Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

DEPLOYMENT OF AN ADVANCED A.I. SCHEDULER AT PHOTOLITHOGRAPHY: A SEAGATE TECHNOLOGY USE CASE

Robert Moss,
Dennis Xenos

Flexciton Ltd.
The Bower, 6th floor, 207 Old Street
London, EC1V 9NR, UK

Tina O'Donnell

Seagate Technology
1 Disk Drive
Londonderry, BT48 0LY, UK

ABSTRACT

Photolithography is the cutting edge of semiconductor manufacturing and requires the most complex and expensive equipment. Photolithography tools are usually the bottleneck area because of the complexities around their operation including the allocation of secondary resources such as reticles to the tools. Reticles are fragile and expensive so the minimisation of the reticle movements helps to mitigate the risk of damaging them, but may sacrifice the fab's fundamental objective of increasing throughput. Our advanced A.I. scheduler can adapt and optimise photolithography tools with a variety of constraints and secondary resources. The scheduler has been implemented in various photolithography toolsets at the Seagate facility in Northern Ireland. It has been running 24/7 for more than a year. The first results from production show a 9.4% increase in lot moves, 4.3% reduction in lot queue time and 5.3% decrease in reticle moves at the same time.

EXISTING METHODS FOR SCHEDULING PHOTOLITHOGRAPHY

The problem of scheduling a photolithography toolset (a toolset consisting of multiple photolithography tools that share the same recipes) is known as the dual resource constrained (DRC) problem. Most of the semiconductor companies rely on their dispatch systems (Waschneck, Altenmüller, Bauernhansl, and Kyek 2016) to schedule their photolithography toolsets based on the work by (Chien 1989). This approach can result in poor outcomes compared to optimisation-based approaches. The work by (Chien 1989) presented an integrated solution with linear programming and dispatching rules which showed an average reduction in cycle time by 3.2% and 25.2% reduced counts of reticle moves. The work by (Munoz, Villalobos, and Fowler 2022) proposed a mathematical programming model that is primed by dispatching rules. To the best of our knowledge, semiconductor manufacturers rely on dispatching methods to solve the DCR problem, the more advanced mathematical models presented in the literature are still not used in production for scheduling complex photolithography toolsets. Our optimisation-based method has been scheduling the photolithography area at a Seagate facility in a closed-loop for more than a year.

SEAGATE PHOTOLITHOGRAPHY USE CASE

Seagate Technology has a more than 40% share of the global Hard Disk Drive (HDD) market and its Springtown facility in Northern Ireland produces around 25% of the total global demand for recording heads, the critical component in an HDD.

The Seagate facility involves various types of photolithography toolsets. Table 1 shows these types of toolsets, the methods of how the reticle is loaded and used inside the tool, and type of reticle storage.

Table 1: Types of photolithography toolsets at the Seagate Springtown facility in Northern Ireland.

	Toolset A	Toolset B	Toolset C	Toolset D
Pod/Library	2 pod load ports	3 pod load ports	Internal Library	Internal Library
Reticle Storage type	Reticle Stocker	Reticle Stocker	Reticle Stocker	Reticle Cabinets

The reticle stocker dispatches reticles in pods (a handling system for transferring reticles) which can hold up to six reticles. These are then placed on a load port and either moved into the tool's internal library before use or are used directly from the pod. From a scheduling standpoint we consider this workflow, the wafers that need to be scheduled, their relative priorities and batching rules, what reticles they use, what other reticles can be combined in a pod, with the aim of utilising these tools to the highest degree.

IMPLEMENTATION

The scheduler uses a combination of mathematical programming, decomposition techniques and heuristics to support the computationally expensive models. The solution is hosted in the cloud and it produces schedules every five minutes in a closed-loop. The data are pulled from the MES system and the scheduling decisions are pushed to the dispatching system in an automatic fashion.

The scheduler provides exact decisions of the batching of wafers into lots, reticles allocation to pods, and pods allocations to the photolithography tools. The schedules of the resources and wafers are produced simultaneously. The scheduler has been integrated with the MES system acquiring data about the location of the wafers and the reticles and the status of the tools. The scheduler provides the wafer schedule via Seagate's real time dispatch system. It also provides the reticle allocations to the tools through the reticle stocker where pods with reticles are built automatically. In the case of the reticles in the reticle cabinets, operators follow the instructions of the scheduler and manually transfer the reticles according to the given schedule.

The deployment of the scheduler required several iterations and trials to validate data, adapt the A.I model, integrate with Seagate's systems and agree the objective weights with different stakeholders. The cloud technology allowed for seamless trials and continuous deployments of new releases, which significantly sped up the roll-out. A modern, user-friendly visual interface helped to explain counter-intuitive decisions by the optimisation-based scheduler. Addressing the change management challenges from the beginning and frequent communication with all the parties contributed to a successful roll-out.

RESULTS

We scheduled hundreds of unique reticles across a variety of tools considering complex constraints in a live production environment at the largest HDD head fab in the world. Since deployment, the throughput of Toolset D has increased by 9.4%, lot queue time has reduced by 4.3% and at the same time the number of reticle moves has reduced by 5.3%. The comparison before and after the deployment of the scheduler considers live production data from multiple months for each system used in the fab. Additionally, operators have given positive feedback since the deployment which is aligned with the data.

REFERENCES

Chien, C. 1989. "Small Sample Theory for Steady State Confidence Intervals". Technical Report No. 37, Department of Operations Research, Stanford University, Stanford, California.

Munoz, L., J. R. Villalobos, and J. W. Fowler. 2022. "Exact and Heuristic Algorithms for the Parallel Machine Total Completion Time Scheduling Problem with Dual Resources, Ready Times, and Sequence-Dependent Setup Times". *Computers & Operations Research* 143:105787.

Waschneck, B., T. Altenmüller, T. Bauernhansl, and A. Kyek. 2016. "Production Scheduling in Complex Job Shops from an Industry 4.0 Perspective: A Review and Challenges in the Semiconductor Industry". *SAMI@ iKNOW*:1–12.