

Modeling Re-entrant Manufacturing Systems with Inspections

Y. Narahari

Computer Science and Automation
Indian Institute of Science
Bangalore 560 012 - INDIA

L. M. Khan

Computer Science and Automation
Indian Institute of Science
Bangalore 560 012 - INDIA

Abstract

Inspection stations are now an integral part of any manufacturing system and help track product quality and process performance. In this paper we consider re-entrant manufacturing systems (such as semiconductor fabrication facilities) with inspections at various stages of processing. At the end of each inspection, three possibilities are assumed, namely accept, reject, or rework at some previous stage. We propose re-entrant lines with probabilistic routing as models for such systems and present an efficient analytical technique based on mean value analysis to predict mean cycle times and throughput rates. We can use the method to compare different ways of locating inspection stations, from a cycle time and throughput viewpoint.

1 Introduction

Global competitive pressures are forcing today's manufacturing companies to become more customer focussed in terms of offering high quality products and reduced product lead times. The recognition that product quality is a strategic asset has spurred factory managers to re-examine the role of on-line and off-line quality in product design and manufacturing. In spite of the best process control methods, it is impossible to eliminate defects altogether, hence inspection stations are essential.

Inspection stations now constitute an integral part of any manufacturing system. They help track the product quality and process performance. Two important problems related to inspection are

- how many inspection stations to use?
- where to locate the inspection stations?

There is much literature on the above two topics. For an overview and references, see [1]. Researchers have

so far not addressed the problem of performance analysis in the presence of inspections, except in two early works of Seidmann, Schweitzer, and Nof [2], and Davis and Kennedy [3].

In this paper we consider re-entrant manufacturing systems such as semiconductor fabrication facilities and explicitly model the effect of inspections at various stages of processing. We extend the re-entrant lines [4, 5] model to include probabilistic routing and propose this as a model for re-entrant manufacturing systems with inspections. We also develop an efficient analytical technique based on mean value analysis (MVA) [6, 7, 8, 9] to compute the mean steady-state cycle time and throughput rate of such models under various scheduling policies. The new model and the analysis technique facilitate

- predicting the mean steady-state cycle time and throughputs of re-entrant manufacturing systems in the presence of inspections, reworking, and rejection of parts, under a wide variety of scheduling policies, and
- comparing different ways of locating inspection stations from a cycle time and throughput viewpoint.

Earlier research in the topic of performance of re-entrant lines has not considered inspections. Kumar [4, 5] has looked at the performance of a wide variety of buffer priority based scheduling policies and due-date based scheduling policies using simulation. Lu, Deepa Ramaswamy, and Kumar [10] have investigated the cycle time and throughput performance of a class of scheduling policies called fluctuation smoothing policies. More recently, Narahari and Khan [6, 11] have used an MVA-based analytical method to evaluate the performance of buffer priority based scheduling policies in re-entrant lines, and the effect of high priority jobs or hot lots on the cycle time and throughput of other jobs in a re-entrant line.

This paper is organized as follows. In Section 2, we present a model for re-entrant manufacturing systems

with inspections and outline an analytical methodology based on MVA to compute the mean steady-state cycle time under various scheduling policies. We provide numerical results obtained for a 4 machine, 13 buffer example including simulation results to validate the analytical method proposed. In Section 3, we show how we can compare the performance of different ways of locating inspection stations in re-entrant lines, using the analytical method proposed.

2 A Model and An Analysis Methodology

In this paper, the model we propose is based on a type of non-traditional queueing models called re-entrant lines. Re-entrant lines [4] are appropriate for modeling manufacturing systems with distinct multiple visits to work centers. Examples of such (re-entrant) manufacturing systems include semiconductor fabrication facilities, thin film lines, and systems with rework tasks. In such systems, parts typically visit a given station several times for undergoing either multiple operations or rework operations. The analysis methodology we propose is based on MVA [6, 7, 8, 9].

2.1 Re-entrant Lines With Probabilistic Routing

A re-entrant line can be described as follows. There is a set of service centers $\{1, 2, \dots, m\}$. Service center $i \in \{1, 2, \dots, m\}$ has n_i logical or physical buffers, $b_{i1}, b_{i2}, \dots, b_{in_i}$. For $j \in \{1, 2, \dots, n_i\}$, the buffer b_{ij} contains parts visiting service center i for the j^{th} time, (call it stage (i, j) of service). A part visits these buffers in a given sequence and any service center is typically visited several times in the route of the part. Figure 1 shows a typical re-entrant line with 4 service centers and 13 buffers. Parts enter the system at buffer b_{11} and visit the buffers at various centers according to a deterministic route as shown. When there is an inspection at the end of a particular stage of processing, it is reasonable to assume three possibilities namely accept, reject, or rework. A part that is *accepted* will queue up for the next stage of processing in the deterministic route. A part that is *rejected* will disappear from the system. A part that needs rework may need to be routed to any of the earlier stages of processing.

We make the following assumptions regarding the parts that need to go for reworking.

1. A part may have to go to any earlier stage say,

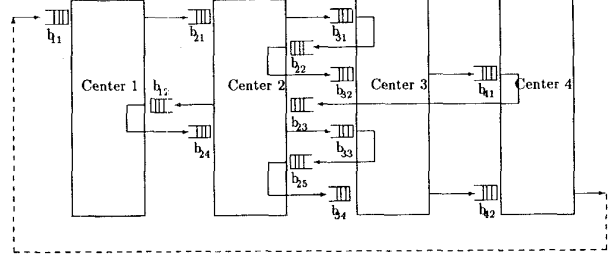


Figure 1: A re-entrant line with 13 buffers

stage (i, j) for reworking. Such a part is indistinguishable from the one going through stage (i, j) for the first time. This implies that the part going for reworking to an earlier stage (i, j) will go through all the stages from (i, j) to the current one.

2. Rework times have identical processing requirement as original processing times.
3. Parts that go for reworking to a particular stage join the tail of the queue at that stage.

After each stage of processing, note that a part may advance to the next stage, come back to the same stage, or go back to any previous stage, or get rejected. The probabilities of each of these events are assumed to be known for all stages of processing. An important assumption we make is that the inspection process is instantaneous. This assumption is relaxed in a forthcoming paper [12].

The inspections, reworking, and rejections detailed above can be described by a re-entrant line with a Markovian routing matrix, P , where the entries indicate the probability of going from a given stage to any other stage. For example consider the matrix P for a 4-buffer re-entrant line shown in Figure 2: Note that P is a square matrix with dimension $B + 1$, where B is the total number of stages. The last entry in each row gives the percentage of parts rejected after an inspection at that stage. Note that the sum of entries in each row, except that corresponding to the last stage i (stage $(2, 2)$ in the example), is unity. In the case of the last stage, it is easy to see that $(1 - \text{sum of entries})$ is the probability that a part exits successfully from the system, after finishing all the processing.

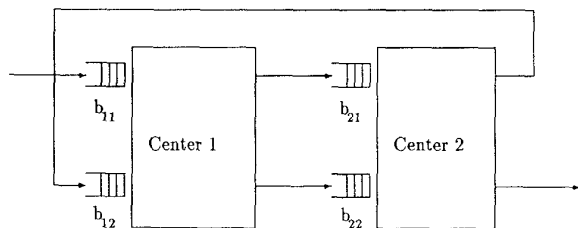


Figure 2: A re-entrant line with 4 buffers

2.2 An Analysis Methodology

The analysis methodology uses ideas from the well known mean value analysis (MVA) technique [7, 8] applied in an approximate way to non-product form queueing networks. In an earlier paper [6], Narahari and Khan have already presented an analysis technique for re-entrant lines with strictly deterministic routing based on MVA. Here this methodology is extended to include inspections. The extended methodology is validated by extensive simulations.

A special feature of this methodology is that it explicitly models any buffer priority based scheduling policy that may be followed at different service centers. That is, when a processing center i finishes servicing a part, it selects the next part for processing from among the buffers $b_{i1}, b_{i2}, \dots, b_{in_i}$, in a fixed priority order that is independent of the state of the system. The analysis assumes that the priorities accorded are non-preemptive and that parts in any given buffer are processed in FCFS fashion.

To apply MVA, we have to assume that the re-entrant line is a closed queueing network. This assumption is valid if the input release policy is a fixed-work-in-process policy (a fresh part is released into the network as soon as a finished part leaves the system) [5]. Also because of inspections, we might reject some parts at intermediate stages, and this will reduce the number of jobs in the system. In order to keep the number of jobs in the system constant, we shall assume that a rejected part is immediately replaced by a fresh part which enters the first buffer in the system.

Let N be the total number of jobs in the system. We shall use the following indices: i denotes a processing center; j denotes a buffer at a given processing center; k denotes a current job population and has the range $1, \dots, N$. Let stage (i, j) as usual correspond to the waiting and the processing of a job visiting center

i for the j^{th} time. Let the performance measures of the network be denoted as follows.

$L_{ij}(k)$: mean steady-state number of jobs in stage (i, j) when the network has k jobs.

$W_{ij}(k)$: mean steady-state delay for jobs in stage (i, j) (mean waiting time in buffer b_{ij} + mean processing time).

$\lambda(k)$: mean steady-state throughput rate of jobs when the network has k jobs.

If $W(k)$ denotes the mean total delay (also called mean cycle time) in the entire network in the steady state, we have

$$W(k) = \sum_{i=1}^m \sum_{j=1}^{n_i} v_{ij} W_{ij}(k) \quad (1)$$

where v_{ij} is the mean number of times a part visits stage (i, j) during its sojourn in the network. We can note immediately that $v_{ij} = 1$ for all i and for all j if there are no inspections in the re-entrant line. If we do have inspections, we can compute v_{ij} 's from the routing probability matrix in the standard way done for product form queueing networks [9, 13].

Using MVA we can recursively compute $W(N)$ and $\lambda(N)$. For details, see [6]. We give an outline of the procedure below. Consider the scenario a job would encounter upon its arrival at a certain buffer b_{ij} and the sequence of activities that occur while it is waiting there. When this *distinguished* part arrives at b_{ij} , it would see a certain number of jobs in various buffers at the service center i . Let S be the set of jobs currently at center i and having higher priority than the distinguished part. Note that S will include all jobs that are ahead of the distinguished job in b_{ij} and all jobs in all buffers having higher priority than b_{ij} . The total mean waiting time of the distinguished job in b_{ij} on each visit can be seen to be the sum of three terms, say Term 1, Term 2, and Term 3 defined as follows.

- Term 1 : Mean total time until all jobs in the set S are serviced and leave center i .
- Term 2 : Mean total time required to process all higher priority jobs which arrive during the stay of the distinguished job in the queue at b_{ij} (i.e. until the commencement of its service).
- Term 3 : Mean processing time of the distinguished job.

Term 3 is easy to compute. The computation of Term 1 and Term 2 is done by presuming that the *arrival theorem* [9] is valid in the given network. In fact

the arrival theorem is not valid for the given network since the network is not product form. However since we are only seeking an approximate analysis, we assume the arrival theorem to be valid for this network and verify the accuracy of the approximation using detailed simulations. The computation of Term 1 and Term 2 is described in detail in [6]. We can thus compute $W_{ij}(k)$ and using (1), we can compute $W(k)$.

Applying Little's Law in the network, we obtain

$$\lambda(k) = \frac{k}{W(k)} \quad (2)$$

Again we use Little's Law to obtain

$$L_{ij}(k) = \lambda(k)W_{ij}(k) \quad (3)$$

Consider the following initial conditions:

$$L_{ij}(0) = 0 \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n_i \quad (4)$$

$$\lambda(0) = 0 \quad (5)$$

Using the initial conditions above and the relations for $W_{ij}(k)$, $\lambda(k)$, and $L_{ij}(k)$, we can compute the above performance measures for all $k = 1, 2, \dots, N$. Thus $W(N)$ and $\lambda(N)$ can be computed.

2.3 A Numerical Example

Consider the re-entrant line shown in Figure 1. This line has 4 machine centers and 13 buffers. The service time for all buffers at a given machine center are assumed to be identical exponential random variables. Let $\frac{1}{\mu_i}$ be the mean service time at each buffer at machine center i . In the analytical and simulation experiments, we have assumed

$$\begin{aligned} \frac{1}{\mu_1} &= 0.5 \text{ hour} & \frac{1}{\mu_2} &= 0.1 \text{ hour} \\ \frac{1}{\mu_3} &= 0.5 \text{ hour} & \frac{1}{\mu_4} &= 0.5 \text{ hour} \end{aligned}$$

In this re-entrant line, machine center 3 is a bottleneck center since there is a maximum service demand on this center.

Assume that there is an inspection at the end of every stage. After inspection at each stage, we have assumed the probability of rejection as 0.05 and the probability of sending for rework to any of the earlier stages as 0.05. Using the proposed MVA based method, and also simulation, we have computed the mean cycle time and mean throughput rate of accepted parts for populations ranging from 1 to 35. Figures 3 and 4 provide a graphical representation of the

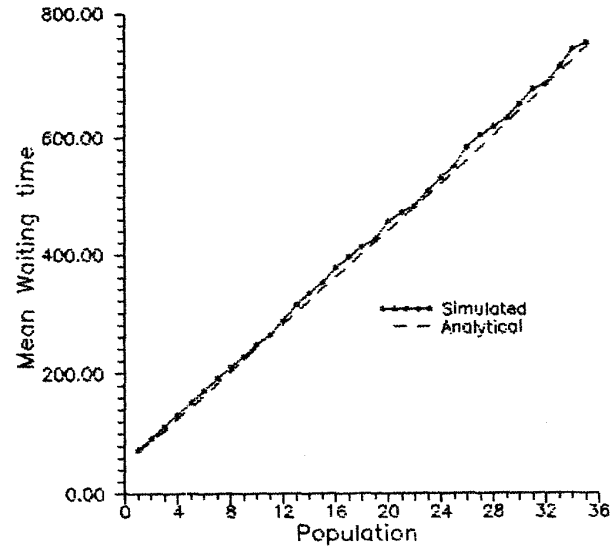


Figure 3: Analytical and simulation results for mean cycle time

numerical results. It can be seen from the graphs that there is a close agreement in the results obtained by the proposed method and simulation. The maximum discrepancy between the simulation and analytical results is about 4% in the case of mean cycle time, and about 7% in the case of mean throughput rate.

3 Comparing Different Ways of Locating Inspection Stations

An important problem in inspection is to determine the optimal number and the optimal location of inspection stations in a given manufacturing system. In this section we shall address the problem of a given number of inspection stations and evaluate different alternatives from a cycle time and throughput viewpoint. We shall illustrate the methodology with an example. Consider the re-entrant line of Figure 2, which has four stages, (1,1), (2,1), (1,2), and (2,2). Call these stages stage 1, stage 2, stage 3, and stage 4, respectively. Assume that an inspection at the end of stage 4 is always required. For $i = 1, 2, 3, 4$, define variable $x_i = 1$ if there is inspection at the end of stage i and $x_i = 0$ otherwise. Call (x_1, x_2, x_3, x_4) as the inspection location vector. If only one inspection is being done, this vector can only take the value $(0, 0, 0, 1)$. If two inspections are being done, the possible vectors are $(1, 0, 0, 1)$, $(0, 1, 0, 1)$, and $(0, 0, 1, 1)$. In the case of three inspections, the possible values are $(1, 1, 0, 1)$, $(1, 0, 1, 1)$, and $(0, 1, 1, 1)$. Finally for four inspections the only possible vector is $(1, 1, 1, 1)$. Thus

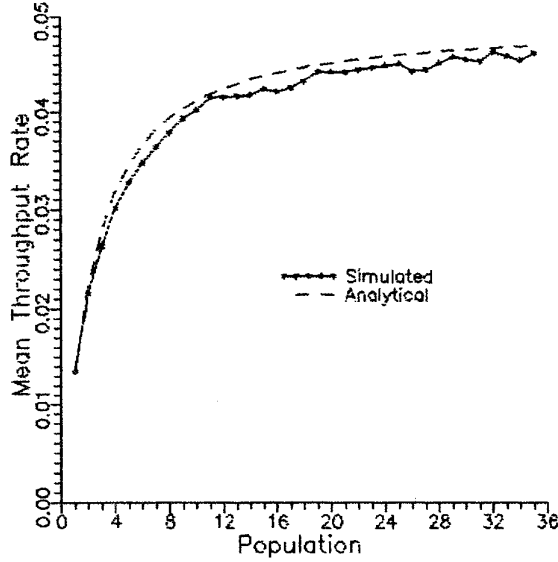


Figure 4: Analytical and simulation results for mean throughput rate

we have eight possible inspection location vectors.

To compare the above alternative inspection location possibilities, we need accept, reject, and rework probabilities in each case. Consider the inspection location vector $(1, 1, 1, 1)$. For $i = 1, 2, 3, 4$, define

r_i = probability of rejection of a part after stage i .
 q_{ij} = probability that a part goes for reworking to stage j after stage i .

Note that $q_{ij} = 0$ for all $j > i$. Define s_i as the probability that a part passes the inspection after stage i . It is easy to see that

$$s_i = 1 - r_i - \sum_{j=1}^i q_{ij}$$

Assuming that we know the probabilities r_i and q_{ij} for all the four stages in the case of the vector $(1, 1, 1, 1)$, we can compute these probabilities for any of the other seven vectors. For example, consider the vector $(0, 0, 0, 1)$. Since there is no inspection at the end of the first three stages, we only need to compute the probabilities for the stage 4. Let these probabilities be $r'_4, q'_{41}, q'_{42}, q'_{43}, q'_{44}$, and s'_4 . It is easy to see that

$$\begin{aligned} q'_{41} &= s_1 s_2 s_3 q_{41} + s_1 s_2 q_{31} + s_1 q_{21} + q_{11} \\ q'_{42} &= s_1 s_2 q_{42} + s_1 q_{32} + q_{22} \\ q'_{43} &= s_1 q_{43} + q_{33} \\ q'_{44} &= q_{44} \\ r'_4 &= r_4 \end{aligned}$$

$$s'_4 = 1 - r_4 - \sum_{j=1}^4 q'_{4j}$$

Similarly, we can compute these probabilities for all possible location vectors.

The analysis methodology of Section 2 can be used for computing the mean steady-state cycle time and throughput rate of accepted, finished parts. Figure 5 shows the mean steady-state throughput rates for all the above inspection location vectors. The mean processing time at each buffer in center 1 is assumed as 1 unit whereas that at each buffer in center 2 is assumed 2 units.

Note from Figure 5 that the maximum throughput rate is obtained in the case of $(1, 1, 1, 1)$, *i.e.*, complete inspection, and the minimum throughput rate is in the case of $(0, 0, 0, 1)$, *i.e.*, no intermediate inspection. This is easy to see since in the latter case, parts can only get rejected after the last stage. It is also interesting to note that the vectors $(0, 0, 1, 1)$, $(0, 1, 1, 1)$, and $(1, 0, 1, 1)$ lead to very nearly the same throughput rates. This is significant because two inspection stations located at stage 3 and stage 4 give out the same performance as three inspection stations located according to $(0, 1, 1, 1)$ or $(1, 0, 1, 1)$. Also the vector $(0, 0, 1, 1)$ outperforms the vector $(1, 1, 0, 1)$, in spite of the fact that former has one inspection station less. This shows that a small number of strategically located inspection stations can perform better than larger number of poorly located inspection stations. The proposed technique can be conveniently used to study comparative merits of different inspection schemes.

4 Conclusions

In this article, we have proposed re-entrant lines with probabilistic routing as an accurate model of re-entrant manufacturing systems with inspections. We have also proposed an efficient analysis method for the new model based on MVA. The proposed analysis method has been validated using simulation.

For a four buffer re-entrant line, we have compared the throughput rates for different inspection location vectors by considering different numbers and locations of inspection stations. Thus the modeling methodology enables different inspection strategies to be compared from a cycle time and throughput rate view point.

In a forthcoming article [12], we have extended the model to include inspection times and contention for inspection stations.

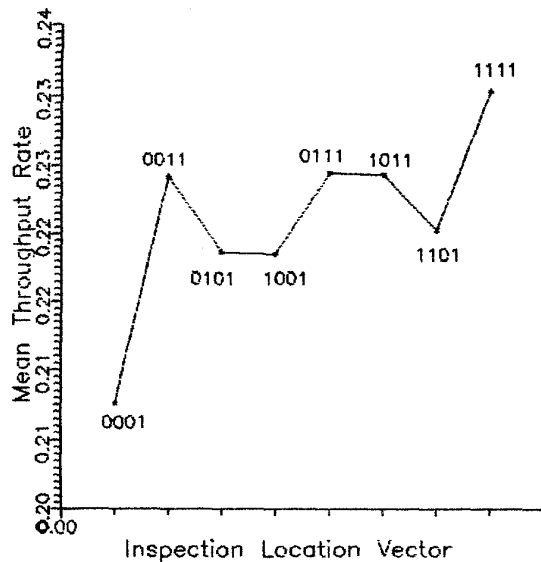


Figure 5: Cycle time and throughput rate for different inspection location vectors

Acknowledgements

This research was supported in part by the Office of Naval Research and the Department of Science and Technology grant N00014-93-1017. We would also like to acknowledge the excellent facilities at the Intelligent Systems Laboratory, Department of Computer Science and Automation, Indian Institute of Science.

References

- [1] N. Viswanadham, S. M. Sharma, and M. Taneja. Inspection allocation in manufacturing systems using stochastic search techniques. Technical report, Department of Computer Science and Automation, Indian Institute of Science, April 1994.
- [2] A. Seidmann, P.J. Schweitzer, and S. Y. Nof. Performance evaluation of a flexible manufacturing cell with random feedback flow. *International Journal of Production Research*, 23(6):1171-1184, 1985.
- [3] R. P. Davis and W. J. Kennedy Jr. Markovian modelling of manufacturing systems. *International Journal of Production Research*, 25(3):337-351, 1987.
- [4] P. R. Kumar. Re-Entrant lines. *Queueing Systems: Theory and Applications*, 13:87-110, 1993.
- [5] P. R. Kumar. Scheduling queueing networks: Stability, performance analysis, and design. In *Proceedings of the IMA Workshop on Stochastic Networks*, July 1994.
- [6] Y. Narahari and L.M. Khan. Performance analysis of scheduling policies in re-entrant manufacturing systems. Technical report, Department of Computer Science and Automation, Indian Institute of Science, January 1994.
- [7] S. Shalev Oren, A. Seidmann, and P.J. Schweitzer. Analysis of flexible manufacturing systems with priority scheduling: PMVA. *Annals of Operations Research*, 3:115-139, 1985.
- [8] R. Suri and R.R. Hildebrandt. Modeling flexible manufacturing systems using mean value analysis. *Journal of Manufacturing Systems*, 3(1):27-38, 1984.
- [9] R. Suri, J.L. Sanders, and M. Kamath. Performance evaluation of production networks. In S. C. Graves, A. G. Rinnoy Kan, and P. Zipkin, editors, *Handbooks in OR and MS, Volume 4*, pages 199-286. Elsevier Science Publishers, Amsterdam, 1993.
- [10] S.H. Lu, Deepa Ramaswamy, and P.R. Kumar. Efficient scheduling policies to reduce mean and variance of cycle-time in semiconductor manufacturing plants. Technical report, Coordinated Systems Laboratory, University of Illinois, Urbana-Champaign, 1992.
- [11] Y. Narahari and L.M. Khan. Modeling the effect of hot lots in semiconductor manufacturing systems. Technical report, Department of Computer Science and Automation, Indian Institute of Science, June 1994.
- [12] Y. Narahari and L.M. Khan. Modeling re-entrant manufacturing systems with inspections. Technical report, Department of Computer Science and Automation, Indian Institute of Science, September 1994.
- [13] N. Viswanadham and Y. Narahari. *Performance Modeling of Automated Manufacturing Systems*. Prentice Hall, Englewood Cliffs, NJ, 1992.