QUEUING THEORY MODEL FOR MANAGEMENT OF SEMICONDUCTOR FABS AND TEST FLOORS FOR OPTIMAL UTILIZATION OF THE EQUIPMENT

Term-paper
Name: B.Ashish
Roll number: 04010209

ABSTRACT:

The term paper gives a brief overview of application of Queuing theory to model the management in the ultra expensive environment of semiconductor fabs and test floors for optimal utilization of the equipment. This paper presents a analysis of Proper staffing in order to increase the throughput of the machines and maximize equipment utilization and overall throughput. This paper is about a more scientific and pro-active approach to the staffing issue than the classic methods used in the industry. The models described in this paper were developed by engineers from Tefen USA. To date, they have been adopted by management teams in over 30 different fabs and test sites over the past two years, both in the US and in Europe.

INTRODUCTION:

The studies involving productivity revealed that three parameters determine the throughput of the equipment-availability (uptime), utilization(how much available time is being used),and run-rate(the speed of the machine - U.P.H.).

Since run rates have not improved, for the most part, over the last 10 years (nor are they projected to improve in the next decade), equipment throughput is, in fact, is a function of the first two. Thus, targeting staffing levels towards improved uptime and utilization will result in better throughput.

The question is how to minimize production loss due to under staffing, without falling into the classic pitfalls of over staffing (reduced alertness, crowding, moderate work quality, etc.).

Traditional approaches to Staffing and their impact on productivity

Current staffing methodologies used in the industry are usually based on “number of activities per operator-hour’’ (operators), and direct labor estimations (maintenance technicians).

Easy and inexpensive as they may seemed, these methods hold significant disadvantages:

- Not responsive to the stochastic nature of machine operation and maintenance
- No distinction between constraint and non constraint equipment
- No knowledge of the impact on productivity (the non labor cost associated with making stifling decisions -"what-if' analysis)
- Built-in inaccuracies because of differences in labor content between "activities" in different equipment types
- Not responsive to changes in layout, automation or work methods
- Poor bench marlung capabilities because of differences in definition of activities in different fabs

QUEUING MODEL CONFIGURATION

An analytical approach to the problem was used first, based on Queuing Theory. Analysis of service frequencies and duration in different machines showed, respectively, close resemblance to Poisson and Exponential processes. For this reason, a multi-server-multi-client model, with finite source of customers (W c) , was chosen to model fab and test work centers. The model was applied in a spreadsheet environment, for cost and upkeep reasons. Two systems were created - one for operators (figure-4), and one for maintenance technicians. Both systems are modular and easy to construct.

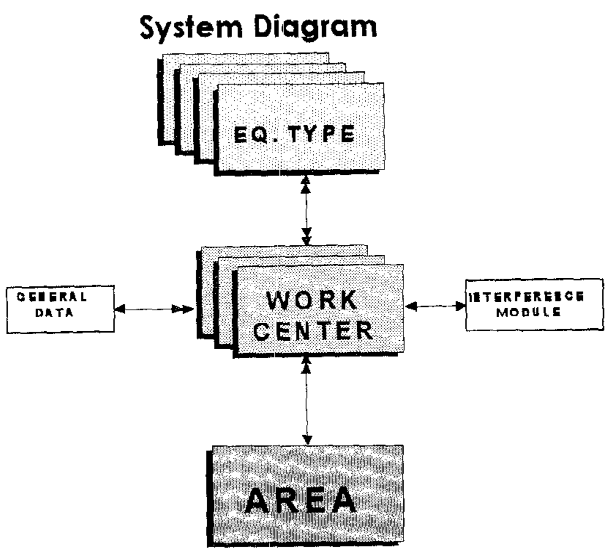

Figure 4

The first level in each system consists of the database - a separate spreadsheet for each equipment type, detailing all the labor operations associated with running or maintaining this equipment. Data includes duration, frequency and some other parameters for each operation. 'Tool spreadsheets' are then linked to their respective work centers. The double headed arrow linking tool to work center files stands for a many-to-many relation. Machines from one type can be located in more than one work center, and naturally one work center can contain more than one type of machine. The ‘work center’ files are all linked to the ‘area’ file, where the user can create various production scenarios and examine their effect on the required staffing level in each work center. In the operators’ case an ‘area’ usually pertains to a fab module (e.g. Etch), whereas a ‘work center’ can be a bay or even a part of a bay. In the maintenance technicians’ case, a work center is usually a cluster or a group of cluster tools serviced by a particular group of people.

Before and after - a case study of productivity improvement
The table below shows a comparison between the performances in particular area in a specific fab, before and after headcount figures was adjusted. Both operators and technician models were applied in this area, and as the table shows, a significant reduction in production loss was achieved.

Productivity Improvement - Case Study

Parameter	*Before*	*After*
Operator Related Idle Time	*10-20% across the floor*	*2-3% bottlenecks* *5-7% non-bottlenecks*
Floor Attendance by Operator	*60-70%* *Management Expectation: 80% (unrealistic)*	*75% Comprehesive Staffing Model*
Operators Non-value Added Activities	*50-70%*	*30-40%*
Maintenance Technicians Response Time for Bottlenecks Down Situations	*1-2 hours*	*15-20 minutes*

Summary

Adequate staS3ng of operators and maintenance crews to bottleneck and non bottleneck areas, is crucial to the overall productivity of semiconductor manufacturers. The use of advanced Queuing Theory and simulation models is highly cost-effective because it reduces human-related inefficiencies within the operation and maintenance of fab/test equipment. Two user-friendly packages were developed and successfully applied in over 30 different **fab** areas and test facilities.