Integrationsaspekte der Simulation:
Technik, Organisation und Personal
Gert Zülch & Patricia Stock (Hrsg.)
Karlsruhe, KIT Scientific Publishing 2010

A WIP Balance Study from Viewpoint of Tool Group in a Wafer Fab

Studie zur WIP-Balancierung von Lastverbünden in der Halbleiterfertigung

Zhugen Zhou, Oliver Rose
Dresden University of Technology, Dresden (Germany)
zhugen.zhou@tu-dresden.de, oliver.rose@tu-dresden.de

Abstract: In a wafer fab, work-in-process (WIP) fluctuation occurs quite often due to tool random failure, setup change, batch processing etc. In this paper, a pull/push concept is proposed in order to balance WIP from viewpoint of tool group. We call it pull and push because on one hand a WIP Control Table is used to measure the pull request of downstream tool group dynamically; On the other hand among the lots which fulfill demand from the downstream tool group with pull request, the upstream tool group tries to schedule an optimal lot considering lot status and local tool group constraint. The simulation results demonstrate that the proposed pull/push concept is superior over First-in-First-out (FIFO) and Minimum Inventory Variability Scheduling (MIVS) with regard to average cycle time as well as cycle time variance under a high fab loading.

1 Introduction

In a wafer fab, material flow is generally nonlinear due to different events such as unpredictable tool failure, setup change, batch processing and so on. ROSE (2003, pp. 1401) pointed out that WIP increases in the first critical tool group in the line. When applying Critical Ratio (CR) dispatching with a tight due date under a high fab loading, because old lots are assigned high priority to be speeded up, more and more fresh lots in the first processing steps are passed by old lots and only waiting in queue. When a critical tool group has a breakdown, the situation gets worse. Consequently, fresh lots pile up in front of the critical tool groups with breakdown, which causes line imbalance and has great impact on cycle time as well as on time delivery. From process operation viewpoint, this WIP imbalance taking place in critical tool groups is caused by WIP imbalance of operations. Old lots are accelerated while fresh lots are waiting, which means the WIP of early operations are accumulated, in contrast, WIP of late operations are always processed.

Many WIP balance approaches have been already developed to correct for WIP imbalance. SPEARMAN and ZAZANIS (1992, pp. 521) presented a detailed discussion about lot release approach in pull production system like Kanban and CONWIP (CONstant Work In Process), and concluded that there is less WIP congestion improvement of pull system than push system. DABBAS and FOWLER (2003, pp. 501) proposed a global Line Balance algorithm combined with local single dispatching rules to minimize the deviation of actual WIP to target WIP of operation. Minimum Inventory Variability Scheduling (MIVS; cf. LI, TANG, COLLINS 1996, pp. 1) is a representative approach to balance the WIP of operations. MIVS considers both upstream operation and downstream operation, and tries to keep the WIP of each operation close to the average target WIP level. It gives the highest priority to an operation which has a high WIP and its downstream operation has a low WIP, in order to avoid starvation at downstream operation. In contrast, it gives the lowest priority to an operation which has a low WIP and its downstream operation has a high WIP. Table 1 describes the principle of MIVS.

		Downstream Operation	
		Actual WIP < Target WIP	***Actual WIP >= Target WIP***
Upstream Operation	***Actual WIP >= Target WIP***	Priority 1	Priority 2
	Actual WIP < Target WIP	Priority 3	Priority 4

Table 1: *MIVS principle*

MIVS succeeds in reducing line imbalance like CR performs, through pulling WIP into a low WIP operation, thus reducing WIP variability and cycle time. However, there are several potential problems of MIVS. First of all, although MIVS is able to keep the dynamic WIP close to the target WIP in long-term view, the tool group which is shared by different operations may have a high WIP and be overloaded in a short-term period. If this tool group has a breakdown, lots have to experience longer queue time, which leads to starvation of other tool groups. Secondary, MIVS considers WIP balance in each operation, which indirectly speeds up lots if they are late. It tells us that MIVS only takes lot status into account, and falls short of considering local tool group constraints, e.g. workload, utilization, failure, setup and batching requirement, when making decision which lot is suitable for processing.

In our opinion, considering local tool group constraint is also important when the fab is running under high capacity loading. Because there are hundreds of tool groups in a wafer fab, on one hand they subject to random failure, preventive maintenance, on the other hand appropriate scheduling to meet setup and batching requirement results in reduction in total setup time and capacity loss. What if watching WIP balance from tool group point of view instead of from operation viewpoint? This study is motivated by not only considering lot status like MIVS, but also considering local tool group constraint. Like traditional pull system, we try to pull WIP to a low WIP tool group based on the downstream tool group status, e.g. high WIP or low WIP, in order to balance WIP in tool groups. After determining which downstream tool group has a pull request, the upstream tool group tries to schedule an optimal lot based on lot status and local tool constraint among lots suitable for the downstream tool group with pull request, which performs like a push system with scheduling work based on demand. Therefore, we propose this pull/push concept in order to take advantage of pull while remaining make-to-order character of push.

This paper is organized as follows. In Chapter 2, we describe the proposed pull/push concept in detail. In Chapter 3, we present and compare the simulation results with MIVS and First-in-First-out (FIFO) with regard to average cycle time, cycle time variance. Chapter 4 gives conclusion and further work.

2 Pull/Push Concept for WIP Balance of Tool Group

2.1 WIP Control Table

To deal with WIP imbalance of tool group, a WIP Control Table is proposed for each tool group. Each upstream tool group maintains a WIP Control Table which contains current WIP information of all its downstream tool groups, e.g. target WIP level, actual WIP level, WIP difference and utilization of downstream tool groups. Figure 1 describes an example of WIP Control Table.

Tool Group	*Target WIP (lot)*	*Actual WIP (lot)*	*WIP Difference (%)*	*Utilization (%)*
3	12	8	-33.3	90
4	16	10	-37.5	70
5	10	14	40	50

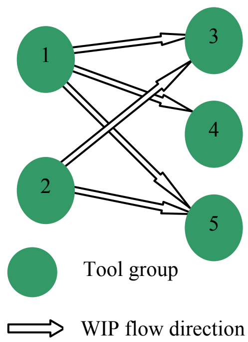

Figure 1: *WIP control table of tool group 1*

The terms in WIP Control Table are described as follows:

- **Target WIP**: the desired WIP level of tool group wanted to be maintained. It is based on simulation model, historical data or engineer's experience.
- **Actual WIP**: the current WIP level of tool group including lots in queue and in process.
- **WIP Difference**: the deviation of actual WIP to target WIP.

$$WIP\ Difference = \frac{Actual\ WIP - Target\ WIP}{Target\ WIP} \tag{1}$$

- **Utilization**: tool group utilization from lot release to current time.
- The Actual WIP, WIP Difference and Utilization will be updated in case of lot move in/out and tool status change (on/off line) of tool group 1.

The objectives of WIP Control Table are:

1. Measuring pull request of downstream tool group;
2. Minimizing deviation of actual WIP to target WIP of downstream tool group.

In case of lot move in/out and tool status change (on/off line), the WIP Control Table will be updated. Through comparison between actual WIP and target WIP, the upstream tool group discovers which downstream tool group runs out of WIP and which downstream tool group has a high WIP dynamically.

2.2 Measure Pull Request of Downstream Tool Group

The following are two steps to distinguish pull request of tool group:

Step 1: For each downstream tool group, if WIP Difference is less than Delta (Delta is predefined negative value, which represents this tool group is running short of WIP), this downstream tool group has a pull request;

Step 2: Compare the utilizations of tool groups which have pull request from Step 1, the one with higher utilization has a stronger pull request. Rank the tool groups with pull request from top to bottom.

2.3 Proportion of WIP Required to be Pushed

After determining which downstream tool group runs short of WIP and has a stronger pull request, the upstream tool group has to push lots to balance the WIP of this downstream tool group. Using the WIP Control Table, the upstream tool group identifies the proportion of WIP needed to be pushed to the downstream tool group.

$$Proportion\ \ of\ WIP = ceiling(\frac{|\,Actual\ \ WIP - T\arg et\ WIP\,|}{Numbers\ \ of\ Upstream\ \ Tool\ Groups}) \qquad (2)$$

Where actual WIP and target WIP are described in the WIP Control Table above, numbers of upstream tool groups is the quantity of upstream tool groups which the downstream tool group has.

In this study, this proportion of WIP ignores capacity availability in the upstream tool groups. We just assume the upstream tool group theoretically has to push average quantity of lots to its downstream tool group, no matter whether the upstream tool group has lots in queue or not.

2.4 Push Optimal Lot to Downstream Tool Group

As we mentioned in Chapter 2.1, the WIP Control Table is only used to measure the pull request of tool group. In order to optimize other wafer fab's performance indicators such as average cycle time and cycle time variance, local tool group constraint, e.g. batch processing, setup change, failure, and lot status, e.g. lot due date, operation time boundary have to be taken into account.

- A lot needs to be accelerated if it is delayed.

- If the time spent by a lot in queue exceeds time boundary, the lot needs to be accelerated. This could avoid:
    1. Upstream tool group processes a lot for a particular high utilized downstream tool group as soon as a lot for this particular downstream tool group arrives. While other lots for the low utilized downstream tool groups could not be processed but only waiting remarkable time in queue;
    2. High variability of time lot spends in operations.
- Try to make batch as full as possible, which can reduce total setup time and capacity loss.

Lots which belong to these 3 categories will be assigned a high priority and pushed to downstream tool group with pull request as soon as possible. In case upstream tool group can no find out lots which fulfill those 3 requirements above. It will push the lot of minimum operation due date to downstream tool group with pull request.

3 Simulation Results and Performance Analysis

3.1 Simulation Model

A wafer fab dataset MIMAC6 from Measurement and Improvement of MAnufacturing Capacities (MIMAC; cf. FOWLER, ROBINSON 1995) is used to test the proposed pull/push concept. MIMAC6 is a typical 200 mm wafer fab model including:

- 9 products, 9 process flows, maximum 355 process steps. 24 wafers in a lot. 2777 lots are released per year under fab loading of 100 %.
- 104 tool groups, 228 tools. 46 single processing tool groups, 58 batching processing tool groups.
- Setup avoidance, rework, MTTR (mean time to repair) and MTBF (mean time between failures) of tool group.

The simulation experiments are carried out with Factory eXplorer (FX) from WWK. The proposed method is not provided by the FX simulation package, but FX supports customization via a set of user-supplied code and dispatch rules. We used a customized FX interface developed by WOLF (2008) to implement the pull/push concept.

3.2 Simulation Results

The simulation length of MIMAC6 was carried out for 18 months. The first 6 months were considered as warm-up periods, and not taken into account for statistic. We simulated the fab with 3 dispatching rules: the pull/push concept, MIVS and FIFO, under 5 different fab loadings of 70 %, 75 %, 80 %, 85 % and 95 %.

At first we consider average cycle time, cycle time variance and cycle time upper pctile 95 % as main performance measures. Table 2 lists the simulation results. For the 75 % and 80 % fab loading case, there is no significant cycle time difference between MIVS and pull/push concept, both of which are better than in the case of FIFO. Moreover, the cycle time variance and cycle time upper pctile 95 % of MIVS are superior over the cases of pull/push concept and FIFO. It seems that using MIVS dispatching is better than the pull/push concept in a low fab loading. Because most of the tool groups have a low WIP, which implies the materials flow quite smoothly through the tool groups and WIP does not need to be balanced in tool group. In this case, considering WIP balance from viewpoint of operation is more efficient, and directly contributes average cycle time and cycle time variance improvement. However, if the fab is running under a high loading, more and more lots enter the fab, due to tools random failure, the situation becomes more complex. Lot piling up in front of the critical tool groups is relatively high. In this case, only considering lot status like MIVS is not enough. As soon as a high WIP taking place in one downstream tool group, the upstream tool group has to stop feeding lots to this high WIP downstream tool group and delivers lots to other low WIP downstream tool groups in time, which can avoid a longer queue time in the high WIP downstream tool group and starvation in the low WIP downstream tool group. As we can see the results, the pull/push concept outperforms MIVS and FIFO in the 85%, 90% and 95 % fab loading cases, with regard to these 3 performance measures. Especially for the 95 % fab loading case, the pull/push concept gets 0.5 and 1.6 days average cycle time improvement compared with MIVS and FIFO cases respectively. Further more, the pull/push concept achieves a better cycle time variance compared with in the case of MIVS and FIFO, which can lead to a better prediction of product completion time.

	Fab loading (%)	*75*	*80*	*85*	*90*	*95*
Avg. Cycle Time (days)	***Pull/Push***	20.3	21.5	22.9	24.8	28.0
	MIVS	20.3	21.4	23.1	25.2	28.5
	FIFO	20.5	21.6	23.3	25.7	29.6
Cycle Time Variance (days^2)	***Pull/Push***	0.97	1.01	1.09	1.18	1.28
	MIVS	0.87	0.98	1.15	1.32	1.71
	FIFO	0.91	1.18	1.20	1.38	1.65
Cycle Time Upper Pctile 95 % (days)	***Pull/Push***	25.4	26.8	28.2	31.8	35.4
	MIVS	24.8	26.5	28.8	32.6	37.2
	FIFO	26.2	27.6	29.8	33.1	39.1

Table 2*: 3 Performance measures of MIMAC6 under different fab loading*

Secondary, we take a close look at tool group behavior under fab loading of 95 %. Table 3 lists the cycle time contribution by top 10 tool groups of product 38090964_B5C. For these 3 different dispatching rules, the furnace tool group 11026_ASM_B2 contributes most cycle time to this product. For FIFO case, 16.5 %

of cycle time is spent in this tool group, which means the line is imbalance. The MIVS is more balanced than FIFO, because this critical tool contributes less cycle time. When the pull/push concept is applied, this critical tool group contributes 3.13 days which is one day less than MIVS. Look at other major contributors, although they contribute a little more compared with MIVS and FIFO cases, the pull/push concept successfully avoids a high WIP occurrence in 11026_ASM_B2 and shifts certain amount of WIP from 11026_ASM_B2 to other tool groups. Therefore, lots do not experience a huge queue time in 11026_ASM_B2. In another word, the pull/push concept balances the WIP among different tool groups, and the line is more balanced than MIVS.

Pull/Push			***MIVS***			***FIFO***		
TG	***CTC (days)***	***PoT (%)***	***TG***	***CTC (days)***	***PoT (%)***	***TG***	***CTC (days)***	***PoT (%)***
11026_ASM_B2	3.13	8.9	11026_ASM_B2	4.17	11.6	11026_ASM_B2	6.54	16.5
20540_CAN_0.43_MII	2.84	8.1	20540_CAN_0.43_MII	2.65	7.3	20540_CAN_0.43_MII	2.76	6.9
12553_POSI_GP	2.15	6.1	12553_POSI_GP	2.16	6.0	12553_POSI_GP	2.46	6.2
13024_AME_4+5+7+8	1.53	4.3	13024_AME_4+5+7+8	1.47	4.1	13024_AME_4+5+7+8	1.52	3.8
15121_LTS_3	1.32	3.7	17421_HOTIN	1.29	3.6	11024_ASM_A4_G3_G4	1.32	3.3
17421_HOTIN	1.26	3.5	15121_LTS_3	1.21	3.3	15121_LTS_3	1.28	3.2
11024_ASM_A4_G3_G4	1.23	3.5	11024_ASM_A4_G3_G4	1.20	3.3	17421_HOTIN	1.26	3.2
16221_IMP - MC _1+2	1.15	3.2	16221_IMP - MC _1+2	1.05	2.9	11027_ASM_B3_B4_D4	1.06	2.6
15627_HIT_S6000	1.04	2.9	11027_ASM_B3_B4_D4	1.02	2.8	16221_IMP - MC _1+2	1.06	2.6
17221_K_SMU236	0.99	2.8	15627_HIT_S6000	1	2;7	15627_HIT_S6000	1	2.5
Where TG stands for tool group, CTC represents cycle time contribution, PoT means percent of total.								

Table 3: *Cycle time contribution by top 10 tool groups of product 38090964_B5C*

4 Conclusion and Further Work

In this study, we proposed a pull/push concept for WIP balance from viewpoint of tool group in a wafer fab. This was accomplished by the WIP Control Table which contains current WIP information of downstream tool groups. Through comparison between the actual WIP and target WIP, the upstream tool group was able to measure the pull request of the downstream tool groups. Apart from that, local tool group constraints and lot status were taken into consideration to push an optimal lot to the downstream tool group with pull request, in order to optimize average cycle time and cycle time variance. The simulation results demonstrated:

1. Under low fab loading case, the pull/push concept is not superior over MIVS, because material flows smoothly through tool groups. In this case, considering WIP balance from viewpoint of operation can directly contribute average cycle time and cycle time variance improvement;
2. Under high fab loading case, the pull/push concept has advantage of shifting WIP among different tool groups, so as to reduce queue time in some critical tool groups, especially during tool random failure period. Therefore, the pull/push concept outperforms MIVS in this case. It tells us that only considering lot status is not enough under high fab loading, local tool group constraint is also extremely important.

Our next step is to apply this pull/push concept in a real wafer fab model, to figure out whether the pull/push concept has good or bad impact on low volume products.

5 Acknowledgements

This work was supported by Infineon Technologies AG, Dresden, Germany.

References

DABBAS, Russ M.; FOWLER, John: A New Scheduling Approach Using Combined Dispatching Criteria in Wafer Fab. In: IEEE Transaction on Semiconductor Manufacturing, New York, 16(2003)3, S. 501-510.

FOWLER, John; ROBINSON, Jennifer: Measurement and Improvement of Manufacturing Capacities (MIMAC): Final Report. Technical Report 95062861A-TR, Austin, TX: SEMATECH, 1995.

LI Shu; TANG, Tom; COLLINS, Donald W.: Minimum Inventory Variability Schedule with Application in Semiconductor Fabrication. In: IEEE Transactions on Semiconductor Manufacturing, New York, 9(1996)1, S. 145-149.

ROSE, Oliver: Some Issues of the Critical Ratio Dispatch Rule in Semiconductor Manufacturing. In: Proceeding of the 2002 Winter Simulation Conference, Eds: YÜCESAN, E.; CHEN, C.-H.; SNOWDON, J. L.; CHARNES, J. M. New York: Association for Computing Machinery et al., 2002, S. 1401-1405.

SPEARMAN, Mark L; ZAZANIS, Michael A.: Push and Pull Production Systems: Issue and Comparisons. In: Operations Research, 40(1992)3, S. 521-532.

WOLF, René: Entwicklung einer Steuerungsschnittstelle für den Simulator Factory Explorer einschließlich ausführlichem Test am Beispiel der Abfertigungsregel "Operation Due Date (ODD)". Dresden: Computer Science Department, Dresden University of Technology, Master Thesis, 2008.