Approximations for Manufacturing Networks of Queues with Overtime

Gabriel R. Bitran*†
Devanath Tirupati**

WP# 2519
    was 1919-89

March 1989

*Sloan School of Management, Massachusetts Institute of Technology
**Department of Management, University of Texas at Austin

Approximations for Manufacturing Networks of Queues with Overtime

Gabriel R. Bitran* and Devanath Tirupati**

Abstract

In this paper we present simple approximations for networks of queues
with overtime operation at some stations. This type of network is
commonly encountered in practice in several manufacturing applications.
We provide bounds on the performance of the approximations for single
and multiple machine stations. Our results suggest that the methods
perform satisfactorily. These approximations can be used in conjunction
with the parametric decomposition methods to analyze queueing networks.
The computational results indicate that the performance of the
decomposition approach does not deteriorate when combined with the
methods proposed in this paper.

 * Sloan School of Management, Massachusetts Institute of Technology
** Department of Management, The University of Texas at Austin.

Approximations for Manufacturing Networks of Queues with Overtime

1.0 Introduction:

Queueing network models have been used to model a variety of
systems in computer and communication networks, service facilities, and
discrete and flexible manufacturing systems. However, exact results for
analysis of such systems is limited to product form networks (see
Jackson 1963; Kelly 1975; Baskett et.al. 1975; and Lemoine 1977). For
more general systems, the focus has been on the development of
approximations and simulation methods. In the analysis of manufacturing
systems modeled by open networks of queues, the parametric decomposition
method has been used with good results, among others, by Shanthikumar
and Buzacott (1981), Whitt (1983) and Bitran and Tirupati (1988a). This
approach generalizes the notion of independence and product form
solutions of Jackson type networks. In this method, the arrival process
at each station is approximated by a renewal process and the squared
coefficient of variation (scv) of interarrival times at each station is
computed approximately. Performance measures such as the mean number of
jobs and queue lengths at each station are estimated based on two moment
approximations. For a detailed description of this approach see Whitt
(1983), Bitran and Tirupati (1988a) and references therein. Recently,
Bitran and Tirupati (1988b) have extended these methods to certain types
of batch operations encountered in manufacturing systems. Segal and
Whitt (1988) describe a new version of the queueing network analyzer
that incorporates several features found in manufacturing lines. These
include machine breakdowns, changing lot sizes, product testing with
partial yields etc.

A feature of manufacturing systems that is not usually considered (at least explicitly) in queueing network models is the use of overtime. It is common to encounter plants where relatively heavily utilized stations work more hours per day than others. Also, in stations with multiple machines, the number of machines that operate during overtime is not necessarily the same as that used during regular time. Job releases to the shop, subsequent movement between stations, and departures from the system usually take place during regular hours.

In theory, such situations can be modeled as nonstationary systems. However, the exact analysis is extremely complex, if not impossible. The focus of the research on nonstationary queues has been on describing the time dependent behavior (for example, see Rothkopf and Oren 1979; Keller 1982; and references therein) and asymptotic behavior of the system (see Lemoine 1981; Wolff 1982; Heyman and Whitt 1984 and references therein). These results do not provide a mechanism to estimate performance measures such as the mean number of jobs, and mean waiting times in networks of queues with features described in the previous paragraph.

In this paper our objective is to present simple approximations that can be used in the context of networks of queues to model stations with overtime. In stations with multiple machines, we consider cases in which only some machines work overtime. We provide bounds on the performance of the approximation for single and multiple machine stations. For multiple machine stations, our results suggest that in most practical situations, the approximation is likely to be robust. These conclusions are supported by computational experiments reported in Section 3. One of these examples is a network with 13 stations and 10

2

products that is based on a semiconductor facility producing specialty devices. In addition, our computational results suggest that the parametric decomposition methods are fairly robust in the analysis of such systems.

The remainder of the paper is organized as follows. In the next section we present the approximations and the related results. In Section 3 we describe the computational experiments. In Section 4 we conclude with comments on the scope of these approximations.

2.0 Approximations to Model Stations with Overtime

2.1 Approximation for a Single Machine Station:

Consider a single machine station (System A) with the following characteristics:

(i) The machine is available for r hours of regular time every day.

(ii) The machine is available for ov hours of overtime every day, i.e., the total availability is (r+ov) hours per day.

(iii) Jobs arrive at the station at a rate λ per day with all arrivals taking place during regular time only.

(iv) Jobs that are completed during regular time leave the station immediately upon completion.

(v) Jobs that are completed during overtime leave the station at the beginning of the following day.

(vi) First come, first serve discipline (FCFS) is observed at the station.

We use the following notation to describe the system:

3

$a_i = (da_i, ta_i)$: arrival time of the ith job; da_i represents the day and ta_i the hour of arrival of the ith job.

b_i: service time of job i (in hours). Let $\bar{b}$ denote the mean service time.

$d_i = (dd_i, td_i)$: departure time of the ith job; dd_i and td_i represent respectively the day and hour of the departure epoch for the ith job.

ρ: utilization $= \lambda\bar{b}/(r + ov)$

We propose an approximate analysis of system A by analyzing a related single machine system A1 with the following characteristics:

(i)   The machine is available for r regular hours per day.

(ii)   The pattern of job arrivals is identical to that of system A, i.e., $a_i' = a_i$ for all i, where a_i' represents the arrival epoch of the ith job in system A1.

(iii)   The service time of the ith job, b_i', is scaled as follows
$$b_i' = kb_i, \text{ where } k = r/(r+ov), \quad i = 1, 2, \ldots$$

(iv)   All jobs leave the system immediately upon completion of service.

(v)   The priority discipline is FCFS.

Let $a_i' = (da_i', ta_i')$, $d_i' = (dd_i', td_i')$, b_i' and ρ' denote the corresponding variables in system A1.

$d_i - x$, $d_i' - x$: denote a time x hours before the departure of job i in system A and A1 respectively.

$d_i + x$, $d_i' + x$: denote a time x hours after the departure of job i in system A and A1 respectively. $|d_i - d_i'| \leq x$ indicates that the departure times of job i in A and A1 are within x hours.

4

The primary motivation for the approximation above follows from the fact that the service or processing rate measured as the number of jobs per day is the same in both systems A and A1. Hence, as long as the systems are busy, the amount of work completed in the two systems during an integral number of days are equivalent. The method has other advantages as well. First, the utilizations in the two systems A and A1 are the same, i.e., $\rho' = \rho$. This is important since, in most queueing systems, utilization is a major determinant of the performance measures. Second, the approximation is easy to apply in a network of queues. By applying this procedure at all stations, we obtain a related network in which all stations work the same hours. This system can be analyzed by conventional methods using either exact or approximate methods or simulation. Proposition 1 below suggests that the approximation is likely to be robust, particularly when the utilization is high.

<u>Definition:</u> A sample path is defined as a sequence of job arrivals with associated service times.

<u>Proposition 1:</u> In every sample path, the departure instances of corresponding jobs in systems A and A1 do not differ by more than k ov. Specifically, this implies the following:

(i)   If job i is completed during regular time, i.e.,

$0 \le td_i \le r$, then $d_i' \in \{d*, d**\}$, where d* and d** are

defined as follows:

d* : $(dd_i - 1, r - k \, ov + td_i)$ if $td_i < k \, ov$

   : $(dd_i, td_i - k \, ov)$ otherwise.

d**: $(dd_i, td_i + k \, ov)$ if $td_i \le r - k \, ov$

   : $(dd_i + 1, td_i + k \, ov - r)$ otherwise.

(ii)   If the ith job is completed during overtime in system A then

either $dd_i = dd_i'$ and $td_i = 0$, and $td_i' < k$ ov;

or $dd_i = dd_i' + 1$ and $td_i = 0$, and $r - k$ ov $\leq td_i' \leq r$.

We denote the above result as $|d_i - d_i'| \leq k$ ov. (The proof

of the proposition above is in Appendix 1.)

When both systems A and A1 are busy, equivalent amount of work is

completed (in the two systems) over any interval spanning an integral

number of days. Proposition 1 describes one consequence of this fact

which suggests that the performance measures in system A may be

estimated from the corresponding measures in A1. For example, the error

in estimating mean waiting time is not greater than k ov hours. (If the

waiting time is measured in days, the error is not greater than k ov/r

days.) The application of Little's law provides a corresponding error

bound on the mean number of jobs. As a result, if the utilization is

high and the average waiting times are large, the relative error in

measuring the mean waiting time and the mean number of jobs is likely to

be small.

Corollary 1: In a system where the interarrival times are independent

and identically distributed (iid), service times are iid and utilization

$\rho < 1$, so that we have steady state, the relative error in estimating

the mean number of jobs tends to zero as $\rho \to 1$.

This is an encouraging result. As we observed earlier in Section

1, we expect stations with relatively high utilization to work overtime.

Corollary 1 suggests that in those cases the approximation is good.

When the utilizations are low and the waiting times are small, we expect

that the absolute error in the estimates of the mean number of jobs will

not be significant. This fact was also observed in computational experiments not reported in this paper for brevity.

2.2 Multiple Machine Station:

Consider a m_1 machine work station (system B) with the following characteristics:

(i) The station operates with m_1 machines on regular time for r hours each day.

(ii) The station operates ov hours of overtime each day with $m_2 \leq m_1$ machines.

(iii) A FCFS discipline is observed.

(iv) Jobs completed during overtime depart from the system the following day, while those that are completed during regular time leave immediately upon completion of service.

As an approximation to system B, we propose a m_1 machine station (system B1) with the following characteristics:

(i) The station operates with m_1 machines during regular time of r hours each day.

(ii) The service times are scaled as $b_i' = k_1 b_i$,

where $k_1 = m_1 r/(m_1 r + m_2 ov)$

(iii) The priority scheme is FCFS and jobs leave the system immediately upon completion.

Observe that the service rate (measured with days as the unit of time) is $(r + ov)/\bar{b}$ jobs per day for the m_2 machines that work over time, while it is $r/\bar{b}$ for the remaining machines that only work regular hours. In system B1 all machines have a processing rate of $r/(k_1\bar{b})$ so that the aggregate processing rate in the two systems is the same. The motivation for the approximation follows from the fact that if all the

7

machines are busy, i.e., the number of jobs is not less then m_1, an equivalent amount of work is completed in the two systems over any interval that spans an integral number of days. The approximation has other advantages similar to those for the single machine case. Both systems B and B1 have the same utilization, and the use of B1 permits analysis of networks using conventional methods. When all machines work over time, i.e., $m_2 = m_1$, the results of proposition 1 hold for this approximation as well.

Proposition 2: Suppose that $m_2 = m_1$ and all machines work overtime in system B. Then in every sample path, the departure instances of corresponding jobs in systems B and B1 do not differ by more than k ov, i.e., $|d_i - d_i'| \leq$ k ov.

The proof of this proposition is in Appendix 1.

Observe that, in general when $m_2 < m_1$, B1 involves approximations in two dimensions -- in the number of machines and the hours worked. Hence, we expect that the approximation will be weaker than that for the single machine station. The following example suggests that in any sample path it is possible for individual jobs to have corresponding departure times (in systems B and B1) arbitrarily apart.

Example: Consider a job that arrives at the station (in system B) when the system is empty. Let the process time for the job be b hours. This represents a processing requirement of b/(r + ov) days. In system B1 the corresponding process time is $b' = k_1 b$ hours, which in turn corresponds to $\{m_1/(m_1 r + m_2 ov)\}$ b days of work.

Note that $\{m_1/(m_1 r + m_2 ov)\}$ b = $\{m_1(r+ov)/(m_1 r + m_2 ov)\}$ [b/(r+ov)]

> b/(r+ov), if $m_2 < m_1$ and ov > 0.

Even if the waiting time for the job is zero in system B1 (same as in B), the difference in departure epochs (measured in days) in systems B and B1 can be made arbitrarily large by choosing a correspondingly large b. This difference will increase further if the job has to wait in B1.

The example suggests that system B1 may not be a good approximation for specific jobs. However, we note that when the number of jobs in the system is not less than m_1 at any point in time, the work content in the two systems would be close. Also in that case the work completion rate on a daily basis will be the same in both systems.

Lemma 1: Consider a multimachine station with $m_2 < m_1$. Assume that the job arrivals and service times are such that the number of jobs in systems B and B1 are at least m_1 at any point in time. Then in every sample path,

$$wk(t) - \frac{(m_1 r + m_2 ov)}{m_1 r} wk'(t) \leq m_2 ov$$

where $wk(t)$, and $wk'(t)$ respectively represent the unfinished work content in systems B and B1. The unfinished work content is measured as the hours of processing required to complete the jobs in process and those in queue in the respective systems.

Furthermore, if t is integral in number of days, $k_1 wk(t) = wk'(t)$.

Lemma 1 provides an error bound on the estimates of the work content $wk(t)$. However, we note that it is not easy to obtain similar bounds for the number of jobs. With $m_2 < m_1$, the processing rates (measured as number of jobs per day) is not identical for all machines in system B, and can result in the assignment of jobs to different machines in the two systems. While the amount of work completed in B

9

and Bl are nearly equivalent, the number of jobs processed depends on sequencing and priority rules used during overtime operation. In Remark 1 below, we provide a parametric bound on the difference in the number of jobs in relatively heavily utilized systems. This bound depends on the ratio of maximum and minimum process times, and in some sense, represents a worst case error bound on the performance of the approximation.

Remark 1: Consider a multimachine station with $m_2 < m_1$. Assume that the job arrivals and service times are such that the number of jobs is at least m_1 in both systems B and Bl at any point in time. Also assume that at time $t = 0$, the processing of first job is initiated in both systems B and Bl. Let $b* = \lfloor b_{max}/b_{min} \rfloor$, where b_{max} and b_{min} respectively denote maximum and minimum process times. Consider systems B and Bl at the start of any day, i.e., at any point in time T, where T is integral in number of days. The number of jobs in the two systems do not differ by more than $(m_1 - 1)$ $(b* + 1) - 1$. (See Appendix 1 for proof of this assertion.)

These results are encouraging. Since in most manufacturing systems overtime is used only when there are waiting jobs, we expect that in most practical cases at stations that work overtime, the mean number of jobs will not be small and the approximation should be adequate. The results of the computational experiments described in the next section support this conjecture.

3.0 Computational Results:

We examined the quality of the approximations proposed in this paper by comparing estimates of the mean number of jobs (at each station) in a series of five experiments. In the first four experiments

we considered single product systems and tandem queues with two and three stations. In two cases we had stations with single machine and in the other two we tested problems with multiple machines. In the last experiment we considered a network derived from a semiconductor facility producing specialty devices. It consists of 13 single machine stations and 10 product classes. A secondary objective of the computations was to evaluate the robustness of the parametric decomposition approach (together with the approximations of this paper) in estimating the mean number of jobs in systems with overtime at some stations. These approximations are based on Shanthikumar and Buzacott (1981), Whitt (1983, 1985), and Bitran and Tirupati (1988a) and are summarized in Appendix 2. Since exact solutions are not available for most of the test problems, we used simulation as a benchmark. This is a common practice in evaluating queueing approximations. (For example, see Albin 1984, Whitt 1985, and Bitran and Tirupati 1988a.) In the design of the test problems the squared coefficient of variation (scv) of interarrival and service times were chosen not greater than one. This choice was motivated by our experience in manufacturing systems in which the corresponding scvs are typically smaller than one.

3.1 Experiments in Networks with One-machine Stations: In the first experiment we considered a single product, two station tandem queue shown in figure 3.1. The work schedules at the two stations were different. While station 2 operated a regular time of 8 hours each day, station 1 was available for additional (overtime) operation each day. The objective of this experiment was two fold. First, we examined the quality of approximations of section 2 in estimating the mean number of jobs at station 1. Second, we examined the impact of overtime operation

11

on the estimates at station 2. We expect that the impact on stations further downstream would be even less significant. In the design of the test problems we considered the following factors:

(i) Overtime operation at station 1: We considered three levels of overtime -- 0, 1 and 2 hours each day. The availability varied from 8 to 10 hours per day. These systems are denoted as (8,8), (9,8) and (10,8) respectively.

(ii) The arrival rate was set at 1 job per day (mean interarrival time 8 hours). Experiments in which the arrival rates are larger than one are described in Section 3.3. Job arrivals take place during regular time hours only. The interarrival times were assumed to have independent and identical distributions (iid). We considered two distributions -- exponential and second order Erlang for interarrival times.

(iii) Service times at each station were assumed to be iid. We considered two distributions for service times -- exponential and second order Erlang. In these problems, the service time distribution was the same at both stations.

(iv) The utilization at station 2 was set at 0.8. At station 1 we considered two levels of utilization; 0.75 and 0.85. Note that station utilization together with the number of hours of operation and choice of distribution (Erlang or exponential) completely specifies the service time distribution at each station.

The number of test problems in this experiment was 24. The simulation estimates of the mean number of jobs is presented in Table 3.1. The table also presents the corresponding estimates by the

decomposition approach using the approximations of this paper. These results are very encouraging. First, note that the variation in the mean number of jobs with overtime level is not significant. Case (a) in Table 3.1 represents the proposed approximation for systems (b) and (c). The estimates of (a) compare very well with those of (b) and (c) which suggests that the approximation adjusting for overtime operation should be adequate in most practical situations. Also note that the impact of overtime on the mean number of jobs at station 2 is also insignificant. The approximations for estimating the number of jobs seem to be very effective. The table also provides, for each problem, the maximum of the absolute relative errors in estimating the mean number of jobs in systems (a), (b) and (c) using the decomposition methods. These errors range between .88% and 7.06% and are typical of the decomposition approach. These results suggest that the performance of the decomposition method does not deteriorate when combined with the approximation of Section 2.1.

The simulation estimates reported in the table are based on six batch runs of 100,000 jobs each and represent the average of 600,000 jobs. We felt that these runs were sufficiently long to provide representative estimates. To illustrate the accuracy of the simulation estimates we present, for two of the test problems, the standard deviation of the six batch estimates for the mean number of jobs. The two examples correspond to the following cases:

Example 1: System (10,8); $\rho_1 = 0.85$; Erlang distribution for interarrival and service times.

Example 2: system (9,8); $\rho_1 = 0.75$, Exponential distribution for interarrival and service times.

13

The standard deviations of the mean number of jobs for the first example were .0512 and .0293 at stations 1 and 2 respectively. The corresponding figures for the second example were .0429 and .0766. These figures are representative of the results obtained. Note that the differences in the mean number of jobs with varying overtime is of the same order as the standard deviation of the simulation estimates. This again suggests that these variations are insignificant.

In the second experiment we considered a three station network shown in figure 3.2. The objective in this test was to examine the quality of the proposed approximation with a wider variation in the operating schedules. We simulated the test problems with the following schedules:

(a) All 3 stations working 8 hours per day (no overtime). This is denoted by (8,8,8).

(b) The first and the third stations work 8 hours each day while the second works 10 hours; (8,10,8).

(c) First station works 6 hours, second 10 hours and the third 8 hours per day; (6,10,8).

(d) The three stations work 8, 10 and 6 hours respectively; (8,10,6).

These cases represent a wide variation in operating hours that accommodate many manufacturing applications. (Note that the ratio of maximum to minimum working hours is 1.67.) Observe that system (a) above is the proposed approximation for systems (b), (c) and (d). For each of these systems we simulated 4 test problems with the following parameters.

14

(i)   The arrival rate was 1 job per day. (Experiments in which

         the arrival rates are larger than one are described in

         section 3.3). The interarrival times were assumed to be

         iid. We considered two alternative distributions --

         exponential and second order Erlang.

(ii)   Service times at each station were assumed to be iid. We

         considered two distributions for service time -- exponential

         and second order Erlang. Again in each problem, the scv of

         service time was the same at all stations.

(iii)   The utilization at stations 1 and 3 was set at 0.8, while at

         station 2 it was set at 0.85.

The computational results for the 16 cases are presented in table
3.2. It is clear from these results that (a) is indeed a very good
approximation for (b), (c) and (d). The variation in the simulation
estimates is small. (It is of the same order as the standard deviation
of the batch results.) The performance of the approximation is also
encouraging. The table presents the maximum absolute relative error in
estimating the mean number of jobs in systems (a), (b), (c) and (d)
using the decomposition method. These errors range between 1.25% and
8.27%. Again, the simulation estimates are based on six batch runs of
100,000 jobs in each run and the accuracy is similar to those in the
first experiment.

3.2 Experiments in Tandem Networks with Multiple-Machine Stations: In
this section we report the results of tests in networks with multiple
machine centers. The network for the first experiment was the two
station system of figure 3.1. The number of machines at stations 1 and
2 was 3 and 2 respectively. Station 2 operated on regular time schedule

15

of 8 hours per day and we considered three levels of overtime at station 1. During overtime operation only 2 machines were available at this station. Thus, we had the following three systems defined by the overtime levels.

(a) No overtime and both stations work 8 hours per day. This is denoted by (8,8).

(b) Overtime of one hour per day at station 1, i.e., system (9,8).

(c) System (10,8) with two hours of overtime at station 1.

For each system we simulated four problems by varying the distribution of interarrival and service times. The alternative distributions considered were similar to the test problems in Section 3.1 and consisted of exponential and second order Erlang distributions. In all test problems the utilization at station 1 was 0.85 and 0.8 at station 2. Observe that system (a) represents our proposed approximation for systems (b) and (c). In these experiments, we implemented the FCFS discipline in the following manner. When the number of jobs in the systems was not greater than m_1 all the jobs are processed during overtime. When the number of jobs was greater than m_1, the machines that were processing the first m_2 jobs (jobs with the highest priority) work overtime.

The results presented in Table 3.3 support conclusions similar to those observed in the single machine experiments of Section 3.1. The mean number of jobs at each station in systems (a), (b) and (c) are very close and well within the accuracy of the simulation estimates. This suggests that the proposed approximation is very reasonable. Table 3.3 presents for each problem, the maximum percentage error of the absolute deviation of the approximation from the simulation results. These

16

errors range from 1.36% to 6.69% which is acceptable in most applications.

The simulation estimates in this experiment are based on twelve batch runs with 50,000 jobs in each run. We felt a total of 600,000 jobs was sufficient to provide good estimates. To illustrate the accuracy of these results, we provide the standard deviation of the batch estimates for the following two cases:

Example 1: System (c); distribution of interarrival and service times are Erlang.

Example 2: System (b); interarrival and service time distributions are Erlang and exponential respectively.

In the first example, the standard deviation of the mean number of jobs was .0716 and .0991 at stations 1 and 2 respectively. The corresponding figures for the second case are .1291 and .0676.

The three station system of Figure 3.2 was the network for the second experiment. The number of machines was 2 at stations 1 and 3, and 3 at station 2. The design of the test problems was similar to the previous experiments. Station 2 had overtime operation while the other two worked during regular time hours only. Also, during overtime only two machines were used. We defined three systems with varying overtime levels. These are (a) (8,8,8) -- No overtime. All stations work 8 hours; (b) (8,9,8) -- overtime of 1 hour at station 2; and (c) (8,10,8) -- overtime of 2 hours at station 2. As in the previous experiments, for each of these systems we tested four problems by varying the distribution of interarrival and service times. The utilizations were set at 0.8, 0.85 and 0.8 at stations 1, 2 and 3 respectively.

The results presented in table 3.4 support the conclusions of the previous experiments. Observe that estimates from system (a) would be a very good approximation for systems (b) and (c). The differences in the simulation figures are not significant. The approximations continue to perform well. The range of the maximum percentage absolute error (relative to the simulation values) was 0.69% to 8.00%. These errors are typical of the parametric decomposition approach in the analysis of general networks. It is interesting to note that differences between estimates in column (d) and columns (a), (b), (c) are comparable. Hence it appears that the performance of the decomposition approach when combined with the approximations proposed in Section 2.2 does not deteriorate.

3.3 A Network Example: In this section we describe an experiment with a network that models the manufacturing facility of a semiconductor company producing specialty semiconductor devices. The facility consisted of 13 single machine stations and processed jobs classified into 10 product families. The routing for jobs in each class was deterministic and described in Table A3.1 of Appendix 3. Observe that the number of operations for each job is between 7 and 13 and a job returns to the photolithography department (station 2) several times. The arrival rate of jobs in each family was 1 per day and the distribution of interarrival times are presented in Table A3.1. The resulting net arrival rates at the stations in the network range from 3 to 25 jobs/day. The station data consisting of operating schedule, utilization and process time distribution is described in Table A3.2. Note that the service time distribution is completely specified by this data together with the net arrival rate computed from product data.

Table A3.2 also specifies an average dollar value (v_j) for each job at station j. The work-in-process (wip) at station j is measured as the product of v_j and the mean number of jobs (L_j). The total wip in the system is obtained by adding the station wips. Observe that the work schedules of the stations in the network vary considerably. For example, stations 1, 4 and 6 work only regular time hours while 8 and 12 operate more than 10 hours. These variations are typical of many manufacturing systems. We refer to this as system (a).

In Table 3.5 we present three estimates for the mean number of jobs at each station and total system wip. Estimates (a) in the table are based on simulation of the network described by Tables A3.1 and A3.2. Estimates (b) are obtained by simulating the network derived by applying the approximations presented in Section 2. In this network all stations work the same schedule. The station utilizations match those in system (a). The third estimate is obtained by using the parametric decomposition method to analyze the network of system (b). The surprisingly close match between the three estimates suggests that the approximations are quite effective. For example the error in estimating the total number of jobs is 2.93%. The error in estimating the total wip is even lower at 1.68%. This limited computational evidence suggests that the performance of the decomposition approach does not deteriorate when combined with the approximations introduced in Section 2.

4.0 Conclusions:

We have considered queueing networks where stations operate on different work schedules resulting in nonstationary product flows. This problem is not usually considered in the traditional approaches to the

19

analysis of queueing networks. In this paper we suggest a mechanism by which we construct a related network where all stations work the same schedule. We also present bounds on the performance of the approximation. We propose that performance measures in the original network can be estimated by analyzing the derived network, either by simulation or by exact or approximate methods. The experimental results of this paper are very encouraging and suggest that the approximation scheme is robust. Also the test problems provide additional evidence regarding the effectiveness of the decomposition methods of Shantikumar and Buzacott (1981), Whitt (1983, 1985), and Bitran and Tirupati (1988a) in estimating network performance.

References:

Albin, S. L., "Approximating a Point Process by a Renewal Process, II: Superposition Arrival Processes to Queues," Operations Research, 32, 5, 1984, 1133-1162.

Baskett, F., K. M. Chandy, R. R. Muntz and F. G. Palacios, "Open, Closed and Mixed Networks of Queues with Different Classes of Customers," J. Assoc. Comput. Mach., 22, 2, 1975, 248-260.

Bitran, G. R. and D. Tirupati, "Multiproduct Queueing Networks with Deterministic Routing: Decomposition Approach and the Notion of Interference," Management Science, 34, 1, 1988a, 75-100.

Bitran, G. R. and D. Tirupati, "Approximations for Product Departures from a Single-Server Station with Batch Processing in Multi-Product Queues," working paper #87-12-1, IC^2 Institute, The University of Texas at Austin, 1988b, forthcoming in Management Science.

Heyman, D. P. and W. Whitt, "The Asymptotic Behavior of Queues with Time Varying Arrival Rates," J. Appl. Probl., 21, 1984, 143-156.

Jackson, J. R., "Jobshop like Queueing Systems," Management Science, 10, 1, 1963, 131-142.

Keller, J. B., "Time Dependent Queues," SIAM Rev., 24, 1982, 401-412.

Kelly, F. P., "Networks of Queues with Customers of Different Types," Journ. Appl. Prob., 12, 1975, 542-554.

Lemoine, A. J., "Networks of Queues - A Survey of Equilibrium Analysis," Management Science, 24, 4, 1977, 464-481.

Lemoine, A. J., "On Queues with Periodic Poisson Input," J. Appl. Prob., 18, 1981, 889-900.

Lemoine, A. J., "On Random Walks and Stable GI/G/1 Queues," Math. Operat. Res., 1976, 159-164.

Rothkopf, M. H. and S. S. Oren, "A Closure Approximation for the Nonstationary M/M/s Queue," Management Science, 25, 1979, 522-534.

Segal, M. and W. Whitt, "A Queueing Network Analyzer for Manufacturing," Proc. 12th Int. Teletraffic Congress, Torino, Italy, June 1988.

Shanthikumar, J. G. and J. A. Buzacott, "Open Queueing Network Models of Job Shops," Int. J. Prod. Res., 19, 3, 1981, 255-266.

Wolff, R. W., "Poisson Arrivals see Time Averages," Operations Research, 30, 1982, 223-231.

Whitt, W., "The Queueing Network Analyzer," Bell System Technical Journal, 62, 9, 1983, 2779-2815.

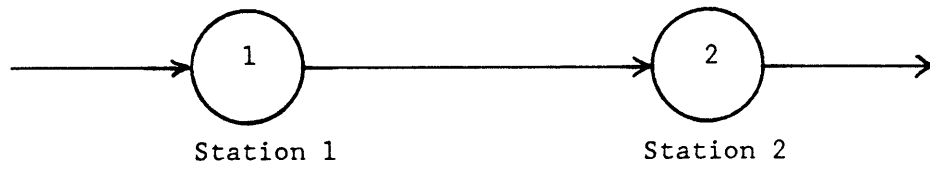Whitt, W., "Approximations for the GI/G/m Queue," Dec. 1985, forthcoming in Adv. Appl. Prob.
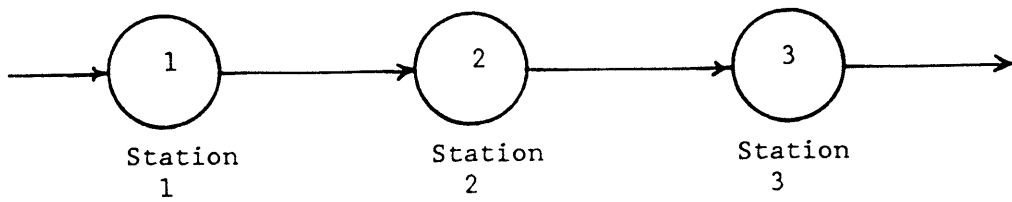
Figure 3.1

Figure 3.2

Table 3.1 Estimates of the Mean Number of Jobs

Utilization at Station 1		0.75				0.85			
SCV of interarrival times		1		0.5		1		0.5	
SCV of Service times		1	0.5	1	0.5	1	0.5	1	0.5
	(a)	2.98	2.45	2.33	1.80	5.68	4.50	4.36	3.16
Station	(b)	3.04	2.47	2.31	1.77	5.64	4.45	4.24	3.09
1	(c)	3.00	2.45	2.36	1.82	5.72	4.40	4.40	3.15
	(d)	3.00*	2.44*	2.37	1.81	5.67*	4.46*	4.39	3.19
	(e)	1.32	1.21	2.60	2.26	0.88	1.36	3.54	3.24
	(a)	3.99	2.81	3.40	2.30	3.96	2.63	3.64	2.30
Station	(b)	3.94	2.83	3.44	2.32	4.01	2.61	3.56	2.30
2	(c)	3.95	2.79	3.42	2.31	4.02	2.55	3.70	2.30
	(d)	4.00*	2.73	3.64	2.33	4.00*	2.59	3.77	2.33
	(e)	1.52	3.53	7.06	1.30	1.01	1.54	5.90	1.30

Note: 1. (a), (b), (c) represent respectively simulation estimates for the mean number of jobs with overtime levels of 0, 1 and 2 hours per day at station 1, i.e, availability of 8, 9, and 10 hours/day.

2. (d) is the estimate of the mean number of jobs based on the decomposition approach (formulae in Appendix 2) using the approximations of Section 2.1.

3. (e) is the maximum % absolute deviation of estimate d from the estimates (a), (b) and (c).

4. * exact result.

Table 3.2 Estimates of the Mean Number of Jobs

	STATION	a	b	c	d	app.	e
ca = 1 cs = 1	1	3.99	3.94	3.97	3.93	4.00	1.50
	2	5.70	5.67	5.86	5.62	5.67	3.24
	3	3.97	3.95	4.02	4.00	4.00	1.25
ca = 1 cs = 0.5	1	3.21	3.26	3.20	3.24	3.20	1.84
	2	3.89	3.89	3.99	3.89	3.66	8.27
	3	2.54	2.55	2.60	2.57	2.43	6.54
ca = 0.5 cs = 1	1	3.10	3.02	3.11	3.13	3.13	3.64
	2	4.93	4.95	5.01	5.00	5.22	5.88
	3	3.71	3.78	3.80	3.76	3.92	5.66
ca = 0.5 cs = 0.5	1	2.28	2.32	2.28	2.31	2.33	2.19
	2	3.13	3.20	3.29	3.23	3.19	3.04
	3	2.29	2.30	2.36	2.31	2.33	1.27

Note: 1. a, b, c, and d represent simulation estimates with the
following work schedules at stations 1, 2, and 3.
a - (8,8,8), b - (8,10,8), c- (6,10,8), d - (8,10,6)

2. app. column represents the estimates based on the parametric
decomposition approach with the approximations of this paper.

3. e is the maximum % absolute error of the approximation
relative to the a, b, c and d estimates.

4. ca is the scv of interarrival times and cs is the scv of
service time.

Table 3.3 Estimates of the Mean Number of Jobs

SCV of interarrival times		1		0.5	
SCV of service times		1	0.5	1	0.5
Station					
	a	6.72	5.67	5.40	4.42
	b	6.81	5.46	5.37	4.40
1	c	6.76	5.54	5.27	4.47
	d	6.69	5.67	5.46	4.46
	e	1.76	3.85	3.61	1.36
	a	4.53	3.21	4.03	2.99
	b	4.37	3.23	4.12	2.99
2	c	4.52	3.24	4.04	3.05
	d	4.44	3.42	4.19	3.19
	e	1.99	6.54	3.97	6.69

Note: 1. a, b and c represent respectively simulation estimates with 0, 1 and 2 hours of overtime at Station 1 (availability of 8, 9 and 10 hours/day).

2. d is the estimate computed from formulae in Appendix 2 with the approximations of Section 2.

3. e is the maximum % absolute error of d relative to estimates a, b, and c.

Table 3.4 Estimates of the Mean Number of Jobs

ca	cs	STATION	a	b	c	d	e
		1	4.55	4.41	4.37	4.44	2.42
1	1	2	6.83	6.65	6.73	6.69	2.05
		3	4.48	4.40	4.46	4.44	0.91
		1	3.67	3.73	3.76	3.74	1.91
1	0.5	2	5.18	5.15	5.23	5.17	1.15
		3	3.17	3.18	3.19	3.32	4.73
		1	3.50	3.55	3.61	3.58	2.29
0.5	1	2	5.81	5.91	5.82	6.23	7.23
		3	4.14	4.14	4.11	4.35	5.84
		1	2.89	2.91	2.88	2.89	0.69
0.5	0.5	2	4.54	4.56	4.43	4.74	7.00
		3	3.04	3.02	3.00	3.24	8.00

Note: 1. ca and cs are the scv of the interarrival and service times
         respectively.

      2. a, b and c are simulation estimates with availability of 8, 9
          and 10 hours respectively at Station 2. At Stations 1 and 3
          the availability is 8 hours.

      3. d is the estimate computed from formulae in Appendix 2 with
          the approximation of Section 2.

      4. e is the maximum % absolute error of d relative to estimates
          a, b and c.

Table 3.5 Results of Tests with 10 Product, 13 Station Network

| | mean number of jobs | | |
Station	(a) Simulation of network with overtime	(b)	(c)
1	2.19	2.20	1.92
2	2.63	2.55	2.44
3	2.51	2.46	2.46
4	1.57	1.56	1.53
5	1.35	1.34	1.33
6	1.20	1.17	1.09
7	1.19	1.19	1.17
8	1.44	1.38	1.46
9	1.54	1.51	1.52
10	1.18	1.19	1.22
11	1.64	1.58	1.77
12	1.99	1.88	1.92
13	1.34	1.29	1.32
Total	21.77	21.29	21.15 (2.93)*
Total System WIP ($)	30739	29920	30231 (1.68)**

Note: 1. a is the simulation estimate with the network data as specified in Appendix 3.

2. b is the simulation estimates with the network derived from approximations of Section 2.

3. c is the computed value for the network in b, with parametric decomposition approach (formulae in Appendix 2).

4. *maximum % error of approximation (c) relative to simulation values (a) and (b).

5. ** maximum % error in WIP value of (c) relative to simulation values (a) and (b).

Appendix 1

Proof of Proposition 1:

First, we show that the proposition is true for job 1 and then, by induction, we show that it holds for any job i, $i \geq 2$. Recall that by notation $(d_i - x, \quad d_i + x)$ denotes an interval of length 2x hours, x hours before and after departure of job i in System A. Note that, both A and A1 are empty when job 1 arrives and the corresponding waiting times are zero. Also note that $a_1 : (da_1, ta_1) = a_1' : (da_1', ta_1')$. We make the following observations about departure times d_1 and d_1':

(i)   If the service time b_1 is an integral number of days, i.e.,

$b_1 = n_1(r+ov)$ for some positive integer n_1, then the corresponding service time in A1 is also equivalent to n_1 days of work, i.e.,

$b_1' = n_1 r$, and the departure instants d_1 and d_1' are the same.

(ii)   Let the service time for job 1 in A be equivalent to n_1 days and n_2 hours, i.e., $b_1 = n_1(r+ov) + n_2$, where n_1 is a non-negative integer and $0 < n_2 < (r+ov)$. Then the difference in the departure times d_1 and d_1' is entirely due to n_2.

Let $y = da_1 + \lfloor (b_1 + ta_1)/(r+ov) \rfloor$,

$z_1 = (r - ta_1)/k$, and $z_2 = ta_1 + k(r - ta_1)$,

where $\lfloor x \rfloor$ denotes the largest integer not greater than x, $\lceil x \rceil$ denotes the smallest integer not smaller than x, and k is the scaling factor for process times as defined in Section 2.1. Observe that, depending on the value of n_2, the completion of service for job 1 and its departure occur either on day y or y+1 in Systems A and A1.

We consider the following four mutually exclusive and collectively exhaustive cases to compare the departure times d_1 and d_1'. The reader

may find figure A1.1, which illustrates these cases, useful in understanding the arguments presented below.

Case (1): $0 < n_2 \leq (r - ta_1)$. Service is completed during regular time of day y in both A and A1. In this case,

$dd_1 = y$, $td_1 = ta_1 + n_2$; $dd_1' = y$ and $td_1' = ta_1 + kn_2$.

Since $td_1 - td_1' = (1-k)n_2 = n_2 \, ov/(r+ov) \leq r \, ov/(r+ov) = k \, ov$, we have $|d_1 - d_1'| \leq k \, ov$.

Case (2): $r - ta_1 < n_2 \leq z_1$. In this case the service for job 1 is completed during overtime of day y (departure at the beginning of day y+1) in System A and during regular time of day y in A1.

$dd_1 = y + 1$, $td_1 = 0$; $dd_1' = y$ and $td_1' = ta_1 + kn_2$

By definition of z_1 and z_2, $z_2 \leq td_1' \leq r$.

Since $r - z_2 = (1-k) (r-ta_1) \leq (1-k)r = ov \, r/(r+ov) = k \, ov$, we have $|d_1 - d_1'| \leq k \, ov$.

Case (3): $z_1 < n_2 \leq (r + ov - ta_1)$. In this case service for job 1 is completed during overtime of day y (and departure at the start of day y + 1) in A and during day (y+1) in A1.

$dd_1 = y + 1$, $td_1 = 0$; $dd_1' = y + 1$ and $td_1' = k(n_2 - z_1)$

The bounds on n_2 imply that $0 \leq td_1' \leq k \, ov$ and $|d_1 - d_1'| \leq k \, ov$.

Case (4): $(r + ov - ta_1) < n_2 < (r+ov)$. In this case the job is completed during regular time of day y + 1 in both systems A and A1.

$dd_1 = y + 1$, $td_1 = n_2 - (r + ov - ta_1)$; $dd_1' = y + 1$ and $td_1' = k(n_2 - z_1)$.

Note that $td_1' - td_1 = k(n_2 - z_1) - n_2 + (r + ov - ta_1) = ov[1 - n_2/(r + ov)]$.

Since $n_2 < (r + ov)$, $td_1' - td_1 > 0$, and

$n_2 > r + ov - ta_1$, $td_1' - td_1 < ov \, ta_1/(r + ov) < k \, ov$.

Hence $|d_1 - d_1'| \leq k$ ov and the proposition follows for job 1.

To prove the proposition for the general case, say job i, $i \geq 2$, assume that the result is true for jobs 1, 2, ...i-1. To prove the proposition we consider four mutually exclusive and collectively exhaustive cases.

<u>Case (1)</u>: At time a_i, the arrival epoch of job i, both A and A1 are empty. This case is identical to that of job 1 and the result follows.

<u>Case (2)</u>: At time a_i, A is busy and A1 is idle.

Let j, j < i denote the last job before i which sees an idle server in System A upon arrival. (Clearly such a job exists.) Let s_j and s'_j denote the start time (day and hour) of service for job j in systems A and A1. Let s_i and s'_i be the corresponding times for job i. We make the following observations on the operation of A and A1.

1. $s_j = a_j$ and $s'_j \geq s_j$

2. The machine in A is busy from a_j till the completion of service job i.

3. $s'_i = a_i$ and $s_i > s'_i$.

To derive the bounds on $|d_i - d_i'|$, consider the following alternatives:

(a) delay the start of service for job i in A1 such that the start time is the same in both systems, i.e., it is s_i in both A and A1. Denote the departure time for job i in system A1 so modified as d1'.

Clearly d1' > d_i'. By an analysis similar to that of job 1, we obtain, $d_i - k$ ov $\leq$ d1' $\leq d_i + k$ ov

which implies that $d_i + k$ ov > d_i' (1)

(b) Consider a system A1 in which the processing of job j starts at the same time as in A, i.e., $s'_j = s_j$ and jobs j, j+1,... i are processed without any intervening idle time. Let d2'

denote the departure time of job i in system A1. It is clear that $d2' < d_i'$.

By treating jobs j through i as a single job with process time $b_j + b_{j+1} + \ldots + b_i$, and an analysis similar to job 1, we obtain $d_i - k \, ov \leq d2' \leq d_i + k \, ov$ which implies that

$$d_i - k \, ov < d_i' \tag{2}$$

From (1) and (2), we have $d_i - k \, ov < d_i' < d_i + k \, ov$, and the proposition follows.

Case (3): At time a_i, A is idle and A1 is busy.

This case can be analyzed in a manner similar to case (2). Let j denote the last job before i which sees an idle machine in A1 upon arrival. Let s_j and s'_j (s_i and s'_i) respectively denote the start of service for job j (job i) in systems A and A1. Note that:

1. $s_j' = a_j$ and $s_j \geq s'_j$

2. $s_i = a_i$ and $s'_i > s_i$

3. System A1 is busy from a_j till the completion of service of job i.

Consider the following alternatives:

(a) Delay the start of service for job i in A such that $s_i = s'_i$. Let d1 denote the departure time of job i in A. Then $d1 > d_i$, and $d_i' - k \, ov \leq d1 \leq d'_i + k \, ov$ which implies

$$d_i < d_i' + k \, ov \tag{1}$$

(b) Consider a system A in which j starts at time a_j and jobs j, j+1, ... i are processed without any intervening idle time. Let d2 denote the departure time of job i. Then $d2 < d_i$, and $d_i' - k \, ov \leq d2 \leq d_i' + k \, ov$ which implies that

$$d_i' - k \, ov < d_i \tag{2}$$

From (1) and (2) $d_i' - k\,ov < d_i < d_i' + k\,ov$ and the

proposition follows.

<u>Case (4)</u>: At time a_i both A and A1 are busy.

In this case note that jobs i-1 and i are processed in sequence without

any interruption. The completion of service for job i is determined by

the time taken to process $(b_{i-1} + b_i)$ in System A and $k(b_{i-1} + b_i)$ in

system A1 from the start of service for job i-1 in both systems.

Note that, by assumption $|d_{i-1} - d_{i-1}'| \leq k\,ov$ and this is independent

of process time b_{i-1}.

Replacing the process time of job i-1 by $(b_{i-1} + b_i)$ and repeating the

analysis for job i-1, we obtain $|d_i - d'_i| \leq k\,ov$.

<u>Proof of Proposition 2</u>

Without any loss of generality we assume that

(i)   If more than one machine is available, a job is assigned to the

      machine which became free first, and

(ii) The machines are numbered such that the first m_1 jobs are

      respectively assigned to machines 1, 2, ...m_1.

<u>Lemma A1</u>: Consider a system B in which all m_1 machines work overtime,

      i.e., $m_2 = m_1$. suppose m_1 jobs with process times b_1, b_2, ... b_{m_1}

      are assigned to the m_1 machines and begin service at the same time

      (say t = 0). Assume that in B1 also the jobs begin service at time

      t = 0. Then the order in which the m_1 jobs complete service is the

      same in B and B1. And the sequence in which the machines become

      free is also identical in B and B1.

<u>Lemma A2</u>: In any given sample path defined by a sequence of job

      arrivals and associated service times, let

$J_j(J_j')$ be the set of jobs assigned to the machine j, j = 1, 2, ...m_1 in system B(B1).

Then $J_j = J_j'$, j = 1, 2, ... m_1.

Outline of Proof of Proposition 2:

Observe that for any given sample path, B can be decomposed into m_1 single machine subsystems A_j for each machine j, j = 1, 2, ...m_1. Jobs in J_j represent the sequence of arrivals and associated service times for A_j.

Let $A1_j$ represent the approximation of section 2.1 for system A_j, and let $d'_{j,i}$ represent departure times of job i in $A1_j$, i ϵ J_j, j = 1, 2, ... m_1.

From proposition 1, we have $|d_i - d'_{j,i}| \leq$ k ov, i ϵ J_j, j = 1, 2, ... m_1.   (a)

Also observe that B1 can likewise be decomposed into m_1 subsystems $B1_j$, j = 1, 2, ... m_1.

Furthermore, from Lemma A2 it follows that $B1_j$ and $A1_j$ are identical, i.e., both systems have same machine schedules, with identical job arrivals and service times.

Hence $d_i' = d'_{j,i}$, i ϵ J_j, j = 1, 2, ...m_1 (b)

From (a) and (b), we obtain $|d_i - d_i'| \leq$ k ov and the proposition follows.

Remark 1

Consider a multimachine system with $m_2 < m_1$. Assume that the job arrivals and service times are such that the number of jobs is at least m_1 in both systems B and B1 at any point in time. Also assume that at time t=0, the processing of first job is initiated in both systems B and B1. Let b* = $\lfloor b_{max}/b_{min} \rfloor$ where b_{max} and b_{min}

respectively denote maximum and minimum process times. Consider

systems B and B1 at the start of any day, i.e., at any point in

time T, where T is integral in number of days. Then, the number of

jobs in the two systems do not differ by more than $(m_1 - 1)$ $(b^* +$

$1) - 1$.

<u>Outline of Proof</u>:

Since the number of jobs is at least m_1 in both B and B1, the

amount of work completed during the interval $[0,T]$ are equivalent,

i.e., $WP(T) = (m_1 r + m_2 ov)T$, $WP'(T) = m_1 rT$, and $WP'(T) = k_1 WP(T)$,

where $WP(T)$ and $WP'(T)$ respectively represent the amount of work

processed in B and B1 during the interval $[0,T]$ and k_1 is as

defined in Section 2.2.

Let $N(N')$ denote the set of jobs processed either completely or

partially in $B(B1)$ during $[0,T]$. Let $NC(NC')$ denoted the set of

completed jobs in $B(B1)$ and $NP(NP')$ the set of in-process jobs.

Let $|NC| = n$ and $|NC'| = n'$. Note that $|NP| = |NP'| = m_1$.

Let x_j (x_j'), $j = 1, 2, \ldots m_1$ be the amount of work completed on

job in process on machine j in $B(B1)$.

We make the following observations:

(i) $\quad WP(T) = \quad \sum_{i \in NC} b_i + \sum_{j=1}^{m_1} x_j$

(ii) $\quad WP'(T) = \quad \sum_{i \in NC'} b_i' + \sum_{j=1}^{m_1} x_j' = k_1 WP(T)$

(iii) Since the priority discipline is FCFS, the start of service of job
     i implies that the number of completed jobs is $i - m_1 + 1$.

<u>CLAIM 1</u>: $n' \geq n - (m_1 - 1) (b^* + 1) + 1$

Case (a): $N' \not\subseteq N$: In this case there exists a job i, which is

processed either partially or completely, only in B1. A FCFS
discipline implies that $i > n + m_1$, and from observation
(iii) above, it follows that $n' \geq i - m_1 + 1 > n + 1 > n - (b* + 1)(m_1 - 1) + 1$.

Case (b): $N' \subseteq N$: Consider the schedule for processing jobs in the set
N in system B1 with the following process times:

$b1'_i = k_1 b_i$ if $i \in NC$

$= k_1 x_j$ if $i \in NP$ and is being processed on machine j.

Note that $\sum_{i \in N} b1'_i = WP'(T)$.

A job is considered in process, if it is assigned a machine
and the service is not yet complete at the end of day T. We
note that:

(i)   At most $m_1 - 1$ jobs are in-process.

(ii) At least one of the in-process jobs in B1 is also in-process in B.
This is clearly true for job $(n + m_1)$ which is the last job to
enter service.

(iii) At most $m_1 - 1$ jobs in the set NP can be scheduled to start prior
to those in NC.

(iv) If all jobs in the set NP start service only after the jobs in NC,
then the number of completed jobs in B1 is at least $n - m_1 + 1$.

(v)   For each job j in NP that is scheduled to start before some jobs
in NC it is possible that amount of processing is greater than
$k_1 x_j$. In that case, each such job can delay completion of at most
b* jobs.

(vi) (iii) and (iv) together imply that $n' \geq n - m_1 - b*(m_1 - 1) = n + 1 - (m_1 - 1)(b* + 1)$.

CLAIM 2: $n \geq n' + 1 - (m_1 - 1)(b* + 1)$. The arguments are similar to those in Claim 1. Claim 1 and 2 together imply that $|n - n'| \leq (m_1 - 1)(b* + 1) - 1$.

Since job arrivals are identical in B and B1, the above results implies that the number of jobs in the two systems do not differ by more than $(m_1 - 1)(b* + 1) - 1$.
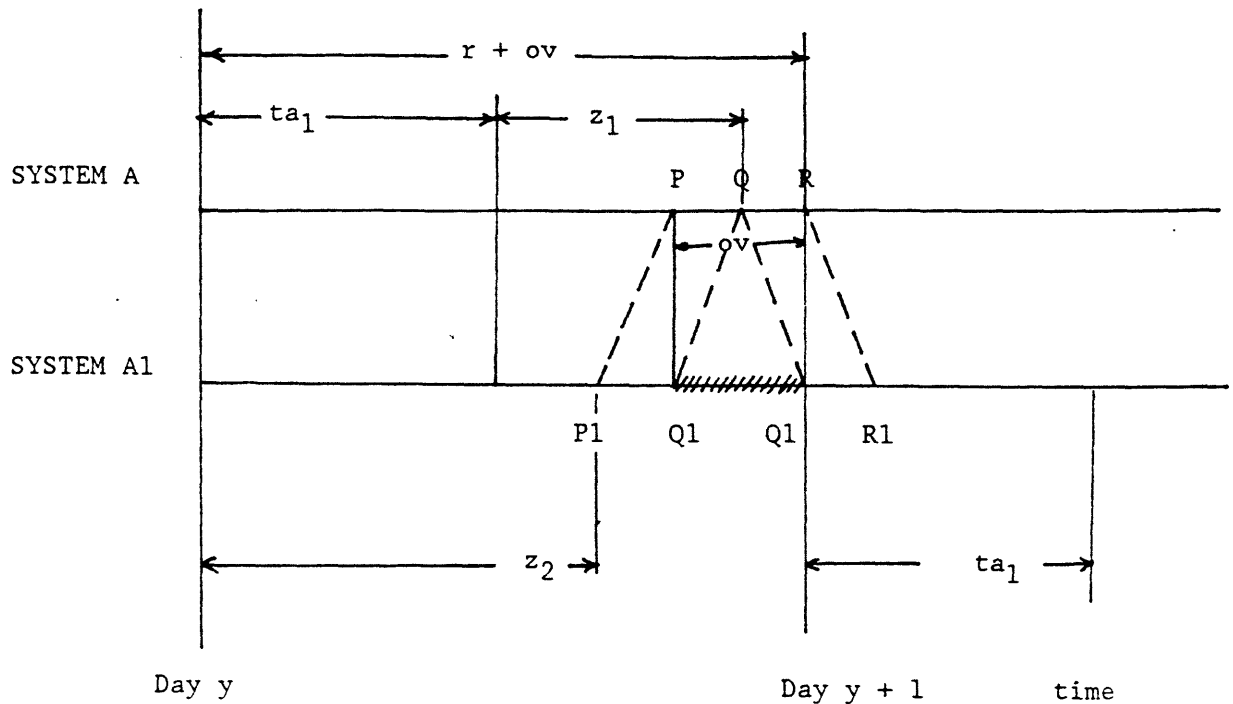
denotes machine not available in A1 (no overtime).

P and R represent start and completion of overtime in A.

P1 and R1 represent points in time in A1 that correspond to P and R in A. (Corresponding points imply equivalent amount of work completion in the two systems.)

Q1 is the corresponding point for Q in A and represents the start of day y+1 and the end of day y.

FIGURE A1.1

Appendix 2

In this appendix, we present a summary of approximations based on the
parametric decomposition approach for estimating mean number of jobs in
open queueing networks with multiple products and deterministic routing.
For details of these results, the reader is referred to Shanthikumar and
Buzacott (1981), Whitt (1983, 1985) and Bitran and Tirupati (1988a).
The approximations involve two steps: (a) determination of the mean and
squared coefficient of variation (scv) of product flows in the network,
and (b) estimation of the mean number of jobs in a $GI|G|m$ queue.

Estimation of the mean and scv of product flows in the network:

Let m = number of products

J = number of stations

a_i = arrival rate of product i

u_i = number of operations required for product i

$u_{i,k}$ = station visited by product i at step k

$cd_{i,k}$ = scv of product stream i after processing at step k

$\qquad$ ($cd_{i,0}$ is the scv of product arrivals to the network.)

λ_j = net arrival rate at station j

μ_j = service rate of each machine at station j

ca_j = scv of arrivals at station j

cd_j = scv of departures from station j

cs_j = scv of service time at station j

m_j = number of machines at station j

ρ_j = utilization at station j = $\lambda_j/(m_j\mu_j)$

$1H(x)$ = indicator function = 1 if $x \epsilon H$ and 0 otherwise.

$$\lambda_j = \sum_{i=1}^{m} \sum_{r=1}^{u_i} a_i 1\{u_{i,r} = j\}, \quad j = 1, 2, \ldots J; \qquad (A2.1)$$

$$cd_j = 1 + (1 - \rho_j^2)(ca_j - 1) + \rho_j^2(cs_j - 1)/\sqrt{m_j} \qquad (A2.2)$$

$$\lambda_j ca_j = \sum_{i=1}^{m} \sum_{r=1}^{u_i} a_i cd_{i,r-1} 1\{u_{i,r} = j\}, \quad j = 1, 2, \ldots J \qquad (A2.3)$$

$$cd_{i,k} = p_i cd_j + cn_i, \quad k = 1, 2, \ldots u_i; \; i = 1, 2, \ldots m \qquad (A2.4)$$

where $j = u_{i,k}$, and $p_i = a_i/\lambda_j$ in (A2.4).

For the case where the arrival distribution of the aggregate product, i.e., the composition of all products arriving at station j except i, can be approximated by a Poisson distribution, (A2.4) reduces to

$$cd_{i,k} = p_i cd_j + (1-p_i)\{p_i + (1-p_i)cd_{i,k-1}\}, \quad k = 1, 2, \ldots u_i \qquad (A2.4')$$

$$i = 1, 2, \ldots m;$$

Other approximations for cn_i are discussed in Bitran and Tirupati (1988a).

Estimation of mean number of jobs in a GI|G|m queue

(see Whitt, 1985 for details)

Additional Notation:

Lq_j = mean number of jobs in queue at station j

L_j = mean number of jobs at station j

Lq_j^* = mean number of jobs in an M|M|m queue with utilization ρ_j.

Then $L_j = Lq_j + m_j \rho_j$, $\qquad (A2.5)$

where Lq_j is given by the following:

<u>Case (1): m = 1</u>:

$$Lq_j = \rho_j{}^2/(1-\rho) \; (ca_j + cs_j)/2 \; g(\rho_j, ca_j, cs_j) \tag{A2.6}$$

where $g(\rho, ca, cs) = \exp [-2(1-\rho)(1-ca)^2/(3\rho(ca + cs))], \; ca \leq 1$

$$= \exp [-(1-\rho)(ca-1)/((1+\rho)(ca + 10cs^2))], \; ca > 1$$

<u>Case 2: m > 1</u>:

$$Lq_j = \phi(\rho_j, ca_j, cs_j, m_j) \; \frac{ca_j + cs_j}{2} \; Lq_j{}^* \tag{A2.7}$$

Where $\phi(\rho, ca, cs, m)$

$$= \frac{4(ca - cs)}{4\,ca-3cs} \; \phi1(m,\rho) + \frac{cs}{4ca-3cs} \; \theta((ca + cs)/2, m, \rho), \; ca > cs$$

$$= (cs-ca)/(2(ca+cs))\phi3(m,\rho) + (cs+3ca)/(2(ca+cs))\theta((ca+cs)/2,m,\rho), \; ca\leq cs$$

$$\theta(\alpha,m,\rho) = 1 \quad , \; \alpha > 1$$

$$= (\phi4(m,\rho))^{2(1-\alpha)}, \; o \leq \alpha \leq 1$$

$\delta(m,\rho) = \min \{0.24, (1-\rho)(m-1)[(4+5m)^{0.5} - 2]/(16m\rho)\}$

$\phi1(m,\rho) = 1 + \delta(m,\rho)$

$\phi2(m,\rho) = 1 - 4\,\delta(m,\rho)$

$\phi3(m,\rho) = \phi2(m,\rho) \exp(-2(1-\rho)/(3\rho))$

$\phi4(m,\rho) = \min \{1, (\phi1(m,\rho) + \phi3(m,\rho))/2\}$

Appendix 3. Data for the Network Example (Section 3.3)

Table A3.1: Product Data

Product	Distribution of interarrival times	Routing Sequence
1	Erlang, order 2	1,2,4,2,9,10,11
2	Erlang, order 2	1,2,5,2,8,9,10,11
3	Erlang, order 3	1,2,6,4,2,9,12,11
4	Erlang, order 3	1,2,7,4,2,9,10,11
5	Uniform	1,2,4,12,2,9,2,13
6	Erlang, order 3	1,2,5,12,2,9,7,13
7	Erlang, order 4	1,2,6,12,2,8,2,13
8	Exponential	1,2,3,7,4,12,2,8,6,9,2,13
9	Exponential	1,2,3,5,4,6,12,2,8,2,10,6,13
10	Erlang, order 3	1,2,3,6,2,4,12,7,2,9,11,5,13

Table A3.2: Station Data

Station	Service Time Distribution	v_j	Work Schedule hours/day	Utilization (ρ_j)
1	Erlang order 2	100	8.00	0.7692
2	Erlang order 4	1612	8.72	0.8284
3	Uniform	733	9.52	0.7979
4	Erlang order 2	1052	8.00	0.7000
5	Uniform	912	8.29	0.6861
6	Erlang order 4	1683	8.00	0.6501
7	Exponential	1662	9.20	0.5797
8	Erlang order 3	1812	10.13	0.7018
9	Erlang order 3	1730	8.96	0.7143
10	Erlang order 3	1600	8.56	0.6547
11	Uniform	1882	9.91	0.7418
12	Erlang order 2	1486	10.05	0.7495
13	Erlang order 2	3250	9.81	0.6522

Note: 1. The distribution of interarrival times for jobs in each product family is completely specified by data in Table A3.1 and the fact that the arrival rate is 1.0 per day.

   2. The service time distribution at each station is completely specified by the data in Table A3.2 together with the net arrival rate computed from routing data (using formulae in Appendix 2) in Table A3.1.

   3. All stations operate on a 5 day/week schedule.

Papers by Gabriel R. Bitran
Nippon Telegraph and Telephone Professor of Management Science

Paper #	Date	Title/Author(s)
3115	1/90	"Analysis of Phi / PH / 1 Queues," Bitran, G.R., and Dasu, S.
3111	1/90	"Distribution-Free, Uniformly-Tighter Linear Approximations for Chance-Constrained Programming," Bitran, G., and Leong, T.
3108	12/89	"Hotel Sales and Reservations Planning," Bitran, G., and Leong, T.
3097	11/89	"Co-Production of Substitutable Products," Bitran, G., and Leong, T.Y.
3090	11/89	"Ordering Policies in an Environment of Stochastic Yields and Substitutable Demands," Bitran, G., Dasu, S.
3071	8/89	"Deterministic Approximations to Co-Production Problems with Service Constraints" Bitran, G., and Leong, T.
3054		"Sequencing Production on Parallel Machines with Two Magnitudes of Sequence Dependent Setup Costs," Bitran, G., and Gilbert, S.
3019		"Ordering Policies in an Environmnent of Stochastic Yields and Substitutable Demands," Bitran, G., and Dasu, S.
3017		"Hierarchical Production Planning," Bitran, G., and Tirupati, D.
2519	3/89	"Approximation for Manufacturing Networks of Queues and Overtime," Bitran, G., and Tirupati, D.
1744	2/86	"Planning and Scheduling for Epitaxial Wafer Production Facilities," Bitran, G., and Tirupati, D.
1764	3/86	"Multiproduct Queueing Networks with Deterministic Routing: Decomposition Approach and the Notion of Interference," Bitran, G., and Tirupati, D.
1635	3/85	"An Optimization Approach to the Kanban System," Bitran, G., and Chang, L.
1486	10/83	"Introduction to Multi-Plant MRP," Bitran, G., Marieni, D., and Noonan, J.
1402	2/83	"A Simulation Model for Job Shop Scheduling," Bitran, G., Dada, M., and Sison, L.
1391	1/83	"Productivity Measurement at the Micro Level," Bitran, G., and Chang, L.

Paper #	Date	Title/Author(s)
1282	2/82	"Analysis of the Uncapacitated Dynamic Lot Size Problem," Bitran, G.R., Magnanti, T.L., and Yanasse, H.H.
1271	11/81	"Computational Complexity of the Capacitated Lot Size Problem," Bitran, G.R., and Yanasse, H.H.
1272	10/81	"Diagnostic Analysis of Inventory Systems: A Statistical Approach," Bitran, G.R., Hax, A.C., and Valor-Sabatier, J.

Please Use Attached Form When Ordering Papers