CASE STUDY

Optimizing Batch Formation for Enhanced Efficiency and Cost Reduction: Insights from Wafer Fab Deployments

flexciton.com

Flexciton

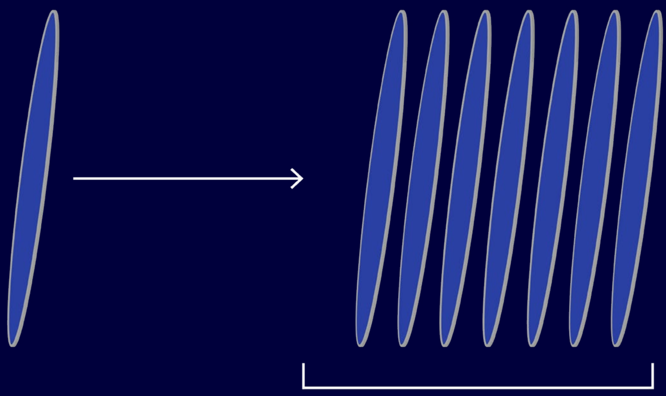
The Complex Challenge of Efficient Batching

Among the myriad factors influencing a front-end fab's production efficiency, batching emerges as a complicated and often overlooked problem that can significantly impact cost reduction—a critical objective for any wafer fab. Batching, in essence, involves efficiently grouping together wafers for processing or transportation. Beyond the immediate goal of minimizing cycle times, batching decisions can reverberate across various objectives within the fab, such as maximizing throughput, increasing tool utilization and reducing timelink violations.

The challenge intensifies as fabs encounter the need to dynamically adapt batching policies, such as 'greedy' or 'full-batch', to influence different objective trade-offs. Full-batch policies require lots to wait until a full batch is available and tend to favor effective capacity utilization and cost factors, while negatively impacting cycle time and variability. Greedy policies don't wait for full batches and instead aim to reduce cycle time. Since certain production KPIs, such as reducing cycle times and maximizing tool utilization, are often at odds with each other, the challenge lies not only in defining a set of batching policies but also in balancing them in real-time to find harmony between conflicting objectives. This adaptability becomes particularly crucial during downturns, when the industry faces the delicate task of reducing costs without sacrificing operational capacities or compromising on customer demands.

Typically, fabs handle batching with a rules-based scheduling approach, where many decisions are still made manually. Rules such as *a minimum batch size of Y lots for dispatch and a maximum wait duration of X hours* are employed in an attempt to avoid underutilization of tools. While rules like this can be beneficial in some circumstances, they have significant drawbacks.

When fab operations change, such as the introduction of new processes or tools or shifts in product mix, the heuristic exhibits a limited capacity to adapt to new process requirements or accommodate changes in demand. The introduction of factors like high-priority lots (also known as hot lots), downstream capacity constraints, and timelink constraints elevates the problem complexity far beyond the scope of rules-based capabilities. Sustaining acceptable performance with a rules-based approach therefore necessitates frequent recalibration by skilled users, a practice that proves highly inefficient in the context of a modern wafer fab, where optimal resource utilization is paramount.



The Solution: Flex Local, the autonomous AI scheduler

In order to effectively form batches, an intelligent and adaptable scheduling technology is required. Flex Local – Flexciton's AI scheduler – is based on a hybrid approach with mathematical optimization at its core. Unlike the fixed rules found in legacy schedulers, Flexciton's optimization-based approach is able to dynamically assess a multitude of variables, including equipment constraints, processing times, and resource availability. This flexibility enables the scheduler to adjust batch sizes in real-time, optimizing efficiency based on the ever-changing conditions of a modern wafer fab.

The AI-driven scheduling solution can decide how long to wait for the next lots, considering the accumulating queuing time of the current lots and the predicted time for new lots to arrive. No predetermined rules are in place; instead, the mathematical formulation encompasses all possible solutions. Flex Local's capacity to consider the interplay of different factors and constraints also enables it to balance complex trade-offs, such as timelink violations and queue time. With a user-defined objective, such as prioritizing tool utilization over cycle time, an optimization solver autonomously identifies the optimal trade-off, eliminating the need for any parameter adjustments. This capability proves critical in scenarios where fab operations evolve, ensuring that the fab is able to keep their batching strategy optimized in real-time.

To showcase the benefits of deploying Flex Local on batch tools, we will examine two real-world case studies.

Case Study: Seagate Technology

Springtown in Derry, Northern Ireland, is home to Seagate Technology's largest wafer fab facility, responsible for producing 25% of the world's read-write heads. They selected Flexciton as their scheduling system in 2021, and highlighted batching efficiency as a problem at certain toolsets. Recent analyses demonstrate that the scheduling system has autonomously optimized batching efficiency, leading to significant reductions in cost.

One of the areas of the fab Flex Local was implemented on contained deposition toolsets that handle valuable materials, such as gold. This meant it was critical to optimize for maximum batch size in order to avoid precious material wastage. In light of this, the scheduler was instructed to minimize batching costs over

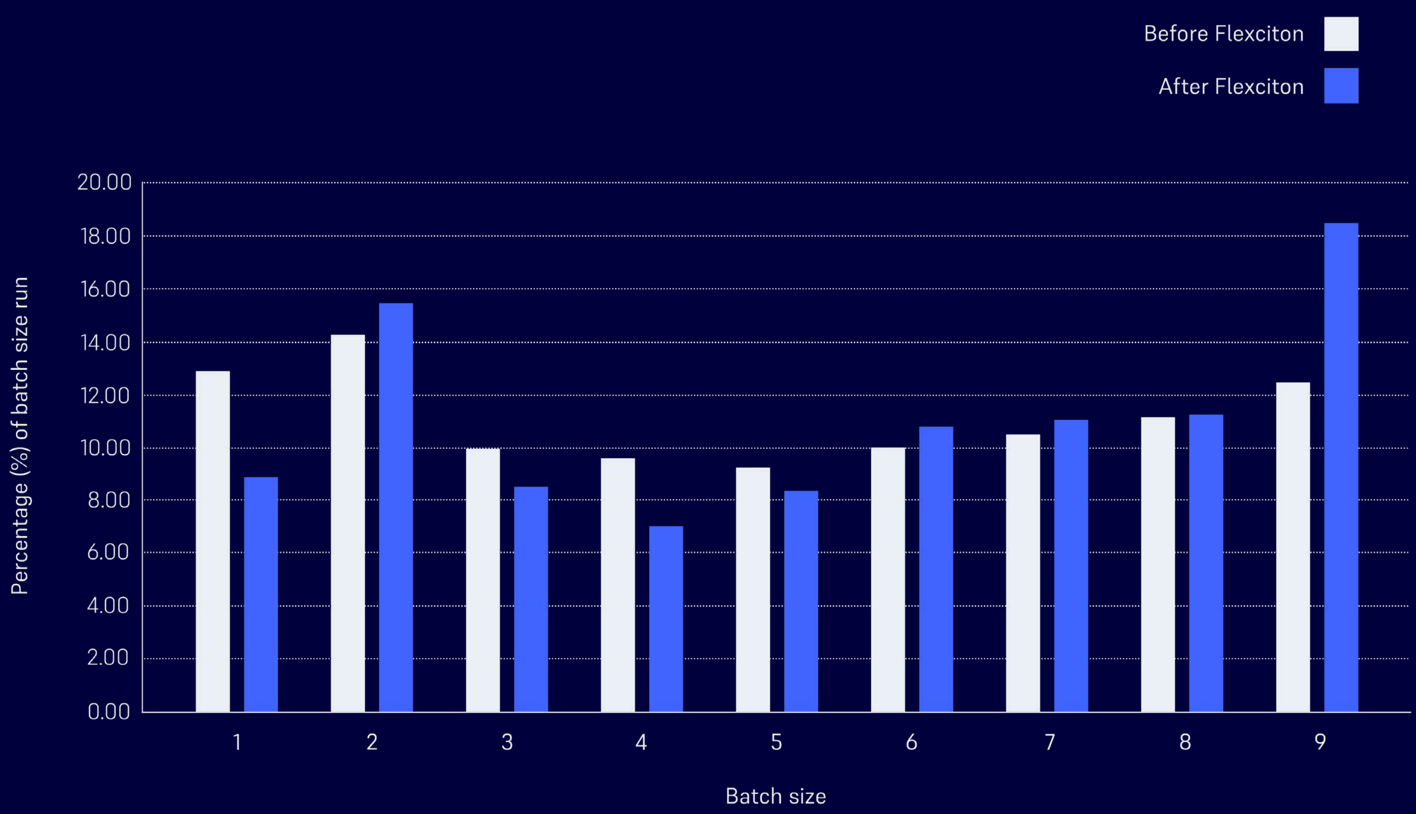


Figure 1 illustrates how Flex Local has increased the average batch size by 8.5% for this valuable toolset. The comparison is based on historical data from both the dispatching logic and the scheduler, considering a few months of data for each.

cycle time, which was achieved by the user adjusting just two different business-driven values—namely, the weight for average cycle time compared to the weight for batching cost. The scheduler adapted by increasing batching size with the optimal minimal sacrifice on other production KPIs, such as cycle time. Without manually tuning if-then parameters, this can only be achieved with an autonomous scheduling system.

Figure 1 illustrates how Flex Local has increased the average batch size by 8.5% for this valuable toolset. The comparison is based on historical data from both the dispatching logic and the scheduler, considering a few months of data for each.

Case Study: Renesas Electronics

Renesas Electronics, a global leader in microcontrollers, analog, power, and SoC products, sought to explore the application of Flex Local to enhance specific toolsets in their US facility, particularly those in the diffusion area. Their main goal was to significantly reduce the number of batches (by increasing batch sizes),

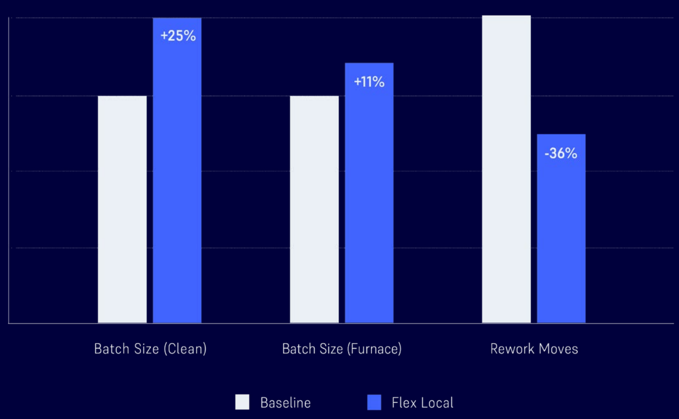


Figure 2 shows the impact of Flex Local on Renesas' diffusion area, helping to boost the average batch size at both furnace and clean, whilst the amount of rework was drastically reduced as a result of minising timelink violations.

while ensuring minimal impact on cycle time. Achieving this posed a scheduling challenge due to the inherent trade-off between batch size and queue time, a critical contributor to cycle time.

This problem was made more complicated by a third element to the trade-off, timelink constraints. Timelinks, also known as time lag constraints, time loops, qtime, or close coupling constraints, refer to a maximum elapsed time a lot can spend between two or more process steps. Their purpose is to minimize the risk of rework or scrap that may have to occur if the wafer is exposed for too long. Since timelinks impose a rigid time limit, it becomes harder to decide how long to wait for batch sizes to increase and, thus more difficult to achieve cycle time targets. This three-way trade-off presented a complex scheduling problem that only a multi-objective optimization approach would be able to solve.

Following the deployment of Flex Local in Renesas' diffusion area, we conducted a comprehensive analysis by comparing the targeted KPIs with the fab's performance during the corresponding period in the previous year. This analysis revealed that the KPI gains correlated with Renesas' objectives, increasing the average batch size in both clean and furnace toolsets by 25% and 11%, respectively. The objective weight given to decreasing the batch size KPI meant that cycle time did rise, but only within the expected range: 3% at the furnace toolset and 9% at the clean toolset. The result from the analysis on timelink violations was similarly significant, with the amount of reworked wafer moves reduced by 36% (see Figure 2). These results were achieved despite an average of 6–10% fewer operators working per shift, demonstrating Flex Local's capability to enhance both operational efficiency and fab performance simultaneously.

About Flexciton

Flexciton provides solutions for semiconductor manufacturers to power their transition towards autonomous factories. Our intelligent planning and scheduling software suite combines advanced optimisation techniques with the power of AI to orchestrate complex fab workflows and achieve critical revenue-to-shop-floor alignment.

www.flexciton.com

About Seagate Technology

Seagate Technology has provided innovative, precision-engineered data solutions for over 40 years and is the leading provider of bytes globally – with over 3 zettabytes1 of data shipped.

www.seagate.com

About Renesas Electronics

A global leader in microcontrollers, analog, power,and SoC products, Renesas provides comprehensive solutions for a broad range of automotive, industrial, home electronics, office automation, and information communication technology applications that help shape a limitless future.

www.renesas.com