Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

DEPLOYMENT OF A NOVEL PRODUCTION PLANNING AND PRESCRIPTIVE ANALYTICS SOLUTION: A SEAGATE TECHNOLOGY USE CASE

Sebastian Steele[1], Semya Elaoud[1], Ioannis Konstantelos[1], Dionysios Xenos[1],
Tina O'Donnell[2], and Sharon Feely[2]

[1] Flexciton Ltd, Floor 7, Linen Court, 10 East Road, London, UK
[2]Seagate Technology, 1 Disk Drive, Londonderry, BT48 0LY, UK

ABSTRACT

An emerging paradigm in semiconductor manufacturing is the "autonomous fab." Companies such as Seagate aim to make decision-making autonomous and aligned at all levels, from supply chain planning to tool scheduling, to cope with ever-increasing complexity and labor supply constraints. One identified gap is between committed deliveries in Enterprise Resource Planning (ERP) systems, and day-to-day execution, where daily production targets and wafer starts are often poorly aligned with shipment targets. Our novel production planning system creates optimized production targets to manage on-time delivery (OTD), line linearity, dynamic bottlenecks and starvation. It also provides proactive, quantified actions for improving fab KPIs. The system thus surpasses existing descriptive and predictive analytics solutions (answering "what happened?" and "what will happen?"), providing a prescriptive solution which "makes it happen."

1 EXISTING METHODS FOR PRODUCTION PLANNING

To achieve OTD, many fabs set lot priority levels using their required average cycle times. This method does not consider the dynamic constraints that apply to each lot, such as recipe enablement or timelinks, nor fab conditions such as WIP levels, which together lead to inaccurate predictions. Further, the impact of priority changes on OTD, linearity and the cycle time of other lots is unclear, as is the ideal number of "hot lots"; if every lot is "hot," then effectively none are. Beyond this, fabs may add heuristics such as putting lots with downstream bottlenecks on hold, and prioritizing WIP to pull it into starved toolsets. These rules also do not assess the impact on the overall line, and can therefore harm predictability and throughput.

Other methods include adjusting lot priorities via (1) line balance algorithms, often based on target cycle times (Leachman et al. 2002; Tsuchiyama and Smith 2022), and (2) discrete-event simulations (Scholl et al. 2016) which have limited ability to search large problem spaces. Recently, Machine Learning methods have improved the cycle time estimates that might be used in priority changes (Borodin et al. 2023). The above methods do not accurately consider the dynamic constraints of the fab, and do not holistically optimize the flow of WIP, considering the impact on every lot, product and tool.

2 SEAGATE USE CASE

Seagate Technology constitutes over 40% of the global Hard Disk Drive (HDD) market, with its Northern Ireland facility producing 25% of global demand for read-write heads, the critical component in an HDD.

Potential improvements in cycle time, line predictability and line balance were identified by replacing the existing system that relies on traditional prioritization methods. This would allow Seagate to commit to delivering higher volumes, better plan new product lines, and further their vision of an autonomous fab by making intelligent planning decisions which adapt to fab conditions. It would also prioritize recipe enablement and tool recovery, alleviating the need to run manual analyses.

3 IMPLEMENTATION

The solution plans the production targets required to achieve the desired trade-off of OTD and throughput whilst preventing bottlenecks and starvation. These targets specify the number of lots of each product and step to process on each toolset until delivery, broken down into periods of a few hours. All targets are calculated holistically using mathematical optimization, which is ideally suited to solving highly constrained problems with large problem spaces and sparse feasible solutions. It enables explicit modeling of fab constraints, such as recipe enablement and timelinks, ensuring dynamic fab constraints and conditions are represented. Endogenous and exogenous uncertainties are addressed by measures including advanced methods for optimization under uncertainty, buffering of lots at key toolsets, and frequent replanning. The model can also be introspected to identify key binding constraints, which allows us to identify, quantify and prioritize actions such as recipe qualification changes.

The system is controlled via trade-offs of business objectives such as cycle time and throughput. By varying the target granularity and optimization horizon, the solution is able to support fabs with very different scales and product mixes. The system fits into Flexciton's hierarchy of control, enabling autonomous and aligned decision-making at all levels of the fab, with each level interfacing via production targets, to which lower-level optimization models include objective terms to penalize shortfall, along with other objectives e.g. batching efficiency.

4 RESULTS

Preliminary results have shown improved throughput and linearity KPIs. The system can autonomously react to excursions (where a large number of lots are scrapped, and tools taken offline), avoiding starvation without manual intervention. The system additionally provides quantified and prioritized actions to be discussed in daily production meetings.

5 FUTURE WORK

The system can be extended to address WIP-aware maintenance planning, which can prevent unnecessary queue time, and wafer starts planning, which will close the gap between ERP commits and daily execution, whilst also preventing bottlenecks caused by starting wafers with conflicting paths in the line. The system may also be used to load balance multiple cleanrooms at one facility, and may be applied across multiple facilities, including assembly and test facilities, to improve supply chain predictability and visibility.

REFERENCES

Tsuchiyama, H. and H. Smith. 2022. "A Novel Approach to Dynamic Line Balance Control and Scheduling with a Digital Twin Production". In *2022 International Symposium on Semiconductor Manufacturing (ISSM)*, December 12-13, Tokyo, Japan, 1-4.

Leachman, R. C., J. Kang, and V. Lin. 2002. "SLIM: Short Cycle Time and Low Inventory in Manufacturing at Samsung Electronics". *Interfaces* 32(1):61-77.

Scholl, W., M. Förster, P. Preuss, A. Naumann, B. P. Gan, and P. Lendermann. 2016. "Simulation-Enabled Development Lot Journey Smoothening in a Fully-Utilised Semiconductor Manufacturing Line". In *2016 Winter Simulation Conference (WSC),* 2559-2567 https://doi.org/10.1109/WSC.2016.7822294.

Borodin, V., N. G. Chembu, A. Khemiri, J. van Heugten, and C. Yugma. 2023. "The Impact of Data Fragmentation on Processing Time Modeling in Semiconductor Manufacturing : *Topic: IE: Industrial Engineering". In *2023 34th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, May 1-4, Saratoga Springs, NY, USA, 1-6.