Improving Service Efficiency in Manufacturing Integrated Circuits

Brian L. Farrell
William Kahan
Jay W. Reightler
Alan C. Woodard

This paper describes a methodology and related system, the Product Flow Database (PFD), that evaluates the performance of facilities that manufacture integrated circuits (IC). This system provides an effective way to detect and analyze degradations in performance. The methodology accounts for the amount of time it should take to process different types of ICs, that is, the expected service times. The system calculates *facility service efficiency*, a measure that normalizes actual service times with respect to expected times. We offer examples of how this system and the PFD are used by AT&T Microelectronics to improve efficiency without impairing quality.

Introduction

There are various continuing efforts in AT&T IC plants to improve manufacturing operations and enhance product quality. One effort, discussed here, focuses on improving *service efficiency*. To understand what we mean by service efficiency, one needs to understand the IC fabrication process.

In IC production, silicon wafers containing hundreds of IC chips per wafer may pass through as many as 250 processing steps in a cleanroom environment. A "recipe" is used to describe the process performed for each individual step, the parameters of the process and the facilities on which it is performed. Such facilities may range from expensive ion implanters and photosteppers to relatively inexpensive wafer-cleaning sinks. ICs are built one layer at a time, each layer developed on the preceding one. Because many of the processes and equipment used to make successive layers are similar, a given wafer is likely to pass through the same facilities many times, each time possibly using a different recipe.

Two very important cleanroom parameters are *capacity* and *cycle time*. Capacity is the maximum rate at which a cleanroom can produce product, usually expressed as units of wafers per week. Cycle time is the interval from the time production starts on a lot of wafers until that lot is finished. Longer than expected service times translate into prolonged cycle times that delay getting product to the customer. Even small increases in service time can significantly inflate product cycle time. While it is critical from an investment point of view to maximize capacity in this capital-intensive business, it is also important, from a marketing perspective, to minimize cycle time to deliver a variety of products to customers who must satisfy rapidly changing market demands. We discuss both these parameters and their relationship to service times.

Capacity, Cycle Time & Service Time. Because cleanroom capacity is extremely valuable, most cleanrooms operate in shifts around the clock. The resulting capacity is determined by wafer processing requirements and the interval required to process each recipe on each facility.[1,2] When service times in a cleanroom take longer than estimated, there is a commensurate reduction in its capacity. However, because of the extremely non-linear way in which cycle time depends upon service time, its degradation is proportionally greater, especially when the capacity utilization is high.

For example, suppose a facility normally takes $1/\mu = 1.0$ hour to process a recipe. Assume, for simplicity, service times are independent exponential random variables and that lots come to the facility with a Poisson distribution rate $\lambda = 16$ lots per day, i.e. about 0.66 lots per hour. We know[3] average facility cycle time, that is, the queueing plus service time, for mean arrival rate λ and mean service rate μ is:

$$\text{Cycle} = \frac{1}{\mu - \lambda}$$

Therefore, for our example facility, we get:

$$\text{Cycle} = \frac{1}{1.0 - 0.66} = 3 \text{ hours}$$

Suppose service time degrades by 10 percent to 66 minutes per lot. Then $1/\mu$ = 66 minutes = 1.1 hours and μ is approximately 0.91 lots/hour. For average cycle time, we now get:

$$\text{Cycle} = \frac{1}{0.91 - 0.66} = 4 \text{ hours}$$

Therefore, a six-minute, or 10 percent, increase in service time per lot will produce an increase of approximately one hour, or about 33 percent, in cycle time per lot. Consequently, a problem at a facility can result in a significant increase in product cycle time. Conversely, even relatively small decreases in a recipe's service time will translate into significant gains in product cycle time. For these reasons, it is important that each shop area ensure its facilities are performing well. We note that high service time variability also adversely impacts cycle time.

Facility Performance. Given the many recipes and product codes involved, the task of measuring IC facility performance is difficult. A given facility may process over 40 distinct recipes spread across many different product codes. Each recipe may take a different amount of time to process. To make it easier to monitor these facilities, we have designed a system to track service efficiency for each facility or how well the facility is performing relative to the service times defined for each recipe. *Service efficiency* is the time it should take to service a given product divided by the time it actually takes. This metric allows shop and engineering personnel to monitor their facilities' performance and provides early warning of service rate degradation. Since the service efficiency is calculated for each facility, each set of engineers, process analysts and operators have ownership of the local metric for the facilities in their area. Tracking this metric and publicizing the results provides a powerful stimulus to improve the facility performance.

Product Flow Database. In 1989 we started development of a system called the Product Flow Database (PFD) to help achieve the goals of reduced work in process inventory and reduced cycle times. This system provides each shop area with timely information needed to track facility performance and to plan and execute meaningful actions to improve service efficiency. To this end, the PFD helps to identify areas needing improvement and provides feedback to shop personnel by quantifying the impact of efforts to improve product flow.

Panel 1. Abbreviations, Acronyms, and Terms

DDL — Device Development Line
IC — Integrated Circuit
`irge` — Interactive Report Generator on Efficiency
IST — integrated service time
PFD — Product Flow Database
`probe` — a menu-driven program to monitor facility or product flow behavior
service efficiency — the estimated time it should take to process a product divided by actual service time

Methods for Gauging Service Efficiency. To run an efficient cleanroom, shop and engineering personnel must monitor many performance measures, including cycle time. However, cycle time can be improved only by improving the parameters upon which it depends, such as arrival and service rates and their variability. There are various efforts underway, such as SmartSelect and SmartGoals, to improve arrival flows to facilities by pulling lots to meet planned cycle times and due dates and balance the workload of highly utilized facilities. Our methodology addresses service times, the other component affecting cycle time. Using information provided by PFD, the shop can define and track meaningful metrics for its facilities based on service times. The shop can easily analyze its facilities' process performance and ensure that service efficiency is maintained at the highest possible levels.

Shop personnel can use these metrics as an integral part of a cyclic paradigm to improve their processes. This procedure is:

- Select appropriate service rate metrics
- Define goals based on chosen metrics
- Observe metrics
- Analyze cases where metrics do not meet goals
- Make improvements that drive metrics toward goals
- Raise goal levels as appropriate

This formula is repeated continuously to achieve high

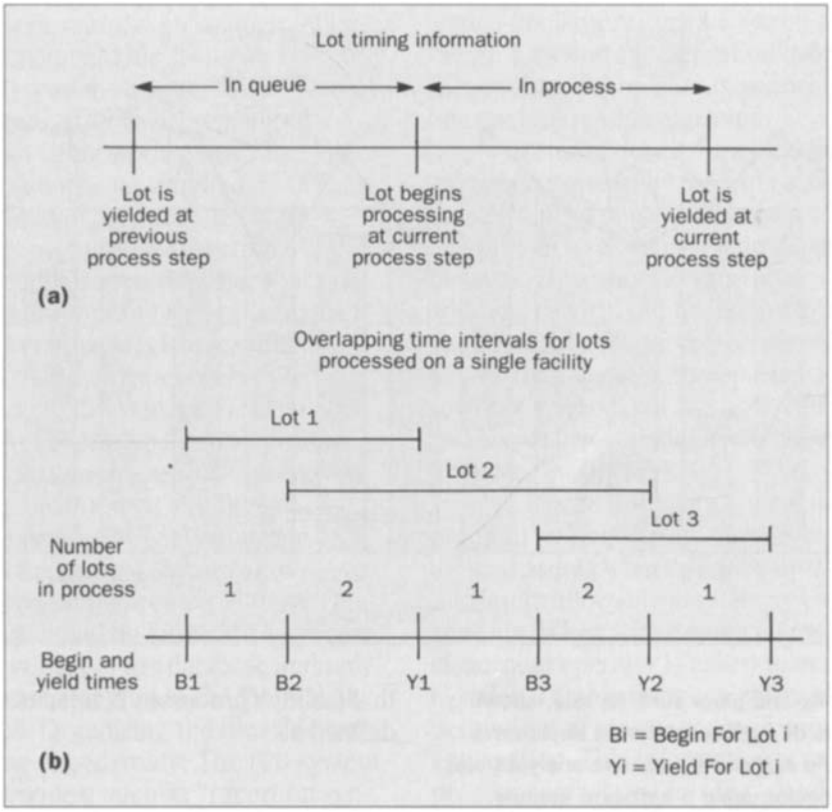

Figure 1. (a) The time a lot spends in a process area is the sum of the times it spends waiting in queue and it actually spends in process. Process time may also include setup time. (b) When several IC lots are processed simultaneously, a timeshare algorithm is used to calculate integrated service time (IST).

performance levels. The PFD system is a critical part of this paradigm. It allows each shop area to define and observe its own metrics and provides direction for making improvements.

Service Efficiency Measures. To evaluate the service efficiency of wafer fabrication facilities, we introduce two related measures of performance, integrated service time (IST) and facility efficiency.[4] Two types of data are required to calculate these measures. The first is a table of *expected processing times* containing the estimated times required for processing each recipe. The table is based on data provided by shop process analysts and process engineers and allocates an estimated time for each particular step and group of facilities. This estimate includes both setup and expected processing times for each lot or wafer on these facilities. Each designated shop area has a table of expected processing times for each set of processing steps and group of facilities on which these can be performed. We will show how the table is used shortly.

The second set of required data are *actual lot processing times* collected from shop process control computers. Figure 1a shows the timing information provided by this system for each lot. The shop's process control computers record the times when processing a lot begins on a specific facility and when that lot is yielded. With this, we can calculate how long the lot was in queue and the actual amount of time it spent *in process*. Other data about the lot, including serial number, product code, technology, and number of wafers in the lot, are also recorded. Later, the data can be retrieved, processed and stored in the PFD.

We now describe how the IST and facility efficiency values are calculated. The IST measures the actual processing time each lot of wafers spends on a given facility. In cases where a facility operates on a single lot at a time, this is simply the *in process* time shown in Figure 1a. However, because many facilities process several lots simultaneously, the data on the lots processed often contain overlapping time intervals, like

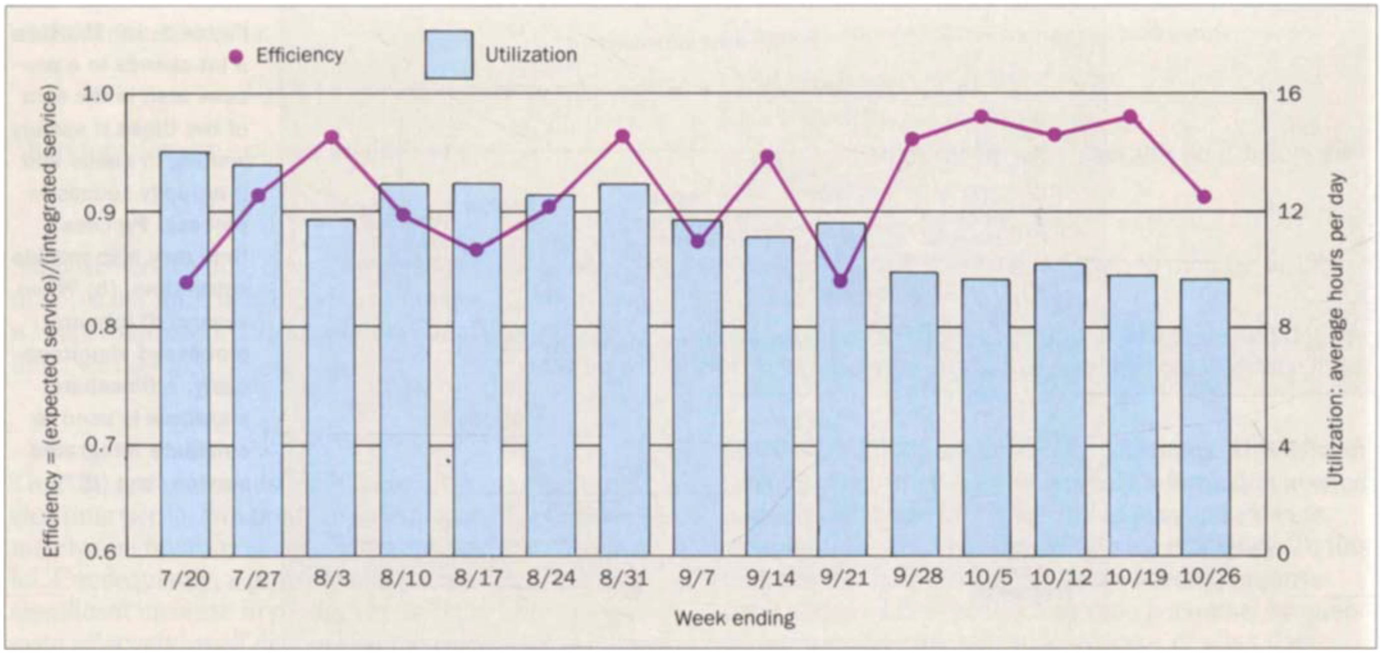

Figure 2. Pre-defined graphs and plots such as this, showing the facility efficiency value of a group of Varian implanters over a 15-week period, help engineers, process analysts and operators monitor line behavior using a software module, called `probe`. Facility efficiency values less than 1.0 show that service times are larger than expected.

those shown in Figure 1b. In these cases, a timeshare algorithm can be used to calculate the ISTs. This formula divides a lot's begin-to-yield time into separate intervals. For each interval, the lot is "charged" the time divided by the number of lots sharing the facility during that period. The IST is the sum of these "charges." For example, suppose the timing data for three lots are as shown in Figure 1b. We calculate the IST for each lot as follows:

Lot 1: (B2 – B1)/1 + (Y1 – B2)/2
Lot 2: (Y1 – B2)/2 + (B3 – Y1)/1 + (Y2 – B3)/2
Lot 3: (Y2 – B3)/2 + (Y3 – Y2)/1

Note that the sum of the lots' ISTs equals the total time the facility was utilized, i.e., (Y3 – B1). This gives us a basis for charging a portion of the processing time to each individual lot.

To calculate facility efficiency, we use the IST calculation and the expected processing time tables. Given that facility f processes N lots, its facility efficiency is defined as:

$$Eff(f) = \frac{\sum_{i=1}^{i=N} E_i}{\sum_{i=1}^{i=N} IST_i}$$

where E_i is the expected processing time of lot i and IST_i is the IST of lot i.

We can also calculate the efficiency for any subset of the lots, such as all lots requiring a particular recipe. In this way, the data can be used to disclose if certain process steps are experiencing problems. The operators can evaluate whether a facility is performing as expected by seeing how close the efficiency measure is to the value one. For example, a facility efficiency value of 0.91 means the facility is taking approximately 10 percent longer than expected to process lots.

The Product Flow Database. The PFD system consists of a database and two interface modules that allow users either to access predefined reports and graphs or to generate custom reports. The PFD system is designed to make available detailed information on product arrival patterns, delays, service time, and cycle time to the wafer fabrication line operators and managers. The system has been installed at AT&T cleanrooms in Orlando, in Allentown for the Device Development Line (DDL) and recently in Madrid, Spain. The PFD system runs on UNIX® Operating System-based computers. (UNIX is a registered trademark of UNIX System Laboratories, Inc.) The PFD is organized in ASCII files and uses a combination of C,[5] AWK,[6] S,[7] and Shell[8,9] language programs to create records and give users access to the information.

The database contains a record for each lot processed per step in the cleanroom. The record contains the lot's identification number, product code, shift on which it was processed, IST, queue time, cycle time through the facility and about 25 other items. The data files are stored in separate directories for each production area of interest. The data are organized this way on the assumption that someone interested in the ion implant area, for example, would probably not want to know about flow through the plasma etch area.

Each week's records are stored separately. This assumes that most queries would be limited to a specific time period covered by a subset of the database and only the files containing records belonging to the interval of interest need be examined. Organizing the files this way gains speed at the expense of generality. The PFD system will respond quickly to a request such as "report on service times of all lots processed in the past 10 days for the *X* group of *Y* machines." It is not designed to track the history of a particular lot as it flowed through the line. However, there are systems available for such queries.

Querying the Database. PFD provides database query capability through two commands. One command, called `probe`, gives access to pre-defined reports and graphs of facility or product flow behavior for the most recent 15-week period. Information about the entire production line is available from `probe` under a "General" category, or for individual work areas under more specific categories. This information is derived from intermediate summary files which are updated weekly. Thus, `probe` can provide information very quickly without having to refer to detailed low-level PFD files.

Process analysts normally use `probe` to monitor or analyze line behavior. It allows them to spot problems, detect trends, and to monitor the effects of actions to improve the performance of a given shop area, group of facilities, or even a single facility. Figure 2 illustrates a typical plot available through `probe`. This particular chart tracks the facility efficiency value for a group of Varian implanters over a 15-week period. It also shows the group's average utilization in hours per day. When these values fall below 1.0, service times for these facilities are longer than expected.

A more detailed method for analyzing performance is provided by the Interactive Report Generator on Efficiency (`irge`) command. `Irge` is menu-driven and provides specialized user-designed reports that can be focused on specific shop areas and time intervals. In addition, reports can be limited to any combination of shift, facility, code, operator, etc. Although `irge` is slower than `probe`, its reports can be focused as narrowly as desired and can be used to supplement `probe` for deeper investigations into behavior that impedes or improves product flow.

Figure 3 shows a typical user-generated `irge` graph. It gives a Pareto analysis of "slow lots," which are defined as lots whose service time exceed a predefined maximum allowable limit set by engineers and process analysts. When a lot's service time exceeds the limit, the cleanroom operator is asked to explain the reason for the slow lot. The operator can either choose from a predefined set of reason codes or provide his or her own explanation. This information is captured by the facility process control systems in Orlando, Allentown and Madrid. Reason codes for all slow lots are stored by the PFD and plots, such as Figure 3, can be produced. This particular plot breaks down lost production hours by reason codes. The `probe` and `irge` commands also offer other analysis options such as user-definable metric graphs, boxplots, histograms and graphical plots of facility activity.

Data Integrity. Given accurate data on expected processing times for each recipe, the PFD system provides useful information on facility efficiency. However, because new recipes are frequently introduced into the cleanroom, it is difficult to keep such data current. To overcome this, PFD has a self-correcting feature. When it does encounter a recipe for which it has no expected time data, it automatically prompts the system administrator, via electronic mail, to obtain the needed information from shop and engineering personnel. This helps to ensure that facility and process data on expected processing times for each recipe are kept current.

Applying PFD at AT&T Microelectronics Plants. The service efficiency methodology described in this paper helps shop and engineering personnel to improve their

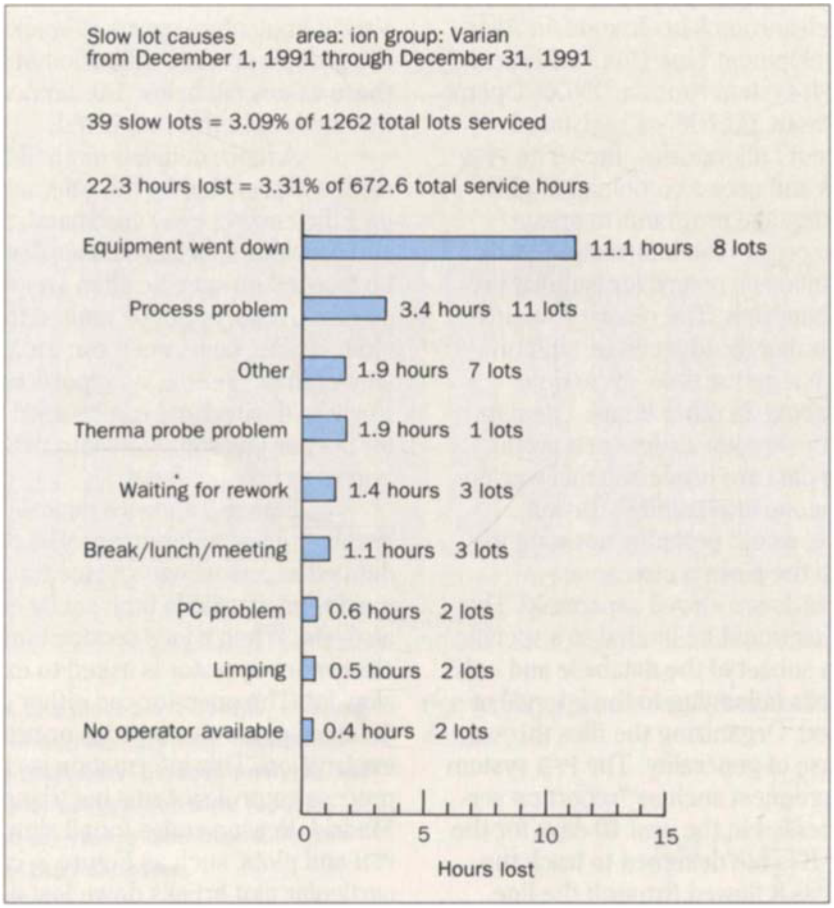

Figure 3. Detailed analysis of facility performance is provided by the `irge` command. This Pareto analysis shows reasons for slow lots. It can be used to identify areas for improving service times.

processes in a structured manner. The PFD system is now used in several AT&T Microelectronics plants. At the Orlando plant, process-based service quality teams are using it to baseline metrics, trouble-shoot, and improve quality. For example, improvements made by the Ion Implant team have increased the capacity of its facilities. During the latter part of 1991, a high-current Nova implanter, a unit costing over a million dollars, had facility efficiency values ranging between 0.6 and 0.7. Service times on this facility were approximately 30 to 40 percent longer than expected. The process quality team discovered that an instrument, called a faraday cup, was malfunctioning. The faraday cup measures beam current at the start of the implant procedure and moves out of the path of the beam when the implant process begins. The cup, though, was not completely moving out of the beam's path, causing the current level to drop and the process to halt. The team worked with the equipment vendor to resolve the problem. However, the facility efficiency values only slightly improved. After further investigation, the team discovered that a malfunctioning dose processor for the atomic mass unit was giving incorrect readings on current beam amperage and was causing the process to shut down. A faulty resistor pack was the cause of the problem and it was replaced.

Once these problems had been isolated and resolved, the facility's efficiency values, as shown in Figure 4, rose to between 0.9 and 1.0, representing approximately a 30 percent reduction in service times and a corresponding 30 percent increase in capacity.

At the DDL in Allentown, engineering and shop personnel have used `probe` and `irge` command options to monitor facility utilizations and service efficiencies, number of wafers processed per week, and the average queue and cycle times per lot. In the Photo Resist Apply area, a quality circle of engineers, operators, and process analysts have used histograms of process service times and slow lot Pareto diagrams to identify areas for improvement.

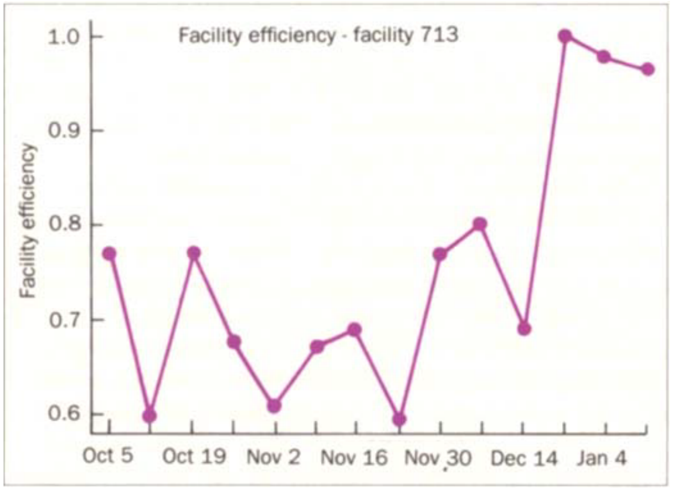

Figure 4. This plot shows the progressive impact of a series of changes made by a quality team that used the PFD system to diagnose long service times on an ion implant unit at AT&T's Orlando plant.

For example, shop personnel devised new procedures to improve production during startup and shutdown periods. Production on the DDL starts at 11 p.m. on Sunday and runs around the clock until shutdown on Fridays. There were significant production drops, on the order of 33 percent of capacity, during the first and the last shifts of the week. To improve output, some personnel now come in three hours early on the Sunday shift to start the line. Similarly, some shop personnel stay after the week's last shift to implement shutdown. These simple procedural changes now allow the facility operators to maximize capacity while the full shifts are operating. Using PFD, shop personnel were able to document the impact of these changes and to identify other problems that were resolved separately.

Other Applications of the PFD System. Because service times are a concern in all manufacturing environments, the PFD system and the methods outlined here can be adapted to other AT&T manufacturing lines. We note that the IST algorithm, as a measure of lot service time, has also been adopted in the SmartSelect and SmartGoals algorithms in the Divisional Shop Control System in Orlando. SmartSelect and SmartGoals were recently implemented in the Orlando cleanroom to pull lots through the production line in a smooth manner to meet planned cycle times and due dates and to balance the workload of highly utilized facilities. PFD is being used to track related metrics such as queue time and throughput to determine the effectiveness of these systems. Another outgrowth of the PFD system in Orlando is a prototype expert system for automated analysis to identify where facility performance problems are occurring. The system currently can analyze PFD records for facilities that are heavily utilized but have low facility efficiencies. The system is written in the rule-based language C5, which makes it easy to add additional capabilities. We are investigating ways to enhance the system to provide in-depth analysis of the underlying causes of performance problems by doing a thorough search through the database and suggesting the most likely causes.

Summary

We have described both a methodology and a system that helps shop and engineering personnel ensure their lines are working well and that their service efficiencies are as high as possible. These objectives play key roles in achieving the ultimate cleanroom goal of short cycle time and high production volume. This work is ongoing at several AT&T plants. We have already seen progress as a result of the focus on metrics and the direction provided by the PFD system. A continued focus such as this will result in lower inventory levels and cycle times and ultimately will increase the profitability of our plants.

Acknowledgments

The authors would like to thank their management for their support during the development and implementation of the PFD system. We would also like to thank those who have played a part in the PFD implementation efforts in Orlando, Allentown and Madrid. They include C. Cirigliano, P. Connor, J. Getz, D. Grube, R. Hnath, F. Hortal, D. Kelly, C. Meeks, R. Mohn, J. Newman, M. J. Nunez, E. Persson, W. Pyka, P. Rice, M. Rico, J. M. Ruiz, E. Searfoss, and L. Snyder. We also acknowledge the efforts of N. Tate and N. Wright in Orlando and J. Swiderski in Allentown, whose efforts on improving service efficiency were used as examples.

References

1. D. Burman, F. J. Gurrola-Gal, A. Nozari, S. Sathaye, J. Sitarik, *"Performance Analysis Techniques for IC Manufacturing Lines,"* AT&T Technical Journal, July/August 1986, Volume 65, Issue 4, pp. 46–57.
2. B. Tullis, V. Mehrotra, D. Zuanich, *"Successful Modeling of a Semiconductor R&D Facility,"* 1990 International Semiconductor Manufacturing Science Symposium, Burlingame, California, pp. 26–32.

3. S. Ross, *"Introduction to Probability Models,"* Academic Press, New York City, 1972.
4. B. Farrell, *"A Production Performance Database and Query Software for Integrated Circuit Manufacturing,"* 1990 International Semiconductor Manufacturing Science Symposium, Burlingame, California, pp. 39–43.
5. B. W. Kernighan, D. M. Ritchie, *"C Programming Language,"* Prentice-Hall Software Series, Englewood Cliffs, NJ, 1978.
6. A. Aho, B. Kernighan, P. Weinberger, *"The AWK Programming Language,"* Addison-Wesley Publishing Company, Reading, MA, 1988.
7. R. Becker, J. Chambers, A. Wilks, *"The New S Language,"* Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1988.
8. S. G. Kochan, P. H. Wood, *"UNIX Shell Programming,"* Hayden Books, Carmel, IN, 1985.
9. M. I. Bolsky, D. G. Korn, *"The Kornshell,"* Prentice Hall, Englewood Cliffs, NJ, 1989.

(Manuscript received January 2, 1992)

Brian L. Farrell *is a member of the technical staff in the Performance Analysis Department of AT&T Bell Laboratories. He currently works on performance analysis and tool development for manufacturing lines and operations support systems. He joined AT&T in 1978. Farrell holds a B.S. in Mathematics and Computer Science, from St. Peter's College, and a M.S. in Operations Research from Columbia University.*

The late ***William Kahan*** *was a supervisor in the Manufacturing Systems Engineering area of AT&T Bell Laboratories. He was working to increase service effectiveness in partnership with the AT&T Microelectronics Orlando works. He joined the company in 1981 and had received B.S., M.S. and Ph.D. degrees in physics from Brooklyn College, Brooklyn, N.Y., and Columbia University, New York City, N. Y., respectively.*

Jay W. Reightler *is a senior technical associate in the Integrated Circuit Reliability Department of AT&T Bell Laboratories. At the time this article was written, he was working on baseline technology development in a Device Development Line, where he was responsible for data analysis of parametric and final test data, Wafer Fab Simulation and monitoring cleanroom product flow. He is currently a device engineer working on qualification of Digital Logic Integrated Circuits. He joined the company in 1984 and has an A.A.S. degree from Northhampton County Area Community College.*

Alan C. Woodard *is a member of technical staff in the Manufacturing Systems Department at the Orlando plant of AT&T Microelectronics. He is responsible for design and implementation of manufacturing systems and is currently working on production scheduling using a starvation avoidance dispatching algorithm. He joined the company in 1984 and has a B.S.E. in Industrial Engineering and an M.B.A. from the University of Central Florida and University of South Carolina respectively.*