Proceedings of the 1997 Winter Simulation Conference
ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson

EXPERIENCES WITH BACKWARD SIMULATION BASED APPROACH FOR LOT RELEASE PLANNING

Sanjay Jain
Stephen Chan

Gintic Institute of Manufacturing Technology
71 Nanyang Drive
SINGAPORE 638075

ABSTRACT

Bottleneck based scheduling approach suggests planning the release of lots based on the capacity of the bottleneck machines in the manufacturing process. Many scheduling systems determine the start of the first operation of the lot based on backward scheduling from its operation on the bottleneck machine. An approach was developed for determining the lot release times based on backward simulation. While conceptually appealing and successful with simple problems, the approach did not lead to expected improvements in a highly complex semiconductor manufacturing scenario. This paper describes the approach, its implementation and the limitations realized in the complex scenario.

1 INTRODUCTION

The manufacturing companies today face stiff competition due to the globalization of the world markets and supplier bases. The semiconductor manufacturing industry in particular is becoming highly competitive as more and more capacity is aggressively being added across the world. Semiconductor industry focuses on some key measures for maintaining its competitiveness (SIA 1994). These include cycle time, overall equipment effectiveness (OEE) and throughput. Organizations need to focus on achieving a balance among the conflicting objectives of reducing cycle time and increasing throughput and OEE.

All of the measures mentioned are strongly influenced by production planning and scheduling mechanisms. To improve competitiveness, organizations need production planning and scheduling approaches that optimize across these multiple objectives. Unfortunately, such approaches are hard to develop. Production planning and scheduling continues to be a challenging problem for researchers and practitioners. Many heuristic, algorithmic and simulation based scheduling approaches have been proposed and developed. However, manufacturers still face the daunting task of identifying the approach best suited for them and customizing it for their unique requirements. Among the many scheduling approaches reported, bottleneck based approaches appear to have been popular among practitioners and scheduling software developers (Morton and Pentico 1993). These approaches propose improvement in the performance of a manufacturing system through understanding and management of the bottlenecks in the system (Goldratt and Cox, 1992). The concept of releasing material into a manufacturing based system based on pull from the bottleneck machine has been proposed (Goldratt 1990) and has been used in production planning and scheduling systems (Fry et al 1992).

The effort reported here was carried out for an assembly and test site of a semiconductor manufacturer. The complexity of semiconductor manufacturing with multiple resource constraints and unique equipment configurations makes it a very demanding scheduling and production planning problem. Simulation based approaches are well suited for such scenarios and have been reported in use in semiconductor industry (Thompson 1995). Some of the simulation based scheduling software provide special constructs and templates for application in semiconductor industry. Such semiconductor industry specific constructs include those for cluster tools in wafer fabs and for multiple headed testers in assembly and test sites. The availability of such features together with its fast execution speed helped the selection of an industry specific simulation software, TestSim/X from TYECIN Systems Inc., for this project.

The paper is organized to successively discuss the concepts, the use of concepts in the proposed approach, the performance of approach and the learnings realized in the effort. The requirements and scope of the proposed system for lot release planning are defined in Section 2. The third section elaborates on the concept of backward simulation. In the fourth section, the proposed lot release planning approach based on backward simulation is described. Section 5 discusses some of the complications faced in implementing the backward simulation using the selected software. The sixth section presents the results achieved using the approach for a semiconductor assembly and test site. Section 7 discusses the limitations of the approach that were realized during the development and the lessons learned. The last section draws conclusions from this study and briefly describes the follow up of this project.

2 LOT RELEASE PLANNING

The quest for improvement on multiple conflicting measures was the prime driver for considering a new approach for controlling the production operations at the semiconductor assembly and test site. Earlier simulation studies had indicated that better control of work in process can lead to substantial improvements in cycle time without an accompanying loss in throughput.

There are several ways to control the level of work in process in a manufacturing system. The options range from Just-in-time operations with manual signals to highly sophisticated detailed finite capacity scheduling software controlling each step of the operation. The suitable options depend very much on the environment and configuration of a manufacturing system. Just-in-time approach was not considered viable due to the large product variety with very high variation in volumes. At the other end of spectrum, a very detailed scheduling software was not considered attractive due to the requirements of real time information infrastructure for the success of such systems.

The particular manufacturing system was part of a worldwide supply chain. While a global central system provided the lots to be released every day, the site had to determine the sequence and timing to be followed during the day. It also had the flexibility to move the release of lots around within the window of a few days. Indeed, at times it was forced to modify the lot releases due to unavailability of required wafers. The current procedure for lot release relied on judgment of the supervisor of the assembly front end operations. The supervisor monitored the work in process inventory at local bottleneck operations and released the lots to keep them from starving. Other than major disruptions downstream from front end, the supervisor did not use information from downstream areas for controlling the flow of lots. Once released, the lots were prioritized by their current cycle time at successive operations. As the lots were released the information was entered into the shop floor tracking system, which then tracked its progress until completion. Figure 1(A) shows the existing flow of information under the current lot release procedures used by the organization.

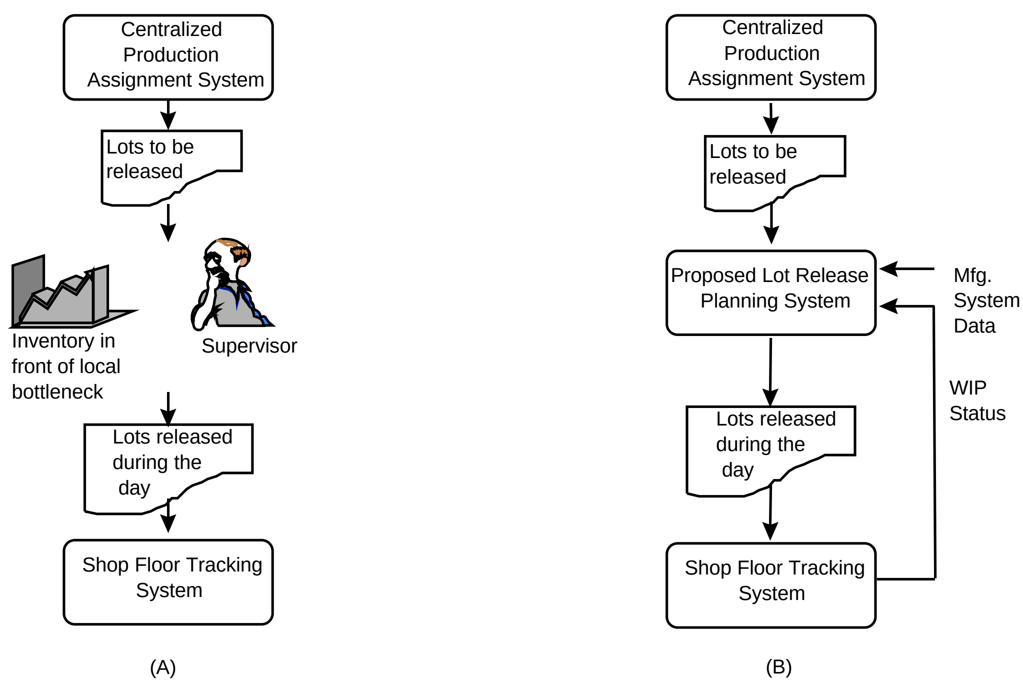

Figure 1: Manufacturing Planning Information Flow (A) under Current Procedures, and, (B) with Proposed Lot Release Planning System

Multiple concepts for lot release control were evaluated and the one based on backward scheduling was selected for development and implementation. Specifically, it was proposed that the release of lots should be controlled based on the pull from the bottleneck resources in the system. This was intended to reduce the queuing of lots at bottlenecks thus reducing the work-in-process and cycle times, while maintaining the throughput roughly at the same levels.

A simulation based system was designed to determine the bottleneck resources. Backward scheduling approach was then used to start from processing of lots at the bottleneck resources and work backwards to determine the lot release times. The resulting lot release schedule was to be provided to the die bank for controlling release of lots into the assembly area. This proposed system was to be integrated in the company information systems as shown in Figure 1(B).

3 BACKWARD SIMULATION

The backward simulation concept focuses on starting from a desired state and moving backwards in time to determine the sequence of events leading to the desired state. The concept of starting from the goal state and moving backwards to determine the path to get to the goal has been used for many years in fields such as artificial intelligence and operations research. In particular it has found application in manufacturing planning and control, at high level planning such as MRP's backward scheduling from due dates to determine the requirements for supply of material and at detailed planning level using backward scheduling from due dates to determine sequence of operations for the lots on machines and the release times for lots into the factory. Backward simulation uses the same principles as backward scheduling, the differences being more in implementation mechanisms used and associated semantics. Indeed, backward simulation can be viewed as an efficient means to implement backward scheduling.

Scheduling approaches typically use multiple passes in time, forward and backward, to develop a good solution to the scheduling problem. This may involve taking one or more lots at a time and developing its complete sequence, backward and/or forward and iterating through all the lots. It may involve developing an initial sequence based on a simple strategy and then iteratively modifying the sequence to improve the quality of the solution. While technically rich, the algorithms usually become quite complicated and use large computing resources to develop schedules. In recent years, (forward) simulation based scheduling approaches have gained popularity due to their ability to generate fast and feasible solutions. The success of these approach is demonstrated by the appearance of many simulation based scheduling software such as AutoSched, FACTOR, TestSim/OnTime, ManSim/OnTime, etc. The simulation based approaches usually execute a single forward pass through time, relying on dispatching rules to develop the sequence. The quality of these solutions is highly dependent on the environment and the dispatching rules used.

Backward simulation based approaches for manufacturing planning applications provide a balance between the solution quality of non simulation based approaches and the execution speed of simulation based approaches. They aim to use the concept of backward scheduling to generate high quality solutions, while utilizing simulation software for speed and feasibility. Similar to non simulation based scheduling algorithms, backward simulation based algorithms use multiple passes. In the earlier implementations reported (Gelders and Van Steelandt 1980, Jain et al 1990), a forward simulation pass is used to determine the sequence for lots already in process and a backward simulation pass(es) is used to determine the release or schedule for new lots. Watson et al (1993) propose using three passes, a backward simulation pass for planning for the new orders, a forward simulation pass to integrate new orders with current work in process, and a forward simulation pass for detailed scheduling if the earlier pass was not at a detailed level.

The implementation of backward simulation in manufacturing is very much like forward simulation except that lots are traveling backwards from final operation to first operation in their routing. Watson et al (1993) describe the backward simulation concept and compare it with forward simulation. The differences highlight the factors to be addressed in implementing backward simulation. These include reversing the routing, reversing assembly into disassembly, reversing the time axis, and designing specific dispatching rules for backward simulations. They also identify the difficulties including the queuing in backward simulation being non intuitive, and the potential infeasibilities due to impact of work in process and going backward further than the current time. Some of these difficulties are further elaborated in Section 5 based on the experience from this project.

Ying and Clark (1994) point out that a more appropriate term for backward simulation is reverse simulation since jobs travel on reversed route in a regular forward simulation. They base this assertion on the view that backward simulation implies a model that should have an exact corresponding representation when simulating forward in time. Using an example with five machines and up to 250 jobs they show that substantial reductions in mean flow time and mean tardiness can be achieved using an algorithm with multiple iterations of forward and reverse simulations.

4 PROPOSED APPROACH

The approach to develop the lot release times based on backward simulation consisted of the following steps:

I. **Downloading of data from company systems and populating of model files**. This step consisted of extracting the information required to describe the current status of production floor from the company systems and providing it in the formats required by TestSim/X.

II. **Capacity Analysis using forward simulation for identification of bottlenecks in the system**. The manufacturing system produced a large variety of products with widely varying requirements on different steps of the production process. There were multiple bottlenecks, that is, several groups of machines had high capacity requirements. The bottlenecks kept changing with the product mix and it was not a trivial task to predict the current bottlenecks. A deterministic forward simulation run was carried out and its outputs analyzed to determine the bottlenecks. The machine groups were ordered from the highest to the least constrained. Only those machine groups that had demand to capacity ratio above a certain defined threshold level were considered to be bottlenecks.

III. **Bottleneck Analysis**. In this step, the lots that needed the bottlenecks were identified and grouped together to reduce the time required for setups. If the same lot required processing across multiple bottleneck machines, it was grouped under the highest constrained bottlenecks used by them. The total capacity requirement at the highest constrained bottleneck machine group was determined for use in the next step.

IV. **Backward Simulation**. This step involved inverting the routing and associated data files of the model to allow backward flow of lots from their operation at the constrained machine to their first operation. The lots were released at the bottlenecks at a time calculated based on the total capacity requirement for the bottleneck machines in previous step. The time was inverted for the backward simulation to start from a future time and end at close to current time or even past time. After the completion of the deterministic backward simulation, the completion times of the first operations of the lots were inverted back on the time axis to determine the start time of the first operation and adjusted for any infeasibility. The results from this step provided the detailed plan for lot release.

The main modules of the Lot Release Planning system developed using the above approach are shown in figure 2.

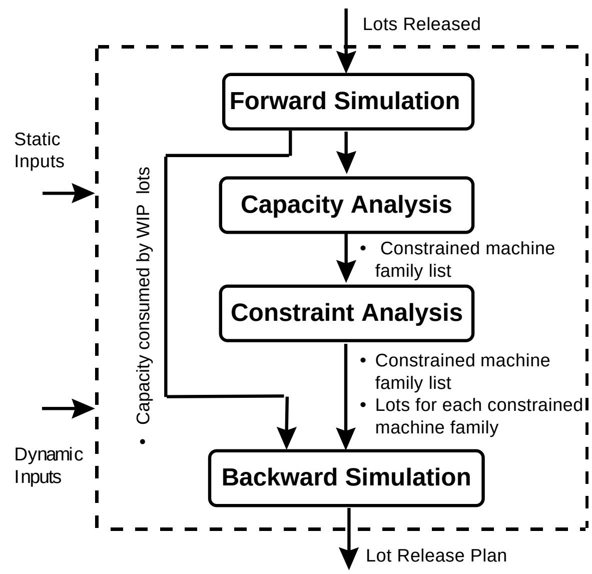

Figure 2: Modules of the Lot Release Planning System

5 COMPLICATIONS IN BACKWARD SIMULATION

The implementation of backward simulation posed some difficulties. The mechanics of the backward flow required some careful thinking and effort to remove oneself from the mindset of forward flow. Dispatching and set ups created some of the major complications.

The dispatching rules in backward simulation have to be reversed to get the sequence to be realized in forward flow. For example, if the forward flow dispatching rule is shortest operation first, the corresponding rule in backward simulation is the longest operation first. Unfortunately, it is not so straightforward with many other dispatching rules. Consider for example, the rule of processing the lot with the highest age (or the longest cycle time so far) first. In backward simulation it is hard to determine a rule that will provide an exact reverse sequence. The closest approximation would be to process the lots with shortest remaining process time first since, under normal circumstances, such lots would have the shortest age in the forward flow and would be processed last. The action of dispatching rules becomes harder to reverse in situations with multiple feeding and consuming machines.

Correct modeling of setups in backward simulation poses the most difficulty, particularly when the simulation is being done through a simulator. Most simulators, including TestSim/X, allow detailed modeling of setups only before the actual process. Facility to define unload times, if provided, usually allows only a specification of time without any consideration of product types, preceding setup, etc. Given these limitations, the setups were defined using the setup feature in the simulator. This resulted in what would amount to setups occurring after the actual processing in forward flow! Using this approach, the capacity requirement of the machines would be quite close to that in forward simulation unless there are widely varying times with sequence dependency. However, the cycle times of the first and last lots for a setup are quite different from that in forward simulation. The first lot in forward flow should incur the setup time at a process step, instead it is the last lot in forward flow that incurs the setup time at a process step when modeled in backward simulation.

Another complication related to setups resulted from the need to control the setups on constrained machine families. Since lots starting at each constrained machine family are separated into setup groups which were then assigned to individual machines, dummy lots were required to influence the setups on the assigned machines. Subsequently, the "same setup" rule was used at the constrained machines so that lots of different setup groups can be processed at corresponding machines only.

In addition to the above major factors, some complications related to inverting of files and time axis had to be dealt with. To facilitate external creation of model files, TestSim/X employs a set of relational tables for model definition. Many of these tables are linked by internal numbers which TestSim/X uses for referencing when loading models. Consequently, when the main routing table is being inverted, many other supporting tables have to be updated (e.g. update of step numbers) in order to safeguard model. Specific changes had to be made for rework loops while inverting the routing files. While inverting time axis, plant shutdowns and equipment off-line schedules (e.g. testers reserved for engineering use) also had to be inverted.

6 EVALUATION AND RESULTS

The Lot Release Planning system was developed in modules described in Section 4. A very large effort was required to interface the system with the company systems to collect and reformat the data. For quantifying the improvement in performance through use of this system, actual lot releases and other data for a month of operation were collected at the company. A base simulation model was built for modeling the current lot release procedures and executed with the data. The results from this base simulation model were used for validation of the simulation model and as the basis for comparison against the results with the new lot release planning approach.

The evaluation of the performance for the month was done by executing the simulation model one day at a time as shown in Figure 3. The lot release plans were generated for each day using the WIP status at the end of the previous day and the list of lots to be released generated on the day. A month long continuous simulation was not feasible since the lot release plan depended on the daily status which could only be determined at the end of the daily simulation. Daily simulations included stochastic factors such as machine failures. Using this scheme, the performance of the Lot Release Planning system was evaluated in a very close representation of the real life system working under the new approach.

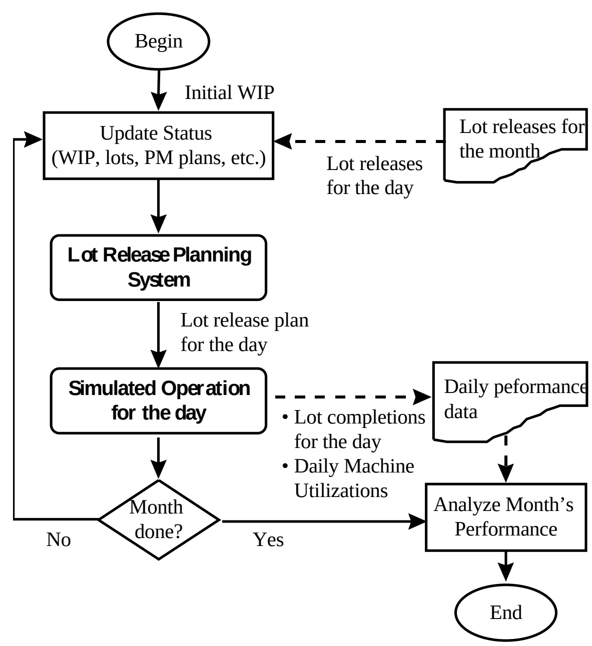

Figure 3: Evaluation of Manufacturing System Performance with the Lot Release Planning System

The results of the evaluation were, however, not encouraging. The performance of the system was not found to improve over the simulated performance of the system in the month's operation using the current lot release procedures. The comparison was carried out against the simulated performance rather than actual system performance to alleviate the impact of modeling inaccuracies. Figure 4 shows the results from different versions of the new approach compared to the simulated performance with the current procedures. The desired objective was to achieve a reduction in cycle time with no reduction or an increase in throughput, leading to points in the “Target Area” shown in the figure.

Several different versions of the original approach were used in attempts to improve the performance. These different versions resulted from detailed study of the events leading to the poor performance and modifications to alleviate perceived limitations. A Gantt chart tool (AESOP 1995) was used to help in the detailed studies. Given the large volume of data and memory limitations of the computer and the software, the Gantt charts had to be built in two or three segments. The different versions of the approach included changes in parameters, and changes to the logic itself. The major parameter changes and enhancements are discussed below:

a) A minimum time was designed in for queuing before the bottleneck machine families. In backward simulation execution, this was implemented by inserting a delay step after the bottleneck machine. The intent of this designed in queuing time was to provide a buffer before the bottleneck machine families to avoid their starvation. Different values of buffer levels were tried.

b) Starting time for the backward simulation was determined using three factors - capacity consumed by WIP lots, the average flow time of new lots to the constrained machine families, and a multiple of the capacity requirement for the new lots. The maximum time among the bottleneck machine families was used as the start time for the backward simulation. Different values for the multiple of the capacity requirement were tried.

c) A threshold value was selected to segregate machine families into bottleneck machine families and non-bottleneck ones. The lower this value the higher the number of bottleneck machine families. Higher number of bottleneck machine families resulted in higher computation effort and larger number of lots controlled by the bottleneck pull.

d) The lots were released at the furthest downstream bottleneck machine family in their route for the backward simulation. An alternate scheme of releasing the lots at the highest constrained machine family in the route was also tested, but did not lead to any significant improvements.

e) Different schemes were used for dealing with lots that did not require the use of any bottleneck machine families on their route. They were released at the beginning based on the pull from the first operation. Another alternate scheme had their release times evenly spread out over the day.

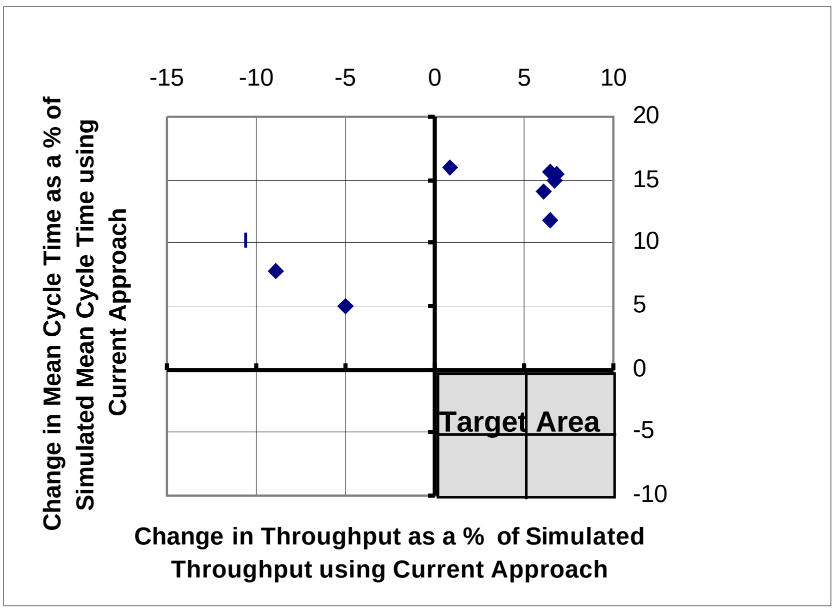

Figure 4: Results from Different Versions of the New Approach Compared to the Simulated Performance with Current Procedures

Lower priority for these "unconstrained lots" was also tested.

None of the above enhancements and parameter changes moved the resulting performance closer to the target area, that is, a reduction in cycle time with no reduction in throughput.

7 LIMITATIONS OF THE APPROACH AND LESSONS LEARNED

The major limitation of the approach was the inability to follow the sequence of operations at successive process steps as predicted by backward simulation. The lots were released using the plan developed by the new approach, however, the flow of lots through their complete route was based on lot age based simple priority rules. Using this mechanism it was found that the sequence of events used by the backward simulation to arrive at the lot release plan was almost never repeated during the forward simulation. This resulted in performance results widely different from that expected. The large impact of minor changes early in lot processing sequences has been discussed recently and doubts have been raised on the usefulness of simulation in such cases (Kempf 1996).

A detailed schedule would help by forcing the lot processing sequences to follow that predicted by the backward simulation. However, that scenario raises the complexity by requiring the use of procedures for responding to real time interruptions that make it difficult to follow a predetermined schedule. Different mechanisms for such response have been studied (Jain 1988), but they require real time information infrastructure for implementation.

The basic bottleneck driven scheduling approach has been shown to work in processes with single major bottleneck (Goldratt and Cox, 1992). In the scenario for this project, with a large variety of products and routes, there were a large number of machine families with capacity demands close to each other. With a minor change in setups or minor machine stoppage, a different machine family could become the highest bottleneck. The bottlenecks also kept shifting based on the product mix. In this complex scenario, the implementation of the approach required many enhancements for dealing with successively lower constrained bottlenecks. The adhoc enhancements made as the limitations were realized did not succeed in handling the complexities. The details of the backward scheduling approach needs to be researched further for implementation in such complex scenarios.

Some project management lessons were also learned during the course of this project. This step of the project was intended to be an implementation step with only an initial part devoted to the proof of concept. However, the team was forced to go into design iterations due to the failure of the designed approach in achieving the desired improvements. The original design was tested on simple scenarios and found to be conceptually acceptable in the previous design phase, but that testing did not turn out to be sufficient. An alternate way would have been to go through the above building of the system and evaluation in a separate phase before getting into implementation. Given the large effort required for building and evaluating the system, the industry partner would have been unwilling to commit to such a phase without having some confidence in its success. It appears to be the chicken and egg situation; it is hard to commit to implementing a new approach until it is expected to lead to a substantial improvement, but it is hard to conclusively determine if it would lead to a substantial improvement without the effort of completely building it and testing it with the complexities of the real life situation.

The use of adhoc enhancements to the approach was the result of a project team struggling to meet the implementation time line. If the team realized upfront about the length of time allowed for successive enhancements, it would have stepped back, identified all the parameters affecting the performance, carried out well designed experiments to determine the major parameters and only then developed the enhancements. The intent has been to develop such complex approaches in research and development projects before attempting an industry implementation, but in this case the complexity was initially underestimated.

8 CONCLUSION

The project provided insights into the performance of a backward scheduling based approach in the complex environment of semiconductor manufacturing. While the development and implementation did not lead to any performance improvements, it did provide valuable learning experience to the project team as well as to the industry partner. This paper aims to share this learning experience with the simulation community involved in similar application areas. Future efforts should focus on overcoming the limitations of this approach and lead to the benefits usually associated with the approach of backward scheduling from the bottleneck.

For the project, the efforts on improving backward scheduling were discontinued and alternate heuristics were built for controlling the lot release. In evaluations, these alternate heuristics were found to lead to improvements on desired performance criteria.

Implementation of these alternate heuristics is under consideration by the semiconductor company.

ACKNOWLEDGEMENTS

Several people were involved at various stages of the project. For the stage described here, the development team from Gintic included Kelvyn Chin, Lieu Ann Horng and Wu Lijun in addition to the authors. An earlier version of approach was designed in the preceding stage with Dr. AI Sivakumar as the project leader. A number of people from the semiconductor company contributed to this project, but are not identified to maintain confidentiality. The project team is grateful to the management of Gintic and the semiconductor company to allow the time for attempts to improve the performance of the approach.

REFERENCES

AESOP, 1995. *GANTT Manual Scheduling documentation*, Version 1.1, AESOP GmbH, Germany.

Fry, T.D., J.F. Cox, and J.H. Blackstone, Jr. 1992. An Analysis and Discussion of the Optimized Production Technology and Its Use. *Production and Operations Management*, Vol. 1: 229-242.

Gelders, L.F. and F.V. Van Steelandt. 1980. Design and Implementation of a Production Planning System in a Rolling Mill: A Case Study. *AIIE Transactions*, March 1980.

Goldratt, E.M. and J. Cox. 1992. *The Goal*, second revised edition, Croton-on-Hudson, NY: North River Press.

Goldratt, E.M. 1990. *The Theory of Constraints*, Croton-on-Hudson, NY: North River Press.

Kempf, K.G. 1996. Simulating Semiconductor Manufacturing Systems: Successes, Failures and Deep Questions. In *Proceedings of the 1996 Winter Simulation Conference*, J.M. Charnes, D.J. Morrice, D.T. Brunner, and J.J. Swain (eds.), 3- 11. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Jain, S. 1988. Decision Framework for Interruption Handling in Flexible Manufacturing Systems, PhD Dissertation, Engineering Science, Rensselaer Polytechnic Institute, Troy, NY.

Jain, S., K. Barber, D. Osterfeld. 1990. Expert Simulation for On-Line Scheduling. *Communications of the ACM*, Vol. 33, No. 10:54-60.

Morton, T. E. and D.W. Pentico. 1993. Chapter 2: Approaches to Scheduling, in *Heuristic Scheduling Systems*, John Wiley & Sons.

Semiconductor Industry Association. 1994. *The National Technology Roadmap for Semiconductors*.

Thompson, M. 1995. Using Simulation-Based Finite Capacity Planning and Scheduling Software to Improve Cycle Time in Front End Operations. 1995 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, 131-135.

Watson, E.F., D.J. Medeiros, and R.P. Sadowski. 1993. Generating Component Release Plans with Backward Simulation. *Proceedings of the 1993 Winter Simulation Conference*, G.W. Evans, M. Mollaghasemi, E.C. Russell, W.E. Biles (eds.), 930-938. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Ying, C.T.C, and G.M. Clark. 1994. Order Release Planning in a Job Shop Using a Bi-directional Simulation Algorithm. In *Proceedings of the 1994 Winter Simulation Conference*, J.D. Tew, S. Manivannan, D.A. Sadowski, and A.F. Seila (eds.), 1008-1012. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

AUTHOR BIOGRAPHIES

SANJAY JAIN is a Senior Research Fellow with the Systems Simulation group at Gintic Institute of Manufacturing Technology, Singapore. His research interests are in the area of using modeling and analysis techniques in development and operation of manufacturing systems, and in improving performance of simulation systems through parallel and distributed execution. Prior to joining Gintic, he worked for several years as a Senior Project Engineer with General Motors North American Operations Technical Center in Warren, MI, USA. He has a Bachelors of Engineering from University of Roorkee, India, a Post Graduate Diploma from National Institute for Training in Industrial Engineering, Bombay, India, and a PhD in Engineering Science from Rensselaer Polytechnic Institute, Troy, New York. He is a member of the editorial board of *International Journal of Industrial Engineering*.

STEPHEN CHAN is a Research Fellow with the Systems Simulation group at Gintic Institute of Manufacturing Technology, Singapore. His research interests include functional modeling, dynamic modeling and simulation as well as development of modeling tools. He received his B.Eng. from the University of London, U. K. and MSc. from the University of Strathclyde, U. K.