*Proceedings of the 2002 Winter Simulation Conference*
*E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, eds.*

# SIMULATING THE TRANSPORT AND SCHEDULING OF PRIORITY LOTS IN SEMICONDUCTOR FACTORIES

Chad D. DeJong
Scott P. Wu

Intel Corporation
5000 West Chandler Boulevard
Mail Stop CH3-113
Chandler, AZ 85226, U.S.A.

## ABSTRACT

As high technology product market becomes more dynamic and competitive, chip manufacturers need to bring products to customers in short periods of time. As a result, semiconductor fabrication plants regularly contain lots with priority status. These lots have several unique characteristics compared to other production lots, both in terms of lot transport and scheduling on tools. These lots consume tool capacity that may impact the factory output rate. Priority lots also have specific policies for transport. The impact of these priority lots on other lots in the fab is not easily quantified, as many factors are involved. Dynamic factory and AMHS simulation models are capable of capturing the variability of a factory, and the interactions of critical constraints that prevent predictable manufacturing. This paper presents a breakthrough modeling approach to study the behaviors of priority lots, and to quantify their impact to manufacturing.

## 1 PROBLEM STATEMENT

Semiconductor fabrication plants (fabs) regularly contain lots that have a priority status. These lots behave differently from other production lots in several ways, both in terms of factory capacity and lot transport. The unique behaviors of these priority lots take away valuable tool capacity in the factory and may impact regular production lot (non-priority lots) output rate. At the same time, there are several automated transport capabilities specific to priority lots that can impact the delivery performance of priority and non-priority lots. The extent of the factory capacity and lot transport impacts must be understood to enable accurate decision support for fab manufacturing. Specifically, manufacturing managers need to make sound decisions regarding priority lot commitments and scheduling, while achieving minimum disruption to the overall fab.

## 2 INTRODUCTION AND BACKGROUND

As the high technology product market is becoming more dynamic and competitive, chip manufacturers are striving to bring products to customers in the shortest period of time. Semiconductor fabrication facilities are running high priority lots to satisfy critical customer demands or to achieve fast process information. The goal is to push these priority lots through the factory as quickly as possible, so that they can be tested and ultimately introduced as new products (Hillis and Robinson, 2002). The impact of these high priority lots on other lots in the fab is not easily quantified, as many factors are involved. For instance, operators, tools, automated material handling system (AMHS), and reticles all have different policies regarding these priority lots. This paper illustrates a breakthrough modeling approach to study the behaviors of priority lots and to quantify their impact to other lots in the line.

In general, the highest priority lot (P1) is the most important lot in the factory. These lots have special tool reservation policies for downstream tools. The second highest priority lot (P2) will break batches and cascade size requirements, enabling them to move faster through the factory than non-priority lots. The last priority classification of lots (P3) will be the next cascade or batch to run on the tool. Table 1 summarizes the characteristics of different levels of priority lots. All other lots in the factory are considered non-priority (NP) lots.

Table 1. Lot Priority Level Description

| Priority Lot Classification | |
|---|---|
| Priority1 | Reserve Tools; Break batches; Break Cascades |
| Priority2 | No Reserve Tools; Break batches; Break Cascades |
| Priority3 | No Reserve Tools; Next batch or Cascade on Tool |
| Non-Priority | Regular Production, FIFO |

There are also several policies for the automated transport of priority lots via AMHS. The priority of a lot is passed between the fab's Material Control System (MCS) and IBSEM-compliant AMHS devices, using a range of integer values. Transport prioritization can be based on the priority of the lot itself, or based on the current or pre-configured status of processing tools. This paper discusses prioritized lot delivery based on the status of the lot.

Of course, material handling affects fab processing, which in turn can impact the scope of movement requests on the material handling system. As a result, integrated factory and AMHS models are required to determine even high level impacts to key success indicators. In order to gain insight into how the components of a factory impact performance metrics, Intel uses an integrated discrete-event simulation modeling approach (DeJong and Fischbein, 2000).

Currently, there are attempts to quantify priority lot impacts using static spreadsheet models. These static models produce quick results, but do not capture the dynamic nature of the factory nor are able to quantify the average and variability impact to factory WIP (Work in Process), cycle time, and weekly wafer outs.

## 3 MODELING METHODOLOGY

Dynamic factory simulation offers a tool to capture the variability existing in a factory, especially the interactions of critical constraints and inhibitors that prevent smooth manufacturing flow of material through the line. Both the fab capacity and AMHS models have a variety of input parameters and complex customization code required to drive the simulation.

### 3.1 Project Approach

The fab capacity simulation model was built to support the analysis and experiments requested by fab customers. This required several meetings with key manufacturing representatives to determine and document general policies for all lot classifications, as relates to tool processing and operator policies.

During lot transport modeling, several simulation studies were done to scope and analyze general automation capabilities and determine which should be built into Intel's architecture. There is concern that prioritizing the delivery of a large number of lots would reduce the ability to expidite the truly important lots, and might even worsen overall fab delivery performance. Closely working with automation personnel allowed the right amount of functionality to be built in without overdesigning.

The next question is always how the tool processing and transport capabilities impact each other. In order to gain insight into how these fab components impact overall performance metrics, such as output and fab velocity, Intel uses an integrated discrete-event simulation modeling approach.

### 3.2 Software

Due to the nature of the hundreds of decisions each priority lot must make in terms of claiming tools and predicting lot delivery performance, Intel requires dynamic discrete-event software in which very complex policies and much customization code could be built in.

### 3.3 Model Logic

In general, representing tool behavior in semiconductor fabrication is extremely challenging. Tool dedication, queuing micro-policies, scheduled and unscheduled downtime events, and reticle allocation must all be accounted for.

Priority lot policies also affect tool behavior. These policies were developed and integrated into Intel's fab capacity model. This new functionality includes operational logic associated with priority lots in a high volume manufacturing (HVM) line. There are two key attributes of the fab capacity model methodology specific to priority lots. First, it has the ability to comprehend different levels of priority (P1, P2, P3, and NP lots). Second, it has the ability to model complex operational rules depending on the lot priority. There are several examples of this complexity. Tools are forced to stay idle to wait for upstream P1 lots. This results in decreased tool availability. User can define how long in advance the tools should be held. Breaking minimum cascade and batch size requirement increases setup, which then induces more variability of the factory WIP flow. Special tool re-entrant policies for priority lots reduce parallel resources that could be used by regular production lots. This reduces available capacity of those tools. The degree to which these operational complexities will impact each toolset is dependent on the flow of WIP through the fab, and each toolset's own utilization vs. availability.

AMHS modeling contains different but equally challenging complexity. AMHS in semiconductor factories can be separated into three distinct classes – software, interbay hardware, and intrabay hardware (DeJong and Fischbein, 2000). Priority lot delivery impacts all three. In terms of software, priority classifications are modeled such that they are comprehended in vehicle assignment logic. Essentially, a lot with a higher priority will always be assigned to the best vehicle before any assignments are made for NP lots. Within equal priority settings, lots are assigned on a FIFO basis. In terms of hardware, similar logic applies to these lots claiming stocker (AS/RS) robots, interbay and intrabay transfer ports, and tool loadports. Whenever there is a queue of lots waiting one of these resources, the higher priority lot will jump to the top of that queue. Incorporating this functionality into the AMHS model allows us to assess various implementations of these policies, and determine the impact of increasing volumes of priority lots.

### 3.4 Model Inputs and Outputs

Both the fab capacity and AMHS models have a variety of input parameters required to drive the simulation. The fab capacity model requires MHS delivery times, loadport configurations, detailed labor schemes, a highly detailed process flow, processing times, tool dedication strategies, etc. The capacity model also has to have a destination table so the simulator knows where to send each lot to be processed for the next step, and how to flag delivery requests for the automation model. As outputs, it tracks throughput time, WIP turns, units (lots, wafers) out, tool starvation rates, and time during which a tool has selected a lot/batch to process but the lot/batch was unavailable (waiting for transportation). The fab capacity model also has the ability to output detailed move requests in a manner that is readable by the AMHS model.

The AMHS model can also be described in terms of its inputs and outputs. Inputs include layout, extensive interbay and intrabay vehicle scheduling logic, tool and stocker bay associations, move requirements from the fab capacity model, material control system and equipment control system software parameters, hardware specifications, (such as component cycle times) reliability metrics (for nodes, vehicles, and stockers), vehicle speed and acceleration, and placement of decision nodes. The outputs of the automation model are interbay wait and travel time for each loop by priority classification, intrabay wait and travel time for each bay by priority classification, vehicle statistics, lot movement rates, and stocker robot and loadport utilization information. Refer to Figure 1 for a high level view of the fab capacity and automation model input and output structures.

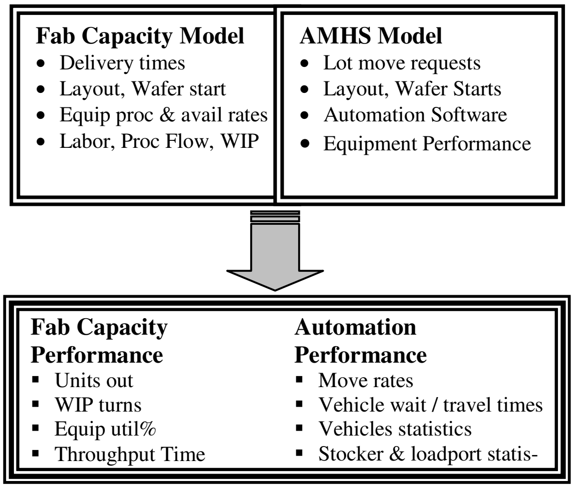

Figure 1: Model Input / Output Structure

### 3.5 Model Verification

Intel spends much energy in model verification and validation. Hours have been spent to verify that priority lots in our simulation models are indeed claiming tools correctly according to their priority classification. Moreover, changing specific model inputs such as varying the number of hours tools are held downstream, or the number of lots in the system, has provided results which indicate model logic accuracy. AMHS logic has been verified through several sets of experimentation, and key model outputs such as how often resource queues are re-sorted due to the entrance of priority lots.

## 4 RESULTS AND IMPLICATIONS

A fab capacity model experiment was setup to evaluate the impact of high priority lots to other production lots. First, a baseline, near constraint factory model was built with no high priority lots introduced to the factory. Using simulation, we determine that the factory is able to produce the expected wafer output, and the factory velocity was recorded. Factory velocity is an indicator of how fast the WIP is moving through the line. Next, high priority lots were introduced to the factory to determine their impact to the velocity of the NP lots. As illustrated in Figure 2, once a single P1 lot is introduced to the factory, there is an immediate impact to the NP lot velocity.

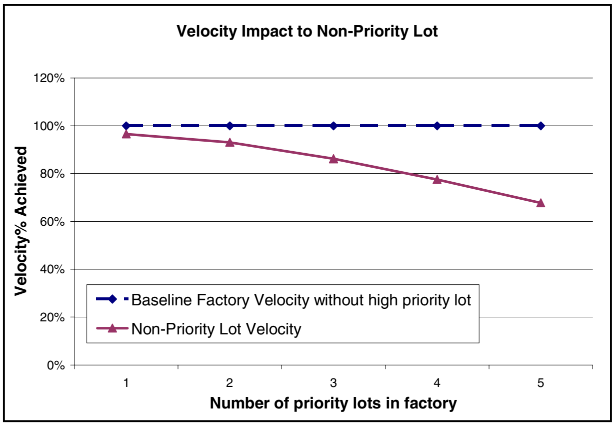

Figure 2: Priority Lot Impact to NP Lot Velocity

The overall weekly outs do not seem to be impacted with a three-hour downstream equipment hold time. Even when there are one to three priority lots in the system, the overall factory outs remain relatively steady when comparing the WIP turns. However, as Figure 3 shows, when there are more than three priority lots in the system, the NP lot velocity and factory outs are both significantly reduced. With four priority lots in the system, outs per week drops to 96%, compared to the baseline factory. When there are five priority lots in the factory, the weekly outs achieved dropped even further to 92% of baseline.

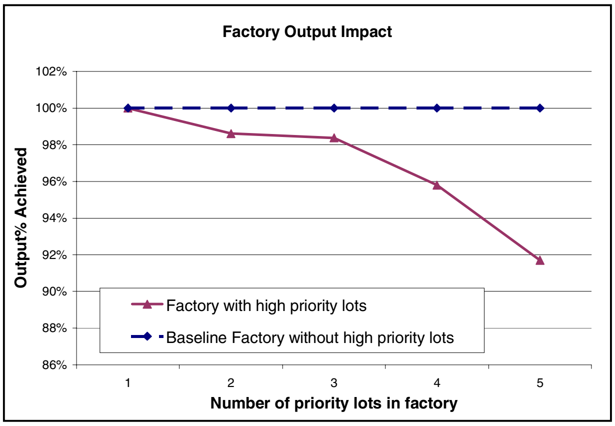

Figure 3: Priority Lot Impact to Factory Output

It is quite clear that, as the number of priority lots is increased to an excessive level, the factory is unable to make the committed output. This lost capacity is due to the effect of priority lot scheduling policies that forces downstream tools to wait for priority lots hours ahead of their expected arrival and lost capacity at certain tool sets due to the breaking of minimum batching requirement.

The primary AMHS experiment simply populated a balanced interbay and intrabay 300mm model (AMHS-A) with a realistically small number of P1 and P2 lots (AMHS-B). The results indicate P1 and P2 delivery times that are significantly better than NP lot delivery times. Moreover, the intrabay delivery time improvement is expectedly greater, as there are more resource types that must be obtained before completing the move, and greater loop move variability. Of course, each bay will have its own unique delivery time differences between NP and priority lots. Bays with high move volumes and move request variability (bays with fast tools or batching tools) will have the greatest differences. Some bays will have no significant difference. Figure 4 shows average delivery time improvements across all bays.

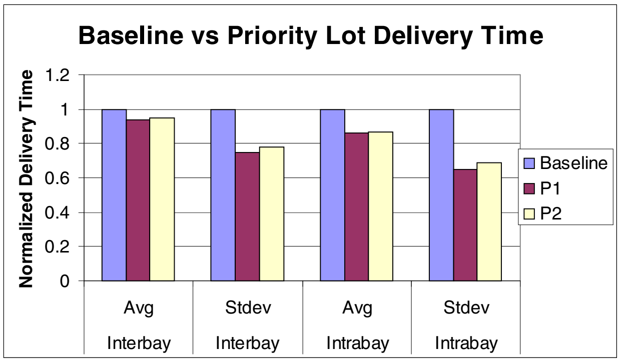

Figure 4: Baseline vs Priority Lot Delivery Time Differences for Interbay and Intrabay

The next set of experiments involved varying AMHS model delivery time improvements in the fab capacity model, to quantify the impact of delivery time improvements to fab velocity for all lot types. First, additional potential AMHS delivery time profiles were created, in terms of percent improvement relative to run AMHS-B. The delivery time profiles are summarized in Table 2.

Table 2: Delivery Time Inputs to Fab Capacity Model

| AMHS run | Description |
|---|---|
| AMHS-A | Baseline |
| AMHS-B | Initial P1/P2 Lot Run |
| AMHS-C | -10% interbay |
| AMHS-D | -10% intrabay |
| AMHS-E | -10% interbay and intrabay |
| AMHS-F | -20% interbay and intrabay |

The AMHS delivery time profiles were modeled as inputs to the fab capacity model. The results are summarized in Figure 5. In all runs, factory outs achieved remained statistically the same. In addition, NP lot velocity remained relatively stable. However, the P1 lot velocity was clearly impacted by delivery time improvements. The largest difference was between the baseline run AMHS-A and AMHS-B.

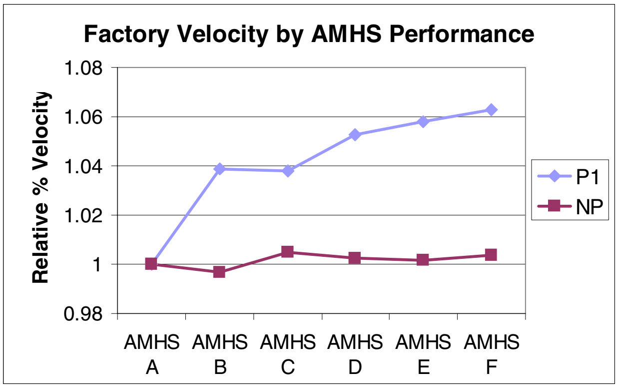

Figure 5: AMHS Performance Impacts on Factory Velocity

Interestingly, the next interbay delivery time improvement of 10% (AMHS-B vs AMHS-C) had virtually no impact to velocity. However, AMHS-D, with an average 10% improvement to intrabay delivery time, did show a marked improvement over runs AMHS-B and AMHS-C. AMHS-E and AMHS-F both showed slight improvements in fab velocity as compared against previous runs, as intrabay delivery time improved by 10% and 20% respectively.

## 5 CONCLUSIONS

Three sets of experiments were completed – fab capacity models which analyzed lot scheduling policies, AMHS models of automated lot transport policies, and finally integrated AMHS and fab capacity models which looked at how AMHS transport of priority lots could impact fab velocity.

The fab capacity runs indicate that the introduction of even one priority lot into a fully loaded factory will negatively impact the velocity of NP lots. This impact appears to grow exponentially as additional priority lots are introduced into the factory. Moreover, output also is also impacted. Our studies indicate that, when holding tools downstream for 3 hours, introducing more than 3 P1 lots into the factory will impact factory outs.

The AMHS model also revealed some very important information about how priority lot delivery can be expidited. When introducing a relatively small number of priority lots to the AMHS system, we were able to significantly improve the delivery performance for these lots. This was done by implementing some relatively simple logic for how these lots claim robots, ports, and vehicles along the way to their final destination. Intrabay delivery times clearly show the most potential for improvement.

Integrating the AMHS and fab capacity models allowed us to quantify the impact of delivery time improvements to overall fab velocity. Initial delivery time improvements led to the greatest increase in fab velocity. Additional delivery time performance upgrades, especially those for interbay, created only marginal fab velocity improvements.

## 6 NEXT STEPS

There are a variety of potential automation transport and lot scheduling policies that lend themselves to further study. For example, there are several ways in which lot transport could be prioritized based on the current status or pre-configured attributes of processing tools. In terms of the fab capacity model, Intel is planning experiments to investigate process step-specific reservation of downstream tools. Of course, the integrated model approach allows further investigation of what transport capabilities would result in actual fab velocity improvements.

## ACKNOWLEDGMENTS

The authors collectively thank the following individuals, without whom this work would not have been possible:

Seth Fischbein, John Braunbeck, Dev Pillai, and Ed Yellig for factory modeling concepts brainstorming and generation. Aidan O’Connor, Casey Swearingen, and Chuck Golla for transport modeling inputs and direction.

## REFERENCES

DeJong, C., and S. Fischbein, 2000. Integrating dynamic fab capacity and Automation models for 300mm Semiconductor Manufacturing. *Proceedings from the 2000 Winter Simulation Conference*, 1505-1509.

Hillis, M., and J. Robinson, 2002. Extremely Hot Lots: Super-Expiditing in a 0.18 Micron Fab. *Proceedings of the MASM 2002 Conference.*

## AUTHOR BIOGRAPHIES

**CHAD DEJONG** is a member of the Operational Decision Support Technology group at Intel Corporation, in Chandler, AZ. He has designed simulation models and various decision support tools for Intel projects for over 5 years. He earned a B.S. in Industrial and Operations Engineering from the University of Michigan, Ann Arbor. He also earned an M.S. in Industrial and Systems Engineering from Georgia Institute of Technology, and completed the Management of Technology certificate program. His previous professional background includes hospital and health care system consulting, automotive component manufacturing, and environmental engineering. Current research and professional interests in the semiconductor industry are in whole factory capacity and operations modeling, supply chain simulation, and model design and execution time reduction. Chad can be reached via email at <chad.d.dejong@intel.com>.

**SCOTT WU** is a member of the Operational Decision Support Technology group at Intel Corporation, in Chandler, AZ. Scott joined Intel in July 2000. Since then, he has designed and developed several key decision support tools, which led to more accurate simulations and improved modeling capabilities. Scott has extensive experience in capacity, and AMHS modeling for both 200 mm and 300 mm factories. Scott is involved in the ongoing effort of developing simulation capabilities to integrate real factory scheduling and operations. Prior to joining Intel, Scott worked for PRI Automation where he used simulation to analyze design layouts of interbay and intrabay monorail systems. Scott can be reached via email at <scott.p.wu@intel.com>.