Proceedings of the 2001 Winter Simulation Conference
B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, eds

SCHEDULING SETUP CHANGES AT BOTTLENECK FACILITIES IN SEMICONDUCTOR MANUFACTURING

Zaid Duwayri

i2 Technologies
909 E. Las Colinas Blvd.
Irving, TX 75039, U.S.A.

Mansooreh Mollaghasemi

Department of Industrial Engineering
and Management Systems
University of Central Florida
4000 Central Florida Blvd.
Orlando, FL 32816, U.S.A.

Dima Nazzal

School of Industrial Engineering and Management Systems
University of Central Florida
4000 Central Florida Blvd.
Orlando, FL 32816, U.S.A.

ABSTRACT

In this paper, a scheduling heuristic was developed to aid the operators in semiconductor fabs in choosing what type of lots to process next on bottleneck facilities and whether to change machine setup in order to reduce cycle time. The scheduling heuristic aims at balancing workload levels for implanters processing lots at different stages of the wafer production lifecycle. This is accomplished by processing lots that contribute most to increasing inventory levels at the bottleneck facility. A whole production line simulation model was used to evaluate the performance of the scheduling heuristic and to compare it against several commonly used scheduling heuristics with respect to mean cycle time, work in process (WIP), and standard deviation of cycle time. Simulation results showed that the heuristic performed better than all other rules in terms of mean cycle time and WIP in all cases, and better in terms of standard deviation of cycle time for most cases tested.

1 INTRODUCTION

There are several industry performance measures used to evaluate the current production policy in a semiconductor fab. These include work in process (WIP) levels, throughput rates, and product cycle times, with the latter being the most frequently monitored measure.

In the semiconductor wafer fabrication process, active circuit elements, such as transistors, are created by selectively introducing impurities (a process commonly referred to as doping) to the wafers. Two commonly used doping techniques are diffusion and (more commonly) ion implantation. In ion implantation, impurities are introduced into unprotected areas of the wafer at room temperature by accelerating dopant ions (atoms stripped of one or more of their electrons) to a high energy so that they are driven into the wafer and become embedded. A full setup change is required when changing from a certain type of dopant to another.

As a part of the complex production line that exists in a semiconductor wafer fabrication facility, implanter workstations are considered to be a bottleneck. The reentrant flow of production passes several times through the implanters at different stages of the wafer production, which may require changes to the current settings of the workstations and thus incurring a significant setup time. Figure 1 depicts the reentrant flow of production that flows, back and forth, between the bottleneck facility and the rest of the facilities in the fab. In this figure, each arrow denotes a loop, where a loop is defined as a set of manufacturing operations between two consecutive implant operations. Material flow in the fab may be smoothed by balancing the workload of each loop. In other words, it would be better if setup changes were performed so that the difference in workloads among the loops is not too large.

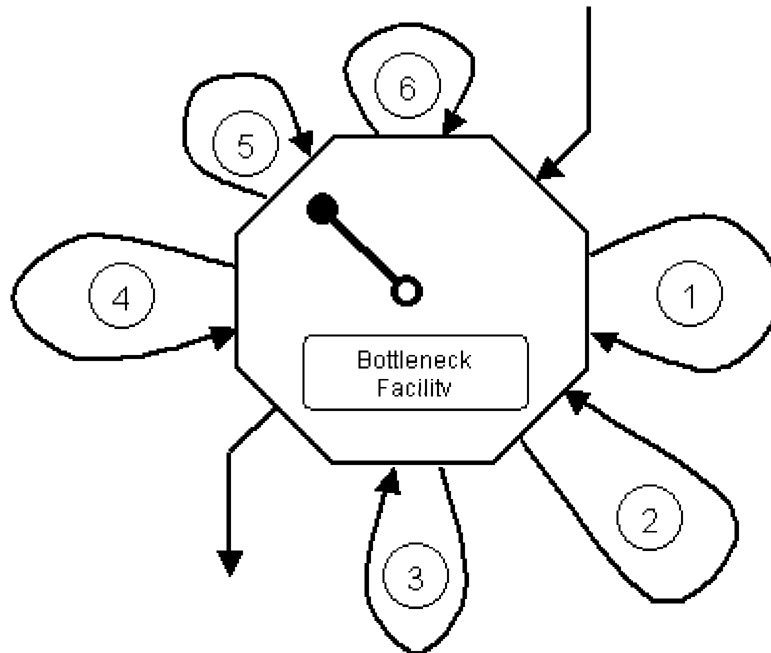

Figure 1: An Illustration of the Reentrant Flow to the Implanter Facility

The purpose of this study is to evaluate a scheduling heuristic to aid the operators in choosing what type of lots to process next and whether a species change is justified with respect to the performance measures monitored by the management. Currently, operators choose lots based on commonly used scheduling rules such as the Same Setup rule. The Same Setup rule recommends choosing the lot that requires no change in the current setup, thus minimizing the number of setup operations required.

Several researchers have studied and compared the use of specially tailored dispatching rules in different areas within the wafer fabrication process. Researchers focused on dispatching rules dealing with inventory levels throughout the fab. Johri (1993) suggested that dispatching rules aid in alleviating long queues in front of bottleneck facilities, reducing cycle-time variance of individual lots, and developing a systematic way to speed up processing of some lots and slow down others to help meet delivery requirements. Li et al. (1996) suggested a scheduling policy that targets reducing inventory levels variability at different workstations. The one-step-ahead policy presented in their paper assigns priority levels to lots in queue for different operations based on the current queue for that operation and the following operation.

Kim et al. (1998) suggested a policy to deal with scheduling production at photolithography operations that require setup time. The authors developed an index to measure the workload at different areas along the production line. The suggested heuristic selects the processing of those lots that would help in reducing the workload within congested areas along the production line.

Chin et al. (1995) introduced a new scheduling policy that takes into account the lots' due date and the required turn rate, and aims at increasing throughput and machine utilization.

In this research, we focused on developing and evaluating a scheduling heuristic for bottleneck facilities that require large setup time and testing this heuristic under a single machine-single product environment.

Section 2 discusses the definition of bottleneck facilities and summarizes their identification process. Section 3 describes the reentrant flow property, the work in process and other performance measures used in this study to evaluate different dispatching rules. In Section 5, the workload calculation procedure is presented and the loop heuristic is discussed. Section 6 will provide a description of the simulation model used throughout the experiments. Sections 7, 8 and 9 present the results of comparing the loop heuristic to other scheduling rules. Sections 10 and 11 present the conclusions and suggestions for future research, respectively.

2 IDENTIFYING BOTTLENECK FACILITIES

A bottleneck facility is, as the name implies, a production facility that constricts the smooth flow of production on the production line. A bottleneck facility can be identified by inspecting queue lengths, machine utilizations, or loading levels.

3 REENTRANT FLOW AND WIP LEVELS

Due to the reentrant nature of the production flow in a semiconductor wafer fabrication facility, the WIP levels at

different stages of production could vary considerably. Schedulers developed several scheduling policies that dealt with the variability of inventory levels and presented evidence that a direct correlation exists between the variability in the inventory levels and the mean and variance of the cycle time (Li et al., 1996 and Kim et al., 1998).

The loop heuristic is based on the principle that if the workloads at different stages of production were balanced, this would lead to a smoother production flow and reduce average delays, thus reducing the average cycle time and the variance of cycle time.

Smooth and efficient operation of a bottleneck facility that processes several stages of the wafer production is very important since it constitutes a major node in the reentrant production flow network. This type of facility can usually process one step at a time, keeping wafers at other stages of production waiting in queue for processing.

The production flows into the bottleneck facility and out again several times during the production lifecycle, forming what is referred to herein as loops. A loop would be a group of processing steps that occur between two consecutive steps on the bottleneck facility. The bottleneck facility would be able to process one loop at a time, moving lots to the next loop. Delaying the switch from processing a certain loop to processing a different one would causes the loop to be congested with lots waiting in queue.

4 PERFORMANCE MEASURES

The performance measure that is mainly monitored in a semiconductor wafer fabrication facility is the cycle time (flow time, as it is sometimes called). Cycle time is the sum of the processing times of all steps required to complete a product plus all the waiting or queue time the product incurs in the process. The semiconductor manufacturers seek to reduce cycle time as a competitive advantage in this highly volatile industry, where prices change quickly and any reduction in cycle time is considered significant. The variance (or standard deviation) of the cycle time is another important performance measure that directly affects customer due-dates and shipping predictions. The lower the variance, the more accurate the ship-date predictions would be. This results in fewer past-due shipments and unsatisfied customers. Other important performance measures are the work in process (WIP), throughput, and machine utilization. In a stable manufacturing environment, WIP and throughput are related to the cycle time using Little's Law:

WIP = Cycle Time × Throughput.

The mean cycle time, the standard deviation of cycle time, and the average WIP level, are the measures to be used to evaluate the performance of the developed scheduling heuristic in this study.

5 ESTIMATING THE WORKLOAD

In order to estimate the workload within each of the loops, a workload index is used (Kim et al. 1998). The workload index is a measure of the workload that the bottleneck facility would have to process from each loop, and is calculated as follows:

$$W_i = \frac{p_i \times n_i}{T_i} \tag{1}$$

where
W_i = is the workload index of loop i.
p_i = is the processing time of the step at the end of loop i (on the bottleneck facility)
n_i = is the number of lots that are in queue or in process at the steps in loop i.
T_i = is the sum of the processing times of all steps in loop i.

5.1 The Loop Heuristic (Single Machine)

In order to balance the workload between the loops, the bottleneck facility has to switch its processing to another loop which has a higher workload index. It should not be overlooked, however, that each time a setup change is performed a certain amount of setup time is incurred, so frequent switching is costly and should be avoided.

In order to be able to determine the appropriate timing for a setup change, a Determination Coefficient (δ) has to be defined and included in the heuristic. The steps for implementing the loop heuristic are as follows:

1. Calculate the value of the workload index for all loops, including the one currently being processed.
2. Calculate the ratio of the highest workload index to the workload index of the loop currently being processed.
3. If the calculated ratio is larger than the Determination Coefficient (δ), then a setup change is justified and the facility should be set up to process the loop with the highest workload index.

The value of δ is dependent on several parameters of the system, such as the setup time, total number of loops involved, and number of products. Simulation experiments would help in determining the optimal value of δ for the system.

6 SIMULATION MODEL

In order to evaluate the performance of the suggested heuristic, a series of simulation experiments were performed. The simulation model that was used is a whole line simulation model for the wafer production fab at Cirent Semiconductor, in Orlando, Florida.

The software used in modeling the fab is AutoSched AP, from AutoSimulations, Inc. This simulation software is a spreadsheet (Excel) based software that defines parameters and distributions using spreadsheet pages and cells that are linked together. This software is optimized for the modeling of semiconductor fabs and is further customized for the use of Agere Systems.

6.1 Main Components of the Model

In order to build a simulation model several data files should be defined, those include

- Station file: This set of parameters defines the main resources in the factory. Each station is a machine or work area. These stations represent constraints on the capacity of the manufacturing system.
- Part file: This set of parameters defines the types of products that are manufactured in the facility. The part file specifies the name and the routing of each part type and the route that it follows to get produced.
- Route file: This set of parameters defines the processing steps that parts must go through to be manufactured. Each step in the route uses a resource.
- Order file: This set of parameters defines the lots and their start time.
- Options file: This file defines the parameters of the simulation, such as the start and end times of the simulation, and the type of reports to generate.
- Operator file: This set of parameters defines the number of operators available at each processing step and their availability.

6.2 Settings

A single product model was used in the simulation experiments. Eight operators were assigned to the implant area. The wafer starts (production starts) associated with this set of experiments are summarized in Table 1.

Table 1: Wafer Starts Settings for the Experiments

Experiment	Product	Number of masks	Weekly Wafer starts
Single Machine – Single Product	Product A	17	840

In this research, a single machine was assumed to be available in the implanters facility group. This provided simplicity in debugging the performance of the loop rule and in conducting the analysis.

The warm up period of the simulation run was set at 100 days. The use of 5 replicates turned out to be an appropriate setting for the simulation model and provided an adequate level of accuracy.

7 OPTIMAL VALUES FOR DETERMINATION COEFFICIENT

At each level of setup time, a series of simulation runs were performed in which the value of δ was varied. The purpose of these runs was to determine the "optimal" value of δ at which the cycle time is lowest, which would then be used in comparing the results of the suggested heuristic against those of other commonly used rules. Figure 2 shows the cycle time values as the coefficient determination is increased from 0 to 100. The experiments were performed at 10, 20 and 30 minute setup times.

Table 2 below summarizes the optimal settings for the Determination Coefficient δ at each setup value.

Table 2: Optimal Values for δ

Setup Time	Optimal δ
10	>=20
20	>=30
30	>=40

8 COMPARING THE RESULTS OF THE LOOP RULE AGAINST OTHER SCHEDULING RULES

In order to measure the overall performance of the loop heuristic, the results of using the suggested heuristic are compared to those of using three commonly used scheduling rules:

- Same Setup rule,
- First Come First Served (FCFS) rule, and
- Earliest Due Date (EDD) rule.

The performance measures that were used for the comparison were the mean and variance of the cycle time, and the work in process (WIP). The value of δ selected for the loop rule was the value of this coefficient at which the cycle time curve stabilizes (Figure 2, Table 2). Table 3 presents the percentage of improvement in three performance measures resulting from applying the loop rule when compared to applying other common rules.

The results show that the loop heuristic has performed better than the three other rules at all levels of setup time. They also show that as the setup time increases, the performance of the loop heuristic, as compared to the performance of the commonly used rules, improves significantly. For demonstration purposes, Figure 3 plots the average cycle time in hours versus the value of δ for a 20 minutes setup time. As the value of δ increases, the performance of the loop heuristic improves, and it surpasses that of the same setup rule at a δ value of 15.

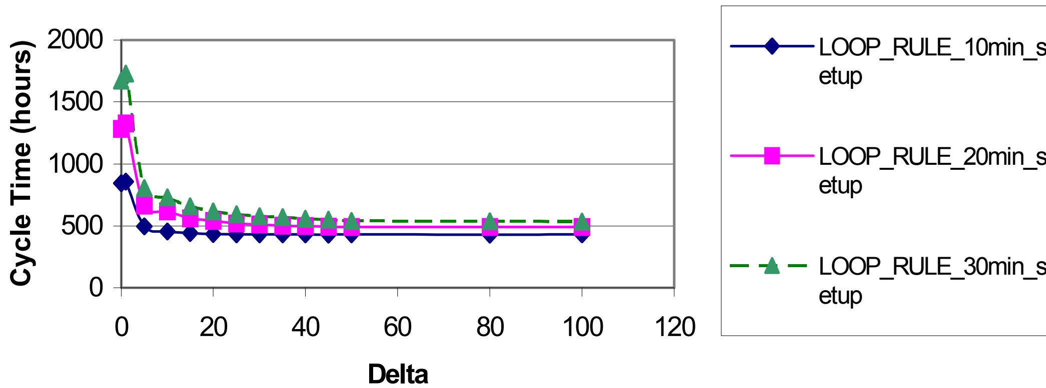

Figure 2: Cycle Time versus δ at 10, 20 and 30 Minutes Setup

Table 3: Percentage Improvement in Performance Measure Resulting from Applying The Loop Rule Compared to Three Common Scheduling Rules

Setup Time	Rule	% Improvement in cycle time	% Improvement in standard deviation of cycle time	% Improvement in work in process
10 minutes	Same Setup	9%	5%	9%
	FCFS	71%	94%	71%
	EDD	56%	88%	56%
20 minutes	Same Setup	15%	3%	15%
	FCFS	79%	93%	79%
	EDD	68%	89%	68%
30 minutes	Same Setup	19%	2%	19%
	FCFS	81%	93%	81%
	EDD	72%	89%	72%

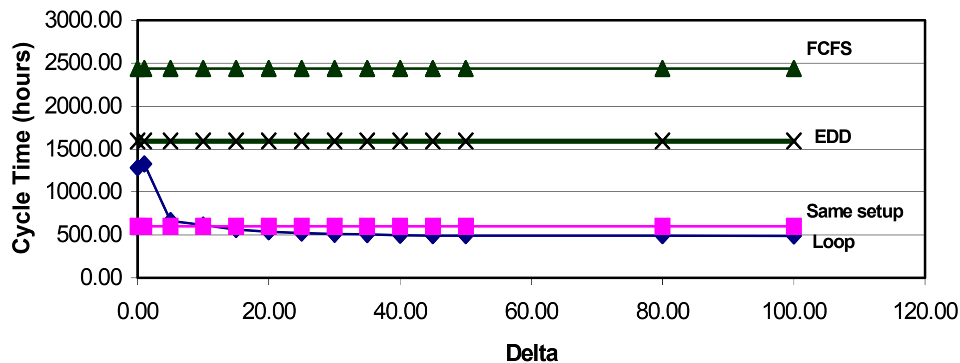

Figure 3: Cycle Time vs. δ for All Scheduling Rules at 20 Minutes Setup Time

9 TEST OF HYPOTHESIS

In order to statistically conclude that the loop heuristic performs better than other scheduling rules, it is necessary to perform a hypothesis test. The null hypothesis states that the mean value of the performance measure when using the loop rule is equal to the performance measure value when using any of the other scheduling rules. The alternative hypothesis states that the mean value of the performance measure when using the loop rule is smaller than that of using any other scheduling rule. Tables 4-6 summarize the results of the test of hypotheses for the three performance measures evaluated in this study.

Table 4: Test of Hypothesis for Mean Cycle Time Comparison

Setup time	Rule	T_o^*	v	$t_{0.05,v}$	Null hypothesis rejected?
10 minutes	Same Setup	50.46	6	1.943	Yes
	FCFS	582.7	7	1.895	Yes
	EDD	367.8	8	1.860	Yes
20 minutes	Same Setup	31.5	10	1.812	Yes
	FCFS	509.2	9	1.833	Yes
	EDD	370.4	10	1.812	Yes
30 minutes	Same Setup	26.2	6	1.943	Yes
	FCFS	620.2	8	1.860	Yes
	EDD	453.5	10	1.812	Yes

Table 5: Results of Test of Hypothesis for Standard Deviation of Mean Cycle Time Comparison

Setup time	Rule	F_0	f_{α,n_1-1,n_2-1}	Null hypothesis rejected?
10 minutes	Same Setup	1.1	6.39	No
	FCFS	238.8	6.39	Yes
	EDD	73.9	6.39	Yes
20 minutes	Same Setup	1.0	6.39	No
	FCFS	216.7	6.39	Yes
	EDD	77.7	6.39	Yes
30 minutes	Same Setup	1.0	6.39	No
	FCFS	191.4	6.39	Yes
	EDD	76.9	6.39	Yes

Table 6: Results of Test of Hypothesis for WIP Comparison

Setup time	Rule	T_o^*	V	$t_{0.05,v}$	Null hypothesis rejected?
10 minutes	Same Setup	49.7	6	1.943	Yes
	FCFS	597.6	7	1.895	Yes
	EDD	349.1	8	1.860	Yes
20 minutes	Same Setup	32.4	10	1.812	Yes
	FCFS	563.9	9	1.833	Yes
	EDD	327.5	10	1.812	Yes
30 minutes	Same Setup	25.5	7	1.895	Yes
	FCFS	785.1	10	1.812	Yes
	EDD	502.5	10	1.812	Yes

The results displayed in Table 4 show that the null hypotheses for all combination of scheduling rules and setup times have been rejected, concluding that the mean cycle time obtained by using the loop heuristic is significantly lower than that when using any of the other scheduling rules.

The results of the hypotheses testing displayed in Table 5 show that the mean standard deviation of the cycle time when using the loop rule is better than that when using the FCFS and the EDD scheduling rules. The test, however, failed to reject the null hypotheses that the performance of the loop heuristic and that of the Same Setup rule, with respect to the standard deviation of the cycle time, are equal.

The results of the hypothesis testing displayed in Table 6 show that the loop heuristic performs better than the other rules with respect to the average WIP level.

10 CONCLUSIONS

The loop scheduling heuristic that was presented in this research dealt with the inventory levels of production at different stages of the reentrant flow. A workload index was used to estimate the amount of workload that the bottleneck facility has to process at different stages of the production life cycle. Lot processing was scheduled for the group of lots with the highest workload index.

A whole line simulation model was used for testing the proposed approach, and the performance measures used to evaluate the scheduling heuristic were the average cycle time, average WIP level, and the standard deviation of the cycle time. The effects of several factors on the performance of the suggested heuristic were evaluated. These include:

- The Determination Coefficient (δ).
- The amount of setup time required.

The performance of the loop heuristic was compared to three commonly used scheduling rules: Same Setup, FCFS, and EDD.

The following points are concluded from this research:

1. The values of the Determination Coefficient (δ) significantly affected the performance of the loop heuristic. Increasing the value of δ reduced the frequency of setup changes performed.
2. The value of δ associated with the best performance of the loop heuristic (optimal δ) varied as the setup time was changed. As the setup time was increased, the value of the optimal δ increased as well.
3. The loop heuristic performed better than any of the other scheduling rules included in the comparison at all settings with respect to average cy-

cle time and the average WIP. The loop heuristic also performed better than FCFS and EDD with respect to standard deviation of the cycle time; however, it did not perform better than the Same Setup rule with respect to the same performance measure.

11 FUTURE WORK

Multiple stations for the bottleneck facility might be used in a future study to evaluate the performance of the loop heuristic with the presence of several identical machines.

The loop heuristic could be modified to account for different setup times by adding weights to the procedure of ranking the loops according to the workload index. High weights can be given to loops that do not require setup changes, while lower weights are given to the loops that require setup changes depending on the associated setup time. This modification would favor processing of loops with little or no setup time over loops with high setup time that could affect the total cycle time of the products.

ACKNOWLEDGMENTS

At Cirent Semiconductor, we would like to thank Jeff Neve for supporting us throughout this research and providing us with all the data that we needed.

REFERENCES

Chin, W., Wang, J., Lin, K. and Huang, S. 1995. New methodology of dynamic lot dispatching required turn rate. *Proceedings of the International Manufacturing Technology Symposium*. 187-189.

Johri, P.K. 1993. Practical Issues in Scheduling and Dispatching in Semiconductor Wafer Fabrication. *Journal of Manufacturing Systems*. 12 (6) :474-485.

Kim, Y., Lee, Y., and Kim, J. 1998. A simulation study on lot release control, mask scheduling and batch scheduling in semiconductor wafer fabrication facilities. *Journal of Manufacturing Systems.* 17 (2) :107-117.

Law, A.M. and Kelton, W.D. 2000. *Simulation Modeling and Analysis*. 3rd Edition. McGraw-Hill.

Li, S., Tang, T., and Collins, W. 1996. Minimum inventory variability schedule with application in semiconductor fabrication. *IEEE Transaction on Semiconductor Manufacturing*. 9 (1) :145-149.

Montgomery, D.C. and Runger, G.C. 1994. *Applied Statistics and Probability for Engineers*. John Wiley & Sons, Inc.

Wein, L.M. 1998. Scheduling semiconductor wafer fabrication. *IEEE Transaction on Semiconductor Manufacturing*. 1 (3) :115-130.

AUTHOR BIOGRAPHIES

ZAID DUWAYRI is an application engineer at i2 Technologies. He received an M.S. degree in Industrial Engineering from the University of Central Florida. He received a B.S. in Industrial Engineering from the University of Jordan, Amman, Jordan. He is a member of ASQ. His email address is <Zaid_Duwayri@i2.com>.

MANSOOREH MOLLAGHASEMI is an associate professor of Industrial Engineering and Management Systems at the University of Central Florida. She received her Ph.D. in Industrial Engineering from the University of Louisville in 1991. She also holds a B.S. and an M.S. in Chemical Engineering from the same University. Her research interests involve simulation modeling and analysis, multiple response simulation optimization, and neural networks. She is a member of IIE and INFORMS. Her email address is <mollagha@mail.ucf.edu>.

DIMA NAZZAL received an M.S. degree in Industrial Engineering from the University of Central Florida. She received a B.S. in Industrial Engineering from the University of Jordan, Amman, Jordan. Her present research interests include simulation modeling and analysis, design of experiments and capacity analysis and planning. She is a member of ASQ. Her email address is <dnazzal@mail.ucf.edu>.