A Planner and Scheduler for Semiconductor Manufacturing

Hugh E. Fargher, Michael A. Kilgore, Paul J. Kline, and Richard A. Smith

Abstract— **The Microelectronics Manufacturing Science & Technology (MMST) project includes two closely related CIM subsystems for planning and scheduling wafer production. The MMST Planner plans all work release into a factory so as to meet stated goals, and predicts work completion dates. The MMST Scheduler operates in real-time to determine the sequence of lot movements and machine loadings that will be performed on the fab floor. Both the Planner and the Scheduler continually maintain plans which are up to date with the factory status by incrementally replanning for unexpected events. The MMST Planner can be used as a decision support tool to rapidly analyze the consequences of various manufacturing decisions. Planning is performed using a modified beam search algorithm, and is based on a time-phased capacity model of the factory. Fuzzy arithmetic is used to model the uncertainty inherent in cycle time data. The MMST Planner is fully distributed, allowing simultaneous users in different parts of the factory. The MMST Scheduler uses a heuristic method called Score Tables to develop schedules of future events. The Scheduler evaluates event prerequisites to determine when to initiate lot transfers and machine loadings, and responds to any failures of execution.**

I. Introduction

The Computer Integrated Manufacturing (CIM) component of the Microelectronics Manufacturing Science & Technology (MMST) project formed an important part of the final project 1000 (1-K) wafer demonstration. This paper describes two closely related CIM subsystems used to manage production control, namely the MMST Planner and Scheduler. The MMST Planner was used to maintain a wafer release plan into the MMST factory, and to predict wafer processing completion dates. However, once released, control of wafer movement passed to the MMST Scheduler. The MMST Scheduler operated in real-time to determine the sequence of lot movements and machine loadings performed on the fab floor. This paper also describes some of the results obtained through using the MMST Planner and Scheduler subsystems during the 1-K wafer demonstration, together with possible future work.

II. MMST Planner

The problem tackled by the MMST Planner is that of determining when to release work into a factory so as to best satisfy customer requests, given the current manufacturing constraints. With this in mind, the MMST Planner has been designed as a *make to order* or *make to forecast* planning system, for use in complex job-shop manufacturing environments. The system maintains a work release plan, which determines when work should be released into the factory and, to a given confidence level, when that work will complete processing. The work release plan is generated so as to avoid starving or overloading bottleneck machines at any projected time in the plan. This in turn helps reduce work-in-process (WIP) and production cycle-times. Plans are incrementally updated in line with a user defined strategy which, for example, could be used to give preference to plans that meet customer requested due dates over those that simply maximize machine utilization. The MMST Planner functions as a decision support tool, continually maintaining an up-to-date plan and providing rapid analysis of user requests. Processing capacity of the factory is represented using a high level capacity model. Plans are generated using an artificial intelligence heuristic search technique which ultimately determines the recommended work release plan. This contrasts with the MMST Simulator [1] which may use a given work release plan as input to determine measures such as resulting production cycle-times. The MMST Planner is typically used to provide a rapid analysis of the consequences of operational decisions (such as when to release a particular order), in terms of how they would affect the factory in its current configuration. Again this contrasts with simulation, which is typically used to determine the consequences of more strategic decisions (such as work release policies or addition of machines) by running suites of simulations on differing factory configurations.

A. MMST Planner Goals

The overall goal of the MMST Planner is to provide decision support for production planning in complex manufacturing environments. Such a tool could enable improved customer satisfaction, while making better use of the production resources available.

There is considerable evidence to show the importance of work release in achieving the typical goals of semiconductor manufacturing [2]. Consequently, an important goal of the MMST Planner is to continually maintain an up-to-date plan, which determines when work should be released into the factory, and to predict when that work will be completed. The plan does not determine precisely when wafers are processed at particular machines, instead, it simply determines that machine processing capacity will be available during the time that work

Manuscript received August 1, 1993; revised December 22, 1993. This work was supported in part by the Air Force Wright Laboratory and the DARPA Microelectronics Technology Office under Contract F33615-88-C-5448.

The authors are with Texas Instruments, Inc., P.O. Box 655012, Dallas, TX 75265 USA.

IEEE Log Number 9400853.

0894–6507/94$04.00 © 1994 IEEE

is planned to be in the factory. All times are determined to within some *granularity*, defined by a time interval duration which is typically a day or shift. A plan horizon is maintained to equal a multiple of the chosen time interval duration. Although the MMST Planner determines work that may be released during each time interval, the MMST Scheduler determines when that work is released into the factory on a minute-by-minute basis. In this way, the MMST Planner can be thought of as shuffling a queue of work *outside* of the factory entrance. Higher priority work is inserted nearer the front of the queue, and is therefore planned for release at an earlier date.

Maintaining an up-to-date plan requires a plan representation which remains consistent with the current factory status and clock-time. To achieve this, the MMST Planner compares the planned and actual work progress in terms of the percent processing complete. Replanning is then performed on work that is ahead or behind plan. In addition, replanning may be performed when machines unexpectedly go down or become available for processing. Furthermore, the plan representation always covers the current clock-time up to the plan horizon.

The MMST Planner must also model uncertainty when planning work release and completion dates. This applies particularly to predicting work cycle-times once released into the factory. Very often, the best available data is in the form of previously observed cycle-time distributions and as a result the MMST Planner associates a confidence level with all planned work completion dates that it calculates.

The MMST Planner requires a strategy, which is used to guide plan generation. A strategy is composed of an ordered list of goals, which are used to define heuristics and constraints. Goals are used to sequence planning decisions so as to attempt to meet given plan measures, such as meeting all due dates or balancing machine utilization. For example, a goal to meet all requested due dates may have a heuristic which sequences work based on slack to due date. Constraints are used to limit plan measures such as planned WIP. A constraint to limit planned WIP would prevent generation of a plan that exceeds the stated WIP at any point in time. At any point in the planning process, decisions are sequenced using the first goal on the strategy's ordered list of goals that is relevant to the decision being made. The user may also define multiple strategies, but only the *active* strategy is used at any one time to sequence alternatives.

Part of the overall goal of the MMST Planner is to allow multiple, simultaneous users of the system, in a distributed environment. This is achieved by providing access to the MMST Planner from any connected workstation in the factory, and by managing the concurrency issues that arise when more than one person is using the system (such as when two users attempt to update the plan at the same time).

Finally, for decision support, the MMST Planner must provide a way of rapidly exploring the consequences of various manufacturing decisions. It achieves this by allowing rapid incremental updates to the existing plan to be explored, without having to necessarily commit to such updates. Consequently, the MMST Planner concentrates on the impact of decisions on the factory as it currently exists.

B. MMST Planner Approach

The planning algorithm used within the MMST Planner is described in detail elsewhere [3], [4]. This section gives only a brief overview of the algorithm.

The plan representation used within the MMST Planner has been devised so as to model the projected work load within the factory, while allowing incremental updates at any time to account for changing circumstances, such as new orders or machine failures. Allowing incremental updates for new orders enables rapid feedback to customers concerning feasible ship dates, without having to wait for a daily or weekly periodic plan update. Allowing incremental updates for machine failures provides an early warning for work that may be late. The plan representation is based on the processing capacity of machine groups within the factory, divided into contiguous time intervals of arbitrary duration. The use of contiguous time intervals is referred to as a *time-phased* representation, and allows a plan to be represented up to some predefined horizon to any level of detail. Consequently, the planning algorithm is used to determine work committed to each machine group during each time interval. In practice plan generation is influenced by the bottleneck machines, a feature shared by other planning approaches [5]. By using a time-phased capacity model, bottlenecks for a particular time interval can always be identified, since bottlenecks may vary over time.

The planning algorithm is divided into two parts, that of determining the sequence of work to be planned (given its requested due-date, customer priority, etc.), and incorporating the required work into the plan (given the current machine group commitments, type of planning request, and any constraints on which time interval the work may be planned for). Ultimately, any update to the existing plan (including replanning due to machine failure) can be tackled in this way. The MMST Planner strategy determines the sequence of work to be planned. To incorporate the required work, the MMST Planner algorithm uses a beam search similar to that used within other artificial intelligence planning systems such as ISIS [6]. However, unlike ISIS the beam width grows with search depth and uses a simple backtracking scheme to search within the beam. The search algorithm is ultimately searching for available processing capacity, over existing time intervals, to incrementally update the plan to accommodate any change in work commitments.

The plan representation must also model the uncertainty inherent in the entire production process cycle-time, since such cycle-times often form the best available data for planning. Uncertainty is modeled by reinterpreting the plan representation in terms of fuzzy sets, an approach which has been previously used within FSS [7]. However, whereas FSS generates an overall cycle-time distribution from all contributing process cycle-time distributions, the MMST Planner performs the reverse operation. Comparisons with simulation have shown the accuracy of this reverse operation [3]. The result can be used to determine the *degree of membership* of work commitment for each machine group within each time interval, which reflects the expected accuracy of the projected work load. Typically, projections become less certain the further they are made into the future. The final result is that planned work completion dates are computed to some confidence level. The overall cycle-time distribution can also be used to determine an earliest and latest expected completion date.

C. MMST Planner Architecture

One of the requirements of the MMST Planner architecture is to operate in a distributed environment. To satisfy this requirement the MMST Planner can be accessed from any connected workstation in the factory, and used to view the current plan, plan new work or perform *what-if* analysis. Clearly, determining the user's authority is important. Few users would have the authority to *accept* updates to an existing production plan, while many may have the authority to view the current plan. Experience has shown that while the main users of the MMST Planner work from within their office, it is important that they be able to use the same features on a workstation on the factory floor.

The MMST Planner is able to run in a distributed environment using two different processes: the server and the User Interface (UI) process. At any point in time there is one server process running and zero or more UI processes, all running on one or more workstations. The server process runs continuously, while UI processes are started and stopped depending on when users need access to the planner. The server process is responsible for maintaining the continuous *presence* of the planning system, and automatically replanning whenever the manufacturing environment significantly changes from the production plan assumptions. Examples of such changes include work release, work completion, or unexpected machine failure. While automated interactions go directly to the server, human interaction with the planner is performed using the UI process. Examples of such interaction include order entry, browsing the current plan, planning new orders, and performing what-if analysis. Since multiple processes can run simultaneously, concurrency issues are handled by preventing users from committing plan changes that violate other committed changes, and informing users of any plan changes that have recently occurred.

The server process also monitors the comparison between planned and actual work progress in the factory, therefore providing some feedback between the MMST Scheduler and Planner. If actual work progress deviates more than some user defined tolerance from planned progress, the work is automatically replanned. This provides a warning to system users that particular work may be deviating from its original planned completion date.

D. MMST Planner User Interaction

The MMST Planner enables the system user to explore the consequences of various manufacturing decisions, without having to necessarily commit to them. Decision support is divided into two types: implicit *what-if* and explicit *what-if* analysis. The two types of decision support reflect two different uses of the system. Implicit what-if is limited to analyzing the consequences of planning new work, or replanning current work, in the factory. Once performed, the updated plan can be either *accepted* by the system user (in which case it becomes the current factory production plan) or *rejected* (in which case all updates are discarded). For this reason, implicit what-if planning requests are limited to those for which the system user has the authority to execute. Explicit what-if analysis is used to analyze the consequences of a variety of operational decisions, such as when to bring a machine down for maintenance, or whether to continue processing particular work over a weekend. However, plans generated using explicit what-if analysis may not be accepted through the MMST Planner user interfaces, since they typically imply production decisions which are outside the authority of the planner system user.

Implicit what-if analysis can be used to plan new work with a variety of commands. The commands either plan new work non-disruptively (which guarantees that no existing planned release or completion dates will be affected) or disruptively (which makes no such guarantee). Disruptive planning requests are those which effectively reshuffle the queue of orders outside the factory door, while non-disruptive requests simply slip a new order into that queue. Furthermore, plan requests either attempt to plan work release so as to complete on a particular day (which invokes a backward planning algorithm) or plan release regardless of completion dates (which invokes a forward planning algorithm). Fig. 1 shows the MMST Planner screen that results from an implicit what-if request. In this case, plans are displayed in tabular form (as opposed to a Gantt chart), showing the current plan (lower), and the modified plan (upper) on the same screen. Notice that the system user can then compare the two plans before deciding whether to accept or reject the modifications. The four smaller lists in between the plan tables show other information such as successfully planned work, work which could not be planned due to capacity constraints, any work dislodged during the planning process (and not planned back in) and all remaining unplanned work.

Explicit what-if analysis is divided further into two types—*brief* and *extensive* analysis. As the name suggests, brief analysis is performed faster, but provides less information to the system user. For example, brief analysis may be used to determine which planned order items would be affected if a particular machine was to go down for 36 hours, but would not determine by how much their planned release or completion dates are pushed back. Extensive analysis, which employs full-scale incremental replanning, would determine precisely by how much the planned dates are pushed back. The advantage in having both brief and extensive analyses available is that system users can quickly browse the overall effects of a wide variety of decisions using brief analysis, before evaluating particular decisions in more detail using extensive analysis.

Fig. 2 shows the MMST Planner screen used to perform brief analysis. Notice that the particular analysis just performed is that of determining which order items are affected by taking the AVP526ChamberSpec machine down for 12 hours, the result being displayed in the scrollable list ("Order Items Affected"). Other brief analysis query types are listed on the screen in Fig. 2.

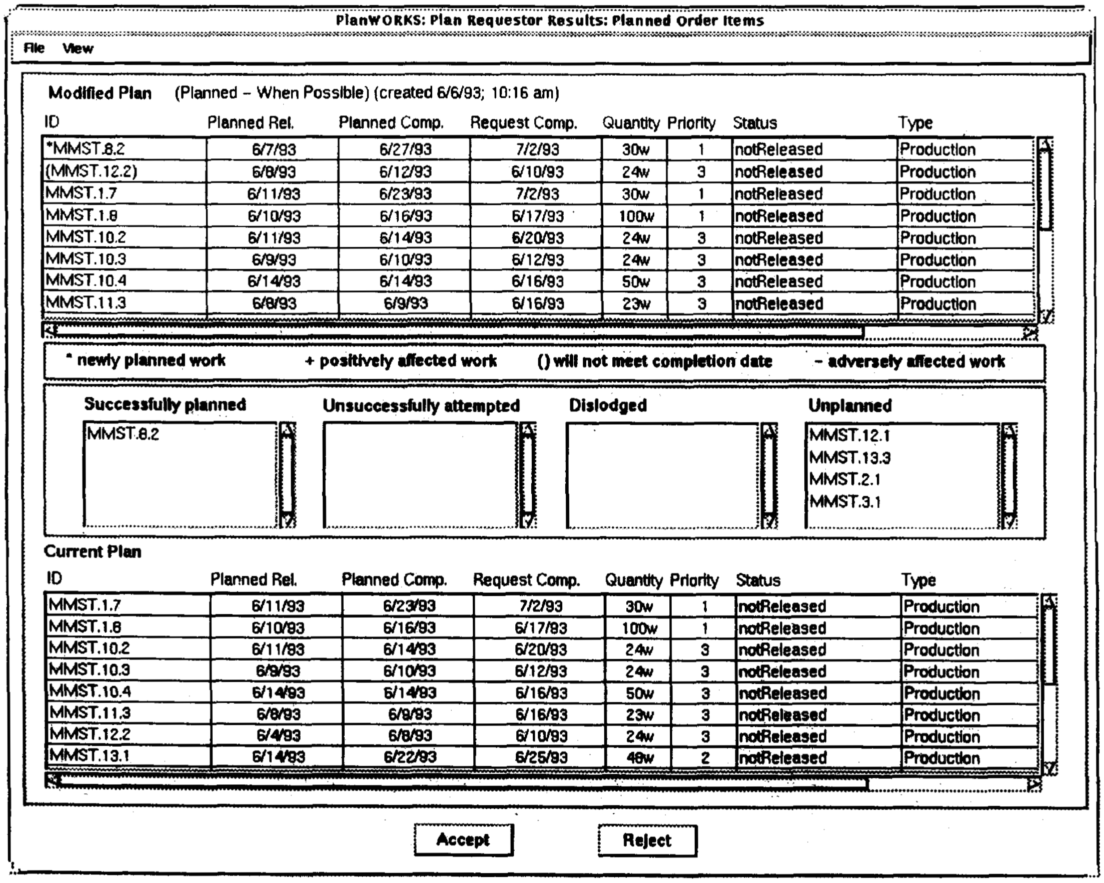

Fig. 1. Implicit what-if results screen.

Performing a single extensive analysis may require more than one screen, since each such analysis may consist of a whole group of planning requests. For example, the analysis may determine the effects of taking one of the Coater and Implanter machines down for maintenance over the next 2 days, while opening the factory to work on all critical order items over the following weekend in an attempt to make up the lost time. Separate MMST Planner screens are used to define the time period to take the machines down, as well as the time period to open the factory. Once all planning requests have been defined, the analysis is actually performed by selecting 'analyze' on the main planner extensive analysis screen. The result of the analysis is a screen similar to Fig. 1, used to compare plans before and after a request has modified the production plan. However, the difference is that the modified plan may not be accepted in this case. It can only be browsed, with a view to warning other personnel of the effects that would result if they decided to place the Coater and Implanter down, together with working on critical order items over the following weekend. If the Coater and Implanter were then actually placed down for maintenance, the MMST Planner would be informed of the change in machine status and would perform an automatic replan so as to maintain consistency with the current factory status.

E. MMST Planner Results

During the MMST 1-K wafer demonstration the MMST Planner was used to perform continuous and off-line planning operations. Continuous planning operations included all day-to-day tasks, such as order entry, planning and work release, as well as planner server execution. Off-line planning operations were performed by the system developers, as and when required, to assist in any decision making.

The day-to-day tasks were typically performed by the factory plan manager. Order creation and planning were generally performed one day prior to the day corresponding to the required release date for the order. One or more orders would be created, planned and prepared for release at a time. Preparing the orders ahead of time ensured that new work was ready to start when the MMST Scheduler was prepared to begin processing the work. In addition to the tasks performed by the users of the MMST Planner, the Planner server was available 24 hours a day. The server monitored progress of work in the factory and supported the work release operation.

As part of the day-to-day tasks, daily reports were produced detailing lot progress. The reports compared planned wafer moves against observed wafer moves, for various subsets of work. Both tabular and graphical versions of the report were

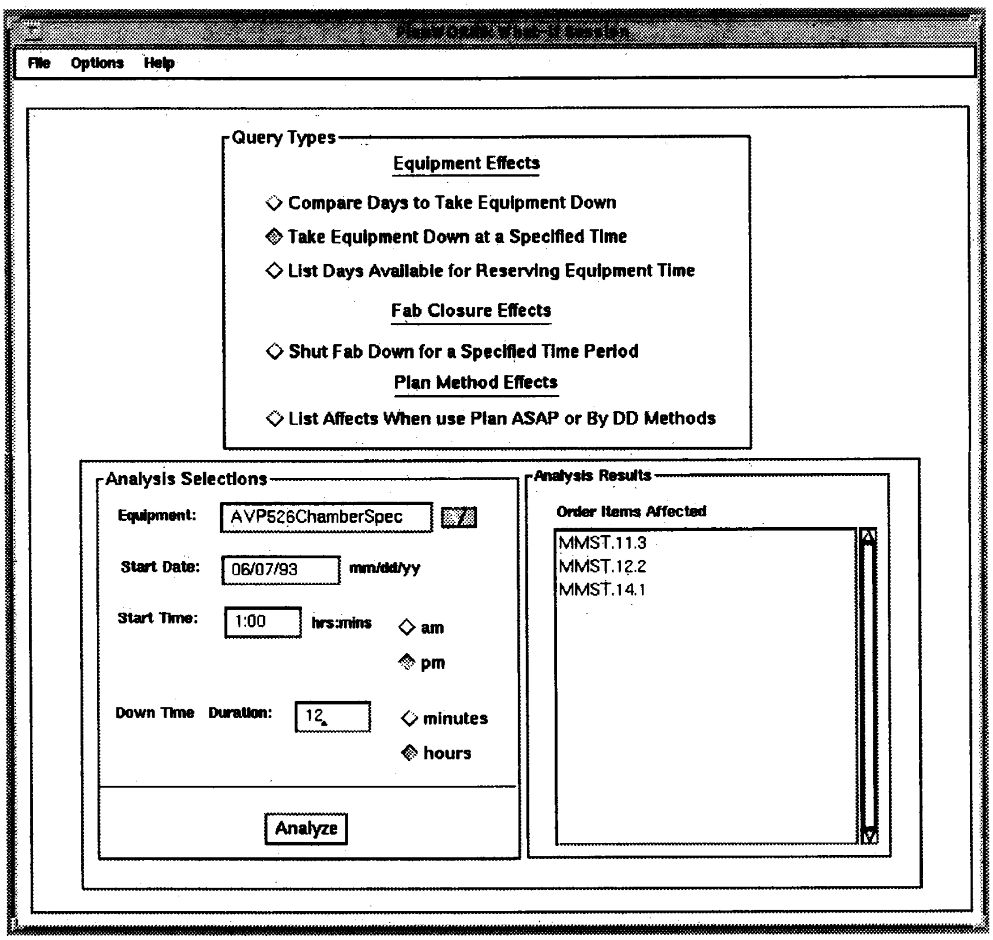

Fig. 2. Brief analysis screen.

generated and were used by managers to determine whether work was progressing adequately.

Historical data provided the opportunity to examine the accuracy of plans generated by the MMST Planner. Fig. 3 shows a graph of planned and actual remaining wafer moves for a subset of lots run during the demonstration, where one *wafer move* corresponds to the completion of one processing step, for one wafer, on one machine. The subset chosen in this case corresponded to larger lot sizes which ran a greater than three day cycle time. The labeled dark line (labeled at each point) shows the planned remaining wafer moves on a day by day basis. The unlabeled grey line shows actual progress of the lots through the factory. By comparing agreement between the observed and planned lines, a measure of the MMST Planner's accuracy can be determined. One measure is to calculate the mean percent difference between planned and actual remaining wafer moves for a given time period. This example results in a measured mean difference of approximately 9%. For simplification, Fig. 3 shows a snapshot of the plan generated on March 24th and does not illustrate the effects of subsequent replanning.

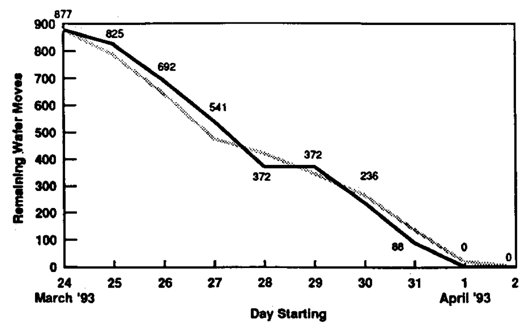

Fig. 3. Planned and actual wafer moves.

In addition to using the MMST Planner as the production planner during the demonstration, it was also used to perform several types of analysis in an off-line planning mode. Off-line analysis resulted in the generation and evaluation of plans that were not intended to be used as production plans; instead, they were generated to answers specific questions

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.