Capacity Planning: Getting Started

Charles Rice
University Systems
The Ohio State University
1121 Kinnear Road
Columbus, Ohio 43212

Introduction

Capacity planning can be defined as planning the future use of your computing resources by what has taken place in the past and what is currently being planned for the future.

There is a proverb that says "any enterprise is built by wise planning, becomes strong through common sense, and profits wonderfully by keeping abreast of the facts." This proverb sums up what capacity planning is all about.

Planning the use of your computing resources is a wise thing to do. One of Murphy's laws states that the demand for your computing resources will always grow to exceed your computing resources. The purpose of a capacity plan is to accommodate the demand for computing resources before those demands exceed the resources available.

Capacity planning requires the use of common sense. You can almost always add another system to your computer, but common sense says "what is the affect of the new system on the other systems?" The new online system could possibly affect the response time of other online systems.

Above all, capacity planning requires you to keep abreast of the facts and rumors. A capacity plan needs to keep track of what happened in the past and what is being planned for the future. Are future versions of a program going to use more resources? Are there plans to add new systems to the current load? As time goes by the future becomes more clear and adjustments are made to previous plans.

Getting Started

The first step to capacity planning is to identify the resource that will limit your capacity. What resource is keeping you from adding more systems? Is the CPU busy 100% or maybe the CPU is paused for disc 50% of the time. As a general rule the resource that limits your productivity is the resource that you want to concentrate most of your effort.

Before MPE started to cache disk IOs in main memory, the number of disk IOs was often the resource that limited activity on our HP. With disk caching though, CPU has become the limiting factor. Therefore at University Systems, we concentrate most of our effort on tracking and predicting CPU usage.

Even though our Capacity Planning Report concentrates on CPU, for historical purposes we track disc IOs, memory activity and global response time. These resources can indicate problems that would not be noticed by tracking just CPU. For example if memory activity seems high, you have a historical record of memory activity to use as a yardstick.

For all of the resources we track we have set a threshold value that we do not want to cross. For instance 85% CPU busy is the typical threshold value for CPU in the industry. If the CPU is busy 85% of the time then you are approaching the capacity of your CPU. Some thresholds that we set are arbitrary i.e. a point where we think we will start to have problems although we have no evidence that would indicate any problems.

There are many tools on the market today that can assist you with capacity planning. The first capacity planning report produced at University Systems' used some tools from the Contributed Library. Before we produced the next report MPE V was installed and the tools we were using were not compatible with the new version of MPE. It is best to use tools that are either supported or supportable i.e. you have the source code and knowledge to support the tools.

At University Systems we use HPTREND from Hewlett Packard and SYSPLAN from Carolian Systems. HPTREND has been very good for tracking resource usage by account and CPU by activity e.g. CPU busy, memory management, paused for disk etc. SYSPLAN has been excellent for identifying CPU activity by process and/or user. SYSPLAN, by allowing us to get at the raw data, offers us a lot of versatility despite its short comings, which I will identify later.

Organizing the Report

At University Systems we produce a Capacity Planning Report quarterly. What follows will be a summary of how our report is organized and then a description of all the sections.

1. Index.
2. Summary.
3. Definition of Business Elements.
4. CPU Graphs:
   a.) Average CPU used during 24 hour weekday.
   b.) Average CPU used during weekend.
   c.) Prime Time CPU Use by Business Element During Quarter.
   d.) Peak Time CPU Use by Business Element During Quarter (see page 7.)
   e.) Table of Projected Growth by Business Element for prime and peak time (see page 8.)
   f.) Graph of Projected Growth by prime and peak time (see page 9.)
5. Graph of Disk IOs during prime and peak time.
6. Graph of Free Disc Space.

7. Miscellaneous Information.
8. Actions taken to improve performance.
9. Recommendations.
10. Glossary.

Report Summary

Most people who read this report will have their own areas of interest. For example the Director of Operations might be more interested in CPU usage on third shift while the Director of Development will be interested in response time the programmers are receiving. The purpose of the summary though, is to direct the readers attention to the areas that need attention. This is where you point out when the current CPU will run out of capacity.

Business Elements

At University Systems we wanted to know what processes were using the CPU. So we divided all processes into eight different categories that we call Business Elements. The eight Business Elements are:

1. Editing Programs (Editor, TDP, SPEEDIT etc.)
2. HPDESK.
3. Office Automation (HPWORD, GRAPHICS etc.)
4. Data Communications (MRJE, DS, NS, IMF etc.)
5. Utilities (MPE subsystems, LIB, TECH etc.)
6. Production (Income producing accounts.)
7. Other (Programs not included in other filters.)
8. Unknown (CPU used but not accounted for by any business element. We have assumed that this is operating system overhead.)

CPU Graphs

Although we are mostly concerned with prime time (8:00 AM to 6:00 PM Monday thru Friday) we do include a graph of CPU usage during the week nights and weekend. This serves two purposes, it gives management an idea of what is happening after hours and records this information for historical purposes. There might come a time when this information will become useful.

We produce two pie charts for prime and peak time that display average CPU used by business element during the quarter. By adding a slice of pie labeled IDLE CPU we force the total of all the business elements to add up to 100% i.e. 100 minus the sum of all business elements equals IDLE CPU.

The table of projected growth by business element is a product of two tables. The first table consists of the percentage that each business element is projected to grow by month (see TABLE ONE.) The second table consists of the actual and predicted percentage of CPU consumed by each business element (see TABLE TWO.) The last month of

actual data is multiplied by the percentage that the business element is expected to grow in the next month. For example if HPDESK is expected to grow 1% in the months of May thru September. The last month of actual data shows that HPDESK used up 15.4% of the CPU. This number (15.4%) is multiplied by 1% to obtain the projected growth of HPDESK for the month of May.

TABLE ONE

	3/88	4/88	5/88	6/88	7/88	8/88	9/88
HPDESK			1%	1%	1%	-25%	1%
EDITORS			2%	2%	2%	2%	2%

TABLE TWO

	3/88	4/88	5/88	6/88	7/88	8/88	9/88
HPDESK	14.0%	15.4%	15.5%	15.7%	15.8%	11.8%	11.9%
EDITORS	15.2%	14.2%	14.5%	14.8%	15.1%	15.4%	15.7%

The table of projected growth by business element is detailed by month and should cover as many months as you think your current CPU will last. This table should reflect your judgment of the expected growth. For example if you know that the performance release of HPDESK is going to be installed in August, then theoretically the amount of CPU used by HPDESK during that month should decline. So the expected growth for that month would be a negative number.

The Graph of Projected Growth is produced from the table of projected growth by business elements. The sum of all business elements are added together and plotted by month in a graph. This is probably the most important graph of the Capacity Planning Report. An example of the table of projected growth and graph of projected growth is on page 8 and 9 of this paper.

You can use this graph to show your track record of predictions versus actual values recorded. By doing this you can determine if the predictions for growth need to be revised. Your first few reports will probably not be very accurate but with experience, you will get a better feel for predicting future growth.

The Capacity Planning Team for our IBM mainframe has found that the margin of error on all the business elements usually averages out to be fairly small. That is they do not have good track record predicting individual business elements, but the margin of error on all business elements averaged together usually comes out to be very small.

SYSPLAN does have a TREND command that does a linear projection of resource usage, but this command has a major flaw. When the TREND command is used, its projection is based on all hours and days of the month with no exceptions. We use our HP very heavily during prime time, but at night and on the weekends, it is just about an idle CPU. CPU usage during prime time is 70% to 80%, however using the TREND command shows the CPU usage around 50%. I don't know about you, but I would have

a hard time justifying the purchase of a new CPU based on the graph produced by the TREND command.

Ideally, the TREND command should let you input what the growth will be from month to month and let you eliminate exceptional days. You would be surprised what a few holidays would do to a linear projection.

DISK IOs

A graph of disk IOs is also included in the report. This is for historical purposes. Since we are not ready to upgrade to a HPPA processor, we needed to extend the life of our current CPU. To do this we are replacing all of the 7933XP disc drives with 7937XP disc drives. Then we will cache IOs at the disc level instead of in main memory. The theory being that the 7937XP disc drives can cache IOs ALMOST as effectively as main memory can. We anticipate freeing up 5% to 10% or our CPU but at the risk of becoming IO bound. Therefore we want to track our average number of disk IOs during the prime time.

Disk Space

Freespace is also included in our capacity planning report. We order another disc drive when 75% of our total disc capacity is full. In general it is hard to offer more services or add new systems when disc space starts to get tight. A lack of disc space can become a performance and/or service inhibitor.

Miscellaneous

The Miscellaneous Information is a good place to put trivial facts that have potential of impressing somebody. For instance this section is a good place to put how many sessions and jobs that occurred on the HP on a average month.

University Systems as a general rule doesn't charge anything for the use of the HP. However we do have a charge back program. We are going to use this program to produce bills to send to our customers with a note explaining that this is the amount of free services we provided them with last month. This produces good will among our customers and conditions them to the day when we might have to start charging for use of the HP. The dollar amount that the HP could bring in if we so chose is good miscellaneous information!

Actions Taken to Improve Performance

You always want to include in your capacity planning report what you did over the last period to improve performance. Sometimes it is hard for upper management to see the monetary value of a system programmer because a system programmer doesn't directly bring in any income. This is a section to let readers of this report know what you do and how valuable you are.

Recommendations

This section is used to make recommendations to improve performance. The recommendations may be as simple as optimizing disk IOs over LDEVs to buying a larger CPU. The idea is to be planning as far into the future as possible. You do not want to wake up one morning and find out you need new CPU. On the other hand you do not want to buy a larger CPU sooner than required.

Conclusion

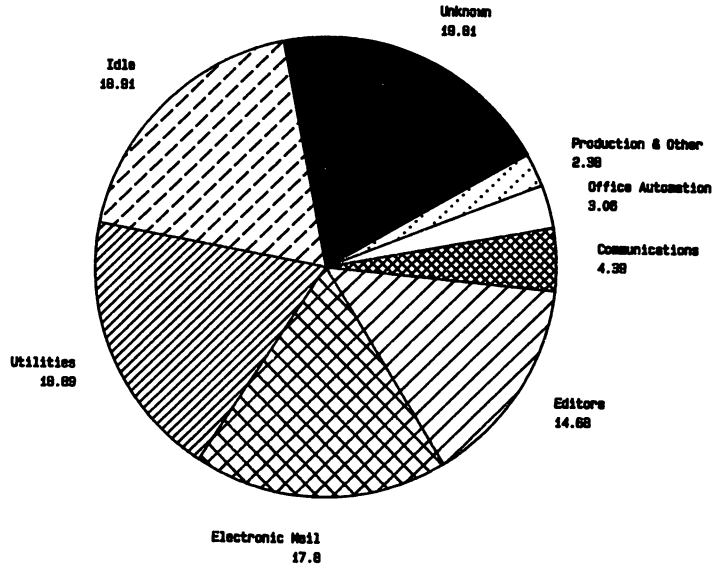
There are several mistakes that can be made in producing the capacity planning report. One mistake is for the systems programming group to become consumed with producing the capacity planning report and have no time for any other activities.

Another mistake is to not expect any surprises. A capacity planning report can only guess what will happen in the future. No matter how good one is at predicting the future, there will come a time when the prediction will miss by a long shot. Usually the cause of the inaccuracy is from underestimating the amount of CPU a system will consume. Be prepared for the worst.

Every computer center should have some idea of its current and future computing load. A capacity planning report can provide management with the necessary information to formulate a plan to accommodate or deny future demands for computing power. The hard part about producing a capacity planning report is getting started. Once started though you will find the report evolving into very useful tool.

UNIVERSITY SYSTEMS HP3000 SERIES 70 US3/RED

Average CPU Used by Business Element during Peak Hour



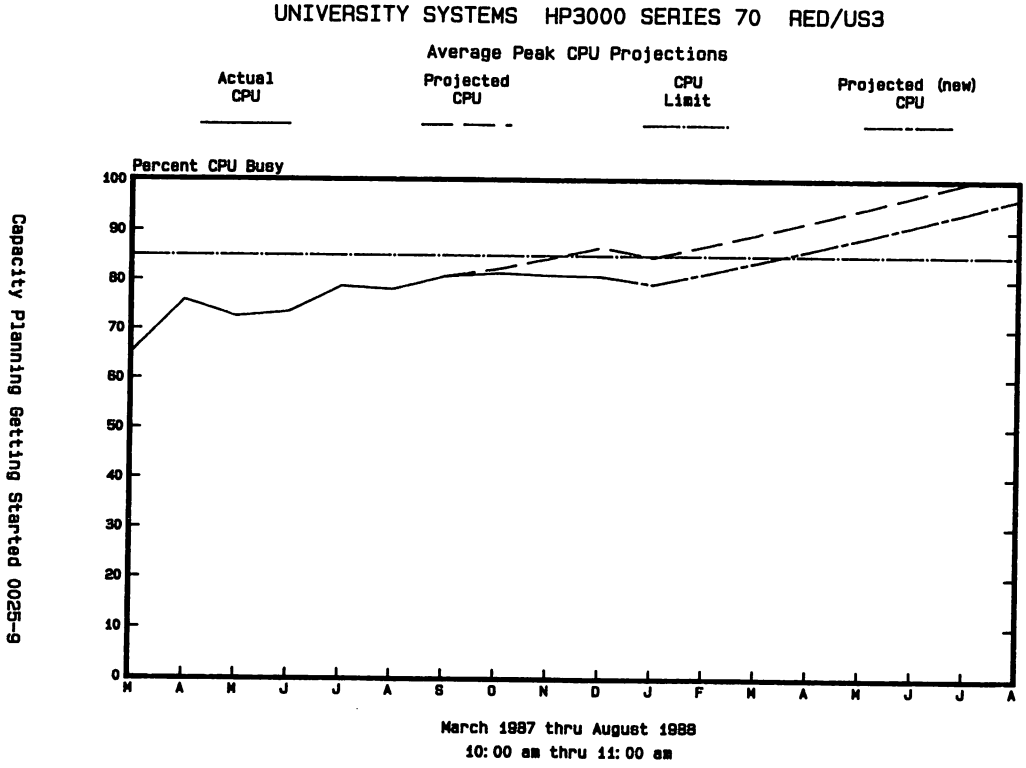
October 1987 thru December 1987
10:00 AM to 11:00 AM

PROJECTED PEAK TIME CPU FALL QUARTER '87

	7/87	8/87	9/87	10/87	11/81	12/87	1/88	2/88	3/88	4/88	5/88	6/88	7/88	8/88
Production	3.6	3.2	1.9	1.1	0.4	0.35	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Office Auto	2.0	1.4	2.2	3.3	3.0	2.76	2.8	2.8	2.9	2.9	3.0	3.1	3.1	3.2
Datacom	4.5	4.5	5.2	4.3	4.4	4.14	4.1	4.2	4.2	4.3	4.3	4.4	4.4	4.4
Utilities	15.5	16.4	15.4	17.8	16.0	22.33	22.7	23.2	23.7	24.1	24.6	25.1	25.6	26.1
Editors	17.2	16.6	17.9	15.1	14.3	14.10	14.3	14.6	14.9	15.2	15.5	15.8	16.2	16.5
HPDESK	16.4	17.0	15.3	20.1	17.6	15.09	11.6	11.8	12.0	12.3	12.5	12.8	13.0	13.3
Other	19.3	18.7	22.5	19.4	24.9	22.00	23.1	24.2	25.4	26.7	28.0	29.4	30.9	32.5
Total	78.7	78.1	80.7	81.4	80.9	80.78	79.2	81.4	83.7	86.1	88.6	91.2	93.8	96.6

PROJECTED GROWTH BY BUSINESS ELEMENT

	7/87	8/87	9/87	10/87	11/81	12/87	1/88	2/88	3/88	4/88	5/88	6/88	7/88	8/88
Production							1	1	1	1	1	1	1	1
Office Auto							2	2	2	2	2	2	2	2
Datacom							1	1	1	1	1	1	1	1
Utilities							2	2	2	2	2	2	2	2
Editors							2	2	2	2	2	2	2	2
HPDESK							-23	2	2	2	2	2	2	2
Other							5	5	5	5	5	5	5	5


UNIVERSITY SYSTEMS HP3000 SERIES 70 RED/US3
Average Peak CPU Projections
Actual CPU
Projected CPU
CPU Limit
Projected (new) CPU
Percent CPU Busy
100
90
80
70
60
50
40
30
20
10
0
M A M J J A S O N D J F M A M J J A
March 1987 thru August 1988
10:00 am thru 11:00 am
Capacity Planning Getting Started 0025-9