1. **INTRODUCTION**

The main emphasis of this talk will be on the uses of SAS itself, rather than explaining our capacity planning methods in detail. We have generally used SAS at Datasolve for at least three years. It is used for the analysis of SMF, RMF, IMS and CICS statistics. My own involvement has been concerned with analysing charging data for overall performance monitoring and capacity planning. We have found that SAS is easy to use, powerful and flexible. It automates regular reporting, and simplifies the work involved in producing one off or special reports. The main facilities used have been tabular reports, SAS/GRAph colour graphics, and linear regression based upon a SAS database updated weekly.

The first diagram (Fig. 1) shows an overview of our SAS based capacity planning system.

- SMF data is processed weekly to update the SAS database.
- Reports on machine loadings etc. for our two machines are produced monthly.
- Every quarter a summary report is produced showing machine usage by major user group.
- This quarterly summary data is then used to detect growth trends and relate usage to hardware capacity.

Once the data is in a SAS database it can be easily manipulated not only to produce the regular reports shown but also for any other special analysis required.

The diagram itself was produced using the SAS/GRAph GSLIDE procedure which allows specification of size, format, colour and position of verbal information using NOTE statements. It also has the facility for drawing lines between pairs of coordinates - with the results shown. No doubt quite complicated drawings could be built up if required.

2. **TABULAR REPORTS**

The Machine Usage and Loading Report is produced monthly. This report is formatted using PUT statements.

- Allows the positioning of heading information by character position along the print line.
- Variable data from the SAS file is handled in the same way.
- Variable data can be formatted to eliminate unnecessary decimal places etc.
The week is divided into prime shift, overnight and weekend periods. The consumption of charge units by Batch jobs and separately for TSO sessions is shown. Data from previous months is also carried forward and shown for comparison. The maximum available value is an estimate of the machine capacity for the time period indicated. This value is calculated by taking into account average resource consumption, peak to average ratios and a number of other factors.

Basically data from each machine is kept in a separate SAS database and usage and loading is reported on monthly in the manner described.

For capacity planning purposes the monthly data from each machine is then combined on a quarterly basis. This is done by sorting, merging and summarising the corresponding records and then outputting to another database. A quarter is normally always two four week and one five week accounting months to give thirteen weeks total. This period of thirteen weeks tends to produce smoother data than does the month which is affected more directly by holiday periods etc.

A further report summarises on a quarterly basis the monthly usage data from the two machines. Once again the report is formatted using the PUT statement.

For capacity planning purposes it is the quarter totals that are utilised. These totals are then related to the estimated machine capacity for the two machines together in order to determine a capacity plan and upgrade schedule. Which is where we utilise the facilities of SAS/GRAPH.

3. USING SAS/GRAPH PROC GPLOT

This rather complicated diagram (Fig. 2) is produced on a quarterly basis. The capacity available from different machine configurations can be depicted as a ratio to our original 158 MP at 6 Mbytes - MP factor.

Although we have now replaced our 158 MP with the 3033 and subsequently installed a 3081G we still relate machine capacity back to this earlier machine. This was originally done to simplify presentation of capacity to management prior to the installation of our Amdahl machine and we have continued to do this ever since.
Total machine usage is calculated from the data in the Quarterly Usage Report, described earlier, and plotted in terms of machine capacity.

At first sight this diagram may appear complicated but we have been producing similar diagrams manually over a number of years so our own management are familiar with this.

The advent of SAS/GRAPH has meant that the production of such diagrams can now be automated. Modifications and variations are easily and quickly incorporated. The hardest part was working out the X, Y coordinates in order to position the notes in the correct places. This was done by trial and error on the IBM 3279 VDU and the final version then printed on our 3287 printer. For reports to management then multiple copies can be produced as required by repetition.

This diagram uses the SAS/GRAPH Procedure GPLOT. Basically it consists of five plots overlaid.

IBM machine capacity in blue points joined
and the additional Amdahl capacity in red using STEPLJ.

The other three lines (usage in green, 13 wk adjusted and GOP) are all plotted with JOIN for linking the points.

In order to project into the future the input data file has to have extra observations added corresponding to the required dates. This is done using PROC EDITOR and ADD; data is supplied for the quarter date and the machine capacity variables. Usage data is not supplied and SAS treats this as missing values which are then not plotted. The missing usage values are then edited in quarter by quarter as the data becomes available.

The descriptive data is generated using TITLE and FOOTNOTE statements. Once again X,Y coordinates are used to position correctly.

The blue vertical line separating the 158 MP from the 3033 is drawn by specifying the required coordinate positions.

Things were slightly more complicated in order to produce the same diagram for overhead projection. The three colours had to be printed separately and photocopied with a foil of the appropriate colour.

The NOAXES parameter was used to suppress the axes for the red and green parts. SAS however then also changed the scale and position of the red lettering when it was separated - I am not entirely sure why. So it was necessary to modify some of the parameters to ensure that it fitted when reassembled.
Part of the power of SAS stems from built in assumptions. If the user requires something different it is not always easy to find the appropriate means to do so. In some cases there is no control available.

There seems to be some conflict between VPOS, HPOS and VSIZE, HSIZE which I have used to produce A4 size output on the printer. Why not allow A4 etc. to be specified. This would be a lot simpler.

The next two diagrams (Figs 3 & 4) are produced from weekly job charge data. This data can be converted in real resource demand as an estimated main storage requirement (Fig. 3) or an estimated total CPU requirement (Fig. 4). In the case of main memory an estimate greater than the 24 Mbytes available is due to allowing a peak to average ratio larger than actual for the calculations involved. Certainly for planning purposes slight overestimates are preferable anyway.

I have now covered the main points relating to our use of SAS for capacity planning. I hope I have brought out the ease with which a variety of things can be done using SAS once the data is available in the right form.

Before I finish I would just like to mention one other SAS procedure that I have played around with - known as G3D.

4. USING G3D

This procedure gives three dimensional perspective plotting.

- Works best with one variable plotted vertically against date and time.
- Results are quite picturesque.

This example (Fig. 5) shows logged on TSO users and the morning and afternoon peaks and the lunch time dip can easily be seen. O.K., it does not convey much more than perhaps we already know, but it looks quite nice on the office wall.

A second example of G3D (Fig. 6) plots estimated main storage demand against date and time. In this case all values below 8 Mbytes have been set to 8 Mbytes so that a flat plateau is obtained, with the higher values rising above it.

The three following diagrams (Figs. 7, 8 & 9) again show the effects using G3GRID with G3D to plot various variables against one another. Each one shows the introduction of spurious negative values caused by the mathematical processes involved. Since in these cases negative values are obviously invalid there is no problem, but in some circumstances such effects could be misleading.
I have shown you some of the results obtained by playing around with G3D in the hope of producing meaningful pictures. At the moment it seems that this process can be used to depict a known viewpoint in quite a useful way. This can be used to improve management reports and enhance the more normal two dimensional plotting.

Of course G3D can be used on artificial data as a result of computing a purely mathematical equation. SAS itself uses this approach to demonstrate G3D with the cowboy hat.

My last diagram (Fig. 10) shows the result of modifying the trigonometric function in the cowboy hat equation. As can be seen this produces quite a spectacular result.
SMF DATA

JOB CHARGES AND RELATED PROCESSING

SAS DATABASE

MACHINE USAGE LOADING ETC.

TOTAL USAGE

CAPACITY

WEEKLY

MONTHLY

QUARTERLY

FIG. 1
3033/3081
ESTIMATED MAIN STORAGE

PEAK TSO
PEAK ONLINE
PEAK BATCH
MVS BASE

22 SEP 82 01 NOV 82 11 DEC 82 20 JAN 83 01 MAR 83
DATE

FIG. 3
IBM3033 ESTIMATED MAIN STORAGE DEMAND

FIG. 6
IBM 3081
(SASDATA.HRLY8482)
G3D9

FIG. 8
IBM 3081
(SASDATA.HRLY8462)
G3DA

FIG. 9