Simulation in Produktion und Logistik
Entscheidungsunterstützung von der Planung bis zur Steuerung
Wilhelm Dangelmaier, Christoph Laroque & Alexander Klaas (Hrsg.)
Paderborn, HNI-Verlagsschriftenreihe 2013

Real life evaluation of a model-based combined dispatching approach at a low-volume high-mix ASIC facility

Praxisbezogene Anwendung und Auswertung eines modellbasierten kombinierten Dispatchansatzes in einer niedervolumigen Halbleiterfabrik mit hohem Produktmix

Mike Gißrau, X-FAB Dresden, Dresden (Germany), mike.gissrau@xfab.com

Oliver Rose, Universität der Bundeswehr München, Neubiberg (Germany), oliver.rose@unibw.de

Abstract: The fabrication of semiconductor devices, even in the area of customer oriented business, is one of the most complex production tasks in the world. A typical wafer production process consists of several hundred steps with numerous resources including equipment and operating staff. Smaller foundries with a high product mix and a low technology volume need reasonable assignments of each limited resource at each time. Several requirements defined by the process requirements, customers and management must be taken into consideration with the objective to find the best trade-off between the different needs. In this paper, we describe the practical assessment of a combined dispatching policy presented in Gißrau and Rose (2012). Besides the facility performance influence, also the human factor is taken into consideration. This includes dispatch compliance parameters and staff surveys.

1 Introduction

A typical application specific semiconductor facility, also called foundry, has a very complex production process. A wide range of different process steps at a characteristic re-entrant material flow makes wafer fabrication and production control very complex. A performance degradation of various facility KPI (Key Performance Indicators) is caused by the large amount of variability during processing like equipment failures or missing operator attendance. Therefore, a reasonable production control is vital for this business.

The wide diversification of different requirements at a typical foundry like a stable on-time delivery and short cycle times necessitate an intelligent approach for controlling the whole factory production process.

In the literature, a wide range of different scheduling and dispatching approaches can be found, often with a focus on mass production or academic examples. Scheduling as described in Pindeo (2002) is quite hard to be implemented and used in a high-mix low-volume facility. Due to the large amount of variability, the schedule's lifetime is very low, which requires a steady recalculation. Grupta et al. (2002) introduced a multiobjective scheduling approach with respect to variable objective targets, which is also not applicable at our case.

In this area, dispatching is quite common. Different techniques beginning with the definition of simple policies like in Rose (2001, 2002, and 2003) are known. However there can also be found more complex ones (like Dabbas et al. 2001), Bagchi et al. (2008)) taking different criteria into account, with focus on automated mass production with a low product mix. The influence of these rules is often not obvious and depends on the field of application (e.g. see Mittler et al. 1999).

In case of high-mix low-volume semiconductor facilities with manual operations, these dispatching approaches often tend to prioritize only one of the interesting facility performance parameters. The fast change of the product mix and a considerable amount of different new technology introductions cause a very unstable WIP (Work In Process). Therefore a dispatching approach is required where the current facility state is taken into account (e.g. current tool downs, operator attendance and qualification). Furthermore optimization targets can be changed on the fly, e.g., by changing the focus from on-time delivery to cycle time.

2 Theoretical background

In this section, we introduce the theoretical background of the combined dispatch policy as a short review of the approach presented in Gißrau and Rose (2012).

2.1 The combined dispatching policy

In the semiconductor foundry business, a wide range of concurrent requirements defined by the technological background, the customer and the management exists. In many cases, the requirements affect each other.

Common facility KPIs at semiconductor manufacturing are the on-time delivery and the cycle timer per mask layer. Focussing on only one parameter and its optimization is not sufficient to fulfil all requirements and finding the best trade-off between the different needs.

The combined dispatching policy consists of a variable set of different single-objective dispatching policies. The goal is to find the best trade-off between the different requirements for the fab. The rule set is combined in a linear way to find the best compromise between the different needs.

The determination of the weights is realized by a detailed facility simulation model. The simulation model is feed by data from the facility data warehouse including different data sources like MES (Manufacturing Execution System) or ERP (Enterprise Resource Planning) system.

A genetic algorithm calculates the weight for each dispatching rule by an objective function. This objective function is defined by the management and the logistics department and changes over time, depending on the needs and requirements of the different customers. The optimization is run periodically at different times of the day.

Different simulation studies (see Gißrau and Rose 2012) show positive effects on various facility performance KPIs like cycle time per mask layer within an improvement up to 8% in comparison to the former applied optimized FIFO policy. The tardiness of the lots, which is one of the most important KPI in the foundry business, could be successfully reduced by about 5%.

The whole approach is illustrated in Figure 1.

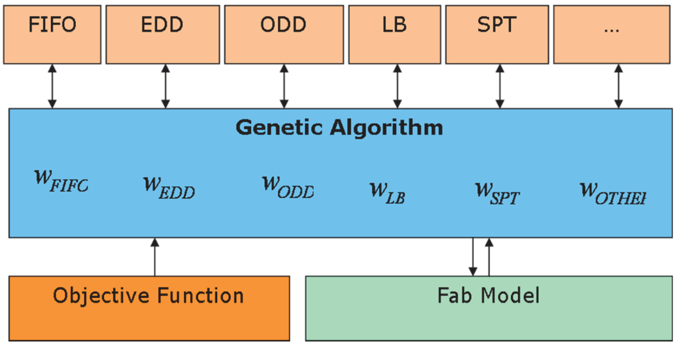


Figure 1: *The combined dispatching approach*

2.2 System implementation

Implementation and introduction of a new system in a running facility is a difficult task. The inhomogeneous IT infrastructure and the different expectations of the operating personal are barriers to be overcome.

In a historically grown IT infrastructure, a wide range of different data sources are in use. For the application of a simulation model and real-time dispatching, these data sources have to be accessed in a stable and reasonable way. Figure 2 illustrates the general environment.

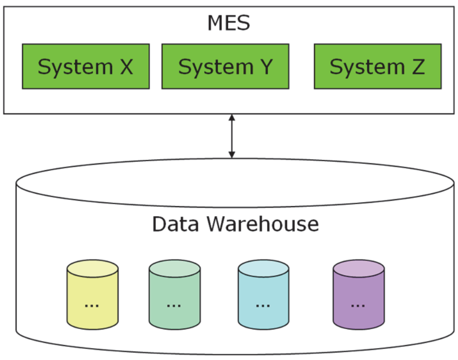


Figure 2: *Data resources*

The steady change of software solutions of the IT landscape caused by software replacements and updates force us to use an independent application structure. For this, the usage of web services is preferred. This allows a flexible and independent access to the required methods of the application. The system is implemented with the programming language JAVA, which leads to a stable solution. The application is divided into two parts:

- *Real Time Dispatcher:* The real time dispatcher is responsible for dispatch list generation for the different equipment and the lot priority update at lot movements. Due to the real time requirements of this operation, the time consumption of the different method calls should be very low.
- *Simulation and Optimization:* The simulation and optimization core is responsible for calculating the new weight combinations according the genetic algorithm. For that, a detailed facility model is generated from the data ware house.

A data structure is established to obtain a standardized access to the required data of the data warehouse. This data structure is filled by various different data sources with online and offline data. Figure 3 illustrates the system overview.

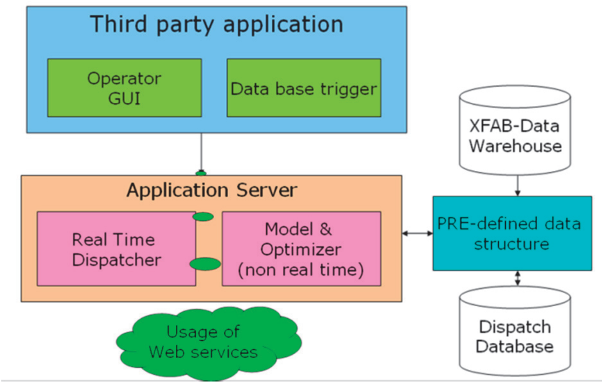


Figure 3: *Application overview*

3 Practical evaluation

This section illustrates the practical assessment of the combined approach at a high-mix low-volume ASIC facility. During the four month test period, the whole dispatching of the fab is realized by the new application using the combined approach. The test period ranges between September 2012 and December 2012.

3.1 Facility KPI results

During the evaluation period, the dispatching system had a positive influence on the facility performance behavior; especially the variance of different facility performance parameters could be successfully reduced.

Figure 4 shows the absolute S-Curve for the cycle time per mask layer without any normalization. The interesting quarters of the year 2012 are Q3 and Q4. In Q1 and

Q2, the technology mix and the facility load changed, therefore only Q3 and Q4 could be compared, because the boundary conditions were the same.

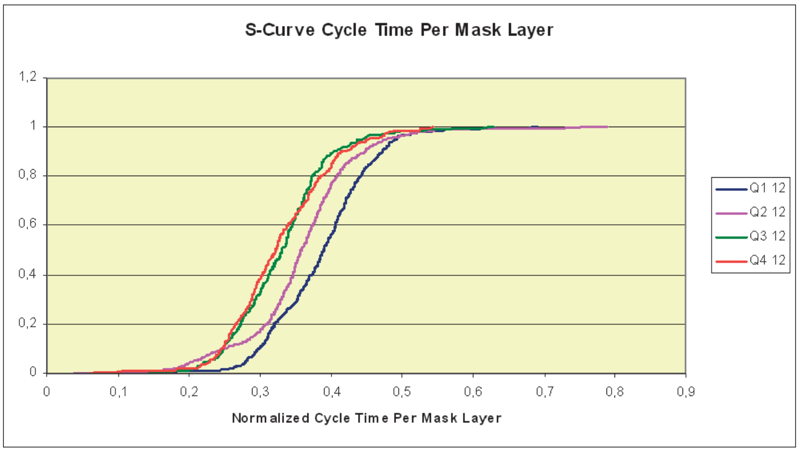


Figure 4: *Cycle time per mask layer over all technologies without normalization*

The positive effect takes place at Q4, after a considerable period without visible changes due to the average processing duration of two to three month for one lot. At the lower border, the increases of the curves in Q3 (first effects of dispatch control) and Q4 are much sharper than in Q1 and Q2. That indicates a better management of faster and slower lots. Before the dispatch algorithm starts, lots are sometimes delivered too early whereas lots with strength due date ware delivered too late.

A further effect can be seen around the median of the curve. The absolute cycle time value could be successfully reduced by 5%. The worse values at the upper end of the curve are caused by production problems rather than by the dispatch control system.

Figure 5 illustrates the S-Curves of the cycle time per mask layer after normalization of the absolute values. Often, different technologies have different cycle time per mask layer distributions due to processing reasons. After elimination of this effect, the positive influence becomes more obvious. Besides a sharper increase, also the distribution values of Q4 are lower than of Q3.

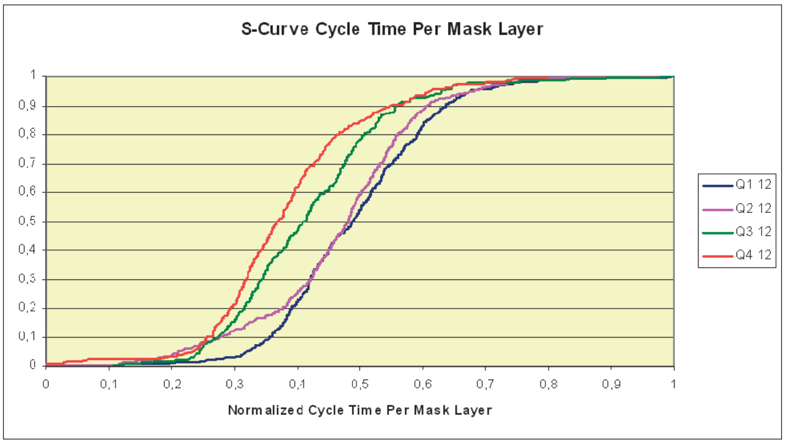


Figure 5: *Cycle time per mask layer over all technologies with normalization*

3.2 System performance results

The performance of the system is vital for the acceptance of the system by the operating personal. To that end several performance parameters were analyzed during the evaluation period. The most vital performance parameter is the time consumption of the dispatch list generation. Because this is a online task, the operator has to wait this time until he can take the next lot for processing. Figure 6 illustrates the statistics for this case.

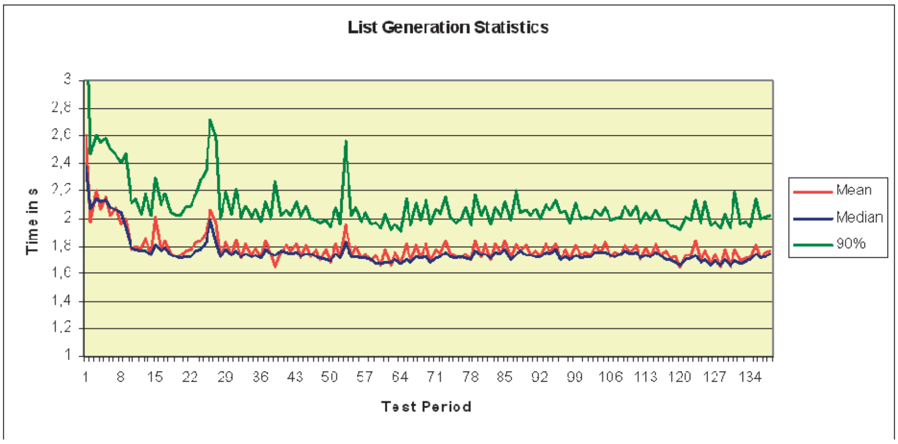


Figure 6: *Dispatch list generation statistics*

The median value is about 1.7 seconds, whereas the 90\% value is around 2 seconds. The 90\% value represents the main statistical output for our needs. Values above 3 seconds are not accepted by the operating staff. The largest amount of time is caused by detecting the right lots for the equipment rather than by the dispatch algorithm itself. This process also includes the interaction with the MES by various system method calls.

3.3 Dispatch compliance

The dispatch compliance is a vital parameter for the acceptance of the whole approach by the operating staff. Due to the huge amount of manual interaction in the production process, the operator also has to choose the next lots for processing. The operating staff is forced to take the first lot or batch from the generated dispatching lists.

Thus, several statistical analyses were carried out to obtain an overview about the compliance level at the facility. The parameter **i** describes the position of each lot or batch in a dispatch list with *N* lots or batches. We define different compliance parameter, e.g.:

- *Compliance Score:* The compliance score P_C is a declarative value for estimating the dispatch compliance. The value ranges between 0 and 1, where the target value is 1.

$$P_C = \begin{cases} 1 & i = 1 \\ \frac{N-i}{N} & 1 < i < N \\ 0 & i = N \end{cases}$$

- *Absolute Dispatch Compliance:* The absolute dispatch compliance P_A describes the absolute fulfillment of the dispatch list.

$$P_A = \begin{cases} 1 & i = 1 \\ 0 & otherwise \end{cases}$$

- *Average Sort Index:* The average sort index P_S describes the average lot or batch position taken from the list.

$$P_S = i$$

- *Correlation:* For evaluation purposes, the correlation P_{CORR} between the waiting time T_{LOT} of a lot and its priority P_{LOT} is an interesting measure how the priority of a lot affects its waiting time.

$$P_{CORR} = CORR(T_{LOT}, P_{LOT})$$

Figure 7 illustrates the range of values of each proposed compliance parameters.

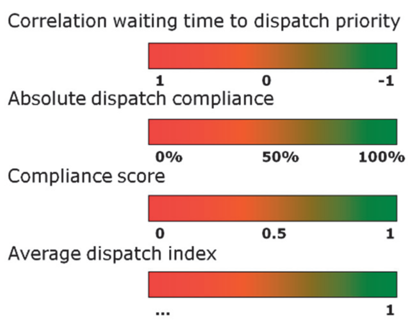
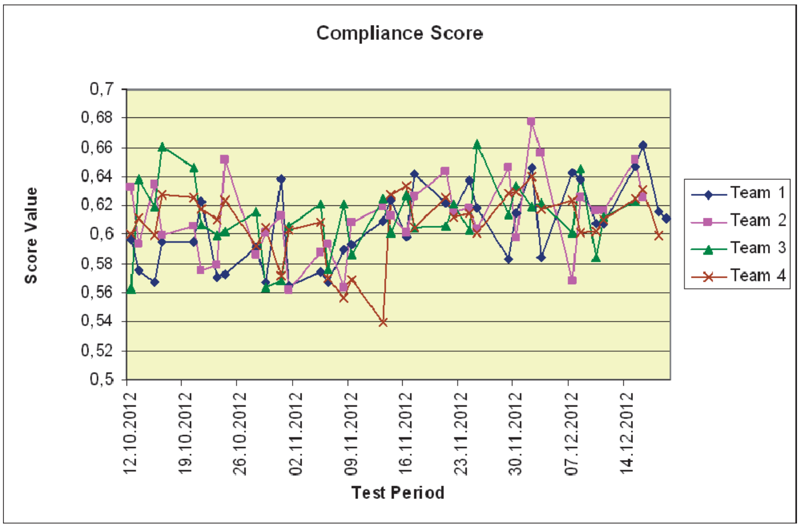


Figure 7: *Compliance parameter overview*

The compliance parameters are collected after each 12 hour work shift. In our case, four teams are available switching between night and day shift. Figure 8 illustrates the compliance score and the correlation value for the four shifts over one part of test period.

The compliance score ranges between 0.55 and 0.68 during the test period, with a positive trend to a higher dispatch compliance. The committed target for this value is P_C=0.75. Several reasons for the partial fulfillment of this target are known:

- Manual batch transport of lots between clusters: Lots are physically not available at the next operation but waiting for transport due to the manual transport system. The information of the current physical lot position is not available.
- Work organization: Several small storage buffers in front of the equipment lead to preference of FIFO processing. Operators tend to use lots in their direct environment rather than taking lots from storage places far away.
- Equipment characteristics: Cluster tools requiring special order of lots according the recipe properties. Often cluster tools are not fully represented in the MES. Information about several process areas in the tool are only partially available.



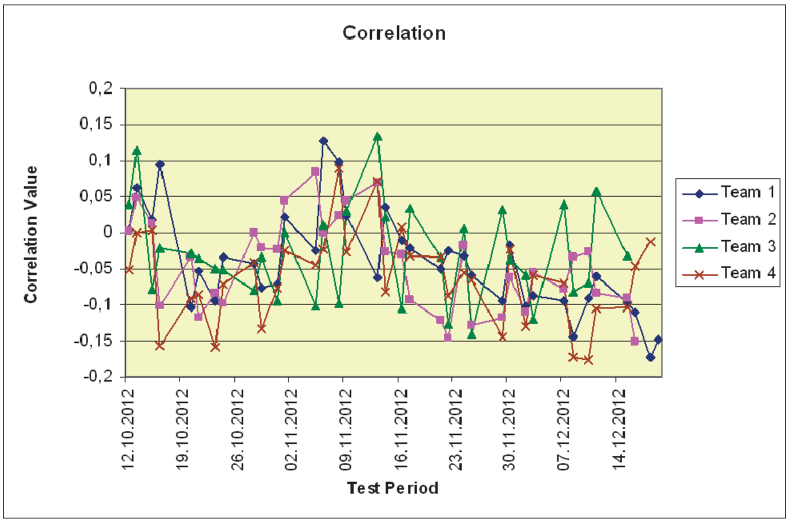


Figure 8: *Compliance score and correlation*

The correlation between the waiting time of the lot and its priority shows a small negative correlation with a negative trend. We expected a higher negative correlation. Reasons for the low negative correlation could be found at low utilized equipment where only a small amount of lots are processed. Often operators are not available at these tools. Therefore single lots with higher priority have to wait.

In general, the compliance analysis offers several improvement possibilities according the work organization and the available data. Several projects are started to increase the dispatch compliance, ranging from operator training to data source improvements.

3.4 Staff survey

After the end of the test period, a web based staff survey was conducted to obtain an overview about the personal opinions from leading staff. The questions were divided into three areas:

1. Evaluation of the current FIFO dispatching procedure
2. Opinions about improvements and required changes for the current dispatching system
3. Evaluation of the new dispatching procedure

The period for the survey was set to 14 days. The survey was completed by 12 persons. These persons included the shift leader and the dispatching personal. The following points show the main results of the survey:

- 11 persons characterize the FIFO policy and the manual priorities as not sufficient for the foundry business with high-mix low-volume characteristics. All persons rate the existing policy worse than 2 (1…perfect, 6…worst case), 5 worse than 3
- All persons rate the need of automated production control systems as very important. 11 persons see the need for a new production control approach
- 9 persons evaluate the dispatching lists as more reasonable. 8 persons see the new approach as a useful improvement. 10 persons rate the new policy better than 3 (1…perfect, 6…worst case).

4 Conclusions

In this paper, we consider the practical assessment of a combined dispatching policy for a typical high-mix low-volume ASIC facility. The combined approach shows an average improvement of the common factory performance parameters like cycle time per mask, the work in process or the on-time delivery of about 3% to 5% compared to former quarters during the test period.

The system performance estimates collected during the test period show sufficiently low time consumptions of the method classes like dispatch list generation.

The dispatch compliance analysis shows improvement potential in the area of operator training, work organization and data source management.

In general, the proposed dispatch approach became an accepted tool for a more efficient production control in our fab.

References

Bagchi, S.; Chen-Ritzo, C.; Shikalgar, S.T.; Toner, M.: A full-factory simulator as a daily decision-support tool for 300mm wafer fabrication productivity, in Proceedings of the 2008 Winter Simulation Conference, S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds., pp. 2021-2029, 2008.

Dabbas, R.M.; H.-N. Chen; J.W., Fowler; D. Shunk.: “A Combined Dispatching Criteria Approach to Scheduling SemiconductorManufacturing Systems”. Computers and Industrial Engineering 39 2001, 307–324.

Gißrau, M.; Rose, O..: “Development and Introduction of a Combined Dispatching Policy at a High-Mix Low-Volume ASIC Facility”. In Proceedings of the Winter

Simulation Conference, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher 2012.

Gupta, A.; Sivakumar, A.: Simulation based multiobjective Schedule Optimization in Semiconductor Manufacturing, in Proceedings of the 2002 Winter Simulation Conference, E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes eds., pp. 1862-1870 2002.

Mittler, M.; Schoemig, A.K.: “Comparision of DispatchingRules for Semiconductor Manufacturing Using Large Facility Models”. In Proceedings of the Winter Simulation Conference, edited by P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 709–713, 1999.

Pinedo, M.: Scheduling Theory, Algorithms, and Systems. 2nd ed. Prentice-Hall, Inc 2002.

Rose, O.: “The Shortest Processing Time First Dispatch Rule and some Variants in Semiconductor Manufacturing”. In Proceedings of the Winter Simulation Conference, edited by B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, 1220–1224., 2001.

Rose, O.: “Some Issues of the Critical Ratio Dispatch Rule in Semiconductor Manufacturing”. In Proceedings of the Winter Simulation Conference, edited by E. Y¨ ucesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, 1401–1405, 2002.

Rose, O.: “Accelerating Products under Due Date Oriented Dispatching Rules in Semiconductor Manufacturing”. In Proceedings of the Winter Simulation Conference, edited by S. Chick, P. J. Snchez, D. Ferrin, and D. J. Morrice, 1346 – 1350, 2002.