ON THE USE OF SIMULATION IN SUPPORT OF CAPITAL UTILIZATION

Adar Kalir
Dean Grosbard

Fab/Sort Manufacturing Division
Intel Corporation
2 HaZoran ST., Qiriat-Gat 82109, ISRAEL

ABSTRACT

This paper provides a modeling approach for dealing with the challenging question of trade-off between capital utilization and production efficiency in semiconductor manufacturing where the ultimate goal is of maximum output at maximum velocity and minimum cost. Full fab simulation is used iteratively between models of "horizontal" and "vertical" simulations in order to rapidly generate results for different possible states of the fab with varying capital costs, matched cycle time (CT), and fixed throughput rate, so as to determine the most efficient operating condition for the fab with respect to cost, CT, and output.

1   INTRODUCTION

The semiconductor manufacturing environment is a unique one. No other manufacturing system has similar properties. It is characterized by an inherent high variability in its process, emanating from equipment availability and a highly re-entrant process flow, along with a very long process typically containing hundreds of sequential process steps performed by shared toolsets subject to unique restrictions such as step dedications, dynamic lock-outs of tools to steps, queue time (also known as process time window) limitations, etc.

   The primary objectives of a high volume semiconductor manufacturing fab are to maximize output with minimal production cycle times (or maximum velocity) and at a minimum wafer cost. These objectives are inherently in conflict with each other. In order to maximize output, more capital which means more equipment is needed, coupled with high equipment utilization. On the other hand, more capital also implies higher wafer cost and high equipment utilization implies higher production cycle times (or lower production line velocity and efficiency). To effectively deal with these conflicting objectives, fabs are challenged hierarchically to first meet required output per demand for a given capital -- both are set at the capacity planning phase -- and then minimize production cycle times. However, in the life cycle of a fab, equipment productivity improvements are realized on the toolsets and the operating curve of the factory changes. What may have been once potentially true at the capacity planning phase no longer holds during the high volume production and maturity phase, not to mention future transitions between technologies and product derivatives. Therefore, the fab has an opportunity to re-position itself at different operating conditions, defined by desired output, velocity, and capital. An illustration of fab CT as a function of Capital and Wafer Starts (or throughput rate in steady-state) is depicted in Figure 1.

   A significant amount of research over the years addresses aspects of the problem described thus far, i.e. of finding the optimal operating condition of a factory. For example, in order to find the most efficient output of the factory, the true constraint(s) has (have) to be identified and exploited. But even the question of identifying the true constraint(s) in a semiconductor manufacturing system is non-trivial, as evident by the extensive literature on this. Moreover, estimating the production cycle time for a given equipment-base with known equipment performance is another challenging area of research with

proposed approximations, all somewhat limited in their accuracy and/or apply only to small scale problems. As noted by Dauzère-Pérès in Chien et al (2008), although semiconductor manufacturing has

been the subject of numerous publications, most of them tackle problems at one decision level, e.g. overall capacity planning or scheduling in various production areas (photolithography, diffusion, etc.) there are many challenges that still arise when considering the integration of two or more decision levels.

In this paper, we propose a simulation as a means for realizing the trade-offs between the conflicting objectives of output, velocity, and wafer cost in an attempt to assist in decision making on the optimal operating condition of a fab throughout its life cycle. We show that via simulation, trade-offs can be outlined by quantifying the marginal gains/losses from each change (i.e. tool reductions; or velocity traded for output and vice versa.)

The paper is organized as follows. In section 2, related literature is reviewed. Survey of the existing literature includes research on capacity planning; capital utilization; bottlenecks and output maximization; production cycle time; and full fab simulation modeling; all in the context of semiconductor manufacturing. Section 3 contains more details about the problem at hand, a description of the simulation in use, and the proposed simulation modeling approach, followed by simulation results and insights in section 4. Concluding remarks are presented in section 5.

2 LITERATURE SURVEY

At the initial phase of designing a new fab, strategic capacity planning is typically employed to address critical questions related to expected fab performance, such as the required capital investment cost; the output at peak loading; the flexibility to transition between products when demand changes; and so forth. Clearly, the development of accurate capacity planning models is of paramount importance in the capital intensive semiconductor industry, where the ability to understand manufacturing capacity requirements can mean the difference between profit and loss ─ as pointed out by Witte (1996). Furthermore, equipment depreciation is by far the largest component in wafer cost (Guo et. al 1999).

Geng and Jiang (2009) review existing solutions and methodologies to capacity planning. They refer to the fact that strategic capacity planning involves trade-offs among finance, throughput, and cycle time but note that the research over the last fifteen years has addressed these aspects separately for the most part. Since capital utilization is an aspect of great importance in this industry due to the fact that huge capital investments are required to build new fabs, investments in fabs are typically related to the primary management goal which is to maximize output (Ovacik and Weng 1995).

Capital utilization and capital management is discussed by Haney and Canter (2005). There, the capital management process is automated via a mathematical optimization model that has been developed to optimize re-assignment of equipment, in order to provide the best possible match of configuration, release date, and need date during technology transitions in the life cycle of a chain of fabs. However, aspects of trade-offs with output and cycle time are ignored. Leu and Yu (2009) further stress the fact that the semiconductor manufacturing industry is a capital-intensive industry which involves tremendous amount of capital investment in production machines as well as in wafer fab construction and therefore an over-investment in fab capacity may cause productivity loss, whereas an under-investment case may lead to a loss of revenue opportunity. They propose a simulation-based capacity optimization model in which both the production output and the capital investment are considered. By this way, the proposed model intends to find an optimal ratio of wafer output to capital investment so that the "production economic" (i.e. capital utilization) can be attained. However, in their work, the simulation model is based on the X-factor theory, implying that the production dynamics is ignored or at the very least assumed static, unlike in this paper.

As far as bottleneck and output maximization, the semiconductor highly re-entrant process poses a real challenge on identifying the true constraint because the true constraint may be virtual, as explained by Morrison and Juen (2008). Therefore, methods have been developed to in an attempt to identify and fully exploit the constraint in order to maximize output. Several methods have been proposed over the past decade. For example, Chung et al (2008) apply data envelopment analysis (DEA) to find a set of product family mix that is efficient for the fab to produce at a given period. To ensure long-term

effectiveness in productivity and in profit gaining, window analysis is adopted to seek the most efficient set of product family mixes for manufacturing efficiency over time.

Fayed and Dunnigan (2007) take a practical approach to manage the trade-off between throughput and cycle-time in order to maximize fab performance. The concept of the operating curve (OC) is deployed. It is suggested that OC can be constructed using simulation or queuing analysis. However in order to drive improvements and to properly grade fabs, they have used interviews with experts, literature and queuing analysis to establish the OC framework. This framework is then utilized to drive continuous improvements in a real fab.

With respect to production cycle time estimation and prediction, numerous approaches have been proposed. One unique approach by Chen et al (2009) involves using a fuzzy back propagation network (FBPN) approach with multiple buckets and partial normalization for lot cycle time prediction in a ramping-up semiconductor manufacturing factory. The proposed methodology is composed of two parts. In the first part, the multiple-bucket approach is applied to consider the ramp-up plan of the semiconductor manufacturing factory. Subsequently, the FBPN approach is applied to predict the cycle time of every lot in the factory.

In Section 3, we provide a brief review of prior uses of simulation for full fab modeling in semiconductor manufacturing.
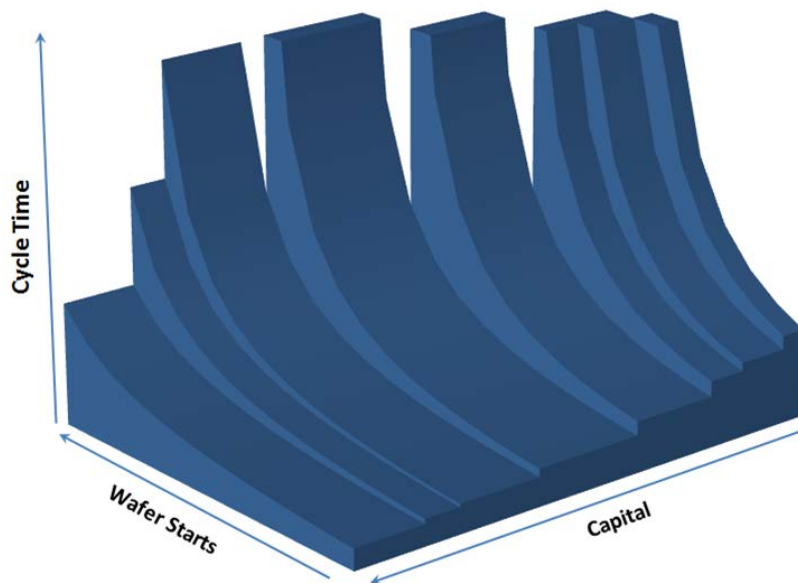

Figure 1: Illustration of Fab CT as a Function of Capital and Wafer Starts

3   PROBLEM DESCRIPTION AND THE SIMULATION MODELING APPROACH

In this work, we use a full factory simulation model, constructed using the AutoSched AP software from AMAT, enhanced by multiple extensions developed by Intel's Decision Support Technology group and described in DeJong and Fischbein (2002). The model in use represents fab operation in detail with regards to fab capacity and contains a full description of each of the tool-sets used in the process including tool downtimes (scheduled and unscheduled), setups, batching, dispatching rules as well as auxiliary resources such as reticles. The model also incorporates general operational elements such as starts plans, hot lots, work in process, tool dedications, rework, test wafer production and time constrained production sequences. The fab's automated material handling system (AMHS) is addressed only be means of the delay distributions of lots moving between different locations in the fab. This approach of extracting the AMHS travel distributions in separate from the capacity simulation model is similar to methods described by DeJong and Fischbein (2002) and Kong (2007), but actual fab data is

used to derive travel time distributions rather than a parallel AMHS simulation model. Extensive analysis of transportation time of AMHS vehicles show that transportation times of the fab are predictable and therefore the use of these distributions does not reduce from the simulations modeling capability, as described in Kiba et al (2009). Following the classification method in Jimenez et al (2008), the model may be classified under the third quadrant with a high level of capacity detail (level 5-6 per their classification) and a low level of AMHS detail (level B per their classification).

The AutoSched AP time phased parameter capability is used to model continuous planned run rate improvement and tool reduction projects. Enabling changes during the simulation run, along with a detailed work-in-process profile, allows to provide a detailed estimate of the cycle time impact of tool reduction projects within the relevant time-frame and under various downtime scenarios. The steady state mean and variance of cycle time after implementation of the projects are estimated by extending the simulation runtime far beyond the relevant time period.

The effects of changes implemented on a single tool-set in a fab are not limited to that specific tool-set and can have an impact on the cycle time (CT) of other tool-sets throughout the production flow. Therefore, if we are to assess whether a project is beneficial by comparing the capital expenditure or avoidance with an expected change in CT, it is necessary to analyze the CT impact on a full fab model and not only with respect to the tool-set itself.

The mechanisms that relate the performance of a single tool-set to those of other tool-sets in the fab are complex and often counter intuitive, making it very difficult to even predict if a change on a given tool-set would increase or decrease the CT of another tool-set. The re-entrant nature of the fab production flow coupled with the complex service policies implemented at the tool-sets cause standard analytical models such as queuing networks and fluid networks to yield poor estimates ,Shanthikumar (2007). However, improved queueing network models specifically designed to manage fab complexities have been used successfully, notably the "EPOS" system presented by IBM, Zisgen et al (2008), Brown et al (2010).

Full fab simulation analysis performed on the Intel "test-bed" to assess the CT impact of potential capital saving projects has provided insight into the relations between the performance of various tool-sets in the fab. Though some of the relations found can strictly be attributed to service polices and operational constraints restricted to a specific fab, these types of interactions tend to affect only a small subset of tool-sets or specific segments in the production while those relations that are the result of the impact that tool-sets have on the WIP flow variability apply to a broader context of re-entrant production lines with a substantial impact on CT.

4   SIMULATION RESULTS AND INSIGHTS FOR TOOL REDUCTION

In this section, we present three cases for which a tool reduction project is simulated, each having a different impact on fab WIP flow variability and subsequently on fab CT. We discuss the mechanisms behind these relations, namely the mechanisms known as "variance generating limiter", "variance canceling limiter", and "mixed variance impact limiter". Finally, we present results of full fab simulation analysis with a fixed throughput rate and varying capital costs to show the trade-offs between cost and CT.

4.1   Variance Generating Limiter

The less tools there are in a tool-set, the greater is the capacity loss or gain when a tool goes down or is brought back up to production. Though a small fleet tool-set may have excess capacity when all the tools are available (N state capacity), it is the gap between the tool-set's required capacity and the tool-set's actual capacity when one of the tools is down (N-1 state capacity) that has the high impact on CT and WIP flow variability. For a tool-set having both substantial N-1 state capacity gap (<-10%) and substantial N state excess capacity (>+10%), each downtime of a single tool has a potential of creating a WIP bubble owing to capacity shortage. This WIP bubble can quickly be moved forward when the tool becomes available again due to the tool-set's N state excess capacity. This situation of course becomes

more extreme when the tool goes down to an N-n state with $1<n<N-1$. This rapid creation and elimination of WIP bubbles creates an increase in the variability of arrival rate to downstream operations in the production flow.

Though the formula given by Hopp and Spearman(2001) for the squared coefficient of variance (SCV) of a departure process from a queue, $Cd^2 = \rho^2 Cs^2 + (1-\rho^2)Ca^2$ predicts that this variance would dissipate rapidly, since tool-sets in the fab tend to have relatively high utilization rates, we find that, as described in Suresh and Whitt (1990) and Whitt (1995), tool-sets with higher utilization rates show an increase in CT regardless of their location. Apparently the impact of the variance in arrival rate that is created by a tool-set can continue to have an impact on tool-sets much further downstream in the production line. Therefore, even if an increase in the tool-sets' CT is not enough to justify the capital of keeping the excess tool, the decision may change when considering the CT impact on the entire fab.

Results of a full fab simulation analysis show the CT impact of the reduction of a single tool from a tool-set having 5 tools, leaving it with 4 tools and a capacity loss of 25% each time one tool goes down. In Figure 2, it can be observed that the WIP bubble is created by the variability in the tool-set capacity over time while in Figure 3 the overall impact the 9 most limiting downstream tool-sets and on the entire production line is depicted.
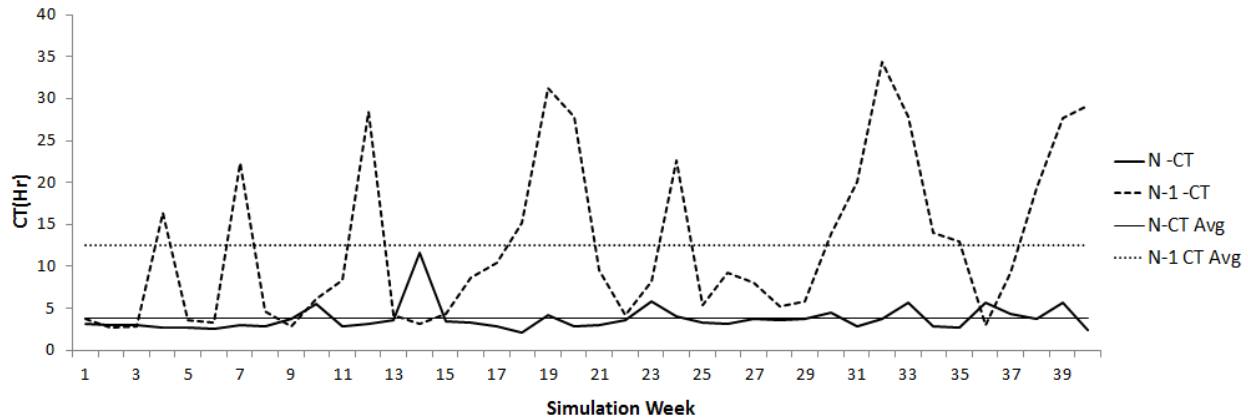

Figure 2: Tool-set Inventory by Time Period With N vs. N-1 Tools

4.2   Variance Canceling Limiter

When a tool-set contains many identical tools with multiple similar operations across the production flow, a reduction of a tool from such tool-set may have an inverse impact on the fab CT. Having many tools within a tool-set normally means that there is very little loss of capacity when a tool goes down and consequently there is much less variability in the tool-set's capacity over time. This is also true when considering preventive maintenance (PM), which is scheduled gradually across the tools to prevent variability of availability. Furthermore, tool-sets with multiple operations and similar process times, spread throughout the production flow, would tend to have relatively low variability in the work arrival rate, despite the fact that arrival rates to individual operations in the production flow are dependent.
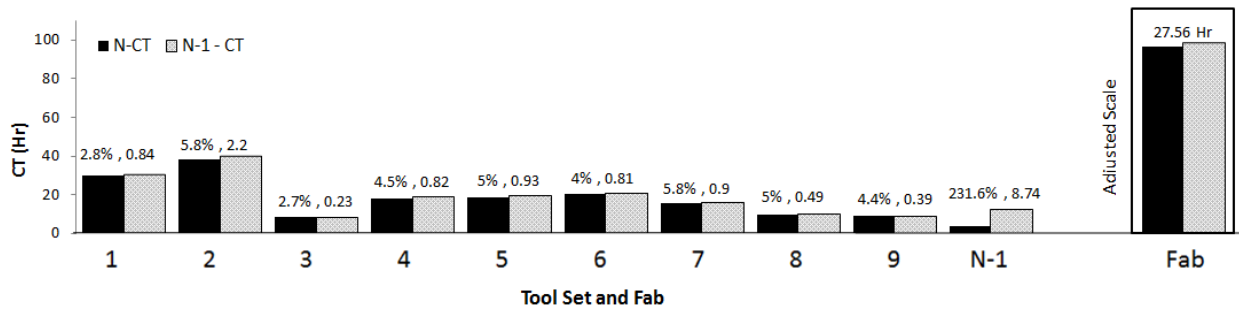
Figure 3: Tool-set CT for high utilization downstream tool-sets with and Fab CT

If such a tool has excess capacity, it tends to be transparent with almost no impact on the variability of WIP flow in the fab, as predicted by the equation for the SCV of inter-departure times from a queue with infinite servers $Cd^2 = Ca^2$. However, when a tool is detracted from the tool-set, the increase in utilization of the tool-set may cause an accumulation of inventory and subsequently an increase in CT. In such cases the relatively low variability of capacity together with the increased utilization tend to fulfill the prediction of $Cd^2 = \rho^2 Cs^2 + (1 - \rho^2)Ca^2$, assuming that $Cs^2 << Ca^2$ due to multiple servers, and cause a decrease in the variability of arrivals to other operations. Furthermore, the low variability allows production planners to manage the tool-set so that WIP bubbles arriving to one of the operations can be released gradually without leaving the tools idle and causing a loss of capacity. This ability to control the movement of WIP bubbles across the production flow causes a decrease in WIP flow variability and thus, even though the a tool-set may show a major increase in CT, there is an accompanying decrease in the CT of other tool-sets due to the reduction in arrival rate variability. Contrary to the example in the previous section, though the increase in CT caused by the reduction of a tool can be substantial, the overall change in fab CT may be small enough to justify the savings in capital. In Figures 3 and 4, the accumulation of inventory and CT owing to a reduction of a single tool is shown. While it does cause an increase in the CT of the tool-set, it is accompanied by a decrease in the CT of other tool-sets, resulting in only a small overall CT increase.
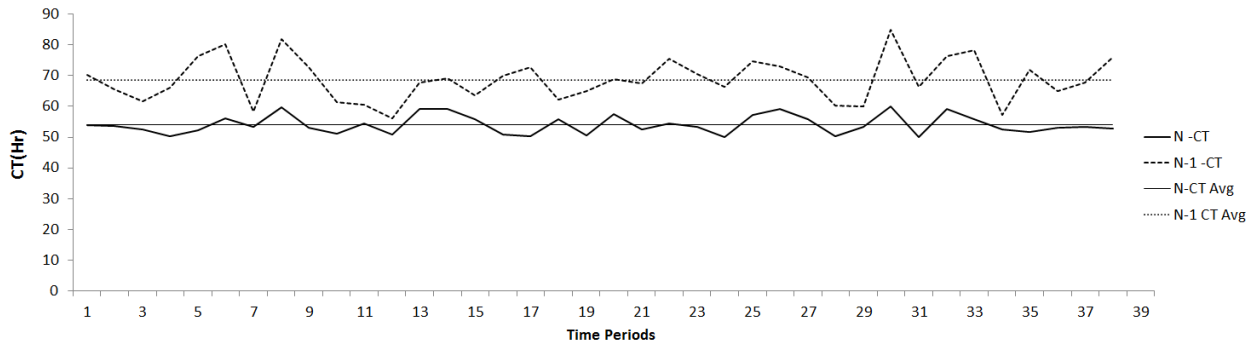

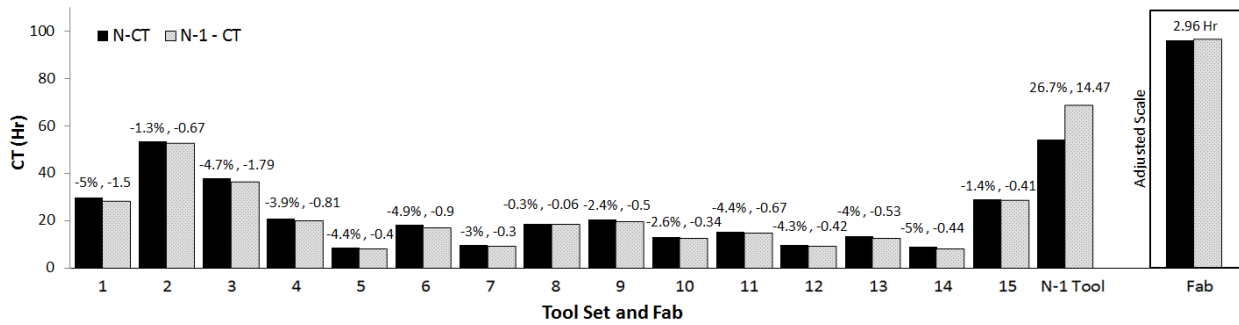Figure 4: Tool-set CT with N vs. N-1 Tools

Figure 5: Tool-set CT for high utilization downstream tool-sets with and Fab CT

4.3 Mixed Variance Impact Limiter

For most tool-sets in the fab capacity of a tool depends not only on the processing times and manufacturing availability but also on the ability to avoid setups and allow overlap between the processing of lots. In many cases, in order to reduce setups and increase the overlap, a tool needs to process lots requiring similar recipes in sequence, and the longer the average sequence length of the tool the higher the tool output rate becomes. The creation of longer sequences, often referred to as cascades, depends on the amount of available lots to process from any specific recipe. While in some cases the tool may be left idle until a sufficiently large cascade if formed in order to reduce setup times, for most tool-sets an idle tool would not wait for a larger cascade. Although the grouping of similar recipe lots throughout the production flow is addressed within the fab dispatching rules, the mean cascade size is still mainly determined by the mean inventory at a tool-set relative to the number of tools and the distribution of that inventory among the different recipes. This subsequent positive correlation between the inventory levels and output rate creates an equilibrium inventory level at which the output rate allows the tool-sets throughput capacity to equal the incoming workload. In Figure 5, we present an example of the equilibrium inventory levels and mean cascade sizes of a tool-set with 144 possible recipes and an average setup to process time ratio of 22 %. This tool-set is simulated with N and N-1 tools in the fleet.
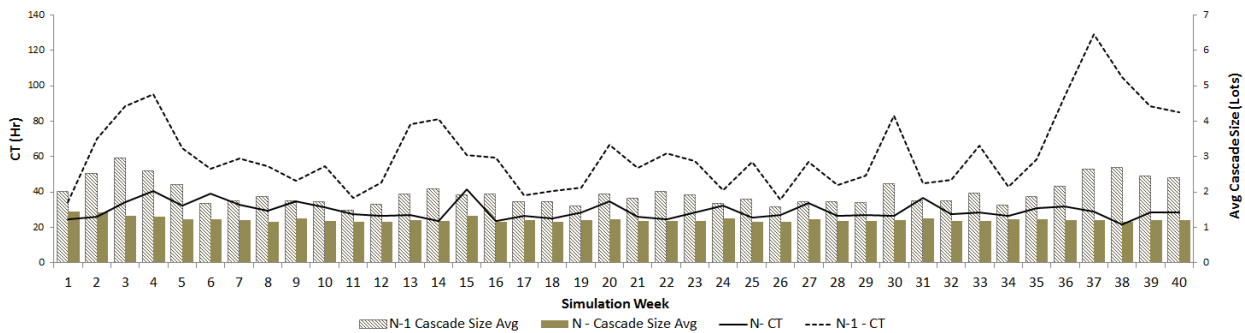

Figure 6: Tool-set CT and mean cascade size with N vs. N-1 tools in fleet

Increasing in the cascade size supports an increase in the tool-set capacity. Yet, in cases where the different recipes refer to different operations in the process flow it also creates an increase in the arrival variability to downstream operations by creating longer alternating epochs of rapid and slow arrivals and therefore increase CT at the operation. However, when the number of tools at the tool-set is reduced, if the tool-set has multiple operations throughout the process this may also lead to a similar impact on arrival variability as described in section 4.2 and cause a reduction in CT by decreasing the arrival variability. In other words, such a tool-set increases the local inter-arrival variability in order to reduce the overall inter- arrival variability .

Although the impact of tool reductions can have conflicting impacts, we can generally predict the type of CT impact a downstream tool-set would experience based on the tool-set utilization. Tool-sets with lower utilization rates have capacity well above the fab loading and would therefore experience relatively short busy periods without accumulation of inventory over time. For tool-sets with low utilization the arrival bursts generated by the upstream tool causes most of the inventory processed during the short busy period to arrive at the beginning of the period and accumulate CT. Therefore, tool-sets with low utilization are impacted to a larger extent by the microscopic inter-arrival variability and would accumulate additional CT owing to the increase in the cascade size of upstream tool-set.

Tool-sets with high utilization rates, on the other hand, have a capacity that is close to fab loading and therefore experience long busy periods in which they accumulate inventory over long periods of time. For such tool-sets, the CT impact of the arrival bursts is negligible in comparison to the CT impact of larger scale variability caused by WIP bubbles. In this case, upstream tool-sets that create short arrival bursts while reducing the arrival variability on a larger scale by dissolving WIP bubble would cause a reduction in the overall CT. This behavior by which tool-sets experience arrival variability differently is consistent with the equations of Whitt (1995) which approximate the inter-arrival variability to each queue in a network based on the utilization of the queue by making separate approximations of the entire network for each utilization rate. In Figure 7 and Figure 8, the CT impact of the N-1 tool reduction used in Figure 5, is shown for cases of low (<70%) and high (>90%) utilization. As already noted, tool-sets with lower utilization show an increase in CT while tool-sets with higher utilization show a decrease in CT.
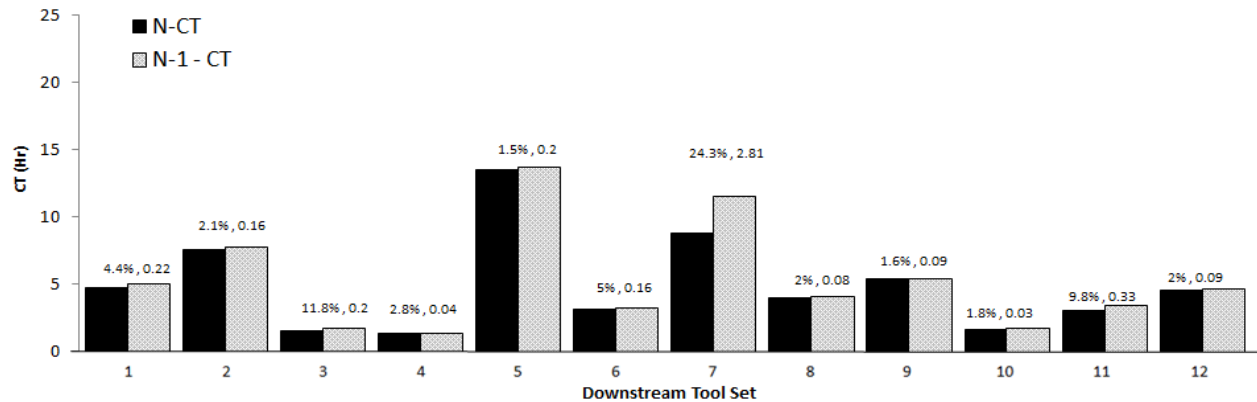

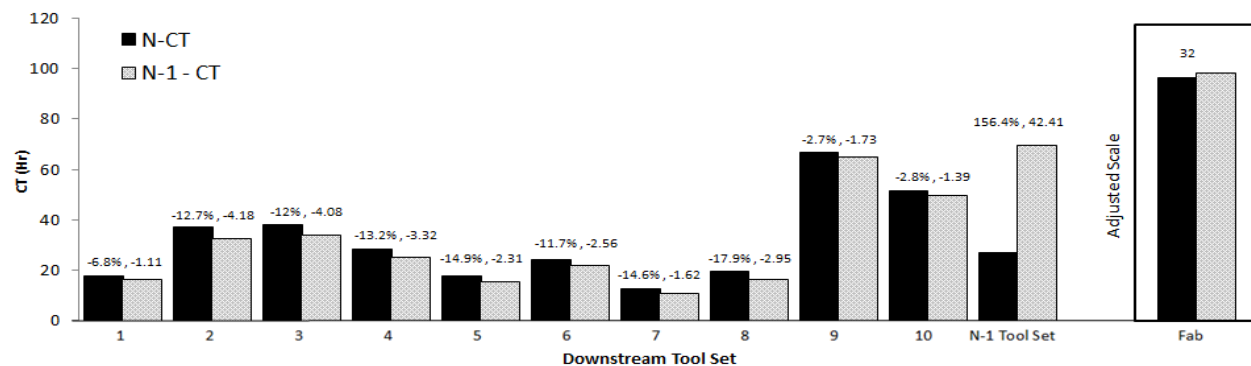Figure 7: CT of downstream toolsets with low utilization by downstream order (<70%)

Figure 8: CT of downstream toolsets with high utilization by downstream order (>90%)

4.4 Complete Fab Tool Reduction Analysis

The models and behaviors described in the previous subsections were constructed in order to analyze the impact of tool reduction from a single tool-set. However, in order to identify multiple potential tool-sets to be reduced and create alternate fab capital and CT scenarios a different approach is required. A full fab simulation, in which each operation is given an individual arrival stream of lots that are terminated after processing (instead of being routed from operation to operation downstream the flow is "horizontal") is used to first identify tool reduction candidates. In this "horizontal" simulation model, unlike running separate simulations for each tool set, a large portion of the run rules between different tool sets may be maintained. A "horizontal" model simulation enables the simultaneous identification of multiple tool-sets that can potentially have a tool reduced from the tool set within a single simulation run. The tool-sets considered for tool reduction are those that, while maintaining sufficient capacity to support fab loading, show either a small CT increase or a substantial capital gain. Once the potential candidates are identified we use a standard, "vertical" full fab simulation model with lots routed from operation to operation down the process flow to estimate the CT impact of the tool reductions. By iterating back and forth between "horizontal" and "vertical" simulations, using the arrival characteristics to operations in the "vertical" simulation to generate the arrival streams of the "horizontal" simulation at each iteration, we can rapidly generate results for different possible states of the fab with varying capital costs and matching CT under a fixed throughput rate. An illustration of the results of these simulation is depicted in Figure 8, where three different fab states all with the same throughput rate but with different cost and CT are presented. Company management can then choose the fab setting that provides a CT sufficient to comply with customer needs with minimal capital cost.
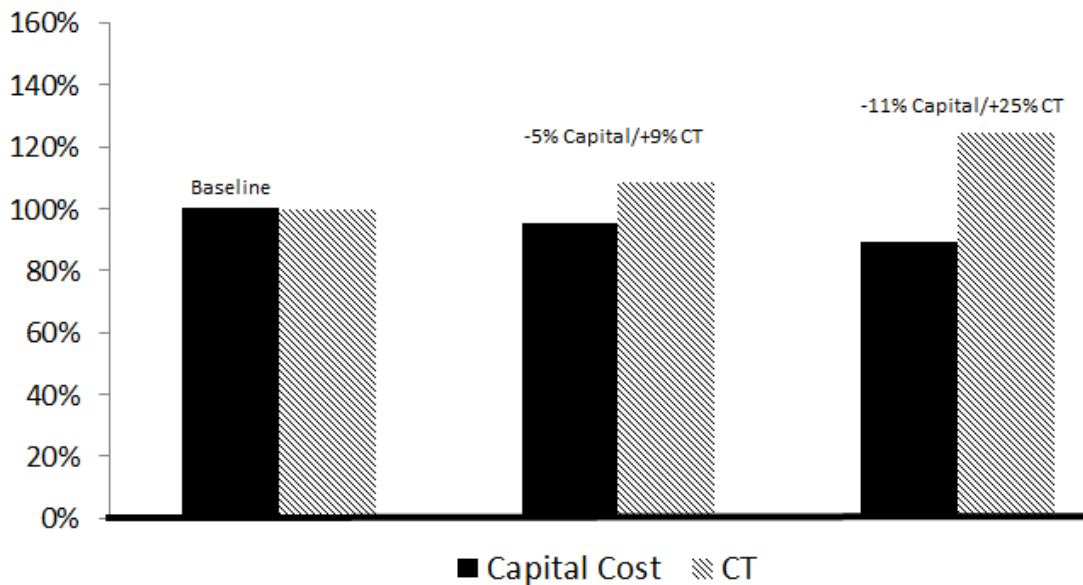

Figure 8: Capital vs. CT scenarios

5 CONCLUDING REMARKS

This paper provides a modeling approach for dealing with the challenging question of trade-off between capital utilization and production efficiency in semiconductor manufacturing where the ultimate goal is of maximum output at maximum velocity and minimum cost. We reviewed three cases for which a tool reduction project has been simulated, each having a different impact on fab WIP flow variability and subsequently on fab CT. We discussed the mechanisms behind these relations, namely the mechanisms known as "variance generating limiter", "variance canceling limiter", and "mixed variance impact

limiter". We also presented results of full fab simulation analysis with a fixed throughput rate and varying capital costs to show the trade-offs between cost and CT.

A full fab simulation has been utilized in an iterative process between models of "horizontal" and "vertical" simulations in order to rapidly generate results for different possible states of the fab with varying capital costs, matched cycle time (CT), and fixed throughput rate, so as to determine the most efficient operating condition for the fab with respect to cost, CT, and output.

More work is definitely needed in this area. We have taken a simulation modeling approach to the problem but clearly there may be viable paths for a mathematical formulation of an optimization problem for this. An optimization framework may consider not only measures at the wafer level (such as wafer cost and output) but at the (good) die level.

REFERENCES

Brown, S, Zisgen, H., Meents, I , Wheeler, B.R , and Hanschke ,T, 2010. "Queuing Model Improves IBM's Semiconductor Capacity and Lead Time", (Wagner finalist), *Interfaces,* Vol. 40, No. 5

Chen, T., Wang, Y., and H. Tsai, 2009. "Lot Cycle Time Prediction in a Ramping-up Semiconductor Manufacturing Factory with a SOM–FBPN-ensemble Approach with Multiple Buckets and Partial Normalization." *Int. J. Adv. Mfg. Tech*., V-42, 1206–1216.

Chien, C.F., Dauzère-Pérès, S., Hans, E., and J.W. Fowler, 2008. "Modeling and Analysis of Semiconductor Manufacturing in a Shrinking World: Challenges and Successes." In *Proceedings of the 2008 Winter Simulation Conference,* edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler, 2093-2099. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chung, S., Lee, A.H., Kang, H., and C. Lai, 2008. "A DEA Window Analysis on the Product Family Mix Selection for a Semiconductor Fabricator." *Expert Systems with Applications,* V-35, 1‑2, 379‑388.

DeJong, D.C., and S.A. Fischbein, 2000. "Integrating Dynamic Fab Capacity and Automation Models for 300MM Semiconductor Manufacturing" In *Proceedings of the 2000 Winter Simulation Conference,* edited by J. A. Joines, R. R. Barton and P. A. Fishwick, 1505-1509. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Fayed, A., and B. Dunnigan, 2007. "Characterizing the Operating Curve – How Can Semiconductor Fabs Grade Themselves? ." In *IEEE International Symposium on Semiconductor Manufacturing (ISSM) Conference*, 1-4.

Haney, B., and M. Canter, 2005. "Lean and Agile Capacity Management." In *IEEE International Symposium on Semiconductor Manufacturing (ISSM) Conference*, 355-358.

Hopp W.J and M.L. Spearman, 2001. Factory Physics: *Foundations of Manufacturing Management.* New York: McGraw Hill.

Jimenez, J. A., G. T. Mackulak, and J. W. Fowler. , 2008. "Levels of capacity and material handling system modeling for factory integration decision making in semiconductor wafer fabs". In *IEEE Transactions on Semiconductor Manufacturing* V-21 600–613.

Morrison, J. R., and J.P. Juen, 2008. "Closed Reentrant Queuing Networks Under Affine Index Policies: Throughput Bounds, Examples and Asymptotic Loss." In *Proceedings of the 17th International Federation of Automatic Control (IFAC),* 3574-3579.

Geng, N., and Z. Jiang, 2009. "A Review on Strategic Capacity Planning for the Semiconductor Manufacturing Industry. " *Int. J. Prod. Res*., V-47, N-13, 3639-3655.

Guo R., Chen, A., Lin, P., and Shih, Y., 1999. "Function-Based Cost Modeling for Wafer Manufacturing and its Application to Strategic Management." In *IEEE International Symposium on Semiconductor Manufacturing (ISSM) Conference*, 31-33.

Kong, S. 2007. "Two-step simulation method for automatic material handling system of semiconductor fab." In *Robotics and Computer-Integrated Manufacturing* V-23, 409–420.

Kiba, J.-E. , Lamiable, G. , Dauzere-Peres, S. and C. Yugma, 2009. "Simulation of a full 300mm semiconductor manufacturing plant with material handling constraints" In *Proceedings of the 2009*

Winter Simulation Conference, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 1601-1609.

Leu, J., and Y. Yu, 2009. "A simulation-based method for the capacity economics of semiconductors manufacturing fabs." In *Proceedings of the International Conference on Computers & Industrial Engineering (CIE),*1793-1797. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ovacik, I.M., and W. Weng, 1995. "A Framework for Supply Chain Management in Semiconductor Manufacturing Industry. ", *IEEE Trans. Semi. Mfg*., 47-50.

Shanthikumar .J.G, S. Ding, and Zhang M.D, 2007. "Queuing Theory for Semiconductor Manufacturing Systems: A Survey and Open Problems" In *IEEE Trans on automation science and engineereing*,V-4 N-4, 512-522.

Suresh, S. and W. Whitt, 1990 " The heavy-traffic bottleneck phenomenon in open queuing networks", *Oper.Res. Letters,* V-9, 355-362.

Witte, J.D., 1996. "Using Static Capacity Modeling Techniques in Semiconductor Manufacturing." In *IEEE/SEMI Advanced Semiconductor Manufacturing Conference (ASMC).*

Whitt.W.,1995 "Variability functions for parametric-decomposition approximations of Queuing Networks", *Management Science* ,V-41, N-2, 1704-1715.

Zisgen, H., Meents, I , Wheeler, B.R , and Hanschke ,T, 2008. "A Queuing Network Based System to Model Capacity and Cycle time for Semiconductor Fabrication". In *Proceedings of 2008 Winter Simulation Conference* edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler, 2067-2074. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

ADAR A. KALIR is a Principal Engineer with the Fab/Sort Manufacturing Division of Intel Corporation, Qiriat-Gat, Israel. He holds a B.S. and M.S. in Industrial Engineering from Tel-Aviv University, Israel, and a Ph.D. in Industrial Engineering and Operation Research from Virginia Tech, USA. He also serves as an Adjunct Professor at Ben-Gurion University, Israel.
His e-mail is: adar.kalir@intel.com.

DEAN GROSBARD is a Manufacturing Engineer with Manufacturing Engineering Department of Intel Corporation, Qiriat-Gat, Israel. He holds a B.S in Mathematics and a B.S and M.S in Industrial Engineering from Ben-Gurion University of The Negev, Beer-Sheva, Israel.
His e-mail is: dean.grosbard@intel.com.