Using Queueing Network Models to Set Lot-sizing Policies for Printed Circuit Board Assembly Operations

Maged M. Dessouky

Department of Industrial and Systems Engineering,
University of Southern California,
Los Angeles, CA 90089-0193

ABSTRACT

The current trend among manufacturers is to reduce WIP by decreasing the lot size. Reducing the lot size in some cases may have an adverse effect on system performance because more setup time is incurred. In this paper, we propose the use of queueing network models to determine the best lot-sizing policy that minimizes the WIP. The queueing network model is validated by comparing its results to the output of a simulation model. Although the focus of this paper is on WIP reduction not demand management, a reduction in the lead time allows a manufacturer to respond faster to new customer orders. This approach is used to set a lot-sizing policy for a printed circuit board assembly process of an electronics manufacturer.

I. INTRODUCTION

The thrust in today's manufacturing environment is to move towards a Just-in-Time (JIT) manufacturing environment. As outlined by numerous studies (e.g., Schonberger, 1986; Voss and Robinson, 1987; Im and Lee, 1989), successful implementation of just-in-time manufacturing principles may lead to reduced inventory costs, improved quality, and increased equipment utilization. A necessary requirement for successful implementation of JIT is the ability to run small lot sizes (e.g., Finch, 1986; Mehra and Inman, 1992; and Handfield, 1993). Numerous approaches have been developed to determine the appropriate lot size. For example, Hill and Raturi (1992) propose an optimization model for determining the lot sizes, and Karmarkar et al. (1992) and Karmarkar et al. (1985) propose a queueing model.

Small lot sizes can reduce work-in-process inventory (WIP) and manufacturing lead time if the setup times are not much larger than the unit run times. A reduction in the lead time allows the manufacturer to respond quicker to new customer orders or any changes in demand and increases the likelihood of meeting the demand on-time. Small lot sizes tend to reduce the WIP because a lot spends less time at a machining center, causing new arriving lots to wait less for the machines to become available. However, reducing the lot size too much can sometimes have the opposite effect by increasing WIP because machine utilization may increase significantly due to an increase in the setup times. Thus, the selection of an appropriate lot size needs to take into consideration the setup time. This consideration is especially important for automatic insertion machines used in printed circuit board (PCB) assembly where there is both a board and component setup time, and up to 35% of the theoretical assembly load of a PCB line is determined by the setup procedures (Feldmann, Franke, and Rothhaupt, 1994).

Carlson, Yao, and Girouard (1994) give an example where a PCB assembly manufacturer produced lots of 100 or more units because the setup time is in the order of three hours per production run. If the lot size is reduced to 5 boards, the setup time translates to around 36 minutes per board. Complicating the problem of determining the appropriate lot-sizing policy in PCB assembly is that the same line produces many different board types, each having its own unique setup time. These factors enhance the need for analytical tools to help determine how low the lot size can be without negatively impacting WIP.

In this paper, a queueing network model is used to determine the lot size that minimizes WIP while considering setup time for a PCB assembly process. Although the focus of this paper is on WIP reduction not demand management, a reduction in the manufacturing lead time allows a manufacturer to respond faster to new customer orders, hence increasing the likelihood of on-time delivery. Queueing network models represent the manufacturing process as a network of queues. Analytical formulas are developed to approximate steady-state performance measures of the manufacturing system such as

average WIP, machine utilization, and average flow time (i.e., the average amount of time a part spends in the manufacturing process). The best lot size is determined by developing operating characteristic curves of the PCB assembly process. Operating characteristic curves of interest include WIP and machine utilization as a function of the lot size. We demonstrate our approach on a high product mix and medium production volume PCB assembly facility located on the west coast.

A queueing network model is used rather than a simulation model because simulation may become tedious in the planning stage due to the numerous alternatives that need to be considered (Suri and Diehl, 1987). Suri (1983) shows the robustness of analytical queueing network models for representation of real systems. The computation time required to solve a queueing network model is significantly less than the simulation run time, greatly facilitating the ability to consider numerous lot-sizing policies. Also, the queueing network models do not require any programming, making them simpler to develop than simulation models. Simulation also requires the analyst to have some knowledge in statistics.

Snowdon and Ammons (1988) survey eight queueing network packages. Some of the queueing network software packages are public domain while others are commercially sold by a software vendor. MANUPLAN (1987) is selected as the queueing software package because of its user-friendly interface, and it has been successfully used in the past in the design of PCB assembly lines. For example, Haider, Noller, and Robey (1986) use MANUPLAN to help identify initial design alternatives for a PCB assembly facility at IBM. Other successful applications of MANUPLAN in the design of PCB assembly lines include studies by Brown (1988) and Garlid, Falkner, Fu, and Suri (1988).

II. PRINTED CIRCUIT BOARD ASSEMBLY PROCESS

The studied PCB assembly process is a high product mix and medium production volume facility. The product mix is approximately 300 board types and the daily production volume is around 600 boards. The assembly facility uses plated through hole (PTH) technology to insert components on a

board. It is common to have both automatic and manual component insertions in PTH assembly lines. For an excellent review of the PTH process, the reader is referred to a book by Kear (1987).

A diagram of the process flow chart of the studied PCB assembly facility is shown in Figure 1. The triangles represent inventory storage locations. This facility uses a mix of automatic and manual operations to insert components on a board. The boards and kits, where the components are stored, are withdrawn from inventory and are sent directly to the auto insertion process. There are two types of automatic insertion machines. One type is axial insertion using a variable center distance (VCD) machine. Sequencing is performed at this step to ensure the components are inserted on the board in the proper order. The other type of automatic insertion operation uses a dual in-line package (DIP) machine. The studied assembly facility has two VCD machines and one DIP machine. All board types do not necessarily have to be processed by both machine types. After the auto insertion process the boards and kits are stored in inventory. There is sufficient demand to justify maintaining a certain amount of inventory for each product type. Later, the boards and kits are withdrawn from inventory and put through a series of manual operations including loading and soldering. The last operation is a manual test and repair station. Since the boards go into inventory storage after the auto insertion process, the auto and manual processes are treated independently and modeled separately. This paper focuses on the auto insertion process since the majority of the setup time is here.

The total time to complete processing of a lot on an automatic insertion machine includes both a *setup time per lot* component and a *run time per board* component for each board in the lot. A setup is required whenever a changeover to a new product type is made. The setup time is incurred once at the beginning of processing of a new lot type. The setup time is independent of the lot size. The run time per board is incurred for each board in the lot.

Let S_{ik} be the setup time of board type i on machine type k ($k=1$ refers to VCD and $k=2$ refers to DIP). The setup time includes a fixed machine handling time and a time to prepare the components for

insertion. The latter time is a function of the number of different component types to insert on the board. The setup time per lot of board type i at machine type k, S_{ik}, is

$$S_{ik} = a_k + m_k h_{ik}$$

where:

a_k fixed prep time on machine k

h_{ik} number of different component types to insert on board type i on machine k

m_k component prep time on machine k

Let B_{ik} be the run time per board of board type i on machine k. The run time includes a board handling time and a time to insert the components on the board The latter time is a function of the number of components to insert on the board. Then, the run time per board type i at machine type k, B_{ik}, is

$$B_{ik} = f_k + r_k p_{ik}$$

where:

f_k board handling time on machine k

p_{ik} number of components to insert on board type i on machine k

r_k insertion time per component on machine k

Let the total time to process a lot of type i on machine k be T_{ik}. Then, $T_{ik} = S_{ik} + q_i B_{ik}$, where q_i is the lot size. Note that the setup time is independent of the lot size.

Table 1 displays representative values for the setup and run times at both the VCD and DIP

machines. The number of components inserted on a board (p_{ik}) varied from a low of 5 to a high of 700 with the number of different types (h_{ik}) typically being 2/3 of that value.

Table 1. Representative setup and run times in minutes.

Machine	Lot Setup Time (a_k)	Component Setup time (m_k)	Board Prep Time (f_k)	Insertion Time (r_k)
VCD ($k=1$)	7.5/lot	0	.25/board	.008/insertion
DIP ($k=2$)	8.5/lot	1/component type	.25/board	.024/insertion

III. MANUPLAN MODEL

MANUPLAN is a data-driven modeling tool. The input file contains data on board and machine characteristics. For each board type the demand rate, lot size, and route need to be specified. The route data contains the setup times and run times for each machine in the route. For each machine type, the capacity, the mean time between failure, and the mean repair time need to be specified.

The queue size in front of the machines is assumed to be unlimited, and the queue dispatching rule is first-come-first-serve. Both the VCD and DIP machines are fairly reliable machines with 98% uptime. The material transfer in the auto insertion process is lot for lot. Since the material handling system is not capacity constrained and has negligible transfer time, the material handling system is not modeled.

The variability of the interarrival time of demand of each board type and the processing time in MANUPLAN are input as a percentage of the mean. Past experience in the facility has shown that the variability is typically 30% of the mean. The outputs of the MANUPLAN model are steady-state performance measures of the manufacturing system including machine utilization, average WIP, and average flow time at each machining center. If all the demand cannot be met within the planning horizon, the output of the MANUPLAN model is an error message indicating that the demand cannot be met with the current machine capacity. The Appendix provides an overview on how the performance

measures such as average WIP and flow time are determined in MANUPLAN.

IV. AN APPLICATION

In this section, we demonstrate the use of queuing network models to help set manufacturing policy for an automatic insertion process of a west coast printed circuit board assembly facility. Currently the lot size for each board type is based on past experience. The purpose of this study is to identify a new lot-sizing policy that reduces WIP and the manufacturing lead time over using the historical lot sizes.

Demand data over a six-month planning horizon is used to perform the analysis. The demand is based on market forecasts. It is assumed that there are 125 working days during the six months. The automatic insertion line operates in two shifts with each shift having 6.5 hours available for manufacturing.

Using the current lot sizes, we first validate the MANUPLAN model by comparing its output to actual values from the assembly process. With the current lot sizes, MANUPLAN estimates the overall system WIP to be around 956 boards. This WIP includes boards in the queue and boards being processed, and is about 20% more than the levels experienced at the auto insertion process. To help identify the discrepancy between the results of the MANUPLAN Model and the real-system, a SLAM II Simulation Model (Pritsker, 1986) is developed to estimate the system performance measurements. Table 2 shows the machine utilization of each type and the average number of lots in the queue waiting for service.

The results show that the DIP machine is busy about 90% of the time while the VCD machine is busy only about 70% of the time. The percent busy time includes both the setup time and run time components of processing a lot. The DIP machine is heavily utilized because it spends about 35% of the time in the setup state while the VCD machine spends only about 8% of the time in the setup state. Due to high machine reliability, the machine utilization estimates from the MANUPLAN and simulation

models are close to the static calculation. The estimates of the average number of lots in the queue differ by about 20% between the MANUPLAN and the simulation model. These results are consistent with the study by Huettner and Steudel (1992) which showed that MANUPLAN tends to slightly overestimate the WIP levels. MANUPLAN is less accurate in estimating the WIP for the DIP machine in absolute terms because approximations for queueing networks are less accurate for heavily utilized machines (Buzacott and Shanthikumar,1993). Nevertheless, we are more concerned with making relative comparisons between different lot-sizing scenarios than with measuring the absolute value of the WIP. That is, typically one is not concerned with the absolute value of the WIP in determining the manufacturing strategy, but rather with how the WIP changes as a function of the strategy. Because MANUPLAN is data-driven and solutions can be found quickly, a large number of scenarios can be compared in a small amount of computation time (roughly 30 CPU seconds per scenario on a Hewlett-Packard workstation).

Table 2. System Performance Measures

Machine	Percent Busy		# in Queue	
Type	MANUPLAN	SLAM II	MANUPLAN	SLAM II
VCD	72.9	72.5	3.2	2.6
DIP	94.3	93.0	23.5	19.0

Currently each board type has a unique lot size and is set based on past experience. The lot sizes vary from a low of 1 board to a high of 747 boards. In order to move to more of a JIT environment, the PCB manufacturer wants to set a limit on the maximum lot size. In this manner, if any of the lot sizes based on historical experience are greater than the maximum limit, the lot size is reset to the maximum limit. Figure 2 plots the total average flow time in hours in the automatic insertion process as a function of the maximum lot size limit. Setting a smaller limit reduces the overall average lot size and increases machine utilization due to more setups. Figure 3 plots the equipment utilization as a function of the

maximum lot size limit. The results indicate that the best operating policy is a maximum lot size between 45-50. As the figure shows, bounding the lot size to a maximum of 50 reduces the average flow time (and subsequently the WIP) by 33%. A limit smaller than 45 increases the average flow time because the setup time starts to dominate. In fact, demand cannot be met with a limit smaller than 25 because the DIP machine utilization becomes greater than 100%. A limit greater than 50 increases the average flow time because a lot spends more time at a machining center due to the increased lot size, causing new arriving lots to wait more for the machines to become available.

The previous analysis helps determine an upper bound on the lot size. However, it still makes use of the historical lot size values for board types below the limit. The company is interested in setting an overall lot-sizing policy in terms of a universal number of lots. For example, if the number of lots to run is set to 5, the lot size for a particular board type is simply the demand over the planning horizon divided by 5 rounded up to the nearest integer. Note that, as the number of lots increases, the lot size decreases. In Figure 4, we plot the WIP (in terms of individual boards) as a function of the number of lots policy. Setting the number of lots for each board type to 5 is equivalent to the current lot sizes in terms of the resultant WIP because both lot size policies yield a WIP level of around 950 boards. A number of lots policy of 9 yields the smallest WIP levels of 756 boards which results in 21% less WIP than the current lot size levels. Increasing the number of lots any further than 9 will increase the WIP because the DIP machine incurs a lot of setups, causing high machine utilization.

V. CONCLUSION

As companies move towards a JIT environment, analytical tools are needed to guide factory managers on appropriate lot-sizing policies. Reducing the lot size too much may have a detrimental effect on WIP due to more setups. In this paper, we demonstrate the use of queueing network models to develop operating characteristic curves on WIP for a studied PCB assembly process. The curves help

identify appropriate lot-sizing policies for WIP minimization. Historically, the facility sets the lot sizes based on past experience. The analysis using MANUPLAN shows by bounding the lot size to a maximum of 50 reduces the WIP by 32% with the same demand levels. The analysis also shows that the demand cannot be met with a bound smaller than 30 due to the frequent number of setups at this level, demonstrating that reducing the lot size will decrease the WIP up to a certain level. Future uses of MANUPLAN include determining an overall number of lots policy. Our initial analysis using the current demand levels shows that a policy of dividing the demand into nine production lots minimizes the WIP. Besides lowering inventory costs, reducing the WIP reduces the lead time which enables a manufacturer to be more responsive to new customer orders or any changes in demand.

Although the results are specific to the studied PCB assembly facility, the same type of analysis can be used to identify appropriate lot-sizing policies for other PCB assembly facilities. Previous to this study, the company had lot sizes mentioned earlier, that varied between 1 and 747 boards. As a result of this study, the company set the maximum lot size to 50 for those products with lot sizes above this limit. Those products with lot sizes less than 50, where maintained at those figures. The advantage of using queueing network models for this type of analysis is that the models are simple to develop and many scenarios can be evaluated because of the fast computation time.

APPENDIX

We now present a brief overview of how the performance measures are determined in queueing networks. Let d_i be the average demand of board type i during the planning horizon and n be the number of different board types to produce during the planning horizon. Let q_i be the lot size of board type i and Q be the vector of lot sizes $Q= (q_1,....,q_n)$. Let c_k be the capacity of machine type k expressed in time units. Then, the utilization of machine type, U_k (Q), as a function of the lot sizes can be estimated as:

$$U_k(Q) = \frac{\sum_{i=1}^{n}\left(\left(\frac{d_i}{q_i}\right)^{+} S_{ik} + d_i B_{ik}\right)}{c_k}$$

MANUPLAN uses node decomposition to estimate the average WIP at each machining center. Buzacott and Shanthikumar (1993) present the details of the node decomposition approach. In essence, each machining center is modeled as a separate **GI/G/s** queue, and the first two moments (mean and variance) of the arrival and service distributions of the machining center are used to approximate the average WIP at each machining center. Then all machining center average WIPs are combined to derive an overall average WIP for the system. Even for a **GI/G/1** queue, it is difficult to exactly determine the average WIP. Therefore, research has focused on developing approximations for the average WIP. Buzacott and Shanthikumar (1993) suggest the following approximation for the average WIP, W_k , where C^2_a and C^2_s are the squared coefficient of variability for arrival and service times, respectively.

$$W_k(Q) = \left(\frac{U_k(Q)^2(1+C_s^2)}{1+U_k(Q)^2 C s^2}\right)\left(\frac{C_a^2 + U_k(Q)^{\,2} C_s^2}{2(1-U_k(Q))}\right) + U_k(Q)$$

Finally, the average flow time is calculated using Little's Law (1961) which states that the average flow time is equal to the average WIP divided by the production rate.

REFERENCES

Brown, E., "IBM Combines Rapid Modeling Technique and Simulation to Design PCB Factory-of-the-Future," *Industrial Engineering,* **20** (1988): 23-36.

Buzacott, J. A. and Shanthikumar, J. G., *Stochastic Models of Manufacturing Systems,* Prentice-Hall, Englewood Cliffs, N.J., (1993) .

Carlson, J. G., Yao, A. C. and Girouard, W. F., "The Role of Master Kits in Assembly Operations," *International Journal of Production Economics*, **35** (1994)**:** 253-258.

Feldmann, K., Franke, J. and Rothhaupt, A., "Optimization and Simulation of the Printed Circuit Assembly," *IEEE Transactions on Components, Packaging, and Manufacturing Technology-Part A*, **17** (1994): 277-281.

Finch, F., "Japanese Management Techniques in Small Manufacturing Companies: A Strategy for Implementation," *Production and Inventory Management*, **27** (1986): 30-38.

Garlid, S., Falkner, C., Fu, B. and Suri, R., "Evaluating Quality Strategies for CIM Systems," *Printed Circuit Assembly*, **May** (1988): 5-11.

Haider, S. W., Noller, D. G. and Robey, T. B., "Experiences with Analytic and Simulation Modeling for a Factory of the Future Project at IBM," *In Proceedings of the 1986 Winter Simulation Conference*, IEEE, Piscataway,
N. J., (1986):, 641-648.

Handfield, R., "Distinguishing Features of Just-in-Time Systems in the Make-to-Order Assemble-to-Order Environment," *Decision Sciences*, **24** (1993): 581-602.

Hill, A. V. and Raturi, A. S., "A Model for Determining Tactical Parameters for Materials Requirements Planning Systems, *Journal of the Operational Research Society*, **43** (1992): 605-620.

Huettner, C. M. and Steudel, H. J., "Analysis of a Manufacturing System via Spreadsheet Analysis, Rapid Modelling, and Manufacturing Simulation," *International Journal of Production Research*, **30** (1992): 1699-1714.

Im, J. H. and Lee, S. M., "Implementation of Just-in-Time Systems in U.S. Manufacturing Firms," *International Journal of Operations and Production Management*, **9** (1989): 5-14.

Karmarkar, U. S., Kekre S. and Kekre, S., "Multi-item Batch Heuristics for Minimization of Queueing Delays," *European Journal of Operational Research*, **58** (1992): 99-111.

Karmarkar, U. S., Kekre S. and Kekre, S., "Lotsizing in Multi-item Multi-machine Job Shops," *IIE Transactions*, **17** (1985): 290-298.

Kear, F. W., *Printed Circuit Assembly Manufacturing* , Marcel Dekker, New York, (1987).

Little, J. D. C., "A Proof of the Queueing Formula: L=8W," *Operations Research*, **9** (1961): 383-387.

Mehra, S. and Inman, R. A., "Determining the Critical Elements of Just-in-Time Implementation," *Decision Sciences*, **23** (1992): 160-174.

Network Dynamics, *MANUPLAN Users Manual,* Cambridge, MA, (1987).

Pritsker, A. A. B., *Introduction to Simulation and SLAM II*, John Wiley & Sons, New York, (1986).

Schonberger, R. J., *World Class manufacturing: The Lessons of Simplicity Applied*, Free Press, New York, (1986) .

Snowdon, J. L. and Ammons, J. C., "A Survey of Queueing Network Packages for the Analysis of Manufacturing Systems," *Manufacturing Review*, **1** (1988): 14-25.

Suri, R. and Diehl, G. W., "Rough Cut Modeling: An Alternative to Simulation,". *CIM Review*, **3** (1987): 25-32.

Suri, R., "Robustness of Queuing Network Formulas," *Journal of the Association for Computing Machinery*, **30** (1983): 564-594.

Voss, C. A. and Robinson, S. J., "Application of Just-in-Time Manufacturing Techniques in the United Kingdom," *International Journal of Operations and Production Management*, **7** (1987): 46-52.