The name SASTRAK was chosen for our system to denote the fact that our Resource Management, Capacity Planning group could use SAS for much more than regression forecasting of future capacity requirements. We planned to reduce the lead time required for forecasting by entering performance data daily. Then the function that we called "tracking" could also be automated. Hence SAS TRACK, or SASTRAK. That original scope has since been extended to several other areas.

When I got the SAS User's Guide, and "discovered" the power of the Data statement, it became apparent to me that we could automate the data entry function. When I read about the SAS Database, it became obvious to me that we could develop a complete system, with confidence that we could add or delete variables as the needs arose. We could publish Performance Reports, and chart and plot to show trends and relationships. From a start as a forecasting tool, we have expanded SASTRAK into a system to automate much of the menial work of Capacity Planning / Resource Management.

As we see it today, SASTRAK is our gopher. SASTRAK gathers the data we need, organizes the data into an easy to use form, and prepares working tools from which we make the decisions. In addition, we have programmed in SAS many of the functions required to keep management advised of the performance of the system, and the longer range resource needs perspective.

To give you an idea of our approach to a Capacity Planning/Resource Management System, let us first look at a system overview, then look over each of the system cycles, with examples of how SASTRAK helps us keep on top of the performance of our Airline Control Program (ACP) based COSMOS (1) system.

SYSTEM OVERVIEW:

The major objectives of SASTRAK were to improve forecasting effectiveness, improve performance reporting, and reduce the amount of manual labor involved in data preparation and manipulation. To achieve those objectives, we developed programs in SAS to extract the resource use data we needed from the ACP Data Collection/Data Reduction package, and place this data on a PDB (Performance Database). We added special datasets to record a "key volume indicator" that is broadly used within Federal Express - actual and forecast Package Volume. The third major component of our PDB is what we refer to as our External Factors datasets. By external, we refer to events external to the system that effect performance, resource use, or the applicability of data recorded during a specific period for use in forecasting or management reporting.

From the PDB, our SAS programs prepare management reports, forecasts of resource requirements, and a several reports, plots, and charts for our own use in tuning the system. We find that inclusion of the key volume indicator data has helped us improve our analysis, forecasting, and reporting. Inclusion of the external data primarily benefits our analysis function - as this data can often explain variances in the resource data that is not explained by volume.

WEEKLY CYCLE:

The Weekly Cycle of the SASTRAK system consists of gathering the resource, volume, and external data; reviewing that data; communicating system status to management; adjusting the system tuning as required; and qualifying the data for future use in forecasting.

Gathering the data consists of running and checking the output of the data gathering programs. This process is designed to assure that we receive and record the data we need.

To assist us in reviewing the prior week's data, we have prepared a set of programs we collectively refer to as the Tracking Subsystem. By tracking, we refer to the fact that the process tracks resource utilization by comparing the daily observations of the week with prior daily history. This system uses reports and plots. In the reports, a column is used for each key resource variable, with a column for our key volume indicator. Quick inspection of the reports and accompanying plots allow us to determine quickly when a particular day's observation is atypical of the trends. We have the capability to "track" one or more "daily peak hours" as needed.

The primary purpose of "tracking" is a quick, cursory review for completeness and reasonableness. The tracking subsystem is also used to highlight apparent aberrations in resource use. When material deviations from the normal trend occur, we determine the cause, and supply that information to management as a part of our weekly Performance Report. Among the features of this subsystem are the ability to include "theoretical limits" of key resources on the plots, and to show the mix of incoming messages during the monitoring run.
To assist in analysis of our performance data, we have established "safe operating ranges" for key system resources. The safe operating range is a tool to remind us that when the resource use reaches a predetermined level, it is time to start planning for expansion, or conservation of that resource. The role of the "safe operating range" is to provide us adequate lead time.

Evaluation of the daily data is a longer-term concern. Using the Tracking subsystem plots, we look for breaks in trends of key resource use - per unit of volume (Package Volume). Breaks in the trend have been found to occur due to changes in programs, changes in the business environment, or changes in policy or practice of the departments that use the system. One way we evaluate our daily data to assure the quality of PDB content is to view a period's observations from a daily peak hour by comparing daily observation means to the grand mean for a period of similar system volume, and the 95 percent confidence limits around that grand mean. This mechanism helps us isolate for further investigation those days where our resource data may have been biased by non-recurring outside factors.

We prepare a Weekly Performance Report for circulation among our technical and management personnel, with some user department reporting features. The user receives message volume and message mix information to assist them in evaluation of their terminal operators, and for their planning functions. Technical and technical management personnel receive a tabular report showing weekly average resource utilization, along with package volume. This report also includes resource use per transaction, to assist Applications Development in evaluation of system changes. Our report form includes space for manually generated comments. In this manner, when we have detected an unusual occurrence via our Tracking function, that fact, and it's effect on the week average, if material, can be shown on the report, often without a complicated cover memo.

MONTHLY:

Our monthly process involves two major functions: performance reporting, and data cycling.

Our Monthly Performance Report consists primarily of plots of forecast resource requirements from past forecasting cycles. On those plots, we have superimposed actual data. With these plots, we furnish a cover memo that states the life expectancy of each resource, in terms of time and volume. The fact that the actual data is following the forecast trend gives management the assurance that our forecast resource life expectancy is on target. A tendency to depart from the trend shows us it is time to schedule a new forecast.

This method approximates a budget variance report, in graphic form. If one views the forecast slope as a "budget", then actual data plotted below the forecast slope can be viewed as a negative variance, and actual observations above the slope as positive variance.

By comparing actual monthly averages with forecasts, we have given management the ability to evaluate our forecasting effectiveness. In those cases where a deviation from the forecast was noted, we have usually been able to identify causal factors that were not anticipated in the forecast. Often this has worked our advantage, by keeping management aware of the extent to which the Capacity Planning function depends on timely, accurate information about changes in design, or changes in the way the system is used.

DATA CYCLING:

This feature has not yet been implemented. The original plan calls for cycling daily detail data to weekly, monthly, and yearly averages by hour of the day. We have yet to experiment with forecasting from weekly average data, or combining weekly average and daily average data into our forecasts. If that proves successful, we will not need to retain daily detail data.

QUARTERLY CYCLE:

The two components of the quarterly cycle are forecasts and a planned Quarterly Performance Report. Nominally, we plan to reforecast each resource quarterly, however, if the Monthly Report shows that we are still tracking with the old forecast, and we know of no upcoming changes, then there is no need to prepare a new forecast.

I have developed a four step approach to forecasting. Following this approach, I have included 3 GLM's, and numerous PRINT, PLOT, and DATA steps in each forecasting program. The purpose of this four step approach is to create an audit trail of each forecast - to show any steps where data from the PDB was omitted for the purpose of the forecast. To effect this, our forecasts begin with plots of the variables to be used, directly from the PDB. Following these plots I have included a "trial regression" step, complete with plots of Actual vs Predicted, Residuals, and Actual vs Predicted with approximate 95 percent Confidence Limits.
When review of the reports and plots show that some observations should be omitted from this forecast, the procedure provides a "cards" data step, where this decision can be recorded, and printed between the trial and final regressions.

A second trial regression is included to "proof" the subsetted data.

The next step of the forecasting program appends our "key volume indicator" (forecast package volume) to the historical record. From that projection, the GLM procedure provides us with an extrapolation through the range of the Package Volume Forecast.

The actual and forecast data are printed and plotted. Residuals are plotted. Confidence Limits are plotted overlaying the Forecast - to show management the "probable accuracy" of the forecast. This plot, showing the Confidence Limits is later used in our Monthly Performance Report to show whether actual experience is falling within the range of probable accuracy of the Forecast.

QUARTERLY MANAGEMENT REPORTS:
At this point, quarterly reports may or may not be implemented. Our probable design, if we choose to implement, will be to show long term trends in resource demand, by Package Volume and by date.

IMPACT ON CAPACITY PLANNING AND RESOURCE MANAGEMENT:
SASTRAK has had a major impact on the work methods we use, and on the effectiveness and efficiency of our department. The automated extraction of ACP resource use data has saved us considerable keypunch type work, and errors. Automated extraction has given us much more data to work with than would otherwise be possible. This has resulted in improved forecasting in two ways: more variables to study and correlate, and more observations to feed the model.

The Tracking Reports and plots are a qualitative and labor-saving improvement. Prior to the development of SASTRAK, we posted key resource data from the Data Collection reports to spread sheets. We referred to this process as Tracking - so we could observe the history of each variable. The new Tracking Reports and plots give us more information, as well as the needed data. They are easier to use. We avoid the onerous clerical task of posting to spread sheets.

SASTRAK and the PDB have helped us improve our forecasting in another way. We have learned to use the various SAS modules, such as PROC MEANS, FREQ, and others as analytic tools - for example to evaluate our raw data for homogeneity. When we do this on a regular, routine basis, we can research the causes of data that varies extremely from the norm prior to the time that the events of that day have vanished from the minds of the people with whom we work.

In our MAINTENANCE subsystem, I have programmed the capability to delete a whole day's data, or to set the value of a variable to null, or to classify a day's data as "suspicions", or "no good". I have set up our forecasting, and some reporting programs to be sensitive to those two classifications. This feature, combined with routine analysis, evaluation, and qualification of data helps us become more effective as forecasters, and in special studies, in that it reduces the lead time we need for data preparation, and use more recent data.

SASTRAK has helped us solve a major analytic problem. We have been running our ACP system under IBM's VM/370. This has resulted in several distortions of the data recorded and reported by the ACP Data Collection/Data Reduction package. Given the raw data on the POB, and some knowledge of the internals of ACP and VM, we were able to devise programmatic adjustment routines to compensate for the distortion. This adjustment greatly reduced the apparent variation in day to day data, resulting in an improvement in the accuracy and reliability of our forecasts.

Now you can see why I earlier referred to SASTRAK as our gopher. It has taken much of the drudgery out of our lives. It has transported us from a group spending a majority of it's time in basically clerical work, to a group that spends most of it's time producing results.

A brief comment about SAS. The SAS Database was a very important factor in this project, as was the SAS language itself. Were it not for the ease of use and self documenting features, and the ease with which variables can be added or deleted, this project would have been much more costly.

(1) COSMOS provides real time dispatching, package tracing, customer record keeping, and other information services.