MODELING FINAL ASSEMBLY AND TEST PROCESSES IN THE SEMICONDUCTOR INDUSTRY

J. David Liljegren

Motorola, Inc.
1303 E. Algonquin Rd.
Schaumburg, IL. 60196

ABSTRACT

Converting wafers to usable products culminates a production process which many consider to be the most complicated type of manufacturing today. The testing process is characterized by long recursive flows through multimillion dollar test equipment. This paper describes two simulation efforts currently underway at Motorola used to evaluate current procedures and policies at the individual tester level as well as the overall system.

1. INTRODUCTION

The process of converting semiconductor wafers into individual useable devices is the task of the final manufacturing group. Wafers are delivered to final manufacturing where they begin the complicated process at a probe step.

The probe process operates on wafers lots of a single product family. A probe indexes between devices on a wafer in a fixed interval, while the probe time is device dependant. A product specific orientation takes place before each wafer is probed. A number of wafers within a lot may be visually inspected for alignment and damage during any portion of the process. The probe handler can be stopped during the inspection process. Probe test mainframes can be shared with device test.

Following probe completion, wafers are sent to the assembly area as logical die. The assembly area transforms the individual die into raw stock. Detailed modelling of the assembly area will be addressed in a later study, but will be represented as a "black box" in the overall system model.

The test area is a major focus for this study. The test area consists of four distinct processes: test, mark, visual inspection, and pack and label. These processes and burn-in are linked one or more times in different routings.

Test systems are made up of a single mainframe and one or two test heads. The mainframe is shared between the two test heads in an alternating fashion. When a device is presented for testing, a signal is sent to the mainframe. If the mainframe is available, the testing

begins, else, the test begins when the second head relinquishes control of the mainframe. At the beginning of each test lot, there is a period of calibration which requires the mainframe. During the calibration period, no other processing takes place as this operation requires the total mainframe resource. A graphical representation is shown in Figure 1.
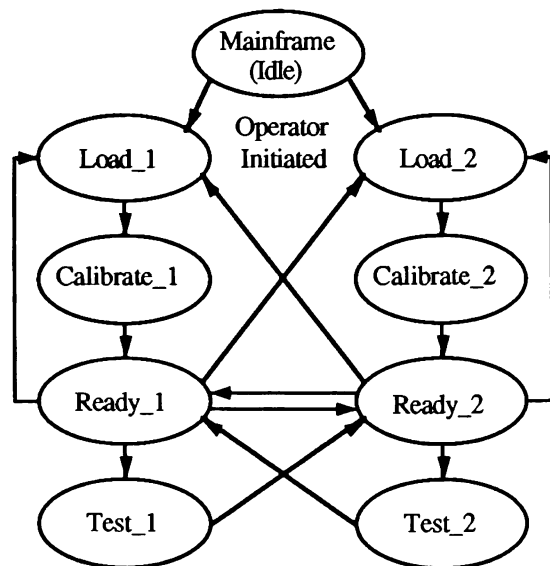

Figure 1. Test System Operation

Each test head can have a different handler for device presentation to the mainframe. Manual handlers require an operator to physically load each device, initiate the test, and unload the device. Manual handlers are only used for room temperature tests.

Tests requiring any change in temperature are tested on universal automated handlers. The automated handlers require an operator to load trays of devices into an input portion of the handler. Automation then takes four devices at a time and presents them to the core of the unit. If the test requires a temperature other than room, the devices are placed into a soaking chamber. When the

devices reach the desired temperature, the group of four devices is presented to be tested. Only one device is tested at a time, allowing control of the mainframe to shift between test heads.

When all four of the devices have been tested, automation carries the devices to the output portion of the machine where they are "binned" in trays according to test results. The operator is then required to unload full trays.

This test operation may be repeated several times in a process route depending on the the product and test specification. Again, systems are extremely capital intensive and very flexible. Because of this, they are a major potential bottleneck.

2. INDIVIDUAL TEST SIMULATION

The first simulation project focused primarily on test systems. Several objectives were identified to be addressed by the simulation.

First, we needed to characterize time multiplexing strategies to evaluate single test head, double test head, single probe head, double probe head, or mixed test and probe heads strategies. The resulting data generated by these runs would be used later in a complete system model.

The model was to be used as a tool to prioritize engineering efforts to increase the effectiveness of the system. And finally, to evaluate various load and unload strategies for the associated material handlers.

The models were built using the WITNESS simulation software. To accurately represent the process, a method was devised to consistently insure that both the individual test head and mainframe operate in parallel, and have the alternation take place, when applicable, between heads.

A single product was introduced with uniformly distributed lot sizes. A single mainframe test unit was used while test heads were changed between experiments. Three different test heads were incorporated, an automated handler, a manual handler, and a handler dedicated to probe operation. Demand for the probe operation was at the wafer level, and held constant across all probe runs. On the automated handler, test temperatures were randomly changed between room and non-room temperatures. A single operator was incorporated to load and unload products. Lastly, re-testing on both the automated and manual test handlers was introduced. Historical data was gathered to generate component reliability.

Six experiments were run to illustrate the possible combinations: 1. two automated test handlers, 2. two manual test handlers, 3. one automated and one manual test handlers, 4. two probe handlers, 5. one probe handler and one automated test handler, 6. one probe handler and one manual test handler.

Typical cycle time and throughput results are shown in figure 2, and figure 3.
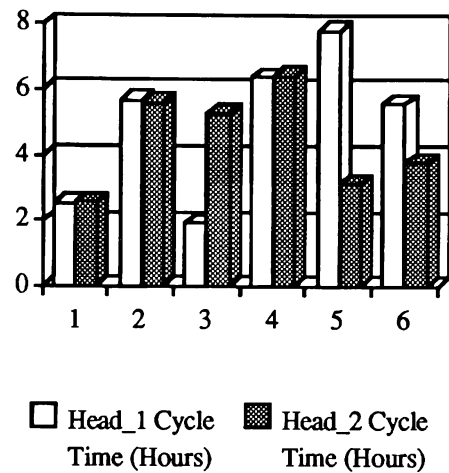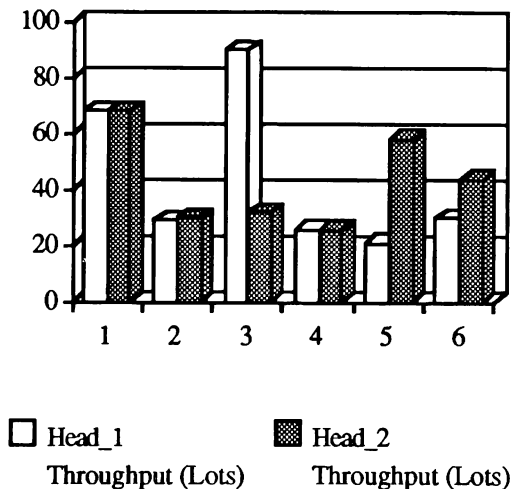

Figure 2. Lot Cycle Times by Head

Figure 3. Throughput by Head

A number of engineering evaluations were based upon the two automated handler model.

The introduction of small engineering lots burdens the system with many additional setups or system calibration. To help understand the impact of small lot sizing on setup and system utilization, a series of lot sizes was introduced to the model. Lot sizes were varied between 50 and 600 devices to force the individual setups. An interesting observation is illustrated in figure 4. in that although the relative throughput of the model increases with lot size increases, the utilization of the

mainframe stays relatively flat. This is largely due to the calibration step which requires the total mainframe resource to complete, and keeps throughput low.
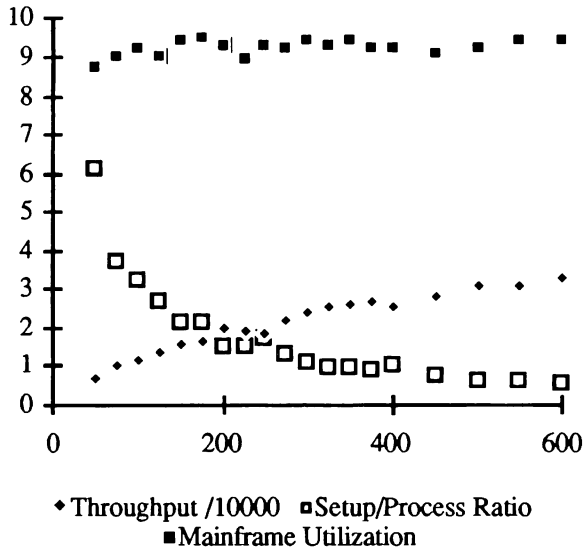

* Throughput /10000   □ Setup/Process Ratio
■ Mainframe Utilization

Figure 4.

The next step in the engineering analysis was to determine logical improvements in current equipment, and possible design changes in future equipment. To be determined was the impact of changing certain time variables on throughput. Targeted were the automated load time, the test times themselves, and the automated unload time.

Additional runs to determine the impact of eliminating a test, and the introduction of parallel testing required the full system simulation to obtain accurate results. These runs will be discussed later.

To model the automated handler correctly, we needed to capture the system operation. A robotic arm loads two device at a time into a 'boat', which holds four. The boat slides into the core of the unit where another robotic arm picks up all four devices and either presents them to the tester, or places them into a soaking chamber for temperature correction. The devices are then prsented to the test head and tested individually. They are then placed into an output boat which slides into the output portion of the handler. The devices are then placed into individual trays by another robot, based upon the test results.

The actual test times themselves are shorter than either the load or unload times, however, because all the devices had to be tested before they could be removed from the core, the sum of all the tests was approximately double the load times. As illustrated in figure 5, varying load times had little impact on system throughput or
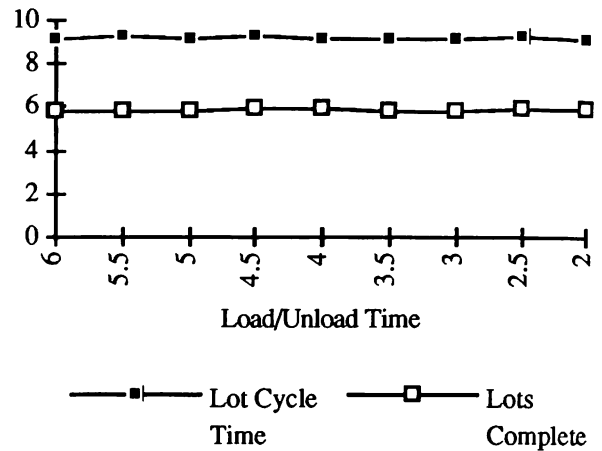
cycle times.

Figure 5.

The only variable which would have a significant impact on either throughput or cycle time is the actual test time itself. Again, the test time was varied over a range to determine what the impact might be. An increase in throughput and a cycle time decrease resulted until we reached the level where load and unload times, held at their initial values, once again pace the operation. The results are illustrated in Figure 6.
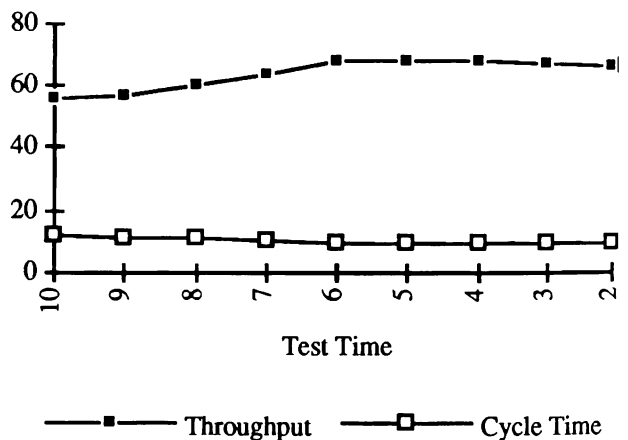

Figure 6.

3. TEST SYSTEM SIMULATION

The test system simulation phase of the project included the major processes which a product must complete before being delivered to a customer. Because the major focus of this project was the test operation, distributed times and capacities were used to represent the all other remaining processes.

Incorporating the individual test and probe system simulation models brought about some unique problems. The most serious being the difficulty in simulating the dynamic factory floor of the test environment. The models developed earlier used a single mainframe with a static pair or combination of test/probe handlers. The factory has several such mainframes where handlers are interchangeable with virtually every mainframe.

The decision was made to investigate a simulation package called TestSim. TestSim is a very easy to use simulator developed by Tyecinb Systems Inc. designed specifically for semiconductor final test operations.

TestSim allowed the input of the additional steps required to complete the product flow illustrated in figure 7. The test flow itself is product specific and may include several test events, multiple burn-in events, a marking event, and a pack and label step.
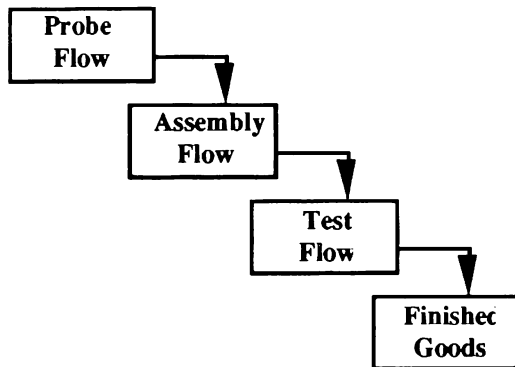

Figure 7. Typical Product Flow

Times were obtained by reviewing and characterizing months of data and placing them into the proper fields in TestSim. Experiments were developed to determine the impact of changing several test parameters.

The first runs used the existing data to establish a baseline which future results were compared to. The version of TestSim used for the models was a beta version which for the first time included the ability to model multiple test heads on a single mainframe. To validate the data obtained from TestSim, benchmarking runs were made from both software packages, and the results compared.

Different from the individual tester models was the introduction of multiple products, and complete test flows. Again, to make the initial validation possible, a single mainframe with two automated handlers were used.

Three experiments were proposed to examine potential policy changes. Earlier, possible cycle time reduction was realized by reducing the individual device test times. Variations of this experiment were run to see the affect on upstream and downstream operations. First, the test times were reduced by half. The elimination of a single test event was examined. And finally, the possibility of

testing multiple devices in parallel was attempted. Unforetunately, multiple device testing is not currently supported by TestSim. This feature is forthcoming in a future release. The results of the runs are indicated in figure 8.

Both runs show a marginal increase in throuput, (approximately 14% and 16% respectively), however, the most impressive gains were made in overall cycle time reduction (73% and 75%), and WIP reduction (69% and 71%).
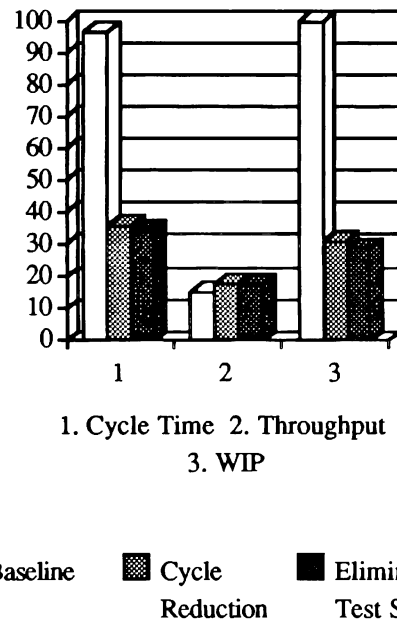

1. Cycle Time 2. Throughput
3. WIP

☐ Baseline   ▨ Cycle   ■ Eliminate
                Reduction   Test Step

Figure 8. System Run Results

CONCLUSIONS

The results used in this paper are no means final. The validation of TestSim is not yet complete. We will continue to work the people from Tyecin Systems to improve TestSim. Efforts are currently underway to introduce parallel testing as well as alternate testing. A similar effort exists to add detail to burn-in constructs.

The individual test cell models will continued to be used for additional equipment experiments, and other engineering projects. These models will also be used to continue the validation of the TestSim Models.

Determining capacity of final test operations is becoming easier with the introduction of TestSim. However, it is not useful to construct detailed models of individual pieces of equipment. For this kind of detail, a more general purpose simulation language or manufacturing simulator like WITNESS will need to be used.

860 *Liljegren*

REFERNCES

AT&T ISTEL Limited, WITNESS User Manual, Version 7.3.0, April 1992.

Tyecin Systems Inc., TestSim User Manual, Version 3.0 (1992)

AUTHOR'S BIOGRAPHY

DAVID LILJEGREN is a Senior Simulation Engineer with Motorola, Inc. Before coming to Motorola, David was a Technical Consultant for a simulation software company. For the last two years he has helped internal organizations throughout Motorola employ simulation as an engineering tool. David holds a B.S. degree in Industrial Technology from Western Illinois University, and did his graduate work in Industrial Computer Applications at Illinois State University.