Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

EFFICIENT SIMULATIONS FOR CAPACITY ANALYSIS AND AUTOMATED MATERIAL HANDLING SYSTEM DESIGN IN SEMICONDUCTOR WAFER FABS

Jesus A. Jimenez
Gerald Mackulak
John Fowler

Industrial Engineering Department
Arizona State University
Tempe, AZ 85287, U.S.A.

ABSTRACT

The Automated Material Handling System (AMHS) must be designed effectively so that it never becomes a limiting factor for the capacity of 300mm wafer fabs. Ideally, a fully integrated fab simulation model (i.e. a model containing detailed modeling constructs for the production operations, the tools, the AMHS, and tool AMHS interactions) should be used in order to design the AMHS. However, the problem is that it takes too much time to simulate and analyze these models. Experimentation has demonstrated that certain capacity models with less detailed AMHS representations can generate accurate system predictions in comparison to the values produced by fully integrated models. Because these less detailed models run faster, we can thus assess efficiently the effects of an AMHS design configuration on equipment capacity. A case study comparing the computational efficiency and the quality of the performance predictions at different levels of detail will be presented.

1 INTRODUCTION

The capacity of semiconductor wafer fabrication facilities (fabs) needs to be planned effectively in order to meet the expected productivity levels with the least amount of resources (i.e. equipment, operators, etc.). As the total costs of building these facilities are currently approaching the $3 Billion dollars (Shelton 2003), effective capacity planning is important for maximizing the bottom line of wafer fabs. Because 300mm fabs are extremely automated, capacity planning models must therefore account for the behavior of key automation components (i.e. the automated material handling system).

The main role of the automated material handling system (AMHS) is to deliver wafers to the right manufacturing step at the right time. Because the costs of equipment within a fab are prohibitively high, the AMHS must never constrain the equipment capacity.

A problem confronting AMHS designers is to determine an adequate number of inter-bay vehicles (i.e. vehicles moving material between two different bays) in the system. An insufficient supply of vehicles can increase delivery times because most of the wafers will be waiting at stockers for unloaded vehicles. On the other hand, too much vehicle capacity can increase delivery times because wafers will spend more time traveling on vehicles due to heavier traffic congestions within the AMHS hallway.

Simulation modeling has been traditionally used in designing the AMHS capacity. A typical simulation approach for determining the inter-bay vehicle capacity consists of using a detailed AMHS model (i.e. a model representing the AMHS operations accurately, but with less fab capacity details) to estimate delivery times across different numbers of AMHS vehicles. The vehicle configuration resulting in the lowest delivery time is then selected. The problem with this approach is that it provides little information about the effects of an AMHS configuration on fab performance.

To detect capacity losses caused by poor AMHS designs, the modeling and analysis of the fab capacity and the AMHS must be therefore integrated. Ideally, a fully integrated fab simulation model (i.e. a model containing detailed modeling constructs for the production operations, the tools, the AMHS, and tool AMHS interactions) should be used in order to determine the AMHS vehicles capacity accurately. However, the problem is that it takes too long to simulate and analyze the desired AMHS configurations with an integrated model.

To reduce the simulation execution time, practitioners build capacity models with less AMHS details. In particular, they use capacity models with “From-To” AMHS delivery time tables (i.e. a table indicating the time the AMHS takes to deliver wafers from a source location to a destination location). Experiments have demonstrated that this modeling approach can produce very accurate performance measures of the fab capacity.

The purpose of this article is to present an accurate and efficient methodology using the capacity model with AMHS delivery times to determine the AMHS vehicle capacity. We then show a case study comparing the computational efficiency and solution quality of the proposed abstract capacity-AMHS model to the corresponding values generated by the fully integrated model.

The remainder of this paper is organized as follows. In Sections 2 and 3, we describe the architectures of the fully integrated model and the capacity model with AMHS delivery times. In Section 4, we present the case study. Finally, in Section 6, we state our conclusions for this work, and propose directions for further research.

2 ARCHITECTURE OF THE FULLY INTEGRATED FAB SIMULATION MODEL

The fully integrated fab simulation model consists of detailed representations of the steps required to manufacture wafers and the AMHS activities involved in moving wafers across the fab.

Simulating the capacity and the AMHS together is very complex. Thus, these systems are modeled by two separate simulation constructs (i.e. the capacity construct and the AMHS construct). These constructs are then integrated so that consistent records of the simulation details (i.e. attributes, simulation clock, variables, etc.) are passed back and forth between them during a simulation run.

The *capacity construct* represents the following fab components [see Jimenez et al. (2005) for more details].

- Tools and tool characteristics (i.e. batching, downtimes, setups, tool dedication policies, etc.).
- Product types (i.e. order frequency, hot lots, initial work-in process, order release policies, etc.).
- Process flow characteristics (i.e. process sequence, type of tool used for each process step, dispatching rules, time-bound production sequences, etc.).
- Factory schedules, operators, reticles, space and layout, etc.

The *AMHS construct* represents the following components [see Jimenez et al. (2005) for more details].

- AMHS layout (i.e. route configurations, control points, modules, vehicles, etc.).
- AMHS operations (i.e. vehicle blocking due to traffic congestions, tool-AMHS interactions, etc.).
- Storage model components (i.e. stocker, stocker robots, input/output ports, etc.).
- Logic controlling the AMHS (i.e. vehicle dispatching, vehicle routing selection, etc.).

Figure 1 illustrates the flow of the simulation data between the capacity construct and the AMHS construct. What this figure indicates is that a wafer completing a manufacturing operation within the capacity construct will trigger a move in the AMHS construct. The AMHS construct then takes control over the simulation and executes the activities involved in transferring the wafer lot to its destination location. When the wafer lot is delivered, the AMHS construct returns the control of the simulation over to the capacity construct and the wafer lot resumes its manufacturing process.

Because the fully integrated model virtually emulates the actual fab, it can be assumed that the capacity performance measures (i.e. throughput, cycle time, work-in process, equipment utilizations, etc.) and the AMHS performance measures (i.e. delivery times, transport times, wafer waiting time for transport, utilization of stocker robots, vehicle utilization, vehicle capacity lost due to congestion, etc.) generated by this model are accurate descriptors of the actual system performance.

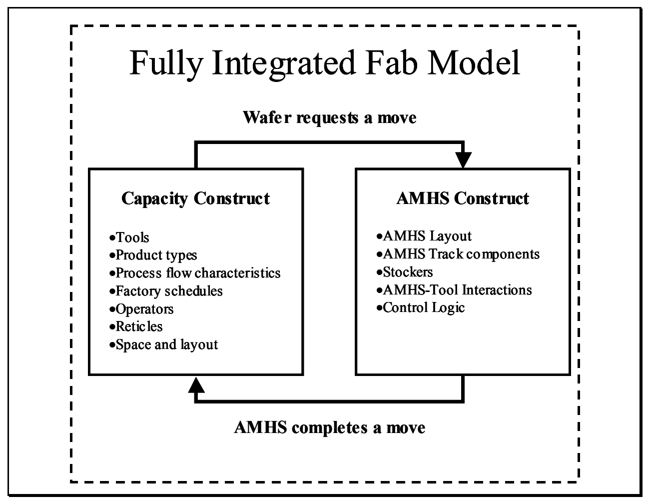

Figure 1: Architecture of the Fully Integrated Model

3 ARCHITECTURE OF THE CAPACITY MODELS WITH AMHS DELIVERY TIMES

Unlike the fully integrated model, the capacity model with “From-To” AMHS delivery times uses one simulation construct only (i.e. the capacity simulation construct). The “From-To” delivery time table thus abstracts the details of the AMHS operations by estimating the average point-to-point delivery times (i.e. the time wafer lots wait at the stocker for unloaded AMHS vehicles plus the transport time) for all the feasible location combinations.

Figure 2 depicts the architecture of capacity models with “From-To” delivery times. This figure shows that after the completion of a manufacturing step, each wafer visits an AMHS simulation block (i.e. a black box) that delays the wafer by the corresponding delivery time. The predictive accuracy of this model abstraction depends on how well the waiting time for vehicle component and the transport time component are represented. These two components can be abstracted accurately from a fully integrated fab model, provided that this model exist. Other methods (i.e. analytical AMHS models) can be also used for generating delivery times, but the problem is that they are less accurate.

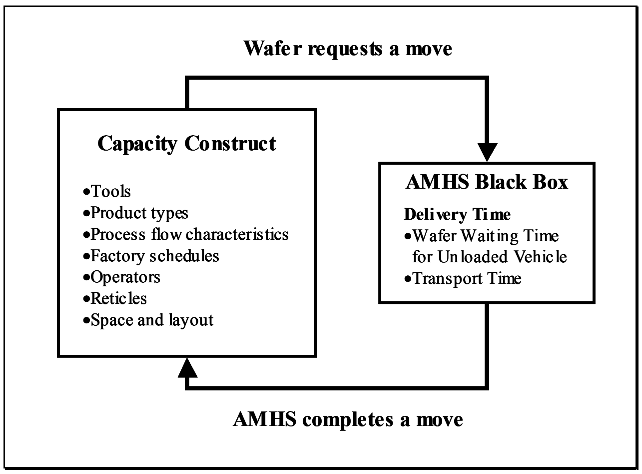

Figure 2: Architecture of the Capacity Model with "From-To" AMHS Delivery Times

Experiments have shown that it takes significantly less time to run the capacity models with AMHS delivery times than what it takes to run the fully integrated model. Clearly, abstracting the detailed AMHS construct and eliminating the overhead required to integrate the two constructs are the two factors that contribute to the significant reduction in execution time.

4 CASE STUDY

4.1 Description

The primary focus of this case study is to demonstrate how to dimension AMHS inter-bay vehicle capacity efficiently by using a methodology based on the capacity models with AMHS delivery times. To determine vehicle capacity, we first produce a curve plotting mean cycle time against different number of vehicles (i.e. 12, 15, 25, 35, 45, 55, and 65 vehicles). We then select the AMHS capacity configuration that gives the lowest cycle time value. To assess the value of this method, we compared solution quality and computational time against the performance of the fully integrated model.

4.2 SEMATECH 300mm Wafer Fab Model

To conduct the proposed study, we used a detailed simulation model of a hypothetical 300mm wafer fab (International SEMATECH 2001). The capacity simulation construct was built in AutoSched AP vs. 7.2 (Brooks 2002). This construct contains the following components:

- 10 types of products (including hot lots).
- A processing route of approximately 316 steps.
- 60 different tool groups.
- Tool downtimes, batching, setups, reticles, and initial work-in process, etc.

The AMHS simulation construct was built in AutoMod vs. 10 (Brooks 2001). This construct contains the following components:

- A layout with 24 bays and 24 tool positions for each bay.
- One inter-bay loop.
- 24 intra-bay loops.
- 48 stockers (two stockers for each bay).
- Two I/O ports for each stocker.
- Detail characteristics of inter-bay and intra-bay vehicles (i.e. speed, load/unload times, downtimes, etc.).

The SEMATECH 300mm wafer fab model can run in a fully integrated mode, or in abstract mode by using the capacity construct and a detailed "From-To" table as the one sampled in Table 1.

Table 1: Sample "From-To" Table of Delivery Times

AMHS System	From Location	To Location	Wait Time for Unloaded Vehicle	Transport Time
Main	b1_out1	b13_out1	87.129	54.975
Main	b2_out1	b3_out1	57.734	56.419
Main	b2_out1	b4_out1	276.195	56.419
. . .	. . .	. . .	. . .	. . .
Main	b24_out1	b9_out1	82.768	56.419
Main	b24_out1	b10_out1	238.298	56.419
Main	b24_out1	b11_out1	90.307	56.419

4.3 Discussion of Results

To estimate mean cycle time at each AMHS capacity point, we generated five (5) replications with the SEMATECH 300mm wafer fab model. Each replicate was run for 100 simulation days; statistics were cleared at 50 simulation days to reduce initialization bias.

Figure 3 shows delivery time estimates at different AMHS capacity levels. This figure indicates that transport times and wafer wait time at stocker for unloaded vehicles change as the number of vehicles increases. This means that the capacity model with AMHS delivery times needs a “From-To” table for each vehicle point (i.e. a total of seven “From-To” tables). For this case study, each “From-To” delivery time table was produced from a long simulation run (i.e. 100 simulation days) of the corresponding fully integrated model.

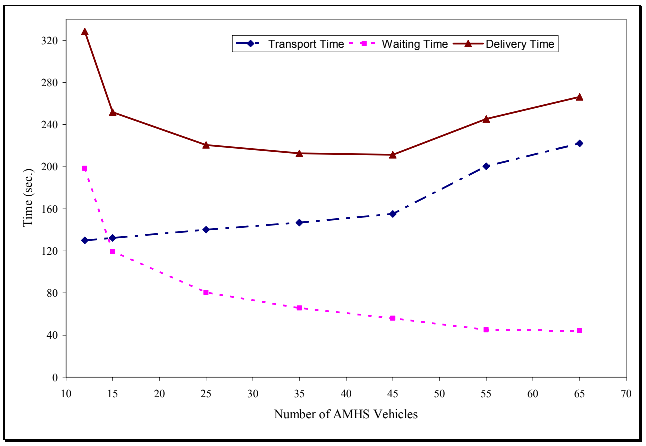

Figure 3: Delivery Time Estimates at Different AMHS Vehicle Capacities Produced by the Fully Integrated Model

Figure 4 depicts the curves plotting cycle time against number of AMHS vehicles. The curve produced by the fully integrated model is represented by the solid line with “♦” symbols, where as the curve produced by the capacity model with delivery times is represented by the dashed line with “■” symbols.

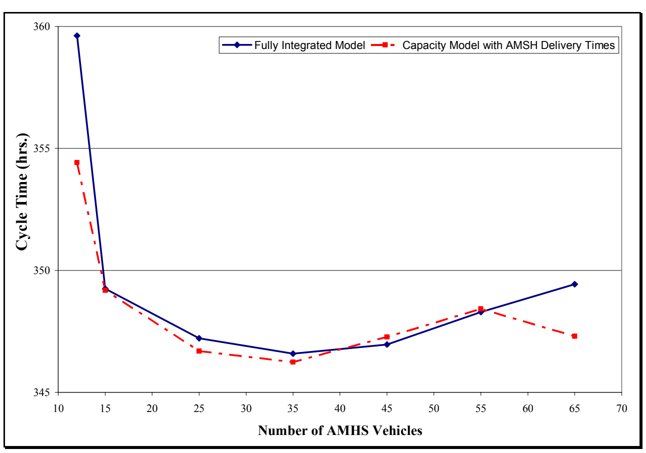

Figure 4: “Cycle Time vs. AMHS Vehicle Capacity” Curves at Different Levels of AMHS Detail

The simulation results generated by the fully integrated model shows that the highest cycle time value is approximately 359 hrs. at the 12-vehicle point (i.e. the minimum number of vehicles). Further experimentation done with this model indicated that using fewer than 12 inter-bay vehicles results in an unstable system because the equipment bottleneck was operating at full capacity.

The significant factor affecting cycle time at the left side of the curve is wafer wait time for vehicles. Cycle time curve thus shows a decreasing pattern between the 12-vehicle point and the 35-vehicle point as the wafer wait time factor decreases in value. Cycle time reaches its lowest value (approximately 346.58 hrs.) at the 35-vehicle point. However, as the AMHS gets more congested, cycle time values start climbing up again.

Our experiments also demonstrate that a similar solution can be obtained by using the capacity model with delivery times. The less detailed model estimates that the lowest cycle time is approximately equal to 346.24 hrs. at the 35-vehicle point. Such cycle time prediction is 0.100% different than the value estimated by the fully integrated model.

The cycle time values predicted by the capacity model with delivery times is, in general, similar to the predictions generated by the fully integrated model. As it can be observed in Table 2, the relative difference in cycle time estimations between these two models is less than two percent. The largest percent relative differences occur at the two extremes of the curve, where delivery time variability is large. One of the reasons why solution quality is different at extreme points is because the capacity model with delivery times can’t explain AMHS variability accurately.

Table 2: Relative Difference in Cycle Time Predictions

Number of Inter-bay Vehicles	**% Relative Cycle Time Difference**
12	1.447%
15	0.020%
25	0.153%
35	0.100%
45	-0.090%
55	-0.038%
65	0.613%

The value of the capacity model with delivery times is its small computational time. Table 3 shows the CPU time required to run one replication of the SEMATECH 300mm wafer fab model for each level of AMHS modeling. What these results suggest is that for each run made with the fully integrated model, approximately 100 runs can be executed with the less detailed model.

Table 3: Computational Time of Fab Simulation Models

Level of AMHS modeling	**Computational Time (seconds)**
Fully Integrated Model	37,056
Capacity Model with Delivery Times	372

5 CONCLUSIONS

This paper shows an efficient method based on the capacity model with "From-To" delivery times. This method enables an integrated analysis of the tool capacity, the AMHS, and the potential capacity losses caused by tool AMHS interactions. The results of our experiments demonstrated that the fidelity of the capacity model with AMHS time delays is similar to the prediction quality of the fully integrated model. For future work, we will define procedures that help us abstract the components of an AMHS accurately so that prediction quality of less detailed capacity AMHS models can be improved.

REFERENCES

Brooks Automation, 2001. *AutoMod User's Manual*, Volumes 1 and 2.

Brooks-PRI Automation, 2002. *AutoSched AP 7.2 User's Guide*.

International SEMATECH. 2001. *SEMATECH Simulation Model: Documentation and User Guide*.

Jimenez, J. A., G. T. Mackulak, and J. W. Fowler. 2005. Levels of capacity and material handling system modeling for factory integration decision making in semiconductor wafer fabs. Working paper.

Shelton, J. 2003. Fabless vision. *Future Fab International* (Issue14).

AUTHOR BIOGRAPHIES

JESUS A. JIMENEZ is a Ph.D. candidate in the Industrial Engineering department at Arizona State University. His research interests are in modeling and analysis of complex manufacturing systems, especially semiconductor manufacturing; manufacturing applications of operations research, computer simulation, and statistics; and the reduction of time required to build, experiment, and analyze complex computer simulation models. His email address is jesus.jimenez@asu.edu.

GERALD T. MACKULAK is an Associate Professor of Engineering in the Department of Industrial Engineering at Arizona State University. He is a graduate of Purdue University receiving his B.Sc., M.Sc., and Ph.D. degrees in the area of Industrial Engineering. His primary area of research is simulation applications within manufacturing with a special focus on semiconductor manufacturing.

JOHN W. FOWLER is a Professor in the Industrial Engineering Department at Arizona State University. Prior to his current position, he was a Senior Member of Technical Staff in the Modeling, CAD, and Statistical Methods Division of SEMATECH. His research interests include modeling, analysis, and control of semiconductor manufacturing systems. Dr. Fowler is the co-director of the Modeling and Analysis of Semiconductor Manufacturing Laboratory at ASU. The lab has had research contracts with NSF, SRC, International SEMATECH, Infineon Technologies, Intel, Motorola, ST Microelectronics, and Tefen Ltd. His email address is john.fowler@asu.edu.