AutoPlan: Planning and Scheduling Semiconductor Assembly/Test Operations Using Simulation

Stephen P. Smith, Naiping Keng
Intel, Components Automation
CH3-68, 5000 W. Chandler Blvd. Chandler, AZ 85226
and
Lito Soto, Cecille Bautista
Intel A2/T9 Assembly Test
T9, 5548 South Superhighway 1233 Makati, Metro Manila., Philippines
and
Joey Skinner, Don Burdett
AutoSimulations, Inc.
655 Medical Dr. Bountiful, UT 84010

Introduction

AutoPlan is a production planning and scheduling system for semiconductor Assembly/Test (A/T) facilities that is being co-developed by AutoSimulations, Inc. (ASI) and Intel. It is currently being piloted at Intel as a decision support planning tool and as a shop-floor lot dispatching aid. It supports the move from weekly production due-dates to daily due-dates while maintaining or improving on-time delivery performance. It is expected to reduce planning cycle time from a few days to less than one day with stable or reduced shop-floor inventory.

This paper will present the need for a system like AutoPlan, outline the key problems to be solved, and describe the basic types of data needed to support A/T planning and scheduling. One key need we identify is the ability to predict day-to-day order performance, which suggests a solution which utilizes a discrete-event simulator with a detailed capacity model of the A/T facility. To support this solution, we will also describe the need for a backward-product mapping (BPM) module which determines die-to-order mappings and inventory release schedules from order due-dates. The paper will then present the basic algorithmic approach and components chosen for the AutoPlan system. Finally, we will explain how the system is being piloted at Intel and identify key learnings and results from initial implementation.

Need and Current Situation:

As a semiconductor manufacturer, Intel must respond quickly and efficiently to changing business conditions. A major business objective is to improve Intel's customer order delivery performance while reducing supporting costs, including direct inventory costs and indirect headcount costs.

Corporate-wide, Intel is implementing the concept of building to customers orders starting at a "die inventory" point mid-way in our internal supply chain. This point is situated between Fabrication (where the basic die are patterned onto raw silicon wafers) and Assembly/Test (where die are separated, packaged, tested, and packed for shipment to the customer). Before that inventory point, the fabrication process, because of it very long cycle times, must build to predicted customer volume, which is based on forecasted, but not necessary firm or placed, customer orders.

Currently, all A/T site planning is done manually on a weekly cycle, with all customer's demand volume for a product rolled up into a single weekly number. This year, Intel will put in place a corporate system which gives each manufacturing site actual customer daily demands for the coming weeks.

Problem(s) Addressed by AutoPlan

A key component of the Build-to-Order concept is the ability of each site to respond to the detailed daily demand set. Manual implementation of this would have resulted in a significant increase in the number of planning staff at each site. AutoPlan provides:

1. The ability of the A/T staff to maximize the performance of the site in responding to daily customer orders,
2. A tool for planning which bases lot launch schedules for all inventory store points on customer orders, and
3. A tool for manufacturing to drive the flow of WIP in the line via actual customer need dates.

Approach

To provide the above capabilities, AutoPlan must be able to predict day-to-day order performance, that is, AutoPlan must be able to accurately predict how lots will move in the A/T factory. The technology which offers the most promise for accurate predictions from simple data (with some downside in the large volume of data required) is discrete-event simulation with a detailed capacity model of the A/T facility. While complex mathematical models [1] are possible, determining and maintaining the key parameters in these type of models is felt to be more difficult that maintaining the "virtual" analogy of the real world in a simulation model [2].

While the simulation gives predictions on lot flow, there remains the determination of:

1. which lots (either in process or in storage) should be used to satisfy each demand (i.e., the lot to order mappings)
2. the time at which lots in storage should be launched to meet the demand(s) to which they are mapped.

To provide these capabilities, AutoPlan contains a backward-product mapping (BPM) module. As designed, AutoPlan's BPM has no view of capacity, so that it provides only an initial solution point while the discrete event simulator checks capacity constraints, modifying lot release times based on when machine capacity is actually available. We will discuss the details of the algorithms in AutoPlan after briefly explaining its input data.

Planning and Scheduling Data

Input data provided to AutoPlan includes both data standard to AutoSimulations' scheduling software, AutoSched, and data specific to AutoPlan.

The standard data includes a description of all equipment used in the Assembly / Test facility. The equipment is grouped into families, which are used in the product routes. Equipment have calendars that define down times, preventative maintenance schedules, shifts, and other availablility constraints.

Product routes are detailed descriptions of the steps required to manufacture specific products or groups of products. Both the finished product and intermediate categorization of product are called part types. Routes are associated with part types. Step data includes equipment and other resource requirements, setup and processing delays, and other step-specific information.

Binning and standard lot size information is used by BPM to create new lots from the existing inventory to fill demand orders.

Data for AutoPlan includes information about storages. Storages are locations where WIP lots are allocated to demand orders. Pieces from several different lots may be combined into a new lot, which follows the routing(s) necessary to become the part type of the demand order to which it is allocated. Storage data include flags for whether or not to combine or split lots when creating new lots. Another flag specifies whether the product in this storage is downgradable. Downgrading is the ability to use alternate part types and bin classes to satisfy customer demand.

The main AutoPlan input file is the backward product mapping table. The backward product mapping table lists finished (target) parts and all of the inventory (source) parts that can fill demand for each target part. Some target parts are final demand parts. Others become source parts for other targets (see the table below):

TARGET	SOURCE
partE	partD
	partC
partD	partB bin1 only
partC	partB bin1
	partB bin2
partB	partA

Assume that no storages are downgradable and there is a demand order for partE. BPM looks for parts of the first source type, partD, then parts that become partD: partB bin1 and then partAs that bin to partB bin1. If not enough of these are found, then BPM looks at partE's second source part, partC, then parts that become partC: partB bin1, partAs that bin to partB bin1 (which don't exist, based on our first search list), and finally partAs that bin to partB bin2.

Downgrading affects the order in which the source parts are searched. For example, if the storage where parts partD and partC are stored is downgradable, BPM would look for partDs first then partCs before searching the table for source parts that become partDs.

Details of BPM

The current implementation of AutoPlan consists of two modes: BPM, which represents the system's capacity through average queue times at each step in a part's routing; and DFSM (Detailed Forward Simulation Module), which uses finite capacity. The BPM mode is designed to be less detailed than the forward simulation. The results from BPM are checked by the simulation and, if BPM's results are desirable and achievable, the lot release schedule from the DFSM is used by planners to drive the facility.

BPM consists of three submodules: rough forward simulation, preallocated demand, and demand allocation.

The rough forward simulation advances all WIP lots through their routings and places them in the appropriate storage for their part type.

Preallocated demand is a file that assigns WIP lots to demand orders prior to BPM's allocation mode. BPM determines a route through the backward product mapping table for the lot to take in order to become the demand part. BPM then moves the lot through that route, similar to the rough forward simulation process, calculating yields, processing times, queue times, etc., in order to calculate whether the lot will be on time for that demand and what quantity of pieces will be available at the end of the routing. If the preallocated lot(s) satisfy that demand, BPM's allocation process will skip that demand order.

The final submodule is the allocation submodule. BPM now tries to satisfy any demand orders that were not filled by preallocated demand by traversing the backward product mapping table. Lots that can meet the demand order are moved through their routings, taking average queuing and processing delays, yielding out scrap and other bin classes.

BPM estimates the completion time and final quantity of pieces that might be allocated to the demand. BPM groups the lots by those that will finish on time and those that will be late.

Ontime lots are sorted by several critiera: the order that their part type was listed in product mapping search; final completion date/time; completion date/time at current store; and lot size. Each subsequent criterion acts as a tie-breaker for the previous criteria. If enough ontime lots are found, late lots are not considered. Late lots are sorted based on final completion date/time.

New lots are created to "consume" the lots needed to satisfy the demand order. BPM records when these new lot(s) need to be released from a storage and which pieces they need to consume from each lot. This information is stored in the release schedule, which is used as input to the forward simulation. Lot size for the released lots are based on the minimum, standard, and maximum lot sizes defined for that lot's part type.

The process of creating the new lots needed for a demand order continues until every piece allocated to it has been moved through the factory to become finished product.

Once one demand has been satisfied or it has been determined that there is insufficient WIP (on time or late) to satisfy it, the next demand order is allocated. This continues for all demand orders. BPM produces the release schedule file for DFSM.

Details of DFSM

DFSM does not use the product mapping table. It relies on the release schedule file to tell it which lots consume other lots. DFSM uses more detail than BPM, such as rules that consider BPM's due dates and equipment preferences for processing parts. DFSM is a fully-featured simulation model that contains capacity constraints and the details to model the user's facility accurately.

If the results from the simulation are satisfactory in terms of how many demands were met on time, late, or not at all, then the output can be used by production people on the floor. If results are not satisfactory, there are several options which can be adjusted to produce better output. Better results could be obtained by changing the order of the demands in the demand file, using more accurate queue times, etc.

Pilot Results

AutoPlan is being piloted in Manila on one of Intel's Flash lines. This line produces about a million units (die) a week and was chosen for pilot because:

1. It has reasonably high volume,
2. It has fairly complex product mapping,
3. It has reasonable stable capacity and customer demand.

The first phase of the pilot consists of generating the current manually determined production goals with the AutoPlan system. Phase 2 will implement detailed inventory store releases using AutoPlan generated release schedules (instead of being determined manually from production goals). Phase 3 will bring the due-dates attached to lots back into Intel's WIP tracking system (Consilium's WorkStream) where WorkStream Rule-Based Dispatch (RBD) module will dispatch all workcenters based on AutoPlan's assigned due-date.

It has taken about ten weeks to go from the completion of initial functional testing of the AutoPlan product to rollout of the system for the planners. During this period we validated that the product was able to generate production goals close to the performance of the manual planners. Figure 1 shows a graph of this performance comparision, which is based on a comparison of the percentage of line item product volumes that both plans (manual and AutoPlan) are able to achieve.

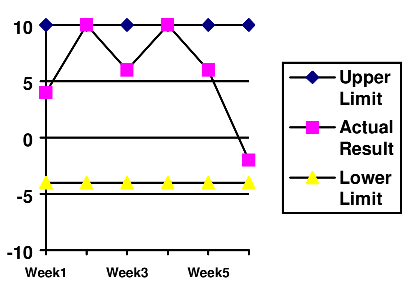

Figure 1: Line Item Schedule Performance of AutoPlan as a percentage of the manual planning result

Before the AutoPlan result could replace the manual result, we determined that we must be within a preset tolerance to the manual result for at least 4 weeks in a row. To reach this entry criteria, we had to perform the following tasks:

1. Fix and improve the software. Some changes included bugs not found during initial functional testing, for instance the release of in-transit lots before they actually arrived. Other major changes included BPM algorithm enhancements which made it more accurate in initial lots selection (Pass1).
2. Reconfirm and revalidated basic input data. Major changes were found to much of the product mapping data, as this data had been maintained manually.

After the entry criteria was reached, we immediately began goaling the line using AutoPlan results and reduced the number of planners from 6 to 3.

References:

[1] Conners, Feigin, Yao, and Kiang, Queueing Network Model for Semiconductor Manufacturing", to be published in IEEE Trans. in Semiconductor Manufacturing.

[2] Leachman and Sohoni, Automated Shift Scheduling As a Tool for Problem Identification and People Management in Semiconductor Factories, in Advances in Production Management Systems, E. Eloranta (ed.), Elsevier Science, North Holland, 1991.

Acknowledgments:

The authors would like to thank the management teams at Intel Components Automation Systems and Intel Philippines for their support of this project.

Comment:

Biography

Naiping Keng. Senior Software Engineer in Intel's Components Automation Systems and Co-manager of AutoPlan project. 1989 Ph.D. Computer Science, Southern Methodist Univerity.

Stephen P. Smith. Senior Software Engineer in Intel's Components Automation Systems and Co-manager of AutoPlan project. 1982 Ph.D. Computer Science, Michigan State University.

Lito Soto. A2/T9 Planning and AutoPlan A2/T9 project leader, Intel A2/T9 Assembly/Test, Operation Planning and Control Department.

Cecille Bautista. Manufacturing Engineering and AutoPlan A2/T9 project leader, Intel A2/T9 Assembly/Test, Operation Planning and Control Department.

Don Burdett. Senior Software Engineer at AutoSimulations, Inc. AutoPlan technical architect.

Joey Skinner. Development Manager at AutoSimulations, Inc. AutoPlan project manager.