n_i , we follow Baskett et al. (1975) to calculate $Y_i(n_i)$. Let I_i be the total number of stations in subnetwork i . We denote the number of jobs in each station j in subnetwork i as k_{ij} . The state of the closed queueing system in Figure 3 is then the array $k_i = (k_{ij}, j = 1, \dots, I_i)$, such that $\sum_{j=1}^{I_i} k_{ij} = n_i$. Let λ_{ij} and m_{ij} be the arrival and service

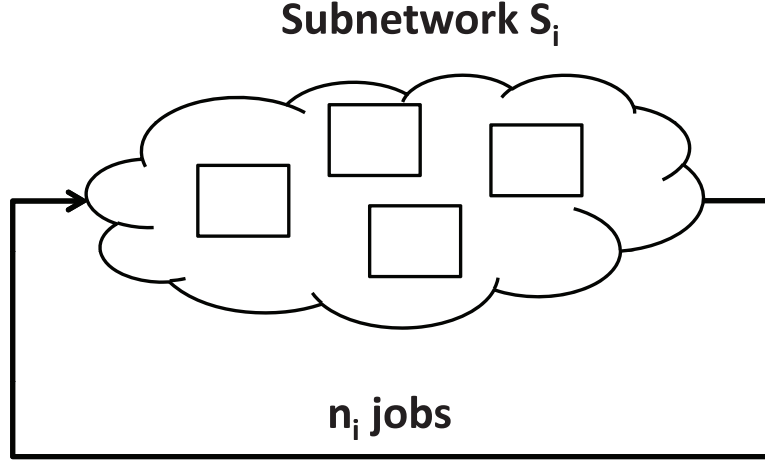


Figure 3: Isolated subnetwork S_i .

rates of station j in subnetwork i , respectively. Denote by $\pi_i(k_i)$ the steady state probability that the system is in state k_i . Basket et al. (1975) show that

$$\pi_i(k_i) = \Gamma_i \prod_{j=1}^{I_i} \gamma_{ij}(\lambda_{ij}, m_{ij}, k_{ij}), \quad (1)$$

where Γ_i is a normalizing constant chosen to make these probabilities sum to 1 and $\gamma_{ij}(\cdot)$ is computed as follows:

$$\gamma_{ij}(\lambda_{ij}, m_{ij}, k_{ij}) := \begin{cases} \left(\frac{\lambda_{ij}}{m_{ij}}\right)^{k_{ij}} & \text{If block } j \text{ is FCFS, PS, or LCFS} \\ \frac{1}{k_{ij}!} \left(\frac{\lambda_{ij}}{m_{ij}}\right)^{k_{ij}} & \text{If block } j \text{ is AS.} \end{cases}$$

One can calculate the throughput of any station j in subnetwork i using the steady state probabilities defined by (1). Then, $Y_i(n_i)$ is obtained by dividing the throughput of station j by its visit ratio (the average number of visits to station j per job).

In the second step of our approximation technique, we analyze the queue length process at the buffer queues B_1 and B_2 using the load dependent service rates of centers M_1 and M_2 computed in step 1. Let $w_1(t)$ and $w_2(t)$ denote the number of jobs in the buffer queues B_1 and B_2 , respectively, at time t . Because two types of entities form the joined single entity as soon as they appear at these buffer queues, it is not possible for both buffers B_1 and B_2 to

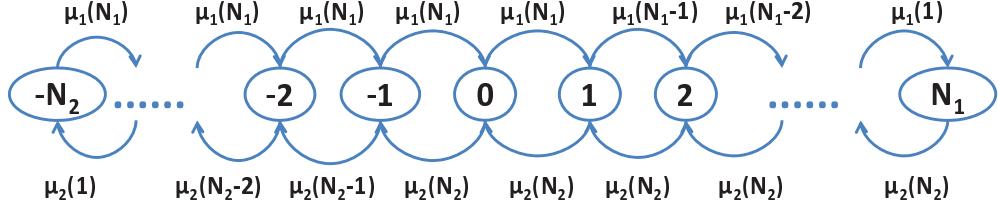


Figure 4: State transition diagram for aggregated fork/join network 2.

be non-empty. For this reason, we denote the number of jobs waiting at these buffer queues at time t as the one dimensional random variable $w(t) := w_1(t) - w_2(t)$; $w(t)$ takes values on $-N_2, \dots, -1, 0, 1, \dots, N_1$. For example, $w(t) = -3$ means that there are 3 type 2 jobs waiting at B_2 and no type 1 job exists at B_1 . There is no job of either type waiting at the buffer queues if $w(t) = 0$. With these definitions, the state of the system in network 2 at any time t is completely characterized by the one dimensional variable $w(t)$ (For convenience, we will omit the index t on the state w , as we will henceforth be referring to stationary quantities). The state space is given by $\mathcal{W} := \{-N_2, \dots, -1, 0, 1, \dots, N_1\}$. Note that this Markov chain has $N_1 + N_2 + 1$ states.

Figure 4 illustrates the state transition rates for the continuous time Markov chain (a simple birth and death process) representing the queue length process in network 2. For each state $w \in \mathcal{W}$ let $p(w)$ be the steady state probability that this continuous time Markov chain is in state w ; these probabilities can be computed as the solution of the following balance equations, also imposing the condition that these probabilities sum up to 1:

$$\begin{aligned}
 p(w)\mu_2(N_2) &= p(w-1)\mu_1(N_1 - (w-1)) \quad \forall w = 1, \dots, N_1 \\
 p(w)\mu_1(N_1) &= p(w+1)\mu_2(N_2 + w + 1) \quad \forall w = -1, \dots, -N_2 \\
 \sum_{w=-N_2}^{N_1} p(w) &= 1.
 \end{aligned} \tag{2}$$

By solving the system of linear equations (2), $p(w)$ can be expressed as

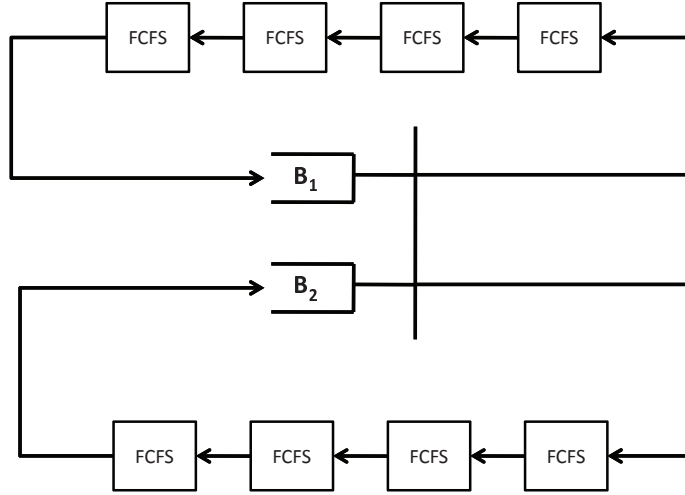


Figure 5: Example closed fork/join network.

$$p(w) = \begin{cases} \frac{1}{A} & \text{If } w = 0, \\ \frac{\prod_{k=0}^{w-1} \mu_1(N_1 - k)}{A(\mu_2(N_2))^w} & \text{If } w \in \{1, \dots, N_1\}, \\ \frac{\prod_{k=0}^{|w|-1} \mu_2(N_2 - k)}{A(\mu_1(N_1))^{|w|}} & \text{If } w \in \{-1, \dots, -N_2\}, \end{cases} \quad (3)$$

where $A = 1 + \sum_{w=1}^{N_1} \frac{\prod_{k=0}^{w-1} \mu_1(N_1 - k)}{(\mu_2(N_2))^w} + \sum_{w=1}^{N_2} \frac{\prod_{k=0}^{w-1} \mu_2(N_2 - k)}{(\mu_1(N_1))^w}$.

Due to the presence of the fork/join station, the throughput of service centers M_1 and M_2 must be equal. The throughput of network 2, X_2 , can then be obtained by calculating the throughput of station M_1 , for example, using $p(w)$ in (3) as follows:

$$X_2 = \mu_1(N_1)(p(0) + \sum_{w=-N_2}^{-1} p(w)) + \sum_{w=1}^{N_1} p(w)\mu_1(N_1 - w). \quad (4)$$

We use X_2 to approximate X_1 , the throughput of the original closed fork/join network.

4. Numerical Study

In this section, we present the results of numerical experiments conducted to quantify the accuracy of our proposed approximation (4). Our focus is to understand how the accuracy

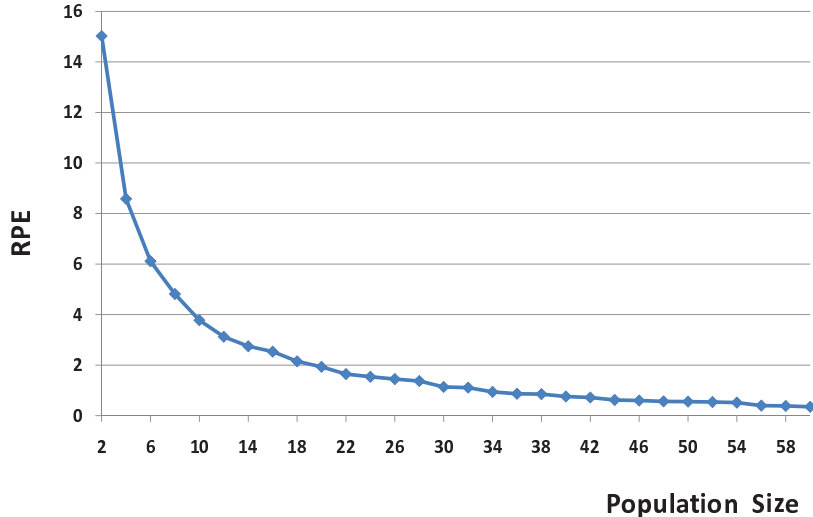


Figure 6: Accuracy of the approximation on the example network: FCFS case.

of our approximation changes with the network size, since the purpose of developing this approximation is to analyze large fork/join networks for which exact numerical methods are not practical. Therefore, we investigate the accuracy of (4) as a function of the population size ($N_1 + N_2$) and the number of stations in the network ($I_1 + I_2$), by comparing our approximation with the throughput estimate of network 1 obtained by using the ARENA Monte Carlo simulation software. We select the simulation run times and number of replications such that throughput reaches the steady state and the half-width of a 95% confidence interval is at most 0.5% of the mean. We then compute the relative percentage error (RPE) of our throughput estimate as

$$\text{RPE} = \frac{X_s - X_2}{X_s} \times 100,$$

where X_s is the throughput computed using Monte Carlo simulation and X_2 is that computed as in (4).

Effect of Population Size. We first analyze the impact of population size on accuracy. To do so, we study the example network illustrated in Figure 5. In this network each input subnetwork consists of four FCFS stations with exponential service rates with mean 0.5 jobs/day. For simplicity, we assume balanced subnetworks, $N_1 = N_2$, but our insights remain qualitatively the same for unbalanced subnetworks. Figure 6 displays RPE as a function of population size. We observe that the gap between the approximated and simulated

throughputs is at most 15%, which occurs when there is only one job in each subnetwork. RPE decreases monotonically with the population size, eventually converging to 0. As the population size increases, both the simulated and approximated throughputs approach the bottleneck capacities. Since our approximation technique matches the bottleneck capacity of the original system in the aggregation step (step 1), RPE must converge to 0 as the congestion in the systems increases.

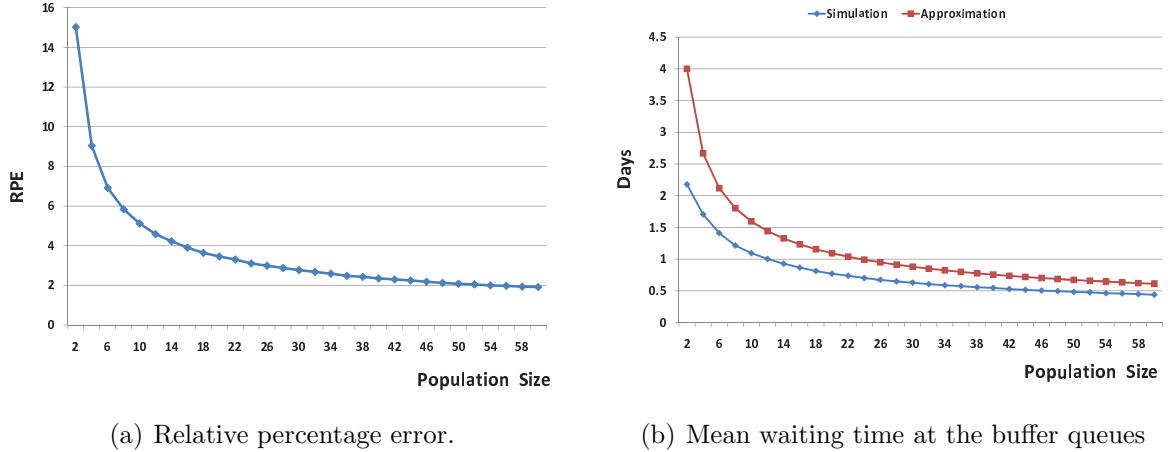


Figure 7: Accuracy of the approximation on the example network: AS case.

In order to exclude the effect of congestion from the analysis of the performance of our approximation method, we also investigate the network in Figure 5 when all the stations in each subnetwork are configured as AS stations. Figure 7(a) illustrates RPE as a function of population size for this case. Once again, RPE decreases monotonically with the population size, as in Figure 6, even though no congestion arises in the subnetworks due to the AS configuration. We discuss the intuition behind this next.

Figure 7(b) plots the mean waiting times at the buffer queues of the fork/join station both in the simulation and approximation models as a function of the population size. These waiting times decrease monotonically with increasing population size. This suggests that the impact of the fork/join station on the throughput of the two subnetworks diminishes as the waiting time in each buffers decreases. Thus, each subnetwork effectively converges to the isolated network shown in Figure 3 as the population size increases. Since for product-form networks the aggregation method used in the first step of our approximation technique yields the exact throughput of the isolated network (Chandy et al. 1975), RPE decreases with population size also in the AS case.

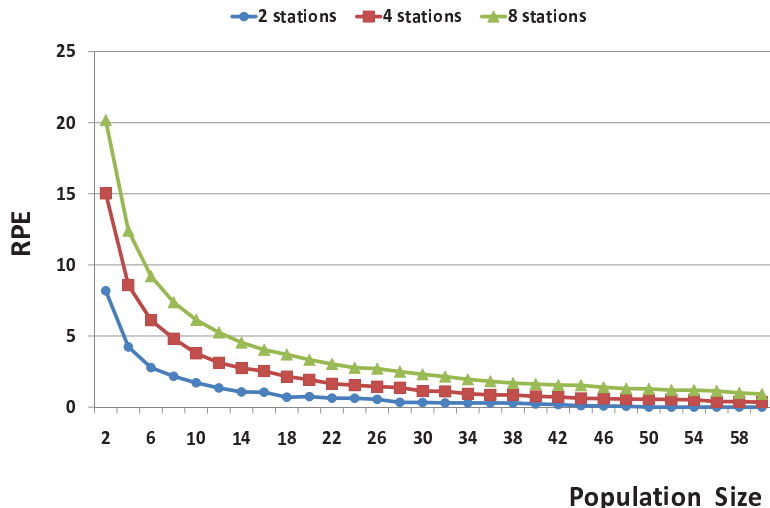


Figure 8: Impact of the number of stations in each subnetwork on RPE: FCFS case.

Effect of number of stations. We now investigate the accuracy of our method as a function of the number of stations in the network. To do so, we change the number of stations in each subnetwork in the example network while keeping their total service requirements constant; we do this to isolate the effect of the service requirement difference on the throughput change. For instance, we set the service rate of each station in a subnetwork as 1 job/day if there are 8 stations in this subnetwork, and 0.5 jobs/day if there are 4 such stations. We first analyze the case of FCFS stations in each subnetworks. Figure 8 plots RPE as a function of population size for subnetworks having 2, 4, and 8 stations. For a given population size, RPE increases with the number of stations. This occurs because increasing the number of stations while keeping the total service requirement constant relieves the congestion in the system by increasing the bottleneck capacity, and RPE increases as the congestion in the system decreases (see Figure 6). All three cases do eventually converge to zero though, due to the congestion effect. This just happens less rapidly as the number of stations increases.

Once again, to exclude the effect of congestion from this analysis, we investigate the case in which all the stations in each subnetwork are configured as AS stations. Figure 9 illustrates RPE in this case. For a given population size, RPE continues to increase with the number of stations in each subnetwork. Notice that, keeping the total service requirement constant, the variability of the arrival rate into the fork/join station decreases as the number

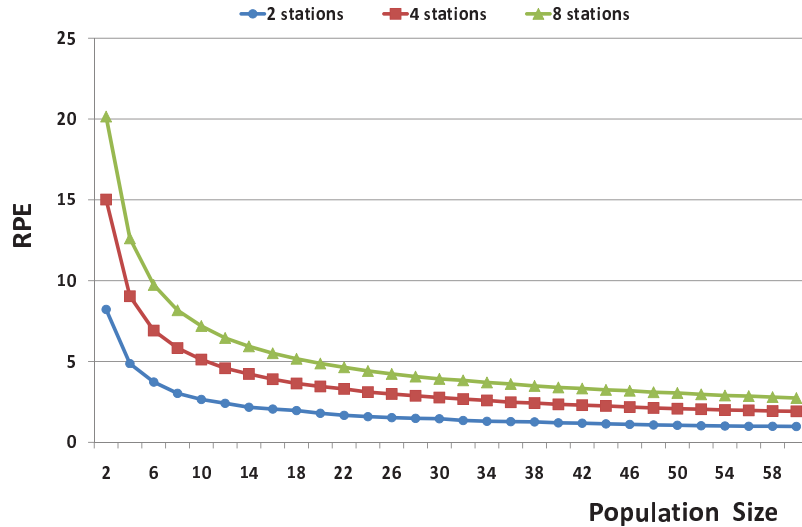


Figure 9: Impact of the number of stations in each subnetwork on RPE: AS case.

of stations in each subnetwork increases. This, consequently, increases the throughput of the system (Krishnamurthy 2004). On the other hand, the throughput of the aggregated network in the approximation model does not change with the number of stations in each subnetwork, since the total service requirement is kept constant (this follows from Baskett et al. 1975). Thus, RPE increases as the number of stations in each subnetwork increases.

5. Conclusions

In this paper, we propose a simple analytical approximation technique to estimate the throughput of a closed queueing network with a single fork/join station and multi-station input subnetworks. Our technique approximates the arrival process into the fork/join station using an aggregation method and approximates the system throughput by relying on an exact analysis of the resulting Markov chain. In contrast to existing approximation methods, we only approximate the arrival process to the fork/join station. This allows us to obtain a closed-form expression approximating the system throughput, which contrasts with the iterative algorithms available in literature. A numerical study shows that our approach is fairly accurate for networks with a large number of jobs, which is when an approximation is most needed.

There are some limitations in our work. (1) We assume that the input subnetworks

are product-form networks, which allows us to calculate the exact steady state probabilities for the isolated subnetworks using a product-form formula. This is no longer possible for subnetworks having a general form. In these cases, the steady state probabilities could be computed exactly by solving the underlying continuous-time Markov chain using appropriate numerical techniques, or approximated by using product-form techniques. (2) We assume that the joined entity splits into its components as soon as it leaves the fork/join station. If the joined entity receives service at some other stations before splitting, these common stations could be removed from the network and their duplicates included in each subnetwork in serial form.

Our work can be extended in several directions. We assume a single fork/join station in the network. However, our closed form expression for approximating the network throughput could be used in decomposition based approaches, as discussed in Baynat and Dallery (2000) and Krishnamurthy and Suri (2006), to analyze settings in which there are multiple fork/join stations in the network. Investigating this possibility is an interesting opportunity for further research. We use an aggregation method to estimate the arrival process into the fork/join station; in fact, the accuracy of the proposed approximation technique depends on that of the aggregation method. It would be of interest to study the case where the arrival process is approximated with other techniques, such as modifications of the aggregation method or Marie's method (Marie 1979).

References

- Bacelli, F., A. M. Makowski, A. Shwartz. 1989. The fork/join queue and related systems with synchronization constraints: Stochastic ordering and computable bounds. *Advances in Applied Probability*. 21 629-660.
- Baskett, F., K. M. Chandy, R. R. Muntz, F. G. Palacios. 1975. Open, closed, and mixed networks of queues with different classes of customers. *Journal of the Association for Computing Machinery*. 22 248-260.
- Baynat, B., Y. Dallery. 2000. An approximation method for general closed queueing networks with fork/join mechanisms. *Journal of the Operational Research Society*. 51 198-208.
- Chandy, K. M., U. Herzog, L. Woo. 1975. Parametric analysis of queueing systems. *IBM Journal of Research and Development*. January. 19-36.

- Di Mascolo, M., Y. Frein, Y. Dallery. 1996. An analytical method for performance evaluation of kanban controlled production systems. *Operations Research*. 44 50-64.
- Duenyas, I., W. J. Hopp. 1992. CONWIP assembly with deterministic processing and random outages. *IIE Transactions*. 24 97-109.
- Duenyas, I., W. J. Hopp. 1993. Estimating the throughput of an exponential CONWIP assembly systems. *Queueing Systems*. 14 135-157.
- Duenyas, I. 1994. Estimating the throughput of cyclic assembly systems. *International Journal of Production Research*. 32 1403-1419.
- Harrison, J. M. 1973. Assembly-like queues. *Journal of Applied Probability*. 10 354-367.
- Hopp, W. J., J. T. Simon. 1989. Bounds and heuristics for assembly-like queues. *Queueing Systems*. 4 137-156.
- Krishnamurthy A., R. Suri. 2006. Performance analysis of single stage kanban controlled production systems using parametric decomposition. *Queueing Systems*. 54 141-161.
- Krishnamurthy A., R. Suri, M. Vernon. 2003. Two-moment approximations for throughput and mean queue length of a fork/join station with general input process. In J. G. Shanthikumar, D. D. Yano, W. H. M. Zijm (ed.), *Stochastic Modeling and Optimization of Manufacturing Systems and Supply Chains*, Kluwer Academic Publishers, 87-126.
- Krishnamurthy A., R. Suri, M. Vernon. 2004. Analysis of a fork/join synchronization station with inputs from Coxian servers in a closed queueing network. *Annals of Operations Research*. 125 69-94.
- Latouche, G. 1981. Queues with paired customers. *Journal of Applied Probability*. 18 684-696.
- Marie R. 1979. An approximate analytical method for general queueing networks. *IEEE Transactions on Software Engineering*. 5 530-538.
- Prabhakar B., N. Bambos, T. S. Mountford. 2000. The synchronization of Poisson processes and queueing networks with service and synchronization nodes. *Advances in Applied Probability*. 32 824-843.
- Rao, P. C., R. Suri. 1994. Approximate queueing network models of fabrication assembly systems: Part I - single level systems. *Production and Operations Management*. 3 244-275.

- Rao, P. C., R. Suri. 2000. Performance analysis of an assembly station with input from multiple fabrication lines. *Production and Operations Management*. 9 283-302.
- Som, P., W. E. Wilhelm, R.L. Disney. 1994. Kitting process in a stochastic assembly system. *Queueing Systems*. 17 471-490.
- Sönmez, E., S. Kekre, A. Scheller-Wolf, N. Secomandi. 2009. Model-based analysis of liquefied natural gas regasification technologies. *Tepper Working Paper 2008-E43*, Carnegie Mellon University.
- Takahashi, M., H. Osawa, T. Fujisawa. 1998. A stochastic assembly system with resume levels. *Asia-Pacific Journal of Operational Research*. 15 127-146.
- Takahashi, M., H. Osawa, and T. Fujisawa. 2000. On a synchronization queue with two finite buffers. *Queueing Systems*. 36 107-123.
- Varki, E. 1999. Mean value technique for closed fork-join networks. *In Proceedings of ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems*, Atlanta, GA, 103-112.