Two Moment Approximations for Fork/Join Stations with Applications to Modeling Kanban Systems

Ananth Krishnamurthy and Rajan Suri
Center for Quick Response Manufacturing, University of Wisconsin-Madison
1513 University Avenue, Madison, WI 53706-1572, USA

Mary Vernon
Department of Computer Sciences, University of Wisconsin-Madison
1210 W Dayton Street, Madison, WI-53706, USA

In queuing models of kanban systems, fork/join stations are used to model the synchronization constraints between parts and kanbans. Efficient analysis of these fork/join stations is quite crucial to analytical performance evaluation of kanban systems. Exact analysis of the fork/join station can be difficult especially if the input processes have general characteristics. We propose a new method for the approximate analysis of the performance of such fork/join stations. Using two moment approximations of the input processes, we study the underlying queue length and departure processes. We develop analytical methods to estimate the throughput, waiting time and queue lengths.

1. Introduction

In recent years there have been several performance analysis studies of kanban systems. The primary aim of these studies has been to understand the impact of factors such as demand, service times and kanban allocation patterns on system throughput, work in process inventory and lead times. Simulation studies such as those presented in Zhou *et al*., 2000, and Krishnamurthy *et al.*, 2000, highlight several interesting tradeoffs in the performance of kanban system for certain environments. They recommend that a more detailed study be conducted to better understand the different tradeoffs in kanban system performance. However, simulation studies become computationally prohibitive for such a trade off study. Analytic queuing models are better suited for such analysis if they can be accurately applied.

Models of kanban systems typically involve solving a set of non-product form closed queuing networks with fork/join synchronization stations. In these models, the fork/join stations model the synchronization constraints between parts and kanbans at different stages of manufacturing. Since, obtaining the exact solutions for these networks is very difficult, approximate methods such as those using product form approximation techniques have been suggested. We have applied the product form approximation approach suggested in Di Mascolo *et al.*, 1996, and our observation from that experience is that the main challenge is the computational complexity resulting from the use of load dependent arrival and service rates for analyzing the different stations in the network. To overcome this difficulty, we propose a new approach based on parametric decomposition and two moment approximation equations. Instead of developing product form approximations for each node in the network, the parametric decomposition approach focuses on approximating the traffic process at the different nodes of the network. Assuming that these traffic processes are renewal processes, simple two moment approximation equations are used to describe the distribution of the inter-renewal times of these traffic processes. These two moment approximation equations form the fundamental basis for the parametric decomposition approach. However in order to be able to analyze queuing models of kanban systems using this approach, new two moment approximations must be developed for the analysis of fork/join stations. Developing these equations is the primary objective of this paper. Although our research is motivated by the analysis of queuing models of kanban systems, the two moment approximations developed in this research can also be used to analyze fork/join stations found in queuing models of other manufacturing systems such as assembly systems.

The outline of the paper is as follows. In section 2 we describe the operation and the principal characteristics of the fork/join station. In section 3 we provide a brief review of the relevant literature and describe some of the challenges present in analyzing fork/join stations. In section 4 we describe our analysis of the fork/join station. In section 5 we present some numerical results. In section 6 we present our proposed approach for analyzing queuing models of kanban systems.

2. Operation of the Join Station

Figure1 shows a typical fork/join synchronization station used during the analysis of queuing models of kanban systems. The join station J has two input queues P and F. In queuing models of kanban control strategies, the join station *J* models the synchronization of parts and free kanbans. Free kanbans queue up at *F* while parts queue up at *P*. If the queue *F*(*P*) is empty, a part (kanban) that arrives in queue *P*(*F*) waits in the queue for the corresponding kanban (part) to arrive. We assume that the waiting units are queued in the order in which they arrive. As soon as there is one entity in each queue, the part from queue *P* and a kanban from queue *F* are joined together and released from the join station. For the purposes of our analysis, we assume that the arrivals to each queue of the fork/join station can be well approximated by an independent renewal process. We characterize the distribution of inter-arrival times at the input queue *P* by its mean λ_P and squared coefficient of variation (SCV) c_P^2 . The distribution of inter-arrival times at the input queue *F* is characterized by its mean λ_F and squared coefficient of variation (SCV) c_F^2 . Further, since these fork/join stations are obtained from the decomposition of closed queuing network models of kanban systems, we assume that that the arrivals to both the input queues are from finite populations, each of size *K*. Operationally, this implies that the arrival process to either queue shuts down when it has *K* units. The arrival process resumes when the respective queue drops below the limit, *K*. For a well designed system, it will be reasonable to expect that the probability that either queue has length near *K* is close to zero.

See Figure 1

For such a fork/join station, we are interested in analyzing the queue length processes at each the input queue and also the output process from the join station. Ideally, we would like to obtain the distribution of the inter-departure times from the join station. However, we approximate the departure process by a renewal process and obtain estimates of the mean, λ_D, and SCV, c_D^2 , of the inter-departure times of this process from our analysis. We also find estimates of the mean queue lengths $\overline{L}_P$ and $\overline{L}_F$ at the input queues *P* and *F* in terms of the input parameters, λ_P, c_P^2 , λ_F, c_F^2 and *K*.

3. Literature Review

Fork/join stations have been analyzed in the context of manufacturing assembly systems and computer networks. Harrison, 1973, showed that assembly systems fed by input streams that are independent renewal processes are unstable when there is no capacity limitation for either stream. Subsequently, he also showed that a sufficient condition for stability of such systems is that the queues be finite. Bhat, 1986, incorporated limited buffer capacities in assembly like queues and derived expressions for the stationary probability vector of the queue length. However, he assumed that the arrival streams are Poisson. Previous studies that are most relevant to our work are the ones by Som *et al.*, 1994, and Takahashi *et al*., 1996. Som *et al.*, 1994, study a fork/join system in their analysis of kitting operations found in assembly systems and give the distribution of the kit completion time intervals. However they also assume that the input streams are independent Poisson processes, i.e., the inter-arrival times are exponentially distributed. Takahashi *et al.*, 1996, extend the study done by Som *et al.*, 1994, and study the effect of different buffer management schemes on kit completion times and sojourn times in the queues. They, too, assume that the inter-arrival times are exponentially distributed.

Our analysis of join stations compares to these previous studies in the following ways. Like the works by Bhat, 1986, Som *et al.*, 1994, and Takahashi *et al.*, 1996, we obtain characteristics for the queue length and the departure process from the join stations. However, our objective is to analyze join stations fed by input processes that are more general than the Poisson process. Di Mascolo *et al*., 1996, study join stations found in kanban systems. They analyze the join station as a double-ended queue with state dependent arrivals. However, obtaining these state dependent arrival rates can itself be quite challenging. Takahashi *et al*., 2000, proposes matrix geometric methods for analyzing fork/join stations fed by 2 input streams, one being a Poisson process and the other a phase renewal process. Their approach involves computations using large matrices. In trying to model the general input processes observed in our application, we are faced with new difficulties. The contribution of this research lies overcoming these difficulties by (1) writing simple equations relating the traffic process parameters, (2) using the theory of semi-Markov processes, to develop detailed equations that approximately describe the queue length and departure processes at the fork/join station. As a result, we obtain performance measures such as queue lengths, distributions of inter-departure times, and steady state probabilities for fork/join stations fed by a wide range of input processes.

4. Join Station Analysis

Our proposed approach, illustrated in Figure 2 consists of two basic steps, the *moment matching* step and the *traffic analysis* step. In the interest of space, in this paper we leave several of the equations in our analysis in their functional form. For detailed derivations and equations we refer the reader to Krishnamurthy *et al.*, 2000.

See Figure 2

4.1 Moment matching step

As mentioned before, the arrival processes are assumed to be renewal processes. Further, the distributions of the inter-renewal times are approximated by their means and SCVs. Assuming that these inputs are given and that the output process is also renewal, we wish to derive expressions for the mean, λ_D, and the SCV, c_D^2 , of the inter-departure times and the mean queue lengths at join station *J*. To do so, knowledge about the mean and the SCV of the arrival processes are not enough. Given the two moment characterizations (λ_P, c_P^2 , λ_F and c_F^2) of the renewal input processes at the input queues, we fit suitable two-phase Coxian distributions, $\hat{G}_P(\lambda_P, c_P^2)$ and $\hat{G}_F(\lambda_F, c_F^2)$ for the inter-arrival times at the input queue *P* and *F* respectively. The specific parameters for the distributions are chosen based on the guidelines discussed in Marie, 1978, and Altiok, 1996. Using $\hat{G}_P(\lambda_P, c_P^2)$ and $\hat{G}_F(\lambda_F, c_F^2)$ to describe the inter-arrival times, we analyze the queue length and departure processes at the join station *J*.

4.2 Analysis of the Queue Length Process

In this section, we derive the steady state queue length distributions at the two input queues of the join station, namely, *P* and *F*.

See Figure 3

Consider the queue length process at queues *P* and *F*. The state of these queues can be described by considering the number of units in these queues. A key observation is that it is not possible for both the queues *P* and *F* to be non-empty for any finite time. Departures occur instantaneously from the join station until at least one of these queues is empty. Further, if the arrival processes were general renewal, the analysis of the join station would be very complicated. Instead, we assume that the arrival processes are phase renewal processes with the underlying distributions having two exponential phases. In this case, the queues can be studied as a continuous parameter Markov process. We solve the continuous time Markov chain to obtain the steady state probability distribution of the queue lengths at *P* and *F*. From these we obtain the mean queue lengths, $\overline{L}_P$ and $\overline{L}_F$.

4.3 Analysis of the Departure Process

Next, we analyze the departure process from the join station *J*. Our analysis proceeds as follows. We first define the output process and recognize it as a marked point process. Then we observe that if the arrival processes at both the input queues are assumed to be phase renewal, the output process is a Markov renewal process. To analyze this Markov renewal process, we derive the corresponding semi-Markov kernel Q. The elements of the semi-Markov kernel consist of the conditional probability distributions for the times between successive departures. To derive these, we first identify the states in the Markov chain embedded at departure instants. Next, we recognize that the semi-Markov kernel has a special structure and use this information to identify the possible sample paths between successive departures. By deriving the conditional probability distributions for each sample path, the elements of the semi-Markov kernel are obtained. From the semi-Markov kernel, we obtain the transition probability matrix *P*, of the Markov chain embedded at instants of departure from the join station. We solve the Markov chain to obtain the stationary probability vector Π.

See Figure 4

Let $\hat{G}_D(t)$ be the distribution function of the inter-departure times. Then, since we assume that the departure process is a renewal process, we have, $\hat{G}_D(t) = \Pr\{D_m \le t\}$, where *m* is any arbitrary inter-departure interval. Then from the property of Markov renewal process, we have

$$\hat{G}_D(t) = \Pr\{D_m \le t\} = \Pi Q \boldsymbol{e} \tag{1}$$

where $\boldsymbol{e}$ is a column vector with all elements equal to 1. From $\hat{G}_D(t)$, we obtain the mean λ_D, and SCV, c_D^2 , of inter-departure times.

5. Numerical Results and Computational Issues

In this section we discuss the results obtained using the analysis presented in the previous sections. To test the accracy of our approach, we compared the results from our analysis to that obtained from simulation. We considered different distributions for the inter-arrival times at the two input queues. We also tested the accuracy of our approach for different values of SCV of the interarrival times at the two input queues. A sample of results is presented in Table 1 and additional details are described in Krishnamurthy *et al*., 2000. As can be seen from Table 1, our two moment approximations provide reasonably good estimates of the mean and SCV of inter-departure times from the fork/join station and the mean queue lengths at the two input queues.

See Table 1

It is interesting to compare the computational burden of our approach against some other approaches discussed in the literature. First, in the two-moment approximation approach for the analysis of join station only two inputs, namely the mean and SCV of the inter-arrival times are required. Further, our choice of the 2-phase Coxian distribution in the moment matching step allows us to model inter-arrivals times that have wide range of means $(0,\infty)$ and SCV $(0.5,\infty)$. The queue length process analysis discussed in section 4.2 involves computing the inverse of a matrix of size $8K$ while the analysis of the departure process involves the computing the inverse of a matrix of size $4K-3$. In the analysis presented by Di Mascolo *et al*., 1996, the join station is analyzed by approximating it as a as a double-ended queue with state dependent Poisson arrival rates. This implies that for the fork/join station illustrated in Figure 1, $2K$ state dependent arrival rates will be required as inputs for the analysis. Given these inputs and for reasonably large values of K their method would approximate the distribution of the arrivals and departures better. However, obtaining these state dependent arrival rates can be quite challenging in itself. Further for most practical applications, information about the first two moments may be more easily available than information about the distribution of the inter-arrival times. Takahashi *et al*., 2000, proposes matrix geometric methods for analyzing fork/join stations fed by 2 input streams, one being a Poisson process and the other a phase renewal process with n phases. However, their approach involves computing the inverse of a matrix of size $n(2K\text{-}1)$. In our approach, by restricting our analysis to phase distribution with two phases, we reduce the computational burden considerably without much compromise on the accuracy of the outputs.

6. Analyzing Queuing Models of Kanban Systems

As mentioned before kanban systems are modeled as closed cyclic queuing networks with manufacturing stations and fork/join synchronization stations. We propose to use parametric decomposition and two moment approximations to analyze these systems. First, the closed queuing network is decomposed into its constituent stations or nodes. Such a decomposition would yield two types of nodes: the *manufacturing stations* and the *fork/join stations*. For each manufacturing stations, *existing* two moment approximations will be used to characterize the mean queue lengths and output process. For the fork/join stations the *new* two moment approximations derived in this research will be used to are characterize the output process and mean queue lengths. Finally, the decomposed nodes are pieced together using the fact that in a cyclic queuing network, the arrival process to each queue is the departure process of the preceding queue in the network. Simple closed form expressions are then derived that equate the mean and SCV of the arrival process at a given node to the corresponding parameters for the departure process at the preceding node in the network. Finally, Little's law is applied to the entire system. This gives rise to a set of non-linear equations in the set of unknowns. These equations are solved to obtain performance metrics such as the system throughput, waiting times, work in progress.

7. Conclusions

Efficient analysis of fork/join stations is quite crucial to developing analytical techniques for performance evaluation of kanban systems. We have presented new two moment approximation equations for the approximate analysis of the performance of fork/join stations with fairly general input characteristics. Our numerical comparisons indicate that the results from our analytical approximations are within 10% of our simulation results for most performance measures. Using these approximation equations we propose to analyze queuing models of kanban systems. Although our research is motivated by the analysis of queuing models of kanban systems, the two moment approximations developed in this research can also be used to analyze fork/join stations found in queuing models of other manufacturing systems such as assembly systems.

References

[Alti 96] Altiok, T., *Performance Analysis of Manufacturing Systems*, Springer, NY, 1996.

[Bhat 86] Bhat, U.N., "Finite Capacity Assembly-like Queues", *Queuing Systems*, 185-101, 1986.

[DiFD 96] Di Mascolo, M., Frein, Y., and Dallery, Y., "An Analytical Method for Performance Evaluation of Kanban Controlled Production Systems", *Operations Research*, 44(1), 50-64, 1996.

[Harr 73] Harrison, J.M., "Assembly-like Queues", *Journal of Applied Probability*, 10(2), 354-367, 1973.

[KrSV 00] Krishnamurthy, A., Suri, R., and Vernon, M., "Push Can Perform Better than Pull for Flexible Manufacturing Systems with Multiple Products", *Industrial Engineering Research 2000 Conference Proceedings*, Institute for Industrial Engineers, Norcross, GA, USA, 2000.

[KrSV 00] Krishnamurthy, A., Suri, R., and Vernon, M., "Two Moment Approximations for a Fork/Join Station", *Technical Report*, Center for Quick Response Manufacturing, University of Wisconsin-Madison, Madison, WI, 2000.

[Mari 78] Marie, R., "Modelisation par Reseaux de Files d'Attente", Ph.D. Thesis, l'Universite'de Rennes, Rennes, France, 1978.

[SoWD 94] Som, P., Wilhelm, W.E., and Disney, R.l., "Kitting process in a stochastic assembly system", *Queuing Systems*, 17, 471-490, 1994.

[TaOF 96] Takahashi, M., Osawa, H., Fujisawa, T., "A stochastic assembly system with resume levels", *Asia-Pacific Journal of Operations Research*, 15, 127-146, 1996.

[IaOF 00] Takahashi, M., Osawa, H., Fujisawa, T., "On a Synchronization Queue with two finite buffers", *Queuing Systems*, 36, 107-123, 2000.

[ZhLu 00] Zhou, X., and Luh, P.B., "The Performance of a New Material Control and Replenishment System: A Simulation and Comparative Study", R. Suri (Ed.) *Proceedings of the Quick Response Manufacturing 2000 Conference* (Society of Manufacturing Engineers, Dearborn, MI), 807-826, 2000.

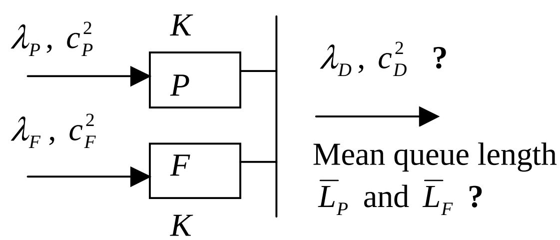

Figure 1. Join station J

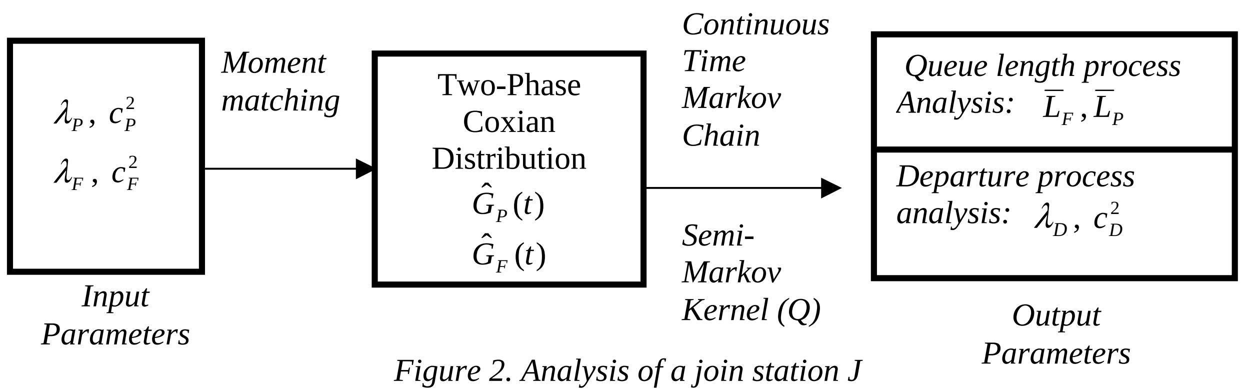

Figure 2. Analysis of a join station J

Figure 3. Queue length process analysis at the join station J

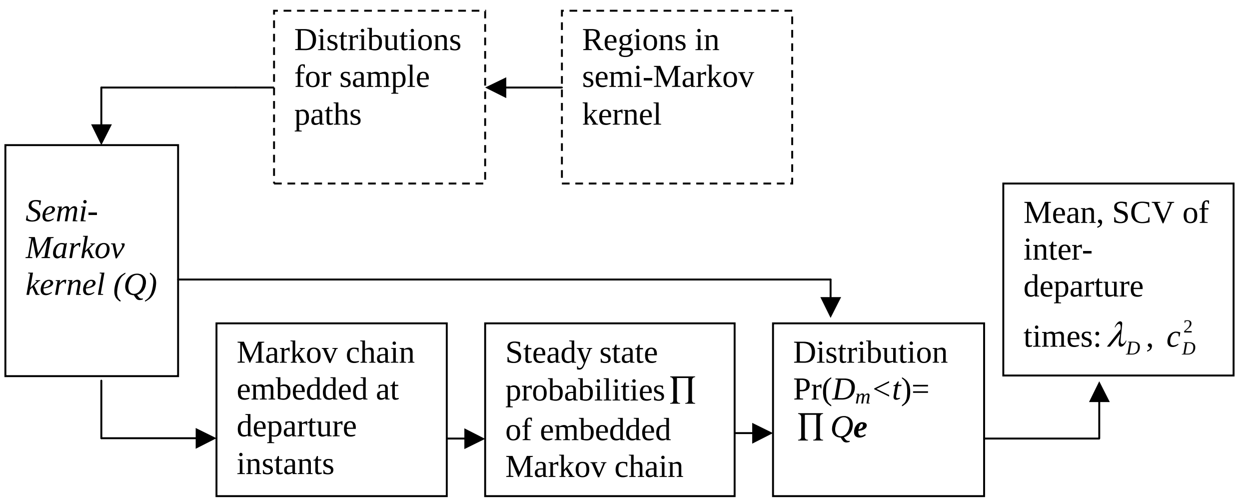

Figure 4. Departure process analysis at the join station J

Table 1. Sample results of join station analysis

Inputs						*Output from Analytical Model*				*Output from Simulation*			
Input Distribution	λ_P	λ_F	c_P^2	c_F^2	K	λ_D	c_D^2	$\bar{L}_P$	$\bar{L}_F$	λ_D	c_D^2	$\bar{L}_P$	$\bar{L}_F$
Coxian	1	1	0.8	0.8	5	1.09	0.72	1.32	1.32	1.09	0.71	1.32	1.32
	1	1	2.0	2.0	5	1.11	1.82	1.54	1.54	1.18	1.80	1.54	1.54
	1.2	0.5	0.6	1.5	5	1.20	0.60	0.01	4.32	1.20	0.72	0.01	4.32
	1.2	0.5	0.6	1.5	10	1.20	0.60	0.01	9.31	1.20	0.72	0.01	9.31
Gamma	1	1	0.8	0.8	10	1.05	0.76	2.56	2.56	1.04	0.92	2.57	2.57
	1	1	2.0	2.0	10	1.06	1.91	2.86	2.86	1.09	1.98	2.84	2.84
	1.2	0.5	0.6	1.5	5	1.20	0.60	0.01	4.32	1.20	0.72	0.01	4.33
	1.2	0.5	0.6	1.5	10	1.20	0.60	0.01	9.31	1.20	0.72	0.01	9.32