PRODUCTION & MANUFACTURING | RESEARCH ARTICLE

Dispatching for order-driven FABs with backward pegging

Yong H. Chung¹, S. Lee¹ and Sang C. Park¹*

Abstract: In this paper, a two-stage dispatching rule for the on-time delivery of an order-driven FAB is presented. Dispatching rules have a significant impact on the performance of FABs. Consequently, many such rules have been created, such as first in first out, earliest due data, operation due date, and critical ratio. Because these conventional dispatching rules make decisions only by considering the due date of each lot, they might not be suitable for dynamic order-driven FABs, for which it is very essential to monitor not only the due date of each lot but also the production rate of each step. In the first stage of the proposed two-stage dispatching rule, the production rate of each step is taken into account, and the due date of each lot is considered in the second stage. In this paper, the production rates of steps are computed by combining forward simulation and backward pegging. A simulation model, based on a wafer FAB data-set, has been developed, and simulation is conducted using commercial software SEEPLAN®. The simulation results clearly show the advantages of the proposed dispatching rule over existing rules.

Subjects: Computer Integrated Manufacturing (CIM); Flexible Manufacturing Systems; Manufacturing & Processing

Keywords: FAB; two-stage dispatching rule; on-time delivery; simulation; pegging

1. Introduction

Nowadays, almost all electronic products require a number of semiconductor chips and these chips are manufactured in FAB facilities. Semiconductor chips are created via a multiple-step sequence of photolithographic and chemical processing steps, during which electrical circuits are gradually created on a wafer made of the pure semiconducting material, silicon. The entire manufacturing
process of chips, from the beginning to the packaging of chips for shipment, takes a couple of months, and is performed in highly capital-intensive production systems. These systems, referred as FABs, require investment of a few billion dollars. In a typical FAB, a large number of product types and variants are produced concurrently. Production in a FAB is considered as one of the most complicated manufacturing processes because of hundreds of steps for a product, re-entrant flows, batch processing, queue time limit, recipe arrangement, and sensitive yield rates of tools (Aaron, Krott, & Doxsey, 2008; Chen, Sarin, & Peake, 2010; Huang & Süer, 2015; Mokhtari & Dadgar, 2015; Tu, Chen, & Liu, 2010). To be globally competitive, full-capacity production for high utilization and just-in-time production for on-time delivery with minimum work-in-process (WIP) are essential in the FAB industry.

Figure 1 shows the relationships between tool groups and steps. To produce a product, a sequence of processing steps is required. Each step requires its own processing time and can be performed by a specific tool group. As shown in Figure 1, “Tool Group A” has three tools (Tool_A1, Tool_A2, and Tool_A3) and two capable steps (Step_I2 and Step_Jk).

In the past, much research has been conducted in the area of efficient operation policies for the re-entrant job shop or FAB. It is possible to classify previous research results according to objective functions, full utilization of critical tools, on-time delivery, WIP balance control, and bottleneck machine detection (Chung & Hsieh, 2008; Mönch & Zimmermann, 2007, 2011; Wang & Liu, 2015; Wu, Chang, & Chiou, 2006; Wu & Hung, 2008; Zhai, Sun, Wang, & Niu, 2011). Although, these key performance indicators (KPIs) focus on different aspects of a FAB, they are all closely inter-related. In Quek, Gan, Tan, and Peng (2007) proposed a dispatching rule to minimize efficiency loss. The rule tried to minimize the investment cost by optimizing the efficiency of critical tools. In Chung and Jang (2009) considered the FAB scheduling problem as a static combinatorial problem and developed linear programming methods for the maximization of throughput.

Among various KPIs, the due date control or on-time delivery is an important performance indicator, especially for an order-driven FAB. To achieve better due date control, it is very essential to solve the WIP imbalance problem. In Leachman, Kang, and Lin (2002) used Little’s formula to translate the target cycle time into target WIP levels. As per the formula, the target WIP level for a step is the product of the target cycle time for the step and the target FAB’s out rate for the device. In Lee and Lee (2003) suggested three different policies—Push, push-pull, and pull type—To control WIP levels. In Kuo, Liu, and Chi (2008) proposed a methodology to determine standard WIP levels by using back
propagation neural networks. In Zhou and Rose (2010) proposed a WIP control table to balance the workloads of tools in a wafer FAB. Although there has been extensive research on the target WIP determination, order-driven FABs have many limitations. Nowadays, FABs must handle changing products and release rates to respond to various customer orders. Under such dynamic situations, it is extremely difficult to determine an appropriate WIP level for a specific operation or machine.

Another approach for better on-time delivery is to develop proper dispatching rules such as first in first out (FIFO), earliest due date (EDD), operation due date (ODD), and critical ratio (CR) (Zhou & Rose, 2010). EDD chooses a lot with the closest global due date at a stage. ODD is a variant of EDD that defines local due date for each lot and stage. The widely used CR rule takes into account the due date of the lot as well as the remaining processing time of the current stage. A common characteristic of these dispatching rules is that they make decisions only by considering the due date of each lot. Conventional dispatching rules ignore the overall production rates of steps, which are very essential to achieve on-time delivery. Although, there are some tardy lots, the on-time delivery still can be met as long as the overall production rate of the step is faster than the schedule. Because of the reason, it is necessary to consider not only the due date of each lot but also the production rate of each step.

The objective of this paper is to develop a two-stage dispatching rule for achieving better on-time delivery for order-driven FABs. In the first stage of the proposed dispatching rule, the production rate of each step will be taken into account, and the due date of each lot will be considered in the second stage. In this paper, the production rates of steps are computed by combining forward simulation and backward pegging. Pegging is a computing process of labeling WIP lots for a target order, which is specified by the due date, quantity, and product specifications (Ko, Park, Yoo, Park, & Kim, 2010). As a result, it determines the out-target for each step and calculates the latest possible start time for each ordered lot. Usually, pegging is used to get the release plan to meet the out-target considering current WIP and machine status. Figure 2 shows the concept of forward simulation and backward pegging for a specific product going through \( i \) steps. The simulation starts at time \( st \), and ends at time \( et \). The simulation can help find the actual production (out-actual in Figure 2(a)) record and WIP for each step. While the time of simulation goes forward, the time of pegging goes backward. At time \( et \), pegging starts with the demand of the product and decides the out-target (out-target in Figure 2(b)) for all steps considering their WIPs.

For the execution of simulation and pegging, commercial software SEEPLAN® developed by VMS solutions was used (Huang & Süer, 2015). The overall structure of the paper is as follows. Section 2 describes the proposed methodology to achieve better on-time delivery. Section 3 presents experimental results for a wafer FAB data-set provided by Fowler and Robinson (1995). Finally, concluding remarks are presented in Section 4.

---

Figure 2. Forward simulation and backward pegging.
2. Approach to achieve on-time delivery

A FAB consists of multiple tool groups, and each tool group has multiple tools capable of assigned steps. As shown in Figure 3, Tool Group B has five WIP lots belonging to different steps. Lot_1, Lot_2, and Lot_3 belong to Step_1, and Lot_4 and Lot_5 belong to Step_2. Lots belonging to the same step may have different due dates. The role of dispatching rules is to determine a WIP lot to be processed next, whenever a tool becomes free.

Before outlaying the detailed explanation of the proposed two-stage dispatching rule, we want to address the overall structure of the rule. As shown in Figure 4, the rule consists of two stages. In the first stage, the production rate of each step is computed by using pegging. All steps are classified into three groups: PR1, PR2, and PR3. The second stage of the dispatching rule computes the tardiness level for each lot. Four tardiness levels are defined: TL1, TL2, TL3, and TL4. As shown in Figure 4, lots belonging to the “PR1-TL1” group have the highest priority and those belonging to the “PR3-TL4” group have the lowest priority. In case of ties, the FIFO rule is followed.

In the first stage of the dispatching rule, it is necessary to use the demand and supply pegging mechanism. For an order-driven FAB, pegging plays an essential role in achieving on-time delivery by labeling WIP lots for a target order, which is specified by the due date, quantity, and product specifications. Backward pegging starts with the amount of demand and the due date for a specific product, requiring multiple steps. For each step, the dispatching rule computes the target output by considering the WIP, raw processing time, and yield rate of the step. Figure 5 shows the pegging mechanism of a process consisting of n steps. To explain the details of our approach, several terms are defined and the pegging algorithm is explained as follows:

Definitions of terms

- WIP(i): Work-in-process of i-th step
- RUN-TAT(i): RUN turnaround time of i-th step (= raw processing time)
- WAIT-TAT(i): WAIT turnaround time of i-th step (= raw processing time × (flow factor – 1))
- Out-target(i): Amount of target production of i-th step by the out-time(i)
- In-target(i): Amount of target input of i-th step by the in-time(i)
- Yield-rate(i): Yield rate of i-th step
- T_{now}: Current time
Backward pegging algorithm

//– Input: Due date (Out-time(n)), Demand (Out-target(n)), WIP(i), RUN-TAT(i) & WAIT-TAT(i) where 0 < i < n

//– Output: Out-target(i) where 0 < i < n

For (i = n; i > 1; i–)
{
    In-time(i) = Out-time(i) – RUN-TAT(i);
    In-target(i) = Out-target(i)/Yield-rate(i);
    Out-time(i-1) = In-time(i) – WAIT-TAT(i);
    Out-target(i-1) = In-target(i) – WIP(i);
}

The pegging algorithm helps compute the out-target for every step. Usually, a FAB operates three shifts a day. For each shift, the out-target can be calculated for every step.
Now, we consider the forward simulation. As simulation proceeds, the out-actual can be found. If the out-actual of a specific step is larger than the out-target, the step produces faster than the schedule. As shown in Figure 4, all steps are classified into three groups based on production rates. The detailed algorithm is described as follows:

**First stage: Production rate of i-th step**

If (Out-target(i) != 0) {
   If (Out-actual(i)/Out-target(i) < 1.0)
      Return PR1; // production rate is slower than schedule
   Else
      Return PR2; // production rate is faster than schedule
} Else
   Return PR3; // undefined schedule

According to production rate definitions, PR1 means the step works slower than the schedule and PR-2 means the step works faster. If the out-target is equal to zero, the production rate becomes PR3, which means that the schedule is not defined yet. Consequently, PR1 is the desirable production rate. Although the production rate can be a good indicator of the priority of steps, it does not explicitly consider the priority of WIP lots. To relieve the problem, the second stage of the dispatching rule computes the tardiness level for each lot. The tardiness level of a lot is determined by considering the due date:

**Second stage: Tardiness level of a lot**

If (Tnow >= due date)
   Return TL1; // already tardy
If (Tnow + Remaining ‘raw’ processing time >= due date)
   Return TL2; // will be tardy
If (Tnow + Remaining processing time including waiting time >= due date)
   Return TL3; // negative on-time delivery
Return TL4; // positive on-time delivery

In the second stage, four tardiness levels are defined: TL1, TL2, TL3, and TL4. As a result, WIP lots are classified into 12 categories: three production rates × four tardiness levels. Figure 4 shows the priority sequence among these categories.

3. Experimental results

For the experimentation of the proposed approach, a simulation model is constructed based on a wafer FAB data-set provided by Fowler and Robinson (1995). The FAB consists of 104 tool groups, constituting 228 tools in all, and produces nine products with different process steps. The total number of steps is 2,241, with 25 steps per product on an average. The product having the longest steps has 212 steps, and the shortest one has only one step. The average processing time of the steps is 3,014 s. Each tool group may have multiple tools, from 1 to 10. A lot consists of 24 wafers, and 2,777 lots are released per year under 100% FAB loading.

To conduct the simulation, the SEEPLAN® engine developed by VMS solutions (Ko et al., 2010) was employed. Figure 6 shows how SEEPLAN® generates the loading schedule for each tool in the FAB. It requires three master data: bill of process (BOP) model, resource model, and dispatching rule.
The current WIP is initialized at the beginning of the simulation. Considering the current WIP, the release plan is used as an input. The simulation results can be analyzed to see KPIs including resource utilization, throughput, and WIP fluctuations.

The BOP model is a network model that combines the bill of material and process routing. It consists of parts, processes, and transitions. The BOP model contains a step sequence, a loadable resource list, tact/flow time for each step, and average transfer time. A resource is characterized by the handling unit, process type, and defect treatment policy. A resource group or tool group indicates the standard step to be processed, jig capacity, setup crew capacity, and a list of unit resources or tools. Each resource has the dispatching rule and tact/flow time. Dispatching rules are used to determine the priority for fulfilling orders.

Table 1 shows simulation results for five dispatching rules. In this paper, we do not consider unpredictable events occurred stochastically, such as rework and machine failure. For that reason, we used deterministic simulation to prove the performance of the proposed dispatching rule. The simulation was conducted for eight months. The first five months were considered as warm-up periods and not taken into account for statistics. The target due date flow factor ranges from 1.6 to 2.6 in steps of 0.2. The performance of the two-stage dispatching rule was compared with conventional dispatching rules. As major performance measures, percent tardy lots (Figure 7), average tardiness of tardy lots (Figure 8), and average cycle time (Figure 9) were chosen. In terms of on-time delivery, the proposed two-stage dispatching rule gives the best results. In addition, the average tardiness of the proposed rule is the smallest. The experimental results show the advantages of the proposed dispatching rule over the conventional ones. As shown in Figure 7, when a flow factor is set as two or more, the proposed dispatching rule is superior to conventional dispatching rules. It means that the dispatching rule is useful when the schedule of a FAB is not tight. Although CR also shows good performance for average tardiness, it causes many tardy lots and long average cycle time. While existing dispatching rules focus only on the due date of each lot, the proposed two-stage dispatching rule also takes into account the production rate of each step.
Table 1. Simulation results for different five dispatching rules

<table>
<thead>
<tr>
<th></th>
<th>1.6</th>
<th>1.8</th>
<th>2</th>
<th>2.2</th>
<th>2.4</th>
<th>2.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tardy lot (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIFO</td>
<td>100.0</td>
<td>97.4</td>
<td>79.0</td>
<td>60.9</td>
<td>12.5</td>
<td>3.0</td>
</tr>
<tr>
<td>ODD</td>
<td>100.0</td>
<td>100.0</td>
<td>65.8</td>
<td>24.3</td>
<td>8.4</td>
<td>0.7</td>
</tr>
<tr>
<td>EDD</td>
<td>100.0</td>
<td>95.0</td>
<td>67.8</td>
<td>30.2</td>
<td>6.5</td>
<td>0.0</td>
</tr>
<tr>
<td>CR</td>
<td>100.0</td>
<td>88.2</td>
<td>43.2</td>
<td>33.4</td>
<td>15.2</td>
<td>6.7</td>
</tr>
<tr>
<td>MI + CR</td>
<td>100.0</td>
<td>100.0</td>
<td>32.1</td>
<td>22.4</td>
<td>4.7</td>
<td>1.3</td>
</tr>
<tr>
<td>Tardiness (day)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIFO</td>
<td>7.8</td>
<td>5.1</td>
<td>3.0</td>
<td>1.8</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>ODD</td>
<td>5.2</td>
<td>3.5</td>
<td>1.6</td>
<td>0.6</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>EDD</td>
<td>6.2</td>
<td>4.2</td>
<td>2.4</td>
<td>1.0</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>CR</td>
<td>4.82</td>
<td>2.05</td>
<td>0.99</td>
<td>0.64</td>
<td>0.39</td>
<td>0.22</td>
</tr>
<tr>
<td>MI + CR</td>
<td>4.5</td>
<td>2.4</td>
<td>0.7</td>
<td>0.6</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Cycle time (day)</td>
<td>30.2</td>
<td>30.3</td>
<td>30.6</td>
<td>31.6</td>
<td>31.0</td>
<td>31.3</td>
</tr>
<tr>
<td>FIFO</td>
<td>27.8</td>
<td>28.7</td>
<td>29.3</td>
<td>29.9</td>
<td>30.9</td>
<td>31.4</td>
</tr>
<tr>
<td>ODD</td>
<td>28.6</td>
<td>29.3</td>
<td>29.9</td>
<td>29.1</td>
<td>29.7</td>
<td>28.8</td>
</tr>
<tr>
<td>EDD</td>
<td>27.22</td>
<td>27.33</td>
<td>28.44</td>
<td>30.11</td>
<td>31.78</td>
<td>33.22</td>
</tr>
<tr>
<td>CR</td>
<td>27.0</td>
<td>27.7</td>
<td>27.6</td>
<td>29.1</td>
<td>30.1</td>
<td>31.0</td>
</tr>
</tbody>
</table>

Figure 7. Comparison of percent tardy lots among dispatching rules.
4. Summary
A wafer FAB is a highly automated manufacturing system. Much research on dispatching rules and scheduling optimization for FABs has been conducted. Among many KPIs of FABs, on-time delivery is becoming more important, especially for order-driven FABs. To achieve better on-time delivery, there are various dispatching rules, like FIFO, EDD, ODD, and CR. Although these rules use different criteria for choosing a lot to be processed next, they share a common attribute—they all make decisions only by considering the due date of each lot. Although, some lots are behind of schedule, the on-time delivery still can be met as long as the overall production rate of the step is faster than the schedule. Because of the reason, it is necessary to consider not only the due date of each lot but also the production rate of each step. This paper proposes a two-stage dispatching rule to achieve better on-time delivery for order-driven FABs. The first stage of the proposed rule takes into account the production rate of each step, and the due date of each lot is considered in the second stage. In this paper, the production rates of steps are computed by combining forward simulation and backward pegging.

For the experimentation of the proposed dispatching rule, a simulation model is constructed by using a wafer FAB data-set provided by Fowler and Robinson (1995). For the execution of the simulation model, commercial software SEEPLAN® developed by VMS solutions is employed. The simulation was conducted for eight months, and percent tardy lots, average tardiness of tardy lots, and average cycle time were chosen as major performance measures. The experimental results show clear advantages of the proposed dispatching rule over existing ones.
References


