

## **Starting a Capacity Plan**

**Vickie Custer  
University of Wisconsin - Milwaukee**

In 1982, The Administrative Information Systems department at the University of Wisconsin - Milwaukee embarked on a new endeavor, the conversion from an IBM 360 model 50 to an IBM 4341. For the first time in years, Administrative Data Processing was going from old technology and batch processing to a newer technology and online systems. Within three years of the conversion, Administrative Data Processing outgrew the 4341 and work became very constrained. The department was unable to upgrade the CPU to a larger, faster model for at least one year. It became evident to the administration that something needed to be done. There was no way that the campus administration would approve funds for a new computer system without any prior planning or budgeting. The budget process usually required two to three years of planning to obtain the funds necessary for large acquisitions. To cap the budget process off, if funds exceeded a certain dollar limit, legislative and governor approval was necessary.

Thus a decision to attempt some capacity planning was made. The beginning was very tough. No one on the staff had any prior expertise in capacity planning or even tuning. A decision was made to acquire the MXG software written by H. W. Barry Merrill to at least decode the SMF and RMF records that were collected. SAS was also purchased at that time, though it was primarily intended to serve the user departments on campus. An IBM ISP for modelling was brought in to give an idea what could be done. Some research into packages was done, but the price proved to be prohibitive for the administrative environment. At this time a person was assigned to try to perform tuning and some capacity planning. The project was given only 25% of the work time, strictly a part time function.

The beginning of developing a capacity plan for Administrative Data Processing was to do tuning on the system and to define the workloads that would be tracked for planning purposes. At the time, online processing consisted of the IS staff and a small number of outside departments using TSO. One CICS application was available, a financial aid system that initially generated less than 10,000 transactions in one week. The rest of the system was devoted to batch processing. Development was being performed on an online registration system that would utilize IMS DB/DC and CICS. This system was a modification of the current system used at Madison. Planned implementation was one year away and no idea as to what kind of workload that would have on the current system. So an attempt to classify the workloads

into the following categories of TSO, CICS, IMS, batch, VTAM and miscellaneous systems.

Once workloads were determined, the next step was to find out how much each workload used the CPU. Performance groups were set up to accumulate the RMF data. This lead into doing some much needed tuning for the system, which extended the life of the CPU for a short period longer. Since the data was available, it was now time to begin creating a capacity plan for Administrative Data Processing.

Using the MXG software to decode the RMF data and also by reading a technical bulletin from the Washington Systems Center, "Capacity Planning A Basic Hand Analysis" by Bronner, we were able to set up formulas to capture an approximation of what each workload utilized the CPU during prime time, which was from 8:00 am until 5:00 pm. By summing the TCB and SRB service units, dividing by the CPU service coefficients and the total time, we were able to determine an approximation of how much CPU was utilized by a particular workload. An example of the SAS code used can be found in figure 1. After the calculations were made, the results were compared to the CPU busy from RMF reports and the data entered into a spreadsheet. The data was collected on a daily basis, summarized monthly and used to produce a historical graph of the CPU utilization by workload. This was a very crude way to begin, but it allowed the use of a pen plotter to produce a graph that could be understood by the administration.

Since the purpose of developing a capacity plan was to determine what the type of workload utilization would be in the future, a method was developed to create projections. The first projections of the workload growth were done by hand. The historical data over six months was evaluated to see what the change was and then that change was applied in three month intervals. When a comparison of the projected growth to the actual growth was made three months later, it was found that the projected growth was not very reliable. There was no way to estimate the growth of the new systems that were being added. After a year of collection and projection, it was much easier to see what the growth was since an entire year had been covered.

Collecting the year of data and plotting out the workload utilizations gave some interesting results. The peaks and valleys were very apparent in the data. As a result, it became necessary to look at the peak periods and base yearly projections of workload growth from that data. Unfortunately having only one main peak of data isn't enough to base a five year projection on. So, data was gathered into a dataset and a simple linear regression was performed upon the data to obtain the slope and the intercept. These values were then used to project utilization over a 60 month period.

The information that was collected to produce the projections including the workload growth was given to IBM. Since we were not sure if the data and the method of projection was reasonable, we needed to make a comparison.

When the output from our projections was placed with IBM's projections from the CP80 product, we were pleased to see that we did agree somewhat. Where the differences occurred was when projections were made for longer than two years. The growth became exponential on the IBM plot, whereas our plot did not increase quite as drastically. We did not compound the growth at the same rate that IBM did.

Our result was a straight line projection of growth based on historical data and gave no indication of what utilization might look like in peak periods. So a modification for the projection had to be made. At this time the historical data spanned three years, so the data was plotted and a linear fit was done to create an average line over the time period. Since the actual data and the linear data was present, a comparison was made on a month basis. Actual January data was compared to the line and the variance was calculated. From the differences, a percentage was derived so that the peaks and valleys could be calculated into the projections. Looking at the data as a variance from the average, gave an insight into the cyclical nature of the administrative processing at the university.

By looking at the variances and the average growth of the individual workloads, it was easy to ascertain the loads that would have little or no growth over the next five years. These were the TSO, VTAM and JES workloads. On the other hand, it was very difficult to predict what batch processing would do. The variance for batch did not give an indication of what would happen at any given time. Our best guess was to say that batch would increase to fill the excess capacity available. So, the online systems were the only workload that we needed to really determine the growth for the next five years. After looking at the variances we could determine the growth without including the effect of new systems.

To make the projections somewhat realistic, the variances for each month were calculated as percentage increases or decreases. Since the data spanned three years, an average increase or decrease was calculated for each month. An example, is that January data increased 5%, 6.2 % and 6.8% in the three years. The average increase used was then 6.0%. For November, the average decrease was usually 6%. So a SAS dataset was built of all the monthly values that would be applied to the monthly projections. This data was input with the projection data as a table so that the monthly projections would have the correct variance applied to it. Thus a projection for January of 1995 would be increased by the appropriate 6% and that the November projections would be decreased by 6 %.

When this data was plotted out along with a straight line projection, the peak processing times were very apparent. This method of projection and graphing made it very apparent when the CPU would be over-utilized. By transferring this data to a spread sheet and by showing the 80% utilization point of the current CPU, this gave management a document that would allow planning and justification of a replacement CPU. We did not need to worry

about what the replacement CPU would be, the philosophy within the management is to double the capacity of the existing CPU.

Over time, changes were made that caused our workloads to change. The major change was that all applications that were IMS DB/DC were converted to CICS and IMS DB. This eliminated a workload, but really increased the CICS workload. As a result, the historical data for CICS showed a large growth when the IMS systems were converted. By the same token, IMS DB/DC decreased, but not by the same proportion. This caused a slight problem in that the projected growth for the CICS workload was far more than the actual.

The projections were checked quite regularly for any variations that might cause a drastic change in growth. The comparison of historical data to the projection was fairly accurate. The projections in most cases were only slightly off of the actual usage. As more historical data was collected, the projections were modified to reflect the changes. This method predicted that we would need to make an upgrade of our CPU during 1989. If we did not, our first peak processing period would have some latent demand and that would just build throughout the year. The CPU was upgraded in March and no problems with the peak processing in August were apparent.

Up until the last year, this method has worked fine for projecting what the CPU utilization will be within 5 years. For the last two years, the only change in the system has been to go from MVS/SP to MVS/XA. We did not add any major applications to our system. In the future we will be implementing a Student Accounting system, a Course system, and possibly a touchtone phone registration system. In addition, we may see modifications to the current registration and transcript printing systems. As a result, our projections for CPU growth are not realistic at all. We foresee that we will outgrow the current CPU before the projections indicate. So a modification is needed to the software to show what impact new systems will have on our current projections. The modification will have to include new growth in the projection of CPU utilization.

To try to include new growth within the software is not difficult. The difficult part of the problem is determining what impact the new systems will have on the system and when will their impact be felt. Through discussions with the systems developers and management staff, a determination of the impact the new systems would have on the current operations was made. In addition, a guess was made as to the probable implementation of the new systems. From this discussion, a table of growth factors for each year in the projection was developed. If the new systems did not have an impact on a month, the value in the growth factor table was a one. If it did have an impact, the growth factor was one plus the percentage of increase for that month. For example, if the system would cause a 10% increase in utilization, the growth factor became 1.10.

To obtain a projection for 5 years, the linear regression was run on the historical data to obtain the slope and the intercept. The monthly variances were applied to the monthly projections to create a dataset that has 60 months of projections in it. The growth factor dataset was merged with the monthly projected data and a new projected CPU utilization was calculated. The new utilization was the projected month times the growth factor. The projected monthly utilizations and the historical file were merged together to create one file for plotting.

Once the data was plotted, management could see the effect of new systems on the utilization projections. This led management to ask questions such as what if touchtone phone registration is not implemented? The table containing the growth factors could be modified easily enough to add new systems or to eliminate systems. This leads well into what-if processing. By changing a table value and recalculating the monthly projections a new plot of CPU utilization is made and management can see the effects of the changes within a short period of time. An example of a graph produced is in Figure 2.

Since the what-if processing has been implemented within the capacity planning software, the next modification will be to incorporate something to recognize when paging is getting out of line and memory must be upgraded. The last upgrade to the current CPU will bring the memory to it's maximum of 64 MBytes. Before the utilization of the 3081-K reaches an 80% level, Administrative Data Processing believes that paging will cause an upgrade in the CPU.

Management is quite pleased with the results of the capacity planning done by the Administrative Data Processing department. The software developed in-house has saved money by providing a good representation of when the CPU must be upgraded. The software works reasonably well within the current environment of MVS/XA. There is no guarantee that the same will hold true when the upgrade to ESA is done. Before that hurdle is taken, Administrative Data Processing will need to upgrade the CPU to a model that will allow ESA to run. The current CPU is a 3081-K and the replacement will not occur until 1993-1994.

#### REFERENCES:

1. SAS, SAS Institute Inc,
2. Merrill's Expanded Guide to Computer Performance Evaluation Using the SAS system, Merrill, H.W. "Barry", 1984
3. "Capacity Planning, A Basic Hand Analysis", Bronner, Washington Systems Center, IBM

```

DATA RMFVDATA (KEEP=RMFVUNIT RMFVOBS RMFVTIME RMFVTRAN);
  SET PDB.RMFINTRV;
  RETAIN RMFVUNIT RMFVOBS RMFVTIME RMFVTRAN;
  FORMAT RMFVUNIT RMFVOBS RMFVTIME RMFVTRAN 12.4;
  RMFVUNIT=VTAMUNIT + VTAMSRBU;
  RMFVTIME = VTAMTCB + VTAMSRB;
  RMFVTAN = VTAMTRNN;
  IF RMFVUNIT > 1000
    THEN RMFVOBS = 1;
    ELSE RMFVOBS = 0;
PROC SUMMARY DATA=RMFVDATA;
  VAR RMFVUNIT RMFVOBS RMFVTIME RMFVTRAN;
  OUTPUT OUT=RMFVSUM
  SUM RMFVUNIT RMFVOBS RMFVTIME RMFVTRAN;
DATA CAPPLAN.VTAMSUM
  (KEEP VTATOTTM TRAMTIME VTAMPG1 VTAMPG);
  SET RMFVSUM;
  RETAIN VTATOTTM VTAMTIME VTAMPG1 VTAMPG ;
  FORMAT VTATOTTM VTAMTIME VTAMPG1 VTAMPG 12.4;

VTATOTTM = (RMFVUNIT / 10) / 391.2;
VTAMTIME = RMFVOBS * 60 * 60;
IF VTAMTIME > 32400 THEN VTAMTIME = 32400;
VTAMPG1 = VTATOTTM / VTAMTIME * 100;
VTAMPG = VTAMPG1 / .9;
IF VTAMTIME = 1 THEN VTAMPG = 0;

PROC PRINT DATA = CAPPLAN.VTAMSUM  SPLIT = '*';
  VAR VTATOTTM VTAMTIME VTAMPG1 VTAMPG;
  LABEL VTAMTOTTM = 'VTAM*TSU*CONSTANT'
  VTAMTIME = 'VTAM*EXECUTION*TIME'
  VTAMPG1 = 'VTAM*PGU*CALCULATION'
  VTAMPG = 'VTAM*FINAL*CALCULATION';

```

FIGURE 1.