A step toward capacity planning at finite capacity in semiconductor manufacturing

Emna Mhiri, Mireille Jacomino, Philippe Vialletelle, Fabien Mangione, Guillaume Lepelletier

Proceedings of the 2014 Winter Simulation Conference

A STEP TOWARD CAPACITY PLANNING AT FINITE CAPACITY IN SEMICONDUCTOR MANUFACTURING

Emna Mhiri
Mireille Jacomino
Fabien Mangione

University of Grenoble Alpes,
G-SCOP, F-38000
Grenoble, FRANCE

Philippe Vialletelle
Guillaume Lepelletier

STMicroelectronics
850, Rue Jean Monnet
F-38926 Crolles Cedex, FRANCE

ABSTRACT

Due to process complexity, technological constraints and constant evolution of the economical context, the management of semiconductor wafer fabs is very complex, especially when operating in the high-mix low-volume model. In this paper, we suggest two different production planning approaches developed by STMicroelectronics and G-SCOP to better control the production of Crolles 300mm production line. At first, a mixed integer programming (MIP) is proposed to determine the start and completion dates for each processing step of the production lots, taking into account the finite capacity of the production line. Moreover, an approach, based on heuristic, is presented to simplify the problem and develop an infinite capacity projection engine that respects lots due date and cycle times' variability.

1 INTRODUCTION

Semiconductor Manufacturing is one of the most complex industrial fabrication processes. The constant evolution in production technologies and products often result in decreasing prices; therefore, it is critical to rapidly install capacity to deliver cutting edge products with short cycle time (Leachman 2012).
To accomplish this, significant attention has been brought to scheduling and dispatching tools. However, the management of capacity and production planning mainly relies on simplistic tools such as spreadsheets as well as heavy methods such as discrete event simulation. Nevertheless, both of these approaches lack precision as they always overestimate capacity and underestimate cycle time (Ignizio 2012). New methods and tools, leveraging on Operational Research techniques and modern computation power, could lead to significant benefits.

The main limitation of current capacity and production planning tools is the evaluation of variability. In "Front-End" production lines, also called wafer fabs, several hundreds of individual process steps are required for the manufacturing of electronic components on silicon wafers. These steps are performed on machines that have various design and operating modes e.g. single wafer, lot or batch, serial or parallel, and sequence dependent set-up or not (Mönch et al. 2011). The number of equipment a.k.a. tools for each type will impact the capacity and downstream inter-arrival times, a first dimension of variability. Production equipment breakdowns or failures are the second dimension of variability. Reentrant flows, a characteristic of the semiconductor manufacturing process, mean that the same tool may be used several times during the manufacturing of a product, or that several products at different stages may be waiting in front of the same tool. A third dimension of variability is hence introduced by dispatching policies.

Semiconductor manufacturers are operating various markets, using various business models. The traditional approach is to consider product mix (High or Low) and the production volume (High or Low).

A production line manufacturing memory components is typically a low-mix high-volume unit; so, equipment redundancy enables significant reduction of the variability. In high-mix manufacturing, the product variety involves options (number of lithography steps, ion implantation and metallization layers) which means high level of variability induced by dispatching policies and tools that may not be qualified for the desired recipe. For fabs operating in the low-volume mode, production tools with complex configurations (heterogeneous clusters) are often limited to minimum and are used at their limits. Low redundancy and processes when pushed to their limits result in high variability.

Most of European wafer fabs operate in the high-mix low-volume mode. Moreover, their capacity has generally been built with incremental investment. Several technological generations are developed in the same line using several types of machines. This further complicates both execution (problem of mix and match for Process Control, heterogeneity of throughputs and qualifications for dispatching) and capacity planning (heavy models, unique machines, important impact of variability). Recently, new challenges linked to the generalization of Time Constraints Tunnels, and the management of small to medium volume products with different customer priorities / expected cycle time, highlights the need for a breakthrough in terms of production and capacity planning.

The ENIAC European Project INTEGRATE aims at the development of new tools and techniques to enable wafer fabs to efficiently manage a high product and technology mix. For capacity and production planning activities, the main challenges to be faced are:

- Choice of technical options when developing / industrializing new technology
- Early identification of blocking points and definition of adjustment / investment strategy
- Validation of start plan / commitment to due date for customer deliveries
- Cycle time challenge with lots / products flowing at different speeds
- Execution / scheduling & dispatching policy for line balancing.

To answer these questions, three different notions and tools are generally involved:

1. **Capacity Planning**: Which capacity do I need to answer the demand?
2. **Production Planning**: What is the best start plan knowing demand and installed capacity?
3. **WIP projection**: What do I need to do to ensure the delivery of the Work In Progress?

While (1) is generally answered using infinite capacity computation, (2) and (3) need to consider the relationship between capacity limitations and cycle time (bottleneck stages increase product cycle time, and delayed products may create 'WIP bubble' conflicts). In this paper, we shall mainly focus on WIP projection where traditional approaches involve two main techniques: projection at infinite capacity and discrete event simulation.

The paper is organized as follows. Standard tools and techniques used today by semiconductor manufacturers are presented in section 2. Then section 3 introduces a brief review of the various approaches proposed in the literature. Section 4 and 5 develop two different approaches (i) MILP and (ii) heuristic. Finally, in section 6, we summarize the conclusions and future perspectives.

2 CURRENT TOOLS AND METHODS USED IN SEMICONDUCTOR FACTORIES

The need for capacity studies has increased with the entrance of semiconductor companies in the industrial era 30 years ago. Capacity studies, also called capacity analyses, gather all the activities aiming at evaluating the ability of a production line to serve its client needs with respect to products quantity, mix and volume. Industrial Engineering is typically in charge of this process of capacity computations.

This process is based on very detailed computer models that take different shapes but always have standardized contents, where each process tool and process step are described with different characteristics including typical tools availability, tools qualification matrix and process steps throughput.

First set of tools that appear to be used for capacity computations are capacity spreadsheets. These are able to evaluate an average loading of tools over time by combining the computer model with a

production plan, expressed as quantities to deliver by period and by product. Main advantages of the spreadsheets are: (1) calculation speed, 'what-if' analysis may be done with instant answer, (2) easiness of bug-identification (3) possibility to compute optimal production plan by adding a solver to the spreadsheet (4) simplicity, which provide answers that are easy to explain to decision-makers. However they do not consider cycle-time aspect of semiconductor manufacturing. It is paramount to consider cycle time figures (up to three months) in production planning to avoid a conflict on bottlenecks. Capacity spreadsheets are of interest for Fab having few different products and stable product mix (Ignizio and Garrido 2012).

An evolution of capacity spreadsheets has given the second set of capacity computation tools. These tools, like CAPACE (acronym for Capacity Explorer) in STMicroelectronics, add cycle-time dimension to capacity spreadsheets. This enables to analyze the production plans made of different products, with significantly different cycle-times and different mix per time bucket. These tools close the gap with spreadsheets considering of cycle-time; however, they work well only if a correct cycle-time is included in the computer model. All the advantages of spreadsheets are affected: (1) calculation speed is reduced, (2) bug are more difficult to analyze due to cycle-times variety which blend data between products, (3) computation of optimal production plan is almost impossible due to the number of "variables", (4) for the same reason, results are more difficult to explain. Furthermore, cycle times must be stable over year-length periods otherwise this second set of tools may lead to improper capacity saturation estimation. Despite all these drawbacks, these tools are mostly used and have overpassed capacity spreadsheets.

These two sets of tools belong to the "infinite capacity projection" category, as they proceed in two steps: (1) projection of products considering or not cycle-time, (2) estimation of toolsets saturation based on the projection results. Other commonly used computation tools belong to the "finite capacity projection" category, as they compute simultaneously cycle-time and toolset saturation, letting the latest impacting the first one. In this category, we find discrete-event simulation tools used for capacity analysis. These tools are used with a similar computer model as capacity spreadsheets. Production plan is input as the 'start plan'. On one hand, the main advantage of these tools is that the modeler doesn't have to model cycle time in detail, as it is an output of the computation and cycle-time will also vary over simulation horizon. On the other hand, their main drawback is intimately linked with their computation principle: discrete event simulation tools just push the Work In Progress and the starts without "goal function". This has important consequences as these tools give feasible projection plans but they cannot answer to the most important question: "Is my production plan feasible?"

3 LITERATURE REVIEW

In addition to tools used in the semiconductor industry, we may find, in the literature, numerous methods and techniques used for capacity planning in the semiconductor manufacturing. Uzsoy et al. (1992, 1994), Gupta et al. (2006) and Mönch et al. (2013) have mentioned in their reviews different capacity planning techniques used in the semiconductor environment which can be divided into infinite capacity planning models and finite capacity planning models.

3.1 Infinite Capacity Planning Tools

Among the methods used for infinite capacity planning, we find classical techniques, considered as the most famous techniques successfully used in many industries, such as Material Requirement Planning (MRP) (Vollmann 2005); Manufacturing Resource Planning (MRPII), a push system (Rondeau and Litteral 2001); the ''pull'' technique Just In Time (JIT) (Golhar and Stamm 1991) and Theory Of Constraints (TOC) (Goldratt 1990).

The application of these traditional tools for capacity planning in semiconductor industry presents some shortcomings. Indeed, it is proven that MRPII method can be inefficient and may produce unrealistic Master Production Schedules (MPS) when applied in a semiconductor manufacturing. It

doesn't take into account the capacity constraints in the system and cannot handle uncertainty caused by machine breakdowns or other unexpected events (Rupp and Ristic 2000).

The JIT technique proves it strengths (Levitt and Abraham 1990); however, it presents some limitations in high-mix low-volume production systems. It seems to be more suitable for a repetitive production environment with stable demand and low mix product (Carlson and Yao 1992).

The TOC seems efficient as a technique of capacity planning in semiconductor industry (Rippenhagen and Krishnaswamy 1998) but it concerns only bottleneck resources.

In addition to these classical methods, new techniques are proposed such as the method of Guan et al. (2008) that is based on a mathematical programming for a dynamic formation of virtual flow paths and a TOC/DBR (Theory Of Constraints / Drum-Buffer-Rope) 'path-specific' mechanism with group scheduling used for the control over each flow path.

There are also capacity planning systems developed to determine lot release time, start fab, and capability of the equipment for multiple semiconductor fabrication lines on the basis of pull philosophy and the assumption of infinite equipment capacity (Chen et al. 2005, Chen et al. 2009). Recently, authors have resorted to linear programming for capacity planning problems. The goal is to find production quantities taking into account demand, capacity restrictions, and current WIP status whereas the objective is to maximize revenue for forecasted orders by minimizing production costs, inventory holding costs, and costs for unfulfilled orders (Habla and Mönch 2008).

Although these techniques present several advantages such as the determination of bottleneck resources or the respect of delivery due dates, they consider a constant average cycle time for processing steps, independent of the workstation utilization.

3.2 Finite Capacity Planning Tools

Under finite capacity planning, the utilization of machine must be lower than its capacity but orders defined in the initial plan may not respect predefined sequences or even meet due dates. The most relevant techniques to establish finite capacity schedules for semiconductor fabs are linear programming models and heuristics.

3.2.1 Linear Programming Models

The linear programming approach is widely applied to the specific issues encountered in capacity planning for the semiconductor industry (Leachman and Carmon 1992, Hung and Leachman 1996).

The objective functions corresponding to the linear programs consist of optimizing strategic resource allocation to maximize long-term profit (Bermon and Hood 1999), minimizing production costs (Catay et al. 2003), minimizing the total weighted lots tardiness (Habla et al. 2007) or maximizing the profit (Ponsignon and Mönch 2012), taking into account capacity constraints.

LPs may require a very long time to generate the input data and huge amounts of memory and disk space to store the data (Sullivan and Fordyce, 1990). Thus, they are generally combined with heuristics (Leachman et al. 1996, Ponsignon and Mönch 2012) or Lagrangian relaxation (Catay et al. 2003, Habla et al. 2007) to reduce the execution time.

3.2.2 Heuristics

Approximate methods have been generally used to develop finite capacity planning systems for semiconductor industry. These systems have considered different inputs and outputs. Some studies have considered as inputs delivery due dates and resource capacity (Rupp and Ristic 2000, Chen et al. 2008) or the WIP status (Liu et al. 2011). As outputs, we find orders release date at bottleneck positions (Horiguchi et al. 2001), the order release time, start date, and equipment capability for each order (Chen et al. 2008) or the mean and the variance of the total cost (excess inventory, unused capacity, penalties…) associated

with a release plan (Liu et al. 2011). Moreover, we find more and more recourse to meta-heuristics especially genetic algorithms (Ponsignon et al. 2008, Ponsignon and Mönch 2009) to solve production planning problems in semiconductor manufacturing.

In the literature review, finite capacity planning techniques found can be divided in long-term strategic capacity planning methods (Bermon and Hood 1999, Catay et al. 2003) and mid-term tactical master planning tools (Ponsignon and Mönch 2012). These different techniques are developed for specific purposes such as minimizing delays, reducing cycle time, defining the completion date of customer orders, etc.

In this study, the mid-term tactical planning at finite capacity is considered. We will base on current techniques mentioned in the literature review to resolve the considered problematic.

4 A FIRST APPROACH

Although the limits of Mixed Integer Programming (MIP) models cited in the literature, we consider this method as the first approach to propose a finite capacity production planning. This technique is used because it is an exact method and it allows us to reformulate the problem mathematically.

The basic approach is quite simple and has been implemented, with some variations and extensions, in several commercial production planning systems. It is also inspired from the work of Habla et al. (2007) but in our case we are not limited on bottlenecks positions. The formulation of the MIP model is as follows:

Indices

l	$\in$	$\{1..L\}$	Lot
s_l	$\in$	$\{1..S_l\}$	Step of lot l
i	$\in$	$\{1..I\}$	Station family
t	$\in$	$\{1..T\}$	Period

Parameters

$p_{sl,i}$	$\forall$	$s_l \in \{1..S_l\}, i \in \{1..I\}$	Process time of step s_l on station family i
q_{sl}	$\forall$	$s_l \in \{1..S_l\}$	Waiting time on the equipment before the start of the process step s_l
o_{sl}	$\forall$	$s_l \in \{1..S_l\}$	Time to process a step s_l to add to the first wafer
Tr	$\forall$	$t \in \{1..T\}$	Transfer time between two successive steps of the same lot
do_t	$\forall$	$t \in \{1..T\}$	Opening of the workshop time during the period t
d_l	$\forall$	$l \in \{1..L\}$	Due date of the lot l
W_l	$\forall$	$l \in \{1..L\}$	Weight of the lot l
M_t	$\forall$	$t \in \{1..T\}$	Number of moves in period t
$K_{i,t}$	$\forall$	$i \in \{1..I\}, t \in \{1..T\}$	Capacity of a station family i in period t

Decision variables

$x_{l,sl,t}$	$\forall$	$l \in \{1..L\}, s_l \in \{1..S_l\}, t \in \{1..T\}$	=1 if lot step s_l is executed in period t
t_{sl}	$\forall$	$s_l \in \{1..S_l\}$	Release date of step $_{sl}$
C_l	$\forall$	$l \in \{1..L\}$	Completion time of lot l
L_l	$\forall$	$l \in \{1..L\}$	Lateness of lot l

MIP

$$\text{Minimize} \quad \sum_{l=1}^{L} W_l L_l \tag{1}$$

$$\sum_{t=1}^{T} x_{l,sl,t} = 1 \qquad \forall \quad l \in \{1..L\},\ s_l \in \{1..S_l\} \tag{2}$$

$$\sum_{l=1}^{L} \sum_{sl=1}^{Sl} p_{sl,i}\, x_{l,sl,t} \le K_{i,t} \qquad \forall \quad t \in \{1..T\},\ i \in \{1..I\} \tag{3}$$

$$t_{sl} \ge t_{sl-1} \qquad \forall \quad s_l \in \{2..S_l\} \tag{4}$$

$$t_{sl} = \sum_{t=1}^{T} x_{l,sl,t} * t \qquad \forall \quad l \in \{1..L\},\ s_l \in \{1..S_l\} \tag{5}$$

$$C_l = t_{Sl} \qquad \forall \quad l \in \{1..L\} \tag{6}$$

$$\begin{cases} L_l \ge 0 \\ L_l \ge C_l - d_l \end{cases} \qquad \forall \quad l \in \{1..L\} \tag{7}$$

$$x_{l,sl,t} \in \{0,1\} \qquad \forall \quad l \in \{1..L\},\ s_l \in \{1..S_l\},\ t \in \{1..T\} \tag{8}$$

Objective function (1) minimizes the total weighted tardiness. Term (2) assures that each step is processed exactly one time. Term (3) is the capacity constraint. Term (4) assures succession of processing steps. Term (5) defines the release date for each step. Term (6) determines the lot process completion time. Term (7) defines the lot lateness. Term (8) is the binary constraint for the decision variable.

The mathematical model presented above was solved by ILOG CPLEX solver ILOG OPL STUDIO. We ran our experiments on an Intel ® Core ™ 2 Duo PC running a 3 GHz processor and 4 GB of RAM. Good results were obtained while testing the MIP on instances of a reduced size: the production schedule obtained respected the capacity of resources, the priority and target delivery dates of the lots. It gave optimal solutions for lots achieved before the end of the planning horizon, taking into account reentrancy. The MIP wasn't implemented for lots with due dates beyond this limit.

Further increasing the size of the tested instances (up to about 1000 steps plan), the resolution of MIP was halted as it required a very large amount of time and computer memory (Figure 1).

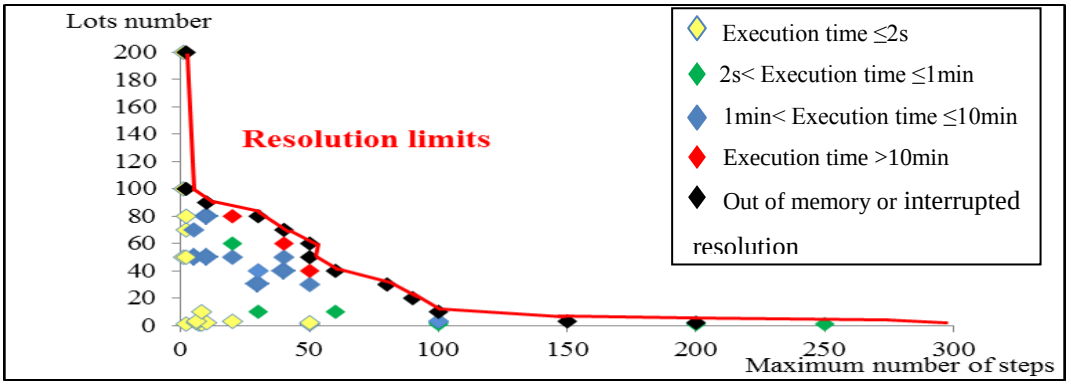

Figure 1: Limits of the MIP solution

Thus, the proposed MIP does not have a feasible solution to the operational perspective for our case study of STMicroelectronics corresponding to the realistic production data set.

Because of this huge size data, even the decomposition approach, consisting in dividing the problem in small sub-problems and solving each sub-problem successively (Ponsignon and Mönch, 2012), couldn't allow reducing the computational burden for this large scale production planning problem.

5 SECOND APPROACH

While applying the first approach of the resolution method, we conclude that the exact method could not afford an optimal solution due to the huge quantity of constraints and the complexity of the problem. We propose to simplify the problem by neglecting, at the first stage, capacity constraints and considering another purpose. Indeed, an infinite capacity planning engine is developed taking into account cycle times variability and target delivery dates. This tool will perform the WIP projection lot by lot, computing by day the activity required to ensure the delivery plan. In the projection, each lot has its own cycle-time model based on the necessary and sufficient speed to achieve its due date and a common fine-tuned reference cycle time curve. Hence, this tool is different from the actual engine, CAPACE used in ST for WIP projection which is presented above.

The proposed tool has, as inputs, the current WIP at lot level, the lots due dates and a new model to compute the objective cycle time per step ($CTobj_{step}$) based on a semi empirical formula named Zachka's formula in STMicroelectronics. In this formula, the data needed are the theoretical cycle time of the step ($CTTH_{step}$) corresponding to the processing time, the theoretical cycle time of route ($CTTH_{route}$) which corresponds on the sum of the steps processing times, the objective cycle time of each route ($Ctref_{route}$) that considers queuing times dependent of equipment saturation rate as following:

$$CTobj_{step} = CTTH_{step} \times Xfactor_{step}^{Zachka}$$

With $$Xfactor_{Step}^{Zachka} = \frac{CTTH_{route} \times (\frac{CTobj_{route}}{CTTH_{route}} - 1)}{\sum \sqrt{CTTH_{route} \times CTTH_{step}}} + 1$$

This tool provides as outputs (Data projection) the number of moves per day (*MovesIn* for wafers entering an operation, *MovesOut* for those completing it), the quantity of WIP at the beginning and at the end of each day and the cycle time coefficient per lot considering the due date for each lot.

This new projection engine is based on an iterative algorithm. The principle of the algorithm is as follows:

1. **For each route**, compute the objective cycle time for each of its steps using Zachka formula.
2. **For each lot**, from its position in its route :
   - ***2.1. compute parameters***
     - Compute the remaining cycle time (*RemCTobj*) from the current step to the end of the route : $RemCTobj = CTobj_{route} - \sum_{s=1}^{s=id_{step\ en\ attente}-1} CTobj_s$
     - Compute the remaining cycle time from the date of extracting WIP (WIP date) to due date : *RemCT = Due_date – WIPDate*
     - Compute the cycle time coefficient : *C= RemCT/ RemCTobj*
   - ***2.2. Project steps for each lot***
     - Determine the start and the end date of each step of the route. The cycle time for each step is equal to $C \ X \ CTobj_{step}$.
3. Compute the number of *MovesIn*, *MovesOut* and the quantity of WIP in the beginning and the end of a day.

This new projection engine is developed in JAVA. After WIP projection, this tool will use current "recipe balancing" and capacity computation modules of CAPACE to define equipment saturation.

To explain the difference between WIP projection principles using MRP, CAPACE and the new engine, two simple examples are developed.

We consider one lot with three remaining steps A, B and C. Table 1 contains cycle time for each step using the different tools. The cycle time using MRP approach is equal to the sum of process time and the average queuing time derived from historical data. For CAPACE, the cycle time is constant for each step; it is equal to the route's "reference" cycle time divided by the number of its steps. With the Zachka approach, the route's reference cycle time is allocated according to the Zachka Factor, which leads there to significantly longer steps.

Table 1: Comparison of cycle times using various tools

Remaining step	**Cycle Time (MRP)**	**Cycle Time (CAPACE)**	**Cycle Time (Zachka)**	**Cycle Time 1st instance (New engine)**	**Cycle Time 2nd instance (New engine)**
A	1,588	2,969	8,909	5,625	1,534
B	0,181	2,969	0,775	0,489	0,133
C	1,383	2,969	7,738	4,885	1,332

For the first instance, we consider a comfortable margin of 11 days between the date of WIP extraction and the lot's due date. So the cycle time coefficient used by our engine is equal to 11/ (8,909+0,775+7,738) =0,631. Figure 2 illustrates the different projection results. It demonstrates that the new projection engine, for this instance, allows the extension of steps queuing times, when we are far from the due date.

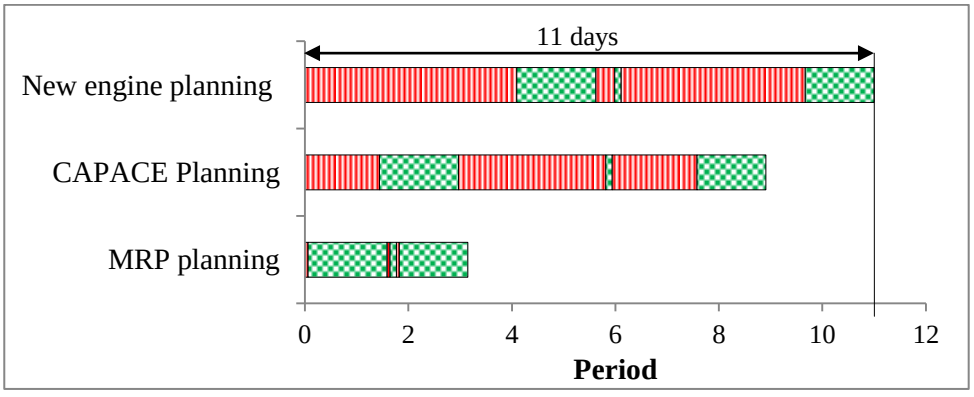

Figure 2: Difference between the different types of planning in case of large margin between WIP date and due date

In the second instance, we reduce the margin between the extraction date of WIP and the lot's due date from 11 to just 3 days. So, the lot's cycle time coefficient becomes equal to 3/(8,909+0,775+7,738)= 0,172.In this case, as it is illustrated in figure 3, projections using both MRP and CAPACE lead to delays

(significant in the case of CAPACE) while the new engine shrinks steps cycle times in order to satisfy the lot's due date.

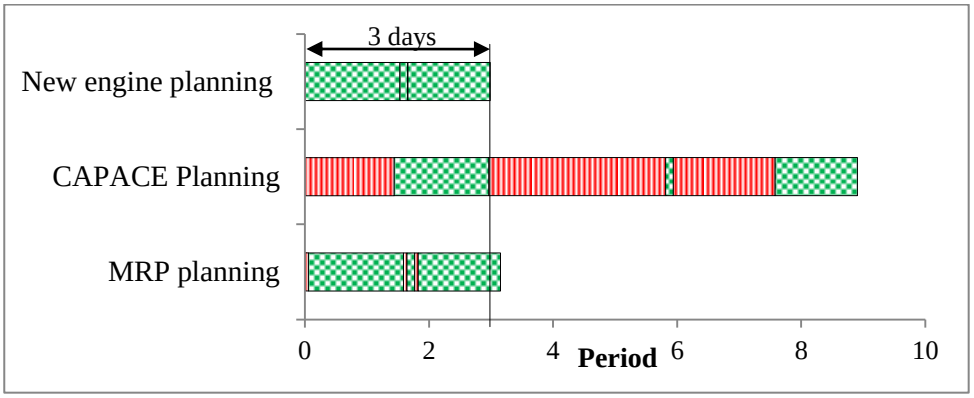

Figure 3: Difference between the different types of planning in case of strait margin between WIP date and due date

Hence, the new engine leads to a projection that is closer to the targeted behavior than current software tools because due dates are a starting hypothesis. In a sense, it delivers the plan that needs to be done in order to fulfill customer delivery commitment. However, the solution still may not be feasible as it runs at infinite capacity: after WIP projection, we may find equipment with workload exceeding capacity.

Nevertheless, the new technique of WIP projection is much faster in terms of execution. It takes approximately 5 seconds to project more than 2000 lots with 200 to 300 remaining steps for each one instead of 5 minutes with the current solution. Knowing that cycle time variability is also taken into account as an input, the new tool looks very promising.

6 CONCLUSION AND PERSPECTIVES

In this paper, we suggested two production planning approaches for semiconductor wafer fabrication facilities. The first approach consists of a mixed integer programming (MIP) that takes finite capacity considerations into account. The MIP's resolution requires a lengthy time. The results of some preliminary computational experiments of the MIP's resolution show that we are very far from the execution time required for the realistic case. So, even the decomposition heuristics couldn't resolve the problem.

In the second approach, we simplified further the problem and developed an infinite capacity WIP projection engine. This tool respects the lots due dates and its design enables to take into account the variability of cycle times. It presents also interesting results in terms of execution time.

There are several directions for future research. First, we can try to resolve this problem using meta-heuristics taking into account all the operational constraints. Second, despite the strengths of the new projection engine in terms of projection specificities and execution time, it is running at infinite capacity. So, after WIP projection, many machines are over saturated. So, in future research, we are interested to add a balancing load capacity module after executing the projection steps. This module sorts bottlenecks, identifies the oversaturation cause and redirects lots to minimize the workload for each period. Third, the new WIP projection engine computes steps cycle time using a semi empirical formula that takes into account the product mix. Although obtained cycle times are approximately near to the realistic case, it is interesting to define other efficient techniques to predict steps cycle times such as the iterative simulation-

optimization scheme, introduced by Hung and Leachman (1996), where a simulation model estimates cycle times that are used in turn as input parameters in an optimization model.

ACKNOWLEDGMENTS

This work is supported by the ENIAC European Project INTEGRATE. The authors also gratefully acknowledge STMicroelectronics for their support on the knowledge of the semiconductor industry.

REFERENCES

Bermon, S., and S. Hood. 1999. ”Capacity optimization planning system (CAPS).” *Interfaces* 29(5):31–50.

Carlson, J. G., and A. C.Yao. 1992. *“*Mixed model assembly simulation.” *International Journal of Production Economics* 26(1-3):161–167.

Catay, B., S. S. Erenguc, and A.J. Vakharia. 2003. “Tool capacity planning in semiconductor manufacturing.” *Computers & Operations Research* 30(9):1349–1366.

Chen, J. C., C. W. Chen, C. J. Lin, and H. Rau. 2005. “Capacity planning with capability for multiple semiconductor manufacturing fabs.” *Computers and Industrial Engineering* 48(4):709-732.

Chen, C.W., J. C. Chen and C. Joe Lin. 2008. “Finite capacity requirements planning with equipment capability and dedication for semiconductor manufacturing.” In *Proceedings of the 9th Asia Pacific Industrial Engineering & Management Systems Conference (APIEMS)*, 1310-1319.

Chen, J. C., Y.-C. Fan, and C.-W. Chen. 2009. “Capacity requirements planning for twin Fabs of wafer fabrication.” *International Journal of Production Research* 47(16):4473-4496.

Goldratt, E. M. 1990. *What is this thing called Theory of Constraints and how should it be implemented*.

Golhar, D. Y., C. L. Stamm. 1991. “The Just-in-time philosophy: a literature review.” *International Journal of Production Research* 29(4):657–676.

Guan, Z., Y. Peng, L. Ma, C. Zhang, and P. Li. 2008. “Operation and control of flow manufacturing based on constraints management for high-mix/low-volume production.” *Frontiers of Mechanical Engineering in China* 3(4):454-461.

Gupta, J. N. D., R. Ruiz, J. W. Fowler, and S. J. Mason. 2006. “Operational planning and control of semiconductor wafer fabrication.” *Production Planning and Control* 17(7):639–647.

Habla, C., L. Mönch, and R. Drießel. 2007. “A Finite Capacity Production Planning Approach for Semiconductor Manufacturing.” In *Proceedings of the 3rd Annual IEEE Conference on Automation Science and Engineering*, 82-87, Scottsdale, AZ, USA.

Habla,C., and L. Mönch. 2008. “Solving volume and capacity planning problems in semiconductor manufaturing: a computational study.” In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson and J. W. Fowler, 2260-2266. Texas, U.S.A.

Horiguchi, K., N. Raghavan, R. Uzsoy and S. Venkateswaran. 2001. “Finite-capacity production planning algorithms for a semiconductor wafer fabrication facility.” *International Journal of Production Research* 39(5):825-842.

Hung, Y.-F., and R. C. Leachman. 1996. “A production planning methodology for semiconductor manufacturing based on iterative simulation and linear programming calculations.” *IEEE Transactions on Semiconductor Manufacturing* 9(2):257–269.

Ignizio, J. P., and H. Garrido. 2012. “Fab Simulation and Variability.” *Future Fab International* 41.

Leachman, R. C. and T. F. Carmon. 1992. “On capacity modeling for production planning with alternative machine types.“ *IIE Transactions* 24(4):62-72.

Leachman, R. C., R. F. Benson, C. Liu, and D. J. Raar. 1996. “IMPRESS: an automated production planning and delivery quotation system at Harris Corporation-semiconductor sector.” *Interfaces* 26(1):6–37.

Leachman, R. C. 2012. "The Engineering Management of Speed." In *Proceedings of the 2012 Industry Studies Association Annual Conference.* Pittsburgh, PA: University of Pittsburgh.

Levitt, M. E., and J. A. Abraham. 1990. "Just-In-Time Methods for Semiconductor Manufacturing." *Advanced Semiconductor Manufacturing Conference*: 3-9.

Liu, J., C. Li, F.Yang, H.Wan, and R.Uzsoy. 2011. "Production planning for semiconductor manufacturing via simulation optimization." In *Proceedings of the 2011 Winter Simulation Conference,* 3612 – 3622.

Mönch, L., J. W. Fowler, S. Dauzère-Pérès, S. J. Mason, and O. Rose. 2011. "A survey of problems, solution techniques, and future challenges in scheduling semiconductor manufacturing operations." *Journal of Scheduling* 14(6):583-599.

Mönch, L., J. W. Fowler, and S. J. Mason. 2013. *Production Planning and Control for Semiconductor Wafer Fabrication Facilities: Modeling, Analysis, and Systems.*

Ponsignon, T., L. Mönch. 2012. "Heuristic approaches for master planning in semiconductor manufacturing." *Computer Operations Research* 39(3):479–491.

Rippenhagen, C., and S. Krishnaswamy.1998."Implementing the theory of constraints philosophy in highly reentrant systems." In *Proceedings of the 1998 Winter Simulation Conference,* edited by D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, 993-996. Texas, U.S.A.

Rondeau, P. J., and L. A. Litteral. 2001. "Evolution of manufacturing planning and control systems: from reorder point to enterprise resource planning." *Production and inventory management journal* 42(2):1–7.

Rupp, T. M., and M. Ristic. 2000. "Fine planning for supply chains in semiconductor manufacture." *Journal of Materials Processing Technology* 107:390-397.

Sullivan. G., and K. Fordyce. 1990. "IBM Burlington's Logistics Management System." *Interfaces* 20(1):4343.

Uzsoy, R., C.Y. Lee, and L. A. Martin-Vega. 1992."A review of production planning and scheduling models in the semiconductor industry part I: system characteristics, performance evaluation and production planning." *IIE Transactions* 24(4):47–60.

Uzsoy, R., C.-Y. Lee, and L. A. Martin-Vega. 1994. "A review of production planning and scheduling models in the semiconductor industry part II: shop-floor control." *IIE Transactions* 26(5):44–55.

Vollmann, T. 2005. *Manufacturing Planning and Control Systems for Supply Chain Management*. New York: McGraw-Hill.

AUTHOR BIOGRAPHIES

EMNA MHIRI is PhD student at Grenoble G-SCOP Laboratory. She is an industrial engineer and she has completed her master in industrial engineering from Grenoble University, in 2013. Her research interests include capacity planning in semiconductor industry. Her email address is Emna.Mhiri@ grenoble-inp.fr.

MIREILLE JACOMINO is professor in G-SCOP Laboratory. She is researcher in combinatorial optimization of systems. Her application fields are manufacturing and energetic systems. She is particularly interesting in execution context of systems. Her works aim at computing decisions of control that guaranty some performance during execution. Robust control and robust decision aided are the key stones of interest that consists in facing uncertainties. Her email address is Mireille.Jacomino@grenoble-inp.fr.

FABIEN MANGIONE is assistant professor in G-SCOP Laboratory. He has a PhD in industrial engineering and works on production planning on industrial cases. His research also deals with lot sizing problems on supply chain modelization. His email address is fabien.mangione@grenoble-inp.fr.

PHILIPPE VIALLETELLE is senior staff member at the Industrial Engineering department of STMicroelectronics Crolles300. He is in charge of the definition and follow-up of collaborative projects in Manufacturing Sciences. His fields of interest cover production planning and management techniques but also Process Control and Big Data. His email address is philippe.vialletelle@st.com.

GUILLAUME LEPELLETIER is Senior Project Leader at STMicroelectronics. He receives a master of sciences in "Advanced Modelling Systems", Brunel University, Uxbridge, UK, 1997; an Engineering degree in Operations and Production Management, INSA de Lyon, France, 1997. He has 15 years of professional experience in Industrial Engineering in the semiconductor industry, working on Capacity Planning, Cycle Time Management, Discrete Event Simulation, Industrial Reporting and Equipment Performance Tracking. His email address is guillaume.lepelletier@st.com.