Data-Driven Strategies for Optimizing Lot Priority Allocation to Reduce Cycle Times in Semiconductor Manufacturing

Adrien Wartelle[1], Quentin Christ[2], Stéphane Dauzère-Pérès[1],
Renaud Roussel[2], Claude Yugma[1]

[1] Mines Saint-Étienne, Univ. Clermont Auvergne
CNRS, UMR 6158 LIMOS, 13120 Gardanne, France
{adrien.wartelle, dauzere-peres, yugma}@emse.fr
[2] STMicroelectronics, 38920 Crolles, France
{quentin.christ, renaud.roussel}@st.com

Mots-clés : *Semiconductor manufacturing, data-driven modeling, lot priority, queuing theory*

1 Introduction

Semiconductor manufacturing companies face a large mix of chip product demands, each with varying urgency levels. This challenge unfolds in a highly complex environment where each wafer lot undergoes hundreds of production steps on hundreds of machines and requires several months to produce. New products must be produced quickly to reduce their time-to-market, whereas established products may have longer cycle times if they meet quality standards and quantity requirements. Consequently, production managers must define an effective queuing discipline, and optimize a mix of priority classes for the lots of the different products in the factory, i.e. a percentage of lots in each class. This approach prioritizes urgent lots in "hot" lot classes over standard lots, as described in [2]. Our study proposes a data-driven statistical tool to analyze the impact of the mix of priority classes on the speed-up of each priority class, essentially their inverse average queuing waiting time in a work center of the factory. The aim is to determine the optimal number of "hot" lots needed to achieve a specified speed-up improvement, acknowledging that increasing the number of "hot" lots inversely affect the speed-up.

2 Material and methods

This study relies on the recorded lot processing data from a 300mm semiconductor factory spanning from January 1, 2023 to April 30, 2023. The analysis focused on one specific work center, which accounted for 301,138 production steps, representing 13% of all production steps in the fab. The factory categorizes lot priorities into six levels : LOW, STANDARD, and MEDIUM constitute the MEDIUM- group with 92.7% of lot arrivals, while HOT, SUPER_HOT, and ULTIMATE (integrated into SUPER_HOT) form the HOT+ group with 7.3% of lot arrivals. Several data-driven queuing models have been developed to study the impact of the ratio of the HOT+ lots in the lot arrivals on the speed-up of each class. These models are characterized by a central probability selection model, which predicts the probability $p_i = \frac{l_i}{\sum_{j \in I} l_j}$, $i \in I$ of a lot i being selected during a queuing departure event, based on a score s_i estimated from the data. In this context, a fluid analytical saturated model was introduced to approximate the simulation results for the speed-up SU_i of each waiting time group (where $\frac{\lambda_{i'}}{\lambda}$ is the arrival ratio) :

$$SU_i = \frac{\sum_{i'} \frac{\lambda_{i'}}{\lambda} \prod_{k \neq i'} s_k}{\prod_{k \neq i} s_k} \tag{1}$$

3 Numerical results

Figure (1) highlights the core findings of this study, depicting the variation of the speed-up of the lots in each priority class in relation to the ratio of HOT+ lots. Notably, reducing the ratio of HOT+ lots from 8% to 4% leads to a 0.89% improvement of the speed-up. Conversely, increasing the ratio of HOT+ lots from 8% to 16% results in a 1.79% decrease of the speed-up. When considering the overall mean waiting time of 4 hours and 38 minutes, these changes equate to an average time reduction of 2 minutes and 25 seconds, and an increase of 5 minutes, respectively. The 5Scores model considered in this study assigns estimated scores directly linked to priority classes, with $s_{LOW} = 0.898$, $s_{STANDARD} = 1$, $s_{MEDIUM} = 1.15$, $s_{HOT} = 1.26$, $s_{SUPER\_HOT} = 1.97$.

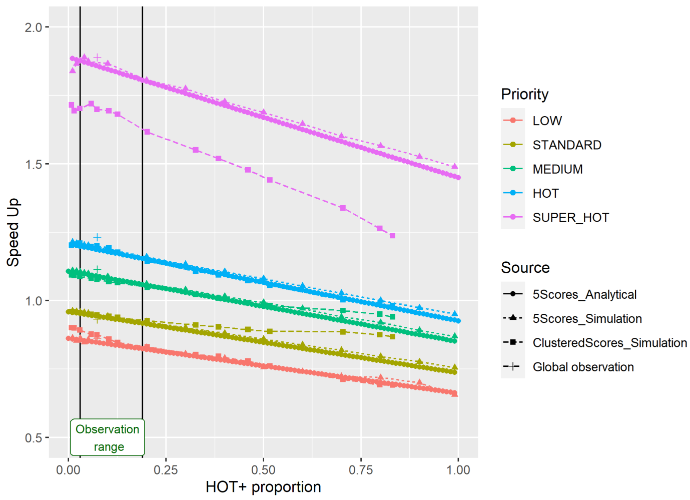

FIG. 1 – Impact on the speed-up of the ratio of HOT+ lots (priority mix) based on 4 months of data

4 Discussion, conclusion and perspectives

Based on a score-based probabilistic priority policy defined from data, this study introduces a new queuing model for predicting the impact of the priority class mix on the speed-ups of lots, using an analytical approach. Despite some biases, the results are in line with the existing literature [1], but we propose a more efficient approach than discrete event simulation. The proposed model, that relies on real behavior of the manufacturing system, offers managers a decision support tool that can quickly provide information to take informed decisions on priority policies.

Références

[1] David Fronckowiak, Andreas Peikert, and Kan Nishinohara. Using discrete event simulation to analyze the impact of job priorities on cycle time in semiconductor manufacturing. In *IEEE/SEMI 1996 Advanced Semiconductor Manufacturing Conference*, pages 151–155, Cambridge, MA, USA, 1996. Institute of Electrical and Electronics Engineers (IEEE).

[2] Kilian Schmidt. Improving priority lot cycle times. In *2007 IEEE/SEMI Advanced Semiconductor Manufacturing Conference*, pages 117–121, Stresa, Italy, 2007. Institute of Electrical and Electronics Engineers (IEEE).