Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

RESERVATION BASED DISPATCHING RULE FOR WAFER FAB WITH ENGINEERING LOTS

Yong H. Chung
Sang C. Park
Dept. of Industrial Engineering
Ajou University.
Woncheon, Suwon, KOREA

Byung H. Kim

VMS Solutions Co., Ltd.
Hanshin S-MECA 611, Gwanpyeong, Yuseong-gu, Daejeon, 305-509, KOREA

Jeong C. Seo
Samsung Electronics Co., Ltd
San #16 Banwol-Dong, Hwasung-City,
Gyeonggi-Do, 445-701, KOREA

ABSTRACT

Presented in this paper is a dispatching rule for engineering lots by a wafer FAB. The proposed rule uses the concept of reservation to ensure greater capacity for the engineering lots. Although a reservation based rule was previously proposed by one of the authors, it has remained difficult to ensure the necessary capacity for the engineering lots because of variations in the processing time for each step of the process. In this paper, we enhanced the previously proposed rule to reflect the fact that batch tools often cause bottlenecks because of their long processing times. We developed a FAB model using the Measurement and Improvement of Manufacturing Capacity (MIMAC) dataset 6, and performed simulations with MozArt®. The simulation results clearly show the advantages of the enhanced reservation based dispatching rule over the previous version of the rule and that it is vital to identify an appropriate reservation range.

1 INTRODUCTION

As the life cycle of electronics products are reduced in the continuously changing marketplace, semiconductor manufacturers have tried to develop new chip-products in the most cost effective manner possible (Park et al 2013; Jia and Mason 2009). To survive and prosper in the modern semiconductor industry, it is necessary to perform consistent engineering research as well as production. Engineering performance is mainly estimated on cycle time, because it is important to process rapidly engineering lots for faster implementation of product improvements and more rapid deployment of new products to the market ahead of one's competitors (Chung and Huang 2002; Crist and Uzsoy 2011). Because of the importance of engineering work, there have been various studies considering the lot (high priority lot, low volume product) that requires a short cycle time similar to engineering lot. In 2011, Rezaei et al. proposed a release and dispatching policy that consider high priority lots. It integrates the theory of constraint (TOC) and the workload control (WLC) to improve the performance (Rezaei et al. 2011). In 2012, Zhou and Rose found that achieving a WIP balance for high volume products negatively affects the on-time delivery of low volume products. To solve this problem, they proposed a dispatching rule for achieving the on-time delivery of low volume products as well as the WIP (work-in process) balance of high volume products (Zhou and Rose 2012). In 2014, Chung et al. proposed a reservation based dispatching rule for the on-time delivery of high priority orders. Here, the reservation means the provisional allocation of capacity to support a guarantee of tool capacity required to produce high priority orders (Chung et al 2014). The proposed

reservation based dispatching rule considers the high priority lots being processed in previous tools as well as the waiting lots in the queue. Once a tool is reserved, then it has to wait until the arrival of the reserved high priority lot. Although there have been various dispatching rules for on-time delivery, still FABs have great difficulties to achieve on-time delivery. To overcome these limitations, it is necessary to carefully observe the attributes of the process whereby semiconductors are fabricated on a wafer. The number of steps can be range of 300 to 600 steps, depending on the products. The processing times can vary depending on the step being performed. In particular, the batch processing steps performed in a wafer FAB takes as long as 10h, while the other steps in the process take only 1- to 2 hours (Sarin et al. 2011; Malapert et al 2012), which possibly leads to the creation of a long queue. Therefore, if the engineering lots are not properly managed, then it might result in long cycle times (the sum of processing times and waiting times) of engineering lots (Bixby and Burda 2008).

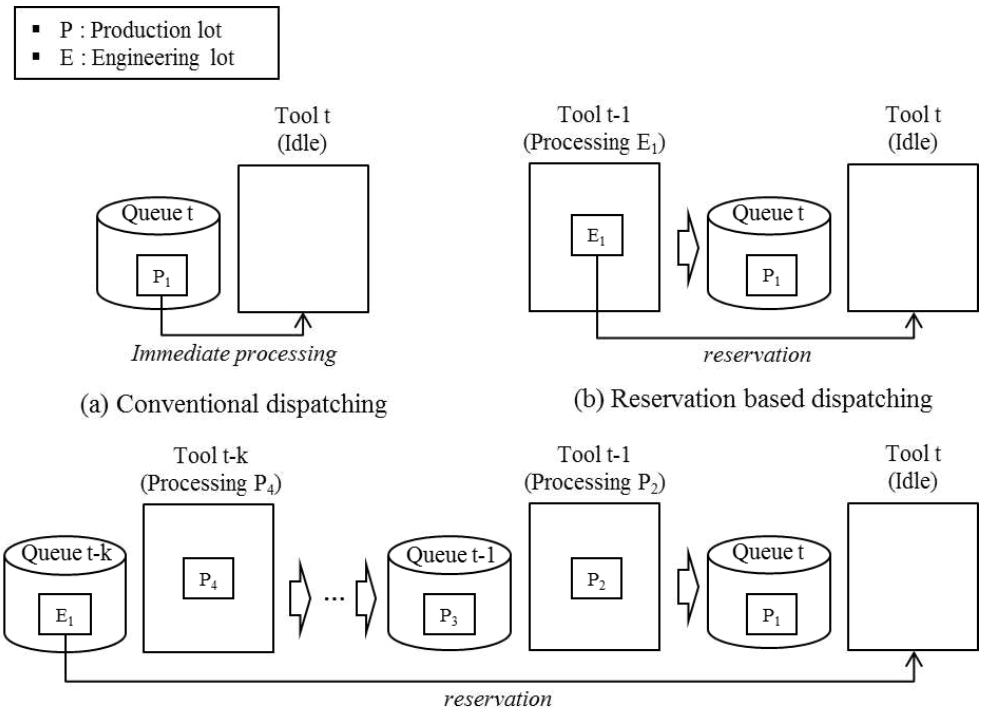


Figure 1: Conventional dispatching, reservation based dispatching and enhanced reservation based dispatching.

The objective of this paper is to minimize the cycle time for engineering lots with the on-time delivery of production lots. In this paper, the previous reservation based dispatching rule has been enhanced. Figure 1 shows the difference among the conventional dispatching rule, the previous reservation based dispatching rule, and the enhanced reservation based dispatching rule. We assumed that Tool t becomes available, and lot P_1 is waiting in Queue t (the queue of Tool t). As shown in Figure 1(a), a conventional dispatching rule prioritizes only those lots waiting in the queue of Tool t. In the situation shown in Figure 1(a), Tool t immediately processes lot P_1, because there is only one lot (P_1) in the queue. To guarantee the capacity required for engineering lots, the reservation based dispatching rule considers the engineering lots being processed in previous tools as well as the waiting lots in the queue of Tool t, as shown in Figure 1(b). If Tool t is arranged to a batch step, then it is necessary to consider wider reservation range. By considering the properties of the batch work, the enhanced reservation based dispatching rule can control

the range of reservation for the engineering lots. The longer a processing time is, the wider a range of reservation is. Therefore, for a batch step, the reservation range is wider than that for the other steps.

To perform the simulation, the commercial software MOZART® developed by VMS solutions was employed. The remainder of this paper is organized as follows. Section 2 describes the enhanced reservation based dispatching rule. Section 3 presents experimental results for the MIMAC6 dataset from Measurement and Improvement of Manufacturing Capacities (MIMAC). Finally, concluding remarks are presented in Section 4.

2 ENHANCED RESERVATION BASED DISPATCHING RULE

This section explains the enhanced reservation based dispatching rule that is designed to minimize the cycle time for engineering lots. To make a reservation, it is necessary to estimate the arrival time of those engineering lots that have not yet arrived, whenever a tool becomes free. The important issue here is how to calculate the arrival time of the engineering lots. In an environment in which the lots arrive dynamically over time, it is very difficult to calculate future lot arrival time. To overcome this problem, it is necessary to carefully observe the attributes of a problem involving an engineering lot. There is an important attribute which distinguish this problem from a general future lot arrival information problem: Engineering lots are given priority over production lots at all steps. This attribute gives rise to the possibility of estimating the accuracy of the arrival time of an engineering lot. As mentioned above, this paper employed the concept of reservation to guarantee the capacity for engineering lots. To minimize the waiting time for engineering lots, it is necessary to reserve engineering lots at all steps. Considering the property, we compute the earliest arrival time of the engineering lots, Figure 2 shows an example of the computation of the earliest arrival time. Engineering lot E_1 is being processed in step k-2, and then moves to the queue of TG_2 at time t_3. TG_2 consists of T_2 and T_3. When E_1 enters the queue of TG_2, T_3 is still processing production lot P_2. To minimize the waiting time of E_1, T_2 preferentially processes E_1 over T_3. As mentioned above, the proposed dispatching rule uses the concept of reservation for the engineering lots. Although P_4 is waiting for processing in the queue of TG_2, it can be more desirable for E_1 to have higher priority than P_4. If E_1 is reserved to T_2 for the period of t_3-t_2 at time t_1, then T_2 has to wait until the arrival of E_1. Then, E_1 can arrive in the queue of TG_1 at time t_3, and the earliest arrival time of E_1 is time t_3. We calculated the earliest time that E_1 is able to move into the queue of TG_1. In other words, the computed arrival time means the earliest arrival time.

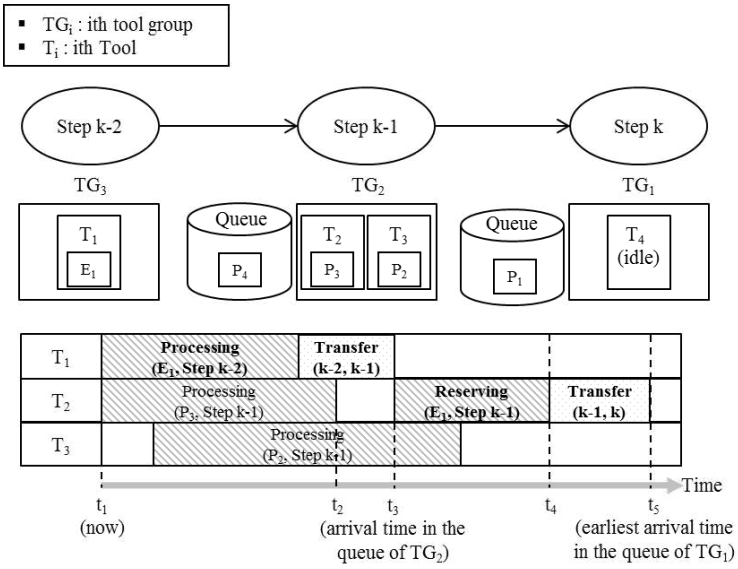


Figure 2: Computation of earliest arrival time for engineering lots.

The overall scheme of the proposed dispatching rule consists of two stages: 1) The computation of the reservation range; and 2) The reservation of the engineering lot that can arrive the earliest. Before explaining the proposed dispatching rule, it is necessary to define several terms as follows.

- T_D : the tool that performs dispatching at current time.
- R-lot : the lot for reservation.
- P-lot : the lot for immediate processing

For the first step, it is necessary to apply the operation due date (ODD) rule for production lots waiting in the queue. The ODD rule can be used to achieve the on-time delivery, and it is calculated in the following way : ODD = Due date – remaining RPT*flow factor where RPT is the raw processing time of the lot, and flow factor (FF) is defined as the target cycle times divided by the raw processing time (RPT). We compute the expected finish time for the lot that has the highest priority by applying the ODD rule. To determine the engineering lot for reservation, the proposed dispatching rule uses the time as a reservation range. The engineering lot to be reserved has to arrive in the queue of a tool, before the expected finish time of the lot selected by the ODD rule.

As mentioned earlier, the proposed rule also considers the on-time delivery for production lots. So, we have to consider an influence from the special handling for engineering lots. The presence of engineering lots negatively affects the flow of the production lots, since engineering lots always have a higher priority than production lots at all steps. To achieve the on-time delivery of production lots, it is necessary to limit the range of reservation. Thus, we have to consider the maximum reservation range as well as the processing time of a production lot for immediate processing. Figure 3 shows the method of determining candidates for reservation. It is assumed that P_1 has the highest priority by ODD rule. As shown in Figure 3, E_1, E_2, and E_3 can arrive in the queue of T_1 before the expected finish time of P_1. However, we have to consider the limitations of the reservation range. If the reservation range is limited to -1 step, E_3 is excepted from candidates for reservation, since it falls outside the reservation range.

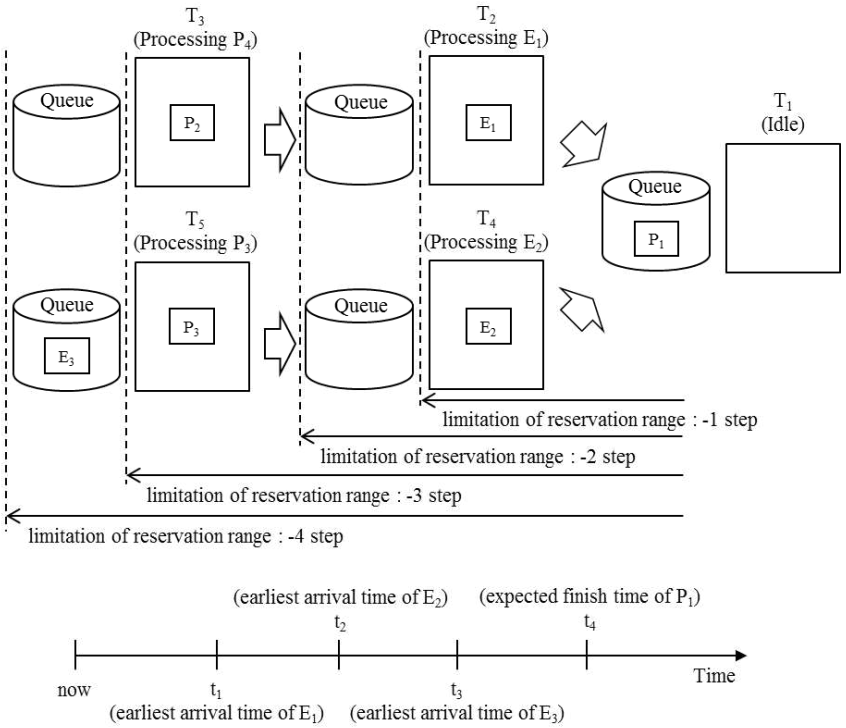


Figure 3: Determination of reservation range and candidate lots for reservation.

In the second step, the proposed dispatching rule reserves the lot which can arrive most quickly in the queue of TD. The method for computing the arrival time of the engineering lots is already mentioned in the earlier. Considering the on-time delivery of production lots, we will say that the best tool for reservation is the one which satisfies the following two conditions;

- Minimizing the waiting time of the R-lot, and
- Minimizing the idle time caused by reservation.

Based on the two stages, the overall dispatching rule can be described as follows.

- Reservation based Dispatching Rule

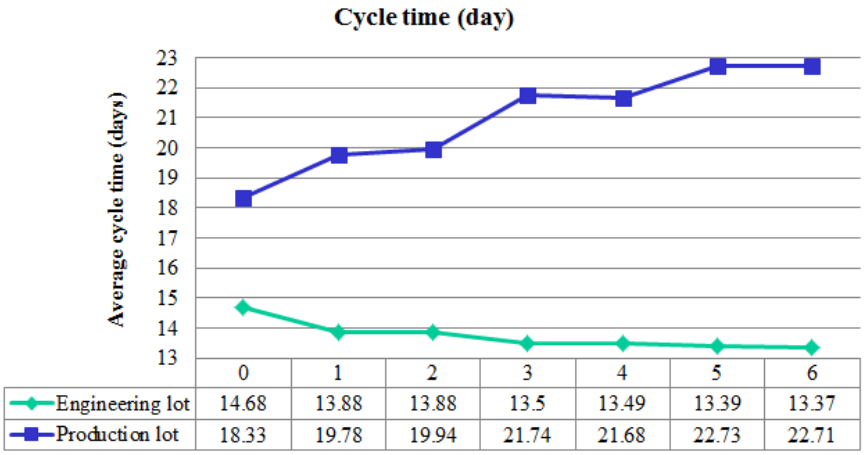
  Step 1) P-lot = Find the engineering lot reserved in the queue of T_D ;
  Step 2) If (P-lot is not null) {
  T_D starts processing of P-lot at current time ;
  Return ;
  }
  Step 3) Compute the reservation range considering the maximum reservation range;
  Step 4) R-lot = Find the engineering lot that can arrive the most quickly among engineering lots within the specified reservation range ;
  Step 5) C-step = current step of R-lot; T-step = target step of R-lot;
  Step 6) Repeat {
  C-step = a next step of C-step;
  Reserve R-lot for the best tool among tools deployed for C-step;
  If(C-step is equal to T-step) Break;
  }

3 SIMULATION RESULT AND PERFORMANCE ANALYSIS

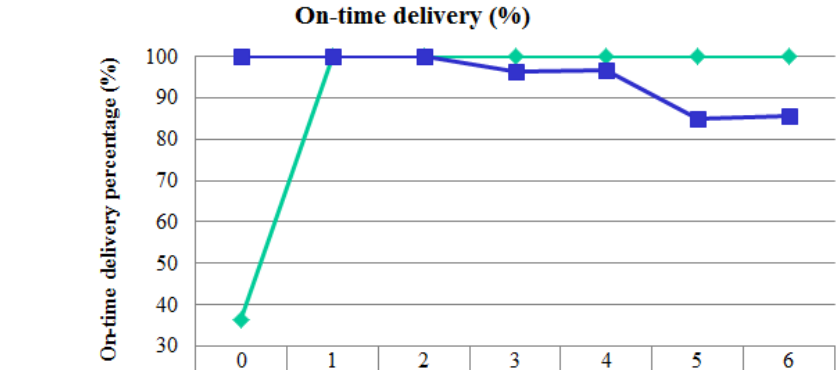
To construct a wafer FAB model, we employed dataset from Measurements and Improvement of Manufacturing Capacities (MIMAC) that consists of seven datasets. Each dataset contains the following information: product, process (flow), operator, and tool. In this paper, we used sixth dataset that presents the small wafer FAB. The small FAB produces nine products having different process, and consists of 93 tool groups. Each tool group may have multiple tools, from 1 to 10. It is necessary to refer to the MIMAC Final Report for the explanation details (Fowler and Robinson 1995). Since the MIMAC dataset does not reflect the nature of a modern FAB, we constructed a FAB model by modifying MIMAC dataset 6. The constructed FAB model does not incorporate any operator information, since modern FABs are automated and controlled by computers. Since a tool group must contains at least two homogeneous tools, it is necessary to increase the number of tools in any tool group with only one tool. Finally, we reduced the excessively large batch size. These changes were performed to enable the application of the nature of modern FAB. However, this results in a decrease in the photolithography tool utilization. To solve this problem, the setup time and processing times of some batch steps were properly changed. We modified the FAB model with the help of related engineers.

For the experimentation of the proposed dispatching rule, we employ the MozArt® engine (Ko et al 2013). MozArt® forward simulation engine based on discrete event simulation (DEVS) mimics the operation of an actual factory. It especially focuses on loading and unloading events, and for each tool, it generates a loading history, step movements and WIP trends. The results can be interfaced with a legacy line or an area scheduler, PM scheduler, or dispatcher. MozArt LSE® enables the planner to test various scenarios, for instance, product mix, PM schedule change and tool dedication modification. It also enables the detection of bottleneck steps or machine groups by analyzing the WIP and movement trends.

The flow factor for production lots is determined as 2.0. Engineering lots are released regularly, with the number of engineering lots set at 10% of the number of production lots. The FAB model was simulated for nine months. Because the first six months constitute a warm-up period, this period is excluded from the results. To show the effectiveness of the level of reservation range, experiments for comparing four different reservation ranges were conducted. Figure 4 shows the simulation results.



	0	1	2	3	4	5	6
Engineering lot	14.68	13.88	13.88	13.5	13.49	13.39	13.37
Production lot	18.33	19.78	19.94	21.74	21.68	22.73	22.71



	0	1	2	3	4	5	6
Engineering lot	36.51	100	100	100	100	100	100
Production lot	100	100	100	96.24	96.6	85.09	85.61

Figure 4: Average cycle time and on-time delivery percentage for different reservation ranges.

We take into account percentage of tardy and average cycle time as major performance measures. To compare the influence by the reservation depth, simulation experiments were performed for seven maximum values of reservation range, which range from 0 to 6 in steps of 1. Figure 4 shows the performance measures for the maximum reservation ranges. The maximum range of reservation is zero means that the rule forces engineering lots to be processed first without any reservations. If there are multiple engineering lots, the FIFO (First-In First-Out) rule may be applied to select the one to be processed. When the maximum reservation range is less than 3, the proposed dispatching rule achieves the on-time delivery of the production lots. However, there is a difference in the on-time delivery percentage of the production lots between the three cases. When the maximum reservation range is 1 or 2, the cycle time for the engineering lots is reduced from 14.7 days to 13.9 days. When the maximum reservation range is 3 or 4, the proposed rule cannot attain the due date for all the production lots. However, it still acceptable, since a 5% tardiness level (the rate of tardy lots) for a production lot is acceptable in practice. According to the results of our experiments, the wider the maximum reservation range is, the shorter cycle time of engineering lots is. Although we can perform dispatching by applying a wider reservation range to minimize the cycle time of an engineering lot, it is not desirable to expand the reservation range, because doing so causes the on-time delivery percentage of the production lots to decrease from 96% to 85%. Therefore, this is not acceptable to apply for real FAB. The simulation experiments show that an optical reservation range can be found. Moreover, the proposed dispatching rule guarantees the minimum cycle time of engineering lots with the on-time delivery of production lots.

4 SUMMARY

To survive in the modern manufacturing industry, it is necessary to perform consistent engineering research as well as production. Engineering lots are required to minimize cycle time for faster implementation of product improvements and more rapid deployment of new products to market. The objective of this paper is to achieve the minimum cycle time of engineering lots assuring the on-time delivery of production lots. The previous reservation based dispatching rule supports minimizing cycle time of engineering lots by considering the engineering lots being processed in previous tools as well as the waiting lots in the queue. However, it is still very difficult to achieve the minimum cycle time for engineering lots, since the rule does not consider the nature of the FAB. In wafer FAB, the processing times can vary depending on the step being performed. In particular, batch processing steps can take as long as 10h to complete as compared with 1 to 2 hours for the other processing steps. Therefore, if the engineering lots are not properly managed, then it might result in long cycle times of engineering lots. In this study, in order to overcome the limitations of the previous rule, it was enhanced. The enhanced reservation based dispatching rule determines the range of reservation for the engineering lots, whenever a tool becomes available. We determined the reservation range by calculating the processing time for the production lot having the highest priority by ODD rule. As the processing time of a lot becomes longer, the reservation range becomes wider. The rule is capable of reserving one of the engineering lots within the reservation range, and can guarantee the required tool capacity for the engineering lots. However, it may lead to significant tardiness of the production lots. To overcome this problem, we defined the maximum value of the reservation range, and it is possible to control the level of the reservations.

To simulate the proposed dispatching rule, we used the commercial software MOZART® developed by VMS solutions and the FAB model constructed by using MIMAC dataset 6. The FAB model is modified to reflect the natures of the real modern FAB. The simulation results show clear advantages of the enhanced reservation based dispatching rule over previous reservation based dispatching rule, and that it is important to identify an appropriate reservation range.

ACKNOWLEDGMENTS

This work was supported by the Defense Acquisition Program Administration (UD110006MD) and Korea Research Foundation (2010-0021040). The authors wish to express sincere gratitude for the financial support.

REFERENCES

Bixby, R., and R. Burda. 2008. “Technology that Upsets The Social Order—A Paradigm Shift in Assigning Lots to Tools in a Wafer Fabricator—The Transition From Rules to Optimization.” *In Proceedings of the Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 2277-2285. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc

Crist, K., and R. Uzsoy. 2011. “Prioritising Production and Engineering Lots in Wafer Fabrication Facilities: A Simulation Study.” *International Journal of Production Research* 49:3105-3125.

Chung, S. H., and H. W. Huang. 2002. “Cycle Time Estimation for Wafer Fab with Engineering Lots.” *IIE Transactions* 34(2): 105-118.

Chung, Y. H., S. C. Park, B. H. Kim, and J. C. Seo. 2014. “Due Date Control in Order-Driven FAB with High Priority Orders.” *In Proceedings of the Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 2544-2551. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc

Fowler, J., and J. Robinson, 1995. “Measurement and Improvement of Manufacturing Capacities (MIMAC): Final report.” Technical Report 95062861A-TR, SEMATECH, Austin, TX.

Jia, J., and S. J. Mason. 2009. "Semiconductor Manufacturing Scheduling of Jobs Containing Multiple Orders on Identical Parallel Machines." *International Journal of Production Research* 47(10): 2565-2585.

Ko K., B. H. Kim, and S. K. Yoo. 2013. "Simulation Based Planning & Scheduling System: MOZART®." *In Proceedings of the Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 4103-4104. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Malapert A., C. Gueret, and L. M. Rousseau. 2012. "A Constraint Programming Approach for A Batch Processing Problem with Non-Identical Job Sizes." *European Journal of Operational Research* 221(3): 533-545.

Park S. C., E. Ahn, Y. Chung, K. Yang, B. H. Kim, and J. C. Seo. 2013. "Fab Simulation with Recipe Arrangement of Tools." *In Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S. –H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 3840-3849. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc

Sarin S. C., A. Varadarajan, and L. Wang. 2011. "A Survey of Dispatching Rules for Operational Control in Wafer Fabrication." *Production Planning and Control* 22: 4-24.

Rezaei, K., H. Eivazy, A. Rezazadeh, and S. Nazari-Shirkouhi. 2011. A Production Planning and Scheduling Model for Semiconductor Wafer Manufacturing Plants, *International Conference on Computers & Industrial Engineering*, 751-756.

Zhou, Z., and O. Rose. 2012. "WIP Balance and Due Date Control in A Wafer Fab with Low and High Volume Products." *In Proceedings of the Winter Simulation Conference*, Edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 2019-2026. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc

AUTHOR BIOGRAPHIES

YONG H. CHUNG received a bachelor degree (2011) in industrial and information system engineering and a master degree (2013) in industrial engineering, Ajou University, Korea. He is now a Ph. D candidate in industrial engineering, Ajou University, Korea. He is interested in simulation-based scheduling and planning, digital manufacturing, and mesh simplification. His email address is yongho1230@gmail.com.

BYUNG H. KIM is the president of VMS Solutions Co., Ltd. since 2000. He received a BS from Sungkyunkwan University in 1993, a MS from KAIST in 1995, and a Ph.D. from KAIST in 2001, all in Industrial Engineering. His main interests are simulation-based scheduling and planning, manufacturing information systems, BPMS, and virtual manufacturing. His email address is kbhee@vms-solutions.com.

JEONG C. SEO is a principal engineer of Samsung Electronics Co., Ltd. since 2008. He received a BS from Chonnam National University in 1997, MS from KAIST in 1999, and a Ph.D. from KAIST in 2008, all in Industrial Engineering. His main interests are simulation-based scheduling and planning, finite capacity planning, lot pegging system, and CPFR(collaborative planning forecasting and replenishment) system. His email address is jungchul.suh@samsung.com.

SANG C. PARK was granted his bachelor (1994), master (1996) and Ph.D. (2000) degrees in industrial engineering, Korea Advanced Institute of Science and Technology (KAIST). He is a professor in Dept. of IE, Ajou University, Korea, since 2004. He is interested in modeling and simulation (M&S), combat simulation for defense, and digital manufacturing system. His email address is scpark@ajou.ac.kr.