Proceedings of the 2008 Winter Simulation Conference
S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds.

DECISION MAKING AND FORECASTING WITH RESPECT TO RISK: A SIMULATION STUDY FOR A SETUP-PROBLEM

Martin Romauch
Walter Laure
Georg Seidel

Infineon Technologies Austria AG
9500 Villach, AUSTRIA

Christian Almeder

Dept. of Business Administration
University of Vienna, Brünner Str. 72
1220 Vienna, AUSTRIA

ABSTRACT

In this paper, the impact of timing a setup operation on the work-in-process and cycle time is investigated. Using a small example of an reentrant production system, it is shown that there is a tradeoff between reducing the average work-in-process inventory and guaranteeing a smooth operation of the system. Furthermore it is shown, that an early setup may lead to an increased variance of the cycle time, but regarding the number of lots with an extreme excess of cycle time the timing decision has only a minor influence.

1 INTRODUCTION

Facing a changing product mix in a semiconductor fab, it is often necessary to reallocate machine capacities to balance the workload. In many situations it is possible to react directly and without delay at low cost by changing dispatching rules or dedications. If these reactions are insufficient, new resources are needed to guarantee a fluent throughput for the critical processes. Avoiding capital investments, it is preferred to attempt a reallocation of resources, e.g. by retooling decisions.

Implementing these decision is usually time consuming and carry some risk. We discuss appropriate criteria to choose a "good" time for the setup by balancing merits of starting early and starting late.

Most of the research literature concentrates on scheduling and setup decisions and their impact on the average work-in-process and throughput, but only a few articles are considering also the variability of these measures, which is essential for running a production system smoothly. For example, Li, Tang, and Collin (1996) developed a special scheduling technique to reduce the variability of the work-in-process. Kim et al. (2007) presented a deterministic work-in-process balancing model where they considered setups necessary when switching between different lots. Duwari et al. (2007) showed that it is important to schedule setup changes for the bottleneck machines at the right time. Samaddar (2001) and Samaddar and Hill (2007) observed that the reduction of setup time in a cyclic production system may increase work-in-process if the variance of the setup time is not controlled.

This article will show for a typical setup decision of an reentrant production system, that the timing of the setup might be considerably different if the objective is to avoid high work-in-process levels instead of trying to keep the average work-in-process inventory low.

2 METHODOLOGY

2.1 Problem description

The problem under investigation is as follows. It is assumed that for two different types of products (e.g. an old and a new version of the same product) there is a bottleneck process which is served by three parallel machines. The machines have to be tooled accordingly to process one of the products. A switch from one product type to the other causes a downtime necessary to retool and reconfigure the machine. It is assumed that the total release of lots of both product types is constant over time, i.e. the total work load of the fab stays fairly constant over the considered time horizon. What is changed is the product mix of the lots released to the fab. It is started with a relatively low fraction of the second product which is increased continuously until it reaches a level twice as high as that of the first product. Initially, most of the machines are tooled for the first product type and only a few are tooled for the second type. As typical for a semiconductor fab, the production process is a reentrant flow, i.e. all the products have to be processed several times on the critical machines. Between these different rounds the products might be processed on other (non-critical) machines in the fab as well.

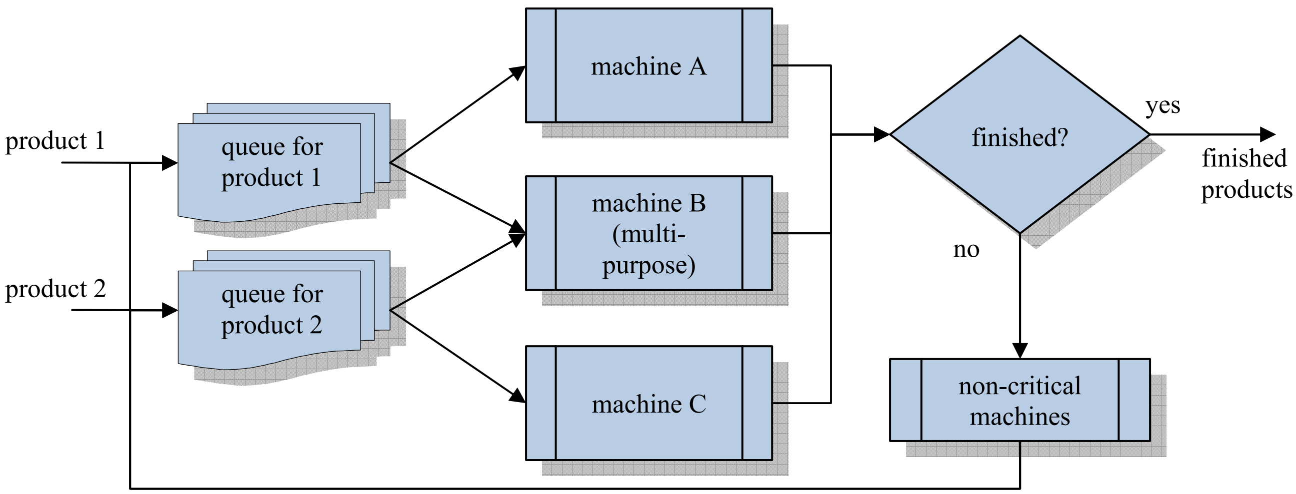

Figure 1: Layout of the simulation model

The research question investigated using this study is to show the impact of the timing of a setup change on the average work-in-process as well as how often a critical work-in-process level is reached.

2.2 Simulation model

In order to answer the above research question, a small simulation model is used (see Figure 1). Lots for 2 different products are generated and stored in the according queues waiting to be processed on the bottleneck machines A to C. Machine A can only process product 1 and machine C can only process product 2, whereas machine B is a multipurpose machine that can process both products. But before the change of the products processed on machine B a setup is necessary for retooling the machine. During this setup the machine is blocked and cannot process any products. After the lots have been processed on one of the machines, they might be processed on some other (non-critical) machines and return to the queues in front of the bottleneck process, or if they have finished all their operations they leave the system.

3 SIMULATION EXPERIMENT

3.1 Parameters

Initially machine B is set up for product 1. The arrival of lots for product 1 is exponential distributed with an arrival rate twice as high as those of product 2. After a lead in of 10 periods (=days) the arrival rate of product 2 is increased and that of product 1 is decreased such that the total arrival rate stays constant. After 50 periods the arrival rate of product 2 is twice as high as that of product 1 and will stay constant. Only once during the simulation run it is possible to switch from product 1 to product 2 on machine B. Further on, the impact of the timing of this switch on the work-in-process and the cycle time will be investigated. Detailed parameters can be found in Table 1.

Table 1: Parameters

Parameter	Value(s)
total arrival rate	0.87
lead in	10 periods
arrival rate ramp up of product 2	50 periods
total number of periods	800 period
processing time on bottleneck machines	triangular distribution (0.9, 1, 1.1)
processing time on non-critical machine	triangular distribution (0.5, 1, 1.5)
capacity on non-critical machine	20 lots
number of loops per lot	3
duration of the setup process	10 periods (fixed) or U[5,15] or U[0,20]

3.2 Experiments

A setup change of machine B was tested for each period between 10 (start of the ramp up of product 2) and 60 (end of the ramp up). For each of these switching times 300 simulation runs have been performed each for a duration of 800 periods. Pretests have shown that only after 600-800 periods the whole system reaches a steady state. For each of the simulation experiments the work-in-process for each period as well as the cycle time for each finished lot has been stored in a database. Three scenarios with an increas-

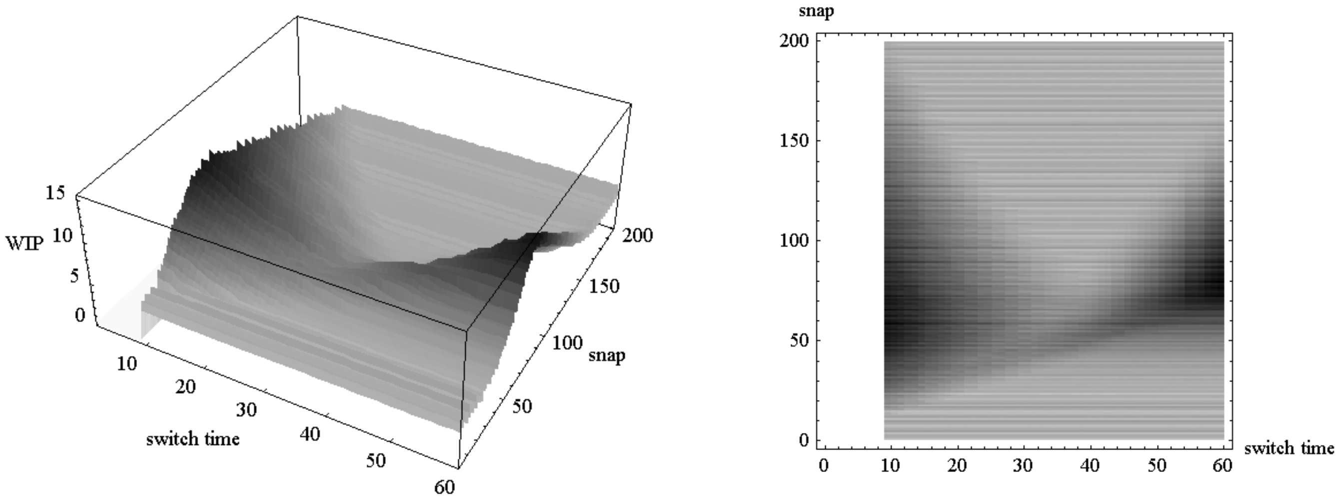

Figure 2: Averarge work-in-process for uniformly distributed setup time between 5

ing amount of stochasticity in the duration of the setup process are generated:

- **(i) 10**: the duration of the setup process is deterministic (i.e. duration of the setup process equals 10).
- **(ii) uniform(5,15)**: the duration of the setup process is uniformly distributed between 5 and 15.
- **(iii) uniform(0,20)**: the duration of the setup process is uniformly distributed between 0 and 20.

Note, the expected value of the duration of the setup process is the same in each scenario.

4 RESULTS

Figure 2 gives a first impression of the influence of the choice of the switching time on the average work-in-process for scenario. On the left side of the plot the dark area represents situations with large values of the average work-in-process. For a switching time near 10 (beginning of the ramp up) one can see that it takes very long to reduce the work-in-process of product 1, and for switching times near 50 the situation is less critical. The reason is that for a low switching time the arrival rate of product 1 is still high while only one machine is dedicated to this product, whereas for high switching times there two machines dedicated to product 2 which provide a fast reduction of the high work-in-process (caused by product 2).

The picture only shows the scenario (ii), but the situation is very similar in all cases. By visual inspection, switching times smaller than 20 or larger than 50 (lines that don't cross the saddle region) aren't favorable. To take a closer look at the reasonable switching times the average work-in-process (holding cost) was set in contrast with the frequency of large work-in-process (risk of a bottleneck situation). The red curves in Figure 3 represent the average work-in-process. Regarding the different scenarios the minimum (around 38) is quite stable. The blue curves represent the frequency of days where the average work-in-process is larger than 10. This is more than twice the minimum average work-in-process. Depending on the level of randomness in the duration of the setup process the minimum moves towards earlier switching times. For all situations a switching time around 30 is optimal for avoiding situations of high work-in-process.

There is a significant tradeoff between the holding costs (determined by the average work-in-process) and the risk to excess work-in-process thresholds, which may cause critical situations in the production system.

Regarding the cycle time Figure 4 shows a comparison of the three scenarios, measuring the percentage of lots with a cycle time larger than 18 (which is equivalent to a flow factor of 3). At switching times between 30 and 40 the cycle times stay at low levels such that it is possible to choose good switching times regarding the work-in-process without delaying lots.

5 CONCLUSIONS

Based on the results of the simulation experiment one might draw the following conclusions regarding the timing of setups:

1. With respect to situations with very high work-in-process, switching very early is more dangerous than switching very late.
2. The level of randomness of the setup time has a stronger impact on the frequency of situations

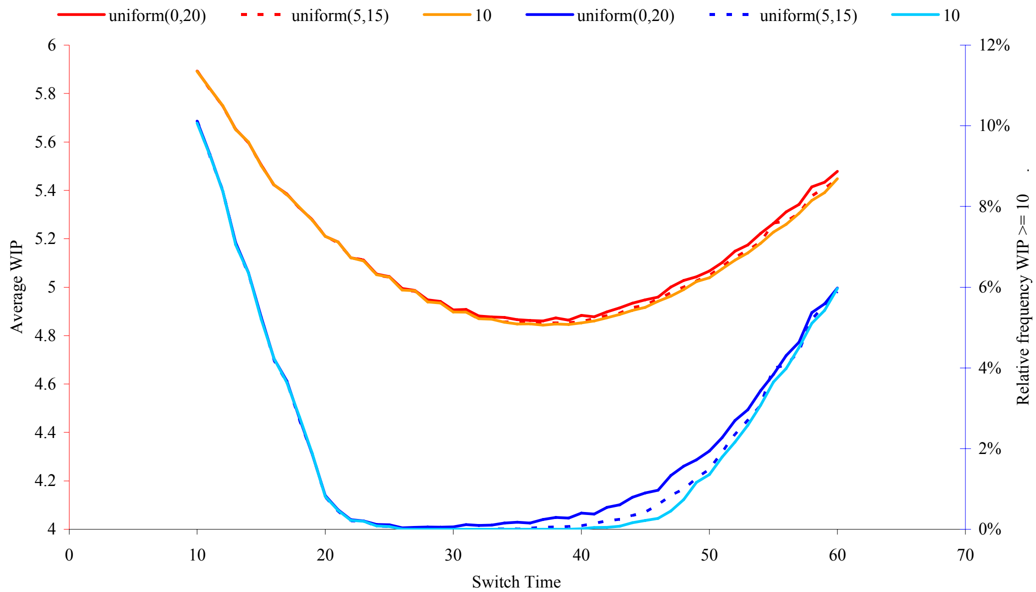

Figure 3: Average work-in-process and relative frequency of snaps with a work-in-process larger than ten.

with a high work-in-process than on the average work-in-process.

A decision maker who wants to avoid bottleneck situations will tend to switch earlier, especially in the case of "large" randomness in the duration of the setup process.

REFERENCES

Duwari, Z., Mollaghasemi, M., Nazzal, D., and Rabadi, G. 2006. Scheduling setup changes at bottleneck workstations in semiconductor manufacturing. *Production Planning & Control* 17:717-727.

Li, S., T. Tang, and D.W. Collins 1996. Minimum inventory variability schedule with applications in semiconductor fabrication. *IEEE Transactions on Semiconductor Manufacturing* 9:145-149.

Kim, S., Y.H. Lee, T. Yang,, and N. Park. 2007. Robust production control policies considering WIP balance and duration of the setup process in a semiconductor fabrication line. *International Journal of Advanced Manufacturing Technology:*doi: 10.1007/s00170-007-1208-4.

Samaddar, S. 2001. The effect of duration of the setup process reduction on its variance. *Omega* 29:243-247.

Samaddar, S., and C.A. Hill,. 2007. Controlling adverse effect on work in process inventory while reducing machine setup time. *European Journal of Operational Research* 180:249-261.

AUTHOR BIOGRAPHIES

MARTIN ROMAUCH is a member of the simulations department at Infineon Austria AG. He received a master's degree in applied mathematics at the University of Klagenfurt and a Ph.D. in the area of Business Administration from the University of Vienna. His main research interests are Complexity Theory, Metaheuristics and mathematical Modeling.

WALTER LAURE is a head of the simulations department at Infineon Austria AG. He received a master's degree in applied mathematics at the University of Graz. He is working on Scheduling and Simulation Topics in Semiconductor Industry.

GEORG SEIDEL is member of the simulations department at Infineon Austria AG. He received a master's degree in applied mathematics at the University of Graz. He is focused on Simulation of Tools in Semiconductor Industry, e,g. Cluster tools.

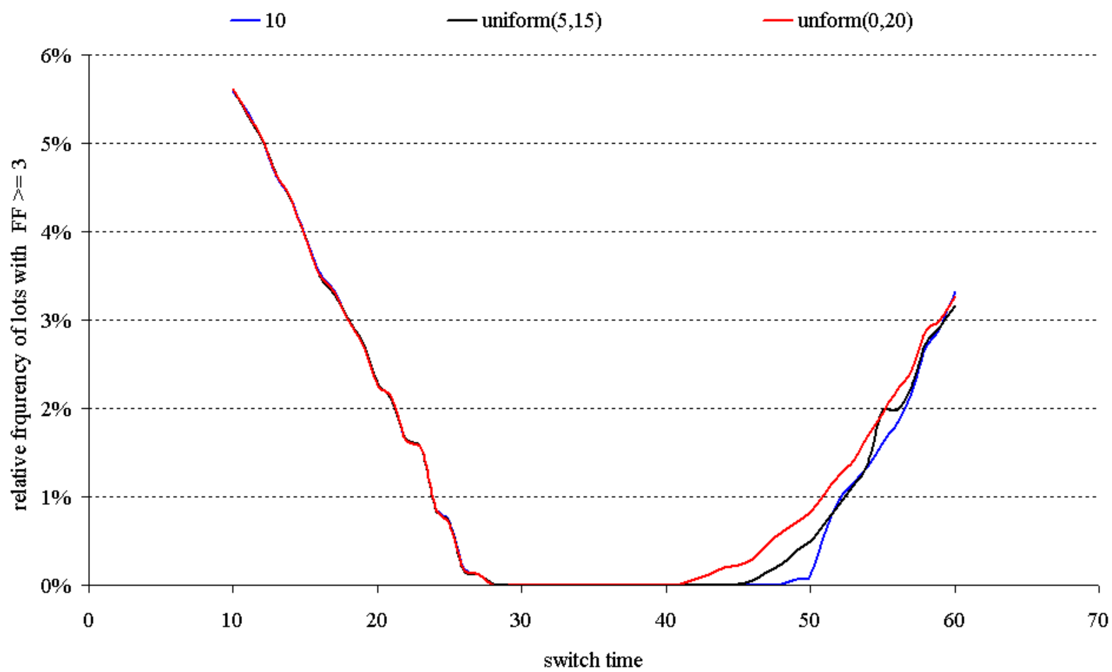

Figure 4: Relative frequency of lots with a flow factor larger than three.

CHRISTIAN ALMEDER is assistant professor at the Department of Business Administration of the University of Vienna. His research interests are discrete-event simulation, combinatorial optimization, metaheurstics, and optimal control models in the context of production and logistic systems. His web page can be found via <www.univie.ac.at/bwl/prod>