Network and Systems Management Services and Support

Balanced Systems and Capacity Planning

P.T. Borchetta
R.J. Wicks

Washington Systems Center Technical Bulletin

GG22-9299-04

Network and Systems Management
Services and Support
Washington Systems Center
Technical Bulletin

Balanced Systems and Capacity Planning

P.T. Borchetta
R.J. Wicks

GG22-9299-04
August, 1993

Fifth Edition (August 1993)

Publications are not stocked at the address given below. Requests for IBM publications should be made to your IBM representative or to the IBM branch office serving your locality.

A form for readers' comment is provided at the back of this publication. If the form has been removed, comments may be addressed to: Ray Wicks, IBM Corporation, Washington Systems Center, 800 N. Frederick Avenue, Gaithersburg, MD 20879-3395.

Contents

Figures

Tables

Special Notices

The information contained in this document has not been submitted to any formal IBM test and is distributed on an "as is" basis **without any warranty either expressed or implied**. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Performance data contained in this document was determined in a controlled environment; and therefore, the results which may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

References in this publication to IBM products, programs, or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM licensed program in this publication is not intended to state or imply that only IBM's program may be used. Any functionally equivalent program may be used instead.

Programming Interfaces

This publication is to help the customer do performance analysis and capacity planning. The information in this publication is not intended as the specification of any programming interfaces. See the PUBLICATIONS section of the IBM Programming Announcements for more information about what publications are considered to be product documentation.

Trademarks

The following are trademarks of the IBM Corporation:

IBM
Hiperspace™
MVS/ESA™
MVS™
MVS/DFP™
ESA/370™
Hipersorting™

Changes and Thanks

Changes for this edition....

- An index was added.
- Processor performance section was revised to expand the application of the Erlang-C formula.
- Added a section to discuss growth.
- Updated the CP90 example to include I/O analysis.
- Expanded the section on Latent Demand.
- Added a section on Sample Selection.

- Correct the relationship between LCU and BCU.
- Added additional formula for predicting the impact on paging when processor storage is changed.
- Expanded the appendix on RIOC.

For this edition, miscellaneous clarifications and corrections suggested in part by reader feedback forms and reviews from:

- Remi Dawalibi of Los Angeles
- Dan Valter of Toronto
- Shelly Weinberg of Los Angeles

All mistakes still contained herein, remain ours.

Abstract

Resources, in a balanced system, are not independent variables in determining system performance. The interrelationship of processing power, I/O capability, and processor storage size will determine system performance and along with performance objectives, their capacity.

This document examines the factors involved in processor and I/O performance and their interrelationship. What do I expect from these resources? How does the performance of one affect the other? And how can one size requirements?

Chapter 1. Performance Topics

What follows in this document is a discussion of system performance and capacity planning. System performance is difficult enough without any philosoph ical discussions. But, we are faced with a chicken and egg problem. How much component performance need one know to handle system problems? And, how can one understand the significance of any component performance issue without some system performance?

I have attempted to establish the concept of balanced systems. It represents a relationship that exists between the individual system resources. This relationship will define system performance. First I will discuss the dependance of system performance on the individual components, then follow that by a discussion of the performance factors of the individual components, and finally put the relationship together in a sample capacity plan.

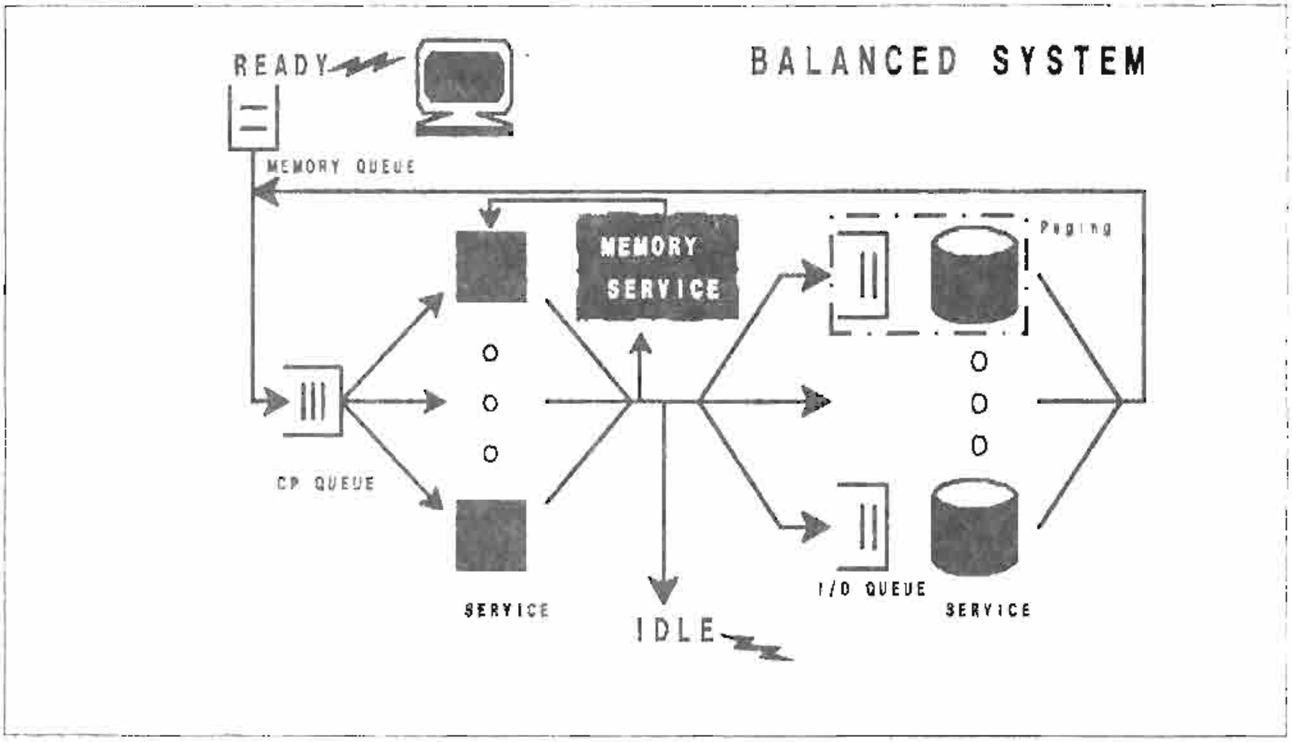


Figure 1. A system Model

The model in Figure 1 represents a common, though simple, system model which is useful when discussing both performance analysis and capacity planning. It will act as my table of contents for comments on processor and I/O performance, and storage guidelines. I will not consider the network resources required. End-user response time is defined here as internal (Host) response time.

From an end user's point of view, none of the model is significant. What is important is the length of time from ENTER to EXIT, from the time between the transaction entering the system until it exits (network excluded). If I am unhappy about that time, it becomes a problem for systems analysts and capacity planning people.

The model serves as a conceptual structure to analyze a performance problem. The model has four basic components.

- A memory queue - here the user acquires memory (frames) in order to execute instructions and I/O commands. The user has to "get into storage".
- The Central Processor (CP) queue and CPs - here the user is ready to execute instructions and queues up for a CP of which there could be more than one. Response time of this server is composed of queueing time and service time.
- The I/O subsystem - device queueing and device service time.
- A server which I have called "memory service". This service is used to replace an I/O request with a request to obtain data from processor storage, either central storage or expanded storage. The technique used may be the use of Hiperspaces in MVS/ESA.

When the user finishes with the I/O service, more CP service is usually required. The amount of CP service required before an I/O request will determine the I/O intensity of the work being done. Eventually, after some CP service, the request will be finished, and the user exits.

The delay for I/O service can be reduced by replacing the need for an I/O operation with a memory service. Placing data in memory reduces I/O time, measured in milliseconds, with a memory reference. Data in memory can be accomplished with access methods such as VSAM by buffering. Access methods such as Data in Virtual (DIV) and DIV using data spaces, will automatically buffer data in memory by design. Hiperspaces permit efficient use of expanded storage. Which ever method is chosen, access to data is improved significantly.

The replacement of an I/O operation with memory service, a replacement of one resource with another, albeit obvious, requires that the resource be available. It makes no sense to attempt to place more data in memory, in a memory constrained system.

The performance of this "system" is determined by all four resources. How long does the user wait before getting all the required storage? How fast is the CP response time? The I/O response time? How many CP-I/O sequences are required? How much I/O can be replaced with a memory service?

There are some general observations about this system which can be made under the "Forced Flow Law".[1]

The forced flow law is the basis of balanced systems. It can be best described as a plumbing problem describing fluids flowing in a set of pipes.

1 See *Quantitative System Performance* by Lazowska et al.

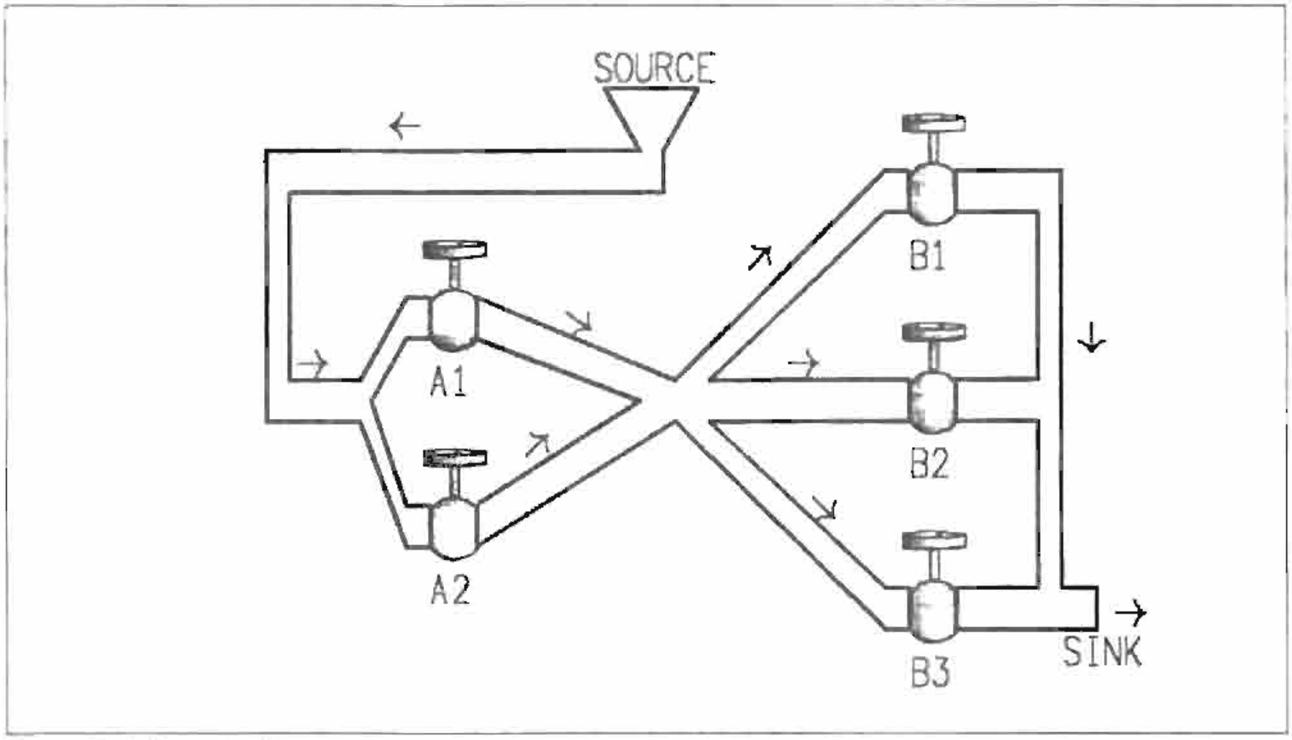


Figure 2. A Plumbing Problem

Let's assume that a fluid is being pumped from the top to the bottom (source to sink in modelling terms) in Figure 2. What could we do to improve the flow through this system? (See how systems analysis sounds like plumbing?)

Could the flow through the B-valves be improved to obtain an improvement in system flow? You should observe that the flow through A-values must EQUAL the flow through the B-valves. We can't have more fluid flowing through the B-valves than through the A-valves and the other way around. Correct? That means to improve the TOTAL SYSTEM flow, we would have to insure that each of the set of valves can handle the desired flow. The forced flow law says that

$$(A1+A2) = (B1+B2+B3)$$

Now look back at the "real" system model. The first set of valves is the CP service or flow. The second set is the I/O subsystem. To improve the flow through the entire system, could we just improve the processor service flow without a corresponding or balanced improvement in the I/O subsystem? Could we forget the amount of storage required by more work?

Like the plumbing problem, system performance and capacity is NOT limited to one resource. System Performance is ALL the Resources in Balance.

The response expectation from such a system is characterized by the graph in Figure 3 on page 5. As the rate at which the system processes work (transactions), the response time increases and it does this non-linearly.

For the capacity planner, the question most given is "How much work can the system do?" From just this chart the answer appears to be "As much as you want". (As long as you don't mind rotten response time.) What happens in fact? As the amount of work increases, some point is reached which constitutes a user's threshold of pain. Then the phone starts ringing.

RRRRRRING.
"Hello, system support speaking."
"What's going on down there? Response time is */& ¬$#¬%*"
"Who is this?"

The last question is significant. The action of the system support people will depend upon the identification of the caller. Is it someone who doing priority work? Priority work requires appropriate action. Priority work runs the business and that work has to get done in a **timely manner**.

This implies either an assumed objective or a specified one for given workload types. It is in terms of this objective that the capacity for a given collection of software and hardware is defined. You can see that with different objectives, the same set of hardware and software can have different capacities. It is not the hardware and software alone that define capacity, but those resources in conjunction with objectives.

When the objectives cannot be met, the system is out of capacity.

There are three ways to gain the resources needed:

- *Buy it.*
- *Create the illusion you bought it. This is known as tuning. Capability to do this implies you have been previously wasting resources. Like purchasing, there is a cost. The cost is people, while it may be higher, it is always less visible than purchase.*
- *Steal it (take it from a less important application). Again there is a cost. Here the cost is lower service to the application from whom the resources were stolen.*[3]

---

3 *MVS Performance Management*, by Siebo Friesenborg and Gary Hall, (GG22-9351) Page 4.

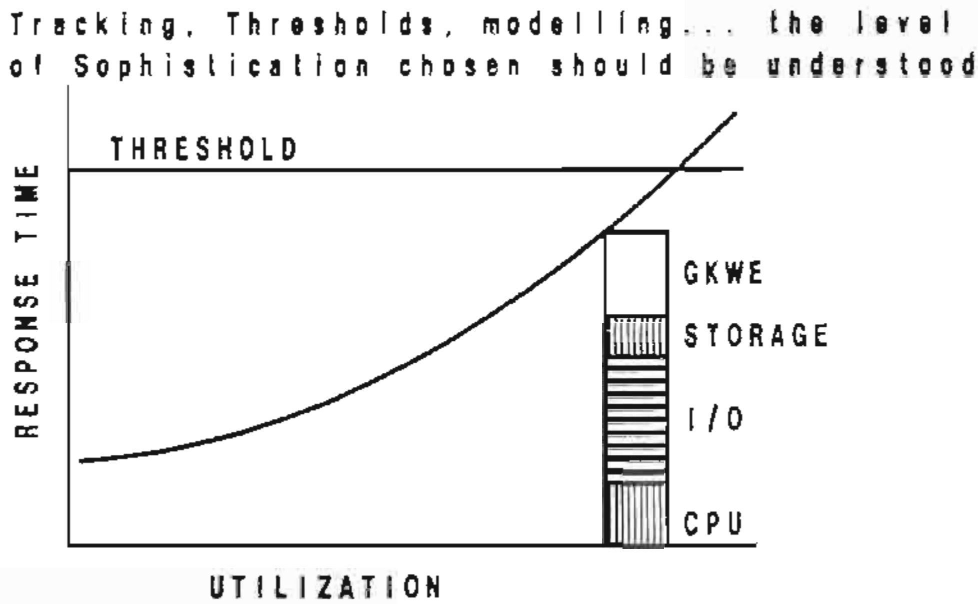


Figure 3. The Resource Components

What to do? From the system model in Figure 1 on page 1, we can see that the components of transaction response time are composed of processor (CP or CPU) and I/O response time. Storage delay, and other components.[3] The other components, in real systems, might be logical delays such as ENQ delays or application queueing. We will be interested in the first three components. Scientific applications aside, I/O time can represent the major portion of transaction time. With the advent of increasing use of Data in Memory (DIM), where the I/O content is dramatically reduced by using larger memory, the time spent in CPU and memory contention is not necessarily less than I/O time.

The shape of the component in the graph is intended to illustrate that as the transaction load increases, the responsiveness of the processor or I/O is affected. It is not that the transaction necessarily uses more processor *service time* (TCB+SRB time), but that the *response time* increases. Response time is queueing+service and the queueing time increases. The transaction has to wait to get at a central processor or DASD device.

[3] God Knows What Else.

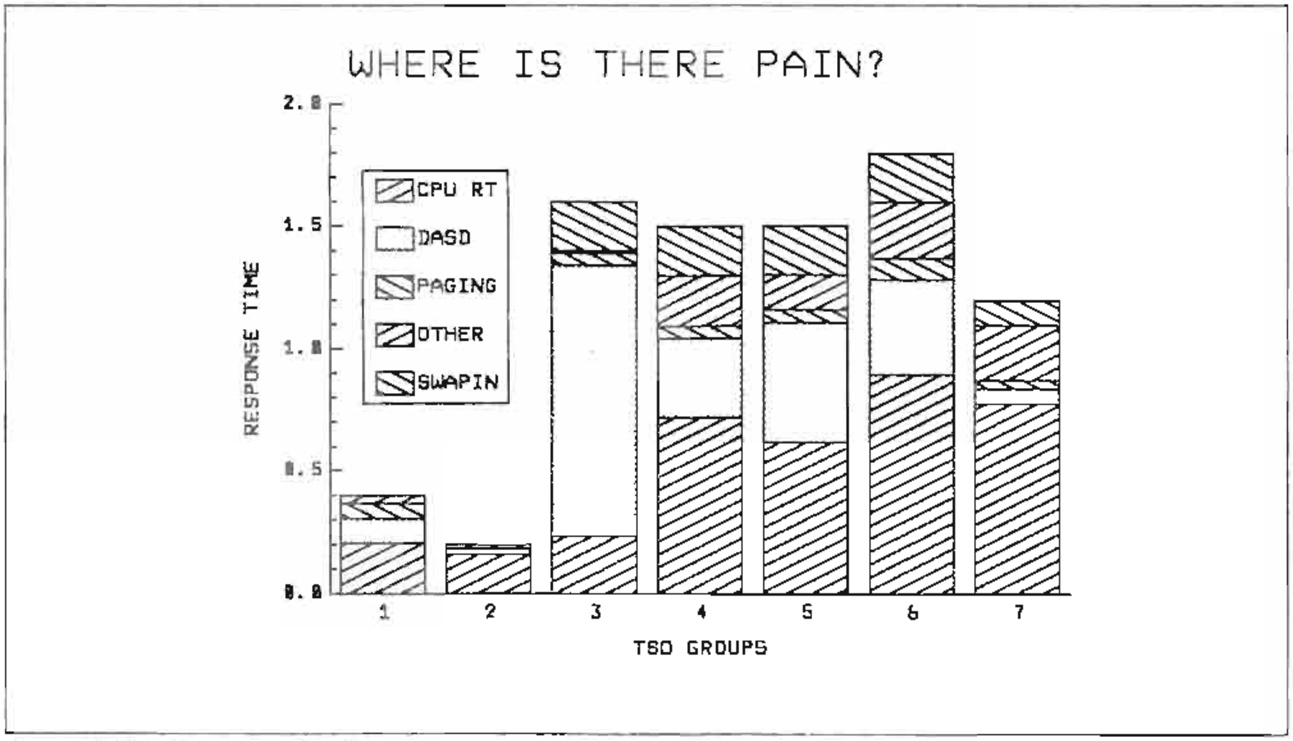


Figure 4. Which resource first?

The message so far indicates that

- System performance, system flow, is defined in terms of all the individual resource performance.
- All work flowing through the system is not equally important.

In this breakdown of TSO time (Figure 4) you can see that all TSO users are not equally happy and the reason for their unhappiness is different. If the response time objective was one second, how can the response time problem be solved? For group 3 the problem is DASD. For group 7 the problem is processor. Groups 1 and 2 are hoping that you don't touch anything: they never call system support. What to do? Well, who is most important? The process, in order, says,

1. Establish Service Level Objectives. The SLO is a negotiated agreement, by business unit, the response time threshold for a given rate - for a transaction rate less than N, the response time will be less than T.
2. Performance management means understanding pain. Where does it hurt when the SLO is not being met? Where is time being spent? Identify the problem response time component.[4]

---

4 RMF Monitor III does some of this automatically. See Figure 21 on page 23.

Balanced Systems Law

THRUPUT IS LIMITED BY THE SLOWEST SERVICE CENTER.

Thruput Limit = 1/Smax

Smax = 250 Ms. service

Limit = 1 / 0.25 = 4 per sec.

Figure 5. Balanced Systems Law

One major principle to remember, before we look at each of the resources in turn, is embodied in the Balanced Systems Law. This law states that the throughput of a forced flow system is limited by the slowest service center. For example, consider an interactive system which required each transaction to do 10 I/O operations to a specific dataset. At 25 ms. per I/O, each transaction would require 250 ms. of I/O service (Smax is the maximum transaction service at this server). How many transactions per second can the *system* do? Four. That's without knowing the size of the processor, the amount of storage, or anything else about this system. Interesting, no?

Let's look at each of the resources in turn and attempt to understand the factors which determine the responsiveness of a resource (or server).

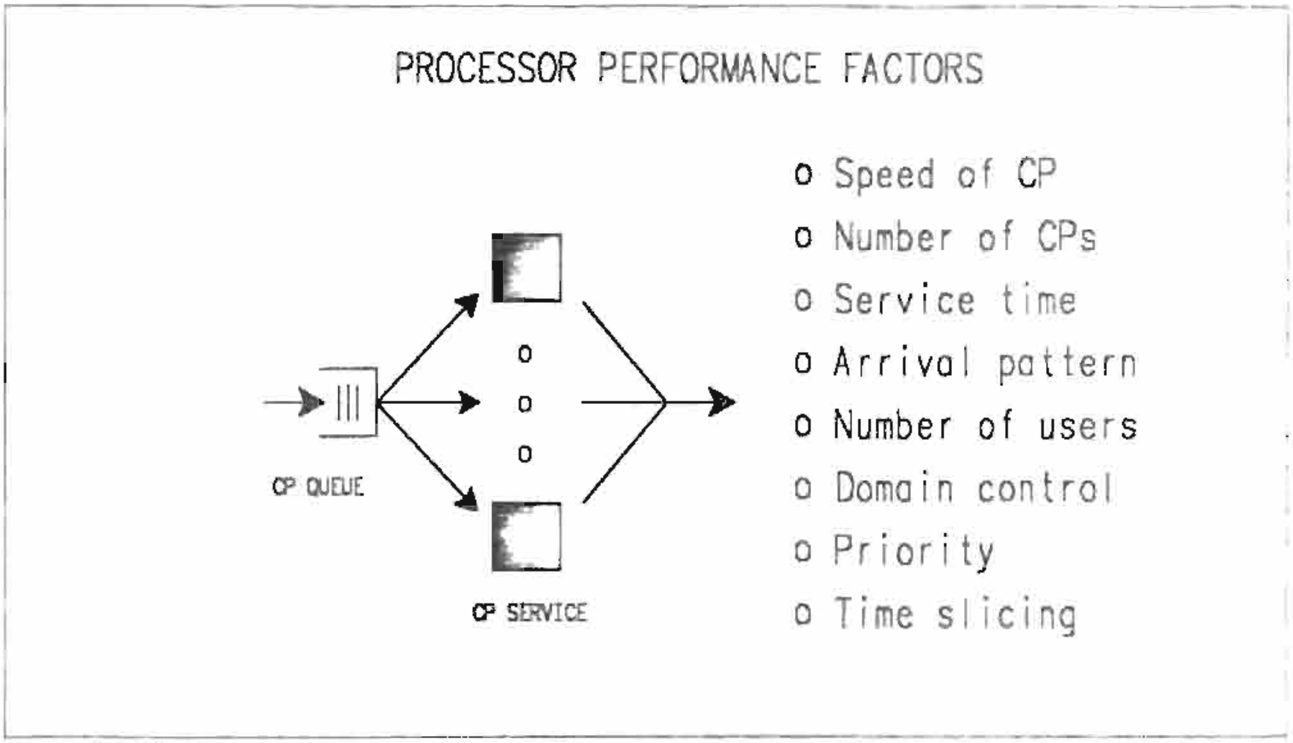


Figure 6. CP Performance

The factors which determine CP responsiveness are the same for many serving mechanisms. Most familiar to us is a bank. What factors contribute to our expectations for "teller" response time?

- How fast are the tellers?
- How fast are customers arriving? (Especially before I get there.)
- How many tellers are serving?
- What kind of transactions are being executed?
- How many customers are there?

The performance analysis for this problem is the same as that of the processor response time. The factors listed in Figure 6 are the same. The factors are generic.

In MVS software, there are controls which can be used to determine who gets access to processor service. To follow the bank analogy, to obtain the illusion of good service in the bank (the processor), we place a guard (the SRM) at the door to stiff-arm people. Only a controlled number get into the bank. Inside the bank, waiting is short and the lobby is not crowded. However, the guard does let his friends in first. Not everyone is equal. The determination of friends in software is done by domain. And even once in line for service, it is done in priority mode. The important go to the front of the line. CP service is not a democracy.

AN UGLY QUEUING THEORY FORMULA

o ERLANG'S C FORMULA (M/M/C) FOR PROBABILITY OF C SERVERS BUSY

$$C(c,u) = \frac{u^C}{c!} \Big/ \left[\frac{u^C}{c!} + (1-\rho) \sum_{N=0}^{C-1} \frac{u^N}{N!} \right]$$

$$C(1,u) = \rho$$

$$C(2,u) = (2\rho^2) \div (1+\rho)$$

$$C(3,u) = (9\rho^3) \div (2+4\rho+3\rho^2)$$

$$C(4,u) = (32\rho^4) \div (3+9\rho+12\rho^2 + 8\rho^3)$$

Where u=total traffic
ρ=Avg. Utilization

Figure 7. Queueing Theory

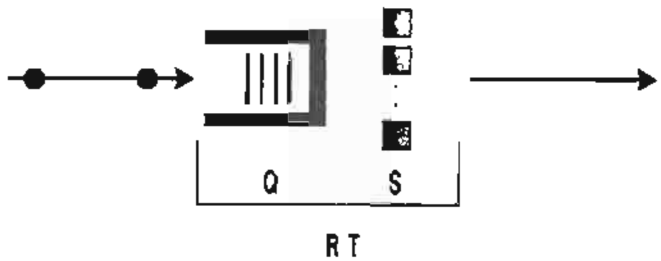
Queueing theory formulae which describe serving behavior can be very simple. The formulae shown in Figure 7 are very menacing since they have to concern themselves with multiple servers. **These formulae and their assumptions** will be the basis of my performance expectations For those weak of heart, I suggest you pass on to the next page while those heartier souls review the formulae.[5]

The first formula estimates the probability that the servers will be busy at any time. That's C(c,u) Where "c" is the number of servers and "u" is the traffic intensity. The Greek letter ρ is the average server utilization (u/c). For a single server (c=1) the formula reduces to C(1,u)= ρ and for two servers (c=2) the formula yields C(2,u)= (2×ρ**2)÷(1+ρ).

If, for a single server, the server is 50% busy, the formula says that, as expected, the probability of finding the server busy is 0.5. Intuition would say that if there were two servers, both 50% busy, the probability of finding both busy would be 0.5*0.5 or 0.25. The formula says it is 0.33. How come? It has to do with the assumptions. How many users are there? How do they arrive? For example, if there were only one user, the probability of finding the 50% utilized server busy would be 0.0! It's not busy when the one user is not using it.

[5] See Table 5 in *Probability, Statistics, and Queueing Theory* by Arnold Allen. Watch the "C" and "c", they mean different things.

QUEUING THEORY M/M/c



$$E[RT] = E[S] + E[Q]$$
$$= E[S] + \frac{C(c,u)E[S]}{c(1-\rho)}$$
$$= E[S]\left(1 + \frac{C(c,u)}{c(1-\rho)}\right)$$

Figure 8. Response Time Calculation M/M/c

The introduction of the C(c,u) formula was necessary to develop a formula for response time.[6] Given the amount of Service (S) requested by a business unit, we want to know the processor response time for this request given the other activity in the system.

We need to know the probability that the business unit will find the server(s) busy when ready to use the server and the average service for the business unit. C(c,u) is used for that. For the expected amount of service $E[S]$ one could use the total CPU service time reported in RMF or SMF divided by the number of transactions. Figure 8 shows the formula which will be used in the next three examples.

Example 1 is in Figure 9 on page 11. Let's assume that one is interested in an estimate of the processor response time seen by the DB/DC workload at various utilization or growth points. In this example, there is some higher priority work which uses an average of 16% of each CP. The DB/DC workload is currently using about 12%.

Note that there is lower priority work in the system, but the DB/DC calculation doesn't bother with lower priority workloads. Why? Because the dispatcher uses a preemptive priority scheme. This effectively eliminates any interference from lower priority work. As mentioned earlier, important work jumps to the appropriate place in line. In this example, not to the front but ahead of lower priority work. This means that the DB/DC workload will see the servers busy any where from 16% to 28%. That's the lower bound of only the higher priority work, to a

6 The use of capital "C" and small case "c" in these formulae occur in the literature. Why? One reason could be that this is another method of torture invented by academics. You know what happens in a chalk talk, don't you? As the talk goes on, the small "c" gets bigger, and the big "C" gets smaller.

higher bound of the sum of higher priority work and DB/DC itself. For this calculation, the higher bound will be used.

Assume a DB/DC service time of .0075 seconds and a growth rate of 25%. Figure 9 shows the expected processor response time versus utilization where the utilization starts around 28% and goes to 95%. The processor is a 9021-640 which has 2 CPs.

There doesn't seem to be much of a concern until the utilization goes over 85%. The service time is plotted at a constant .0075. The true response time is in the shaded area somewhere. Why only *somewhere*? Remember our discussions earlier about the factors affecting multiple server response time? Such as different distributions. The assumptions here are all that goes with the M/M/c theory.

You could use your RMF Monitor III or Omegamon to help calibrate the estimation.

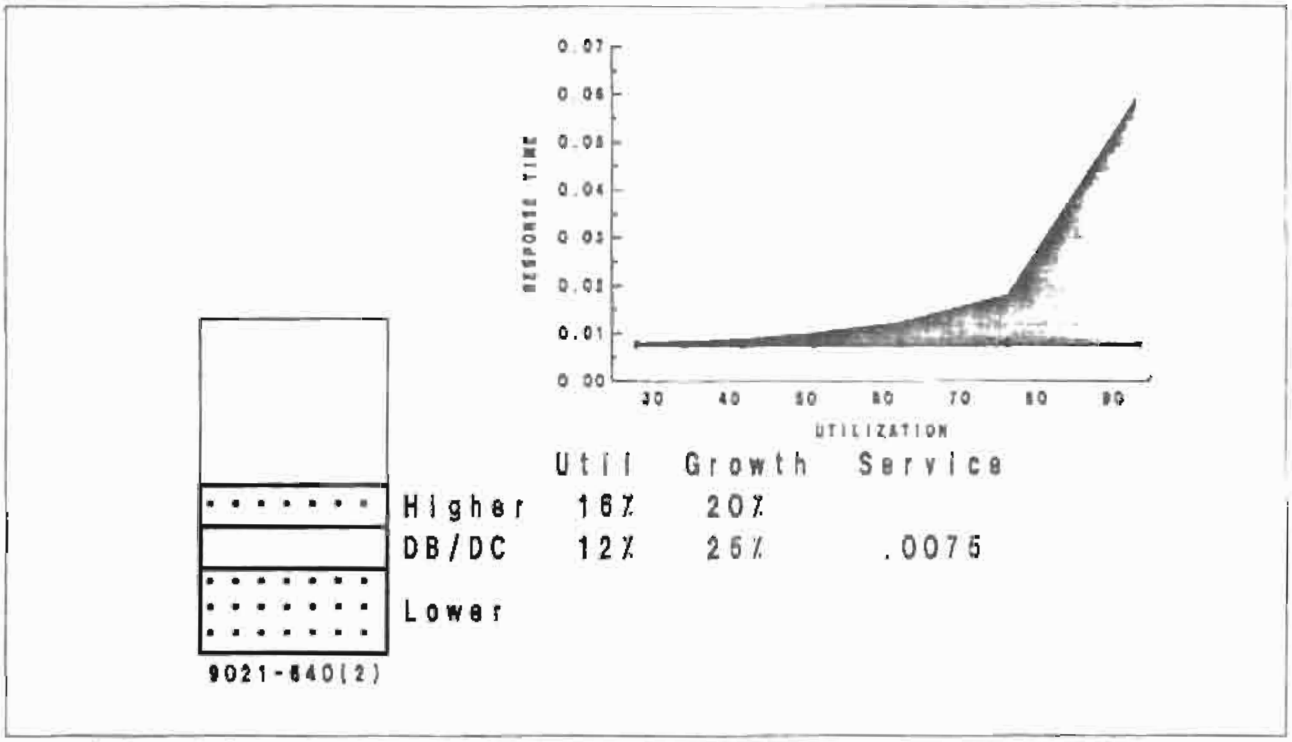


Figure 9. Example 1

Example 2 is in Figure 10 on page 12. Here the formula is used directly to compare three processors. The processors have a different number of CPs and the total power is different. The purpose of this example is to examine the trade-off of response time and throughput. The base processor is the 9021-640. The service time is 0.02 seconds. To adjust the service time for the other processors, I used the average power per CP compared to the base processor. Although the 3090-600S has more power than the 640, the power per processor is less. Hence the service time is longer.

Notice that at low utilizations, the response time is nearly identical to the service time. Also notice that the lines converge at very high utilizations where the amount of queueing is large enough to over come any service time differences. Until very high utilizations, the 3090-600S never catches up. The situation for the

9121-742 with 4 CPs is different. At around 80% the queueing on the 2 CP 640 becomes larger more rapidly than the 4 CP 742. In fact the lines cross (albeit a tiny difference and probably not empirically detectable). This raises the possibility that slower CPs might do better if there's enough of them.

The significance of all of this will depend upon the impact on total response time. Usually, the end user would not detect a change in processor response time in the order of 0.02 to 0.05 seconds. However the difference in total power for these machines could be significant to throughput. That's shown in Figure 11 on page 13. This is essentially the same data as the previous figure except the utilization for each processor has been converted to transaction rate. Here the 9121-742 leads the pack with throughput at higher utilizations even though it was slightly behind the 640 in response time at lower utilizations.

Clearly this set of graphs is dependent upon the models chosen in the comparison.[7]

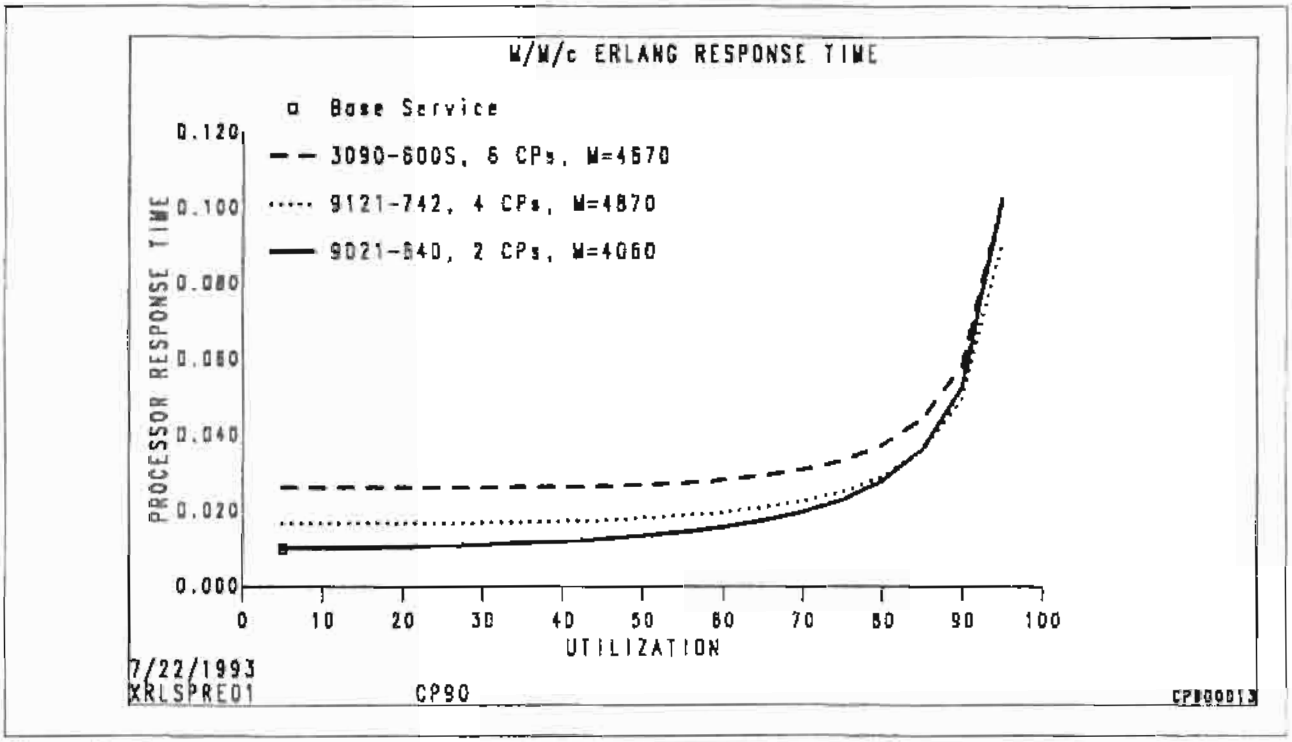


Figure 10. Example 2

[7] A generalized facility can be found in CP90 on HONE for IBM SEs and on IBMLink as a service to customers.

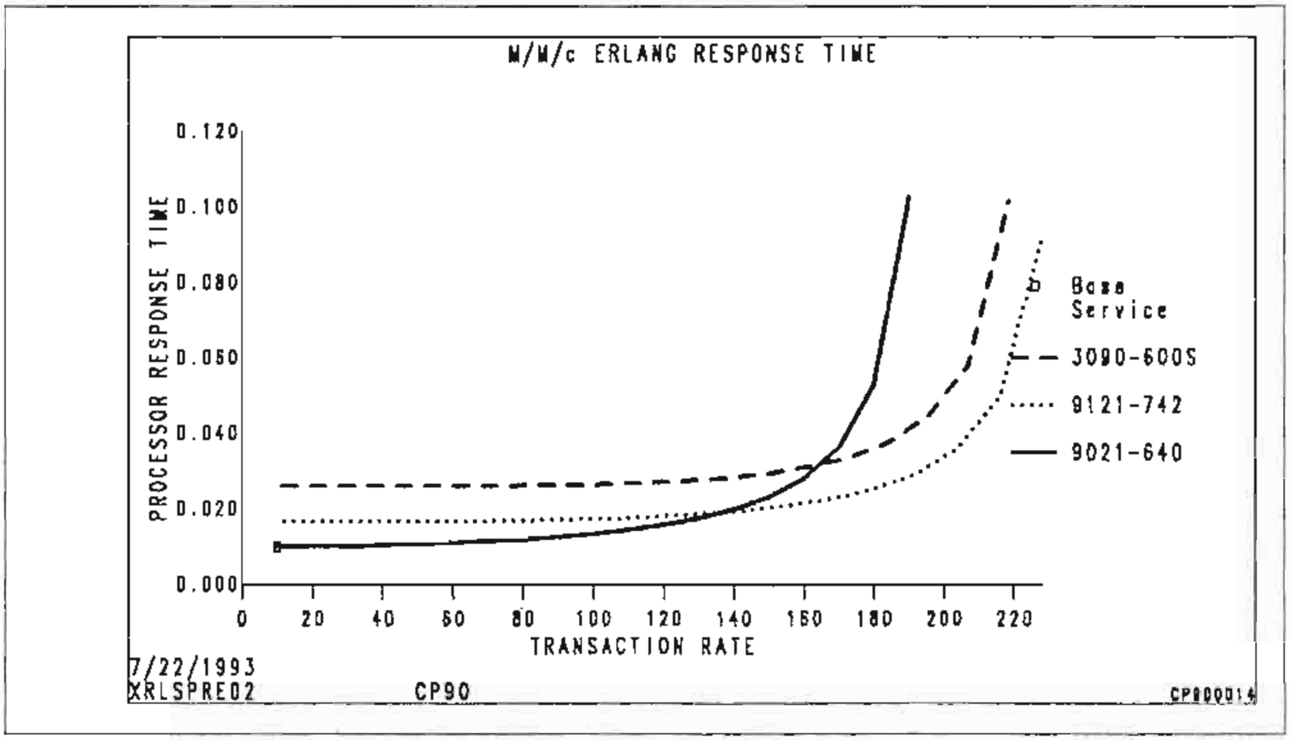


Figure 11. Example 2 Continued

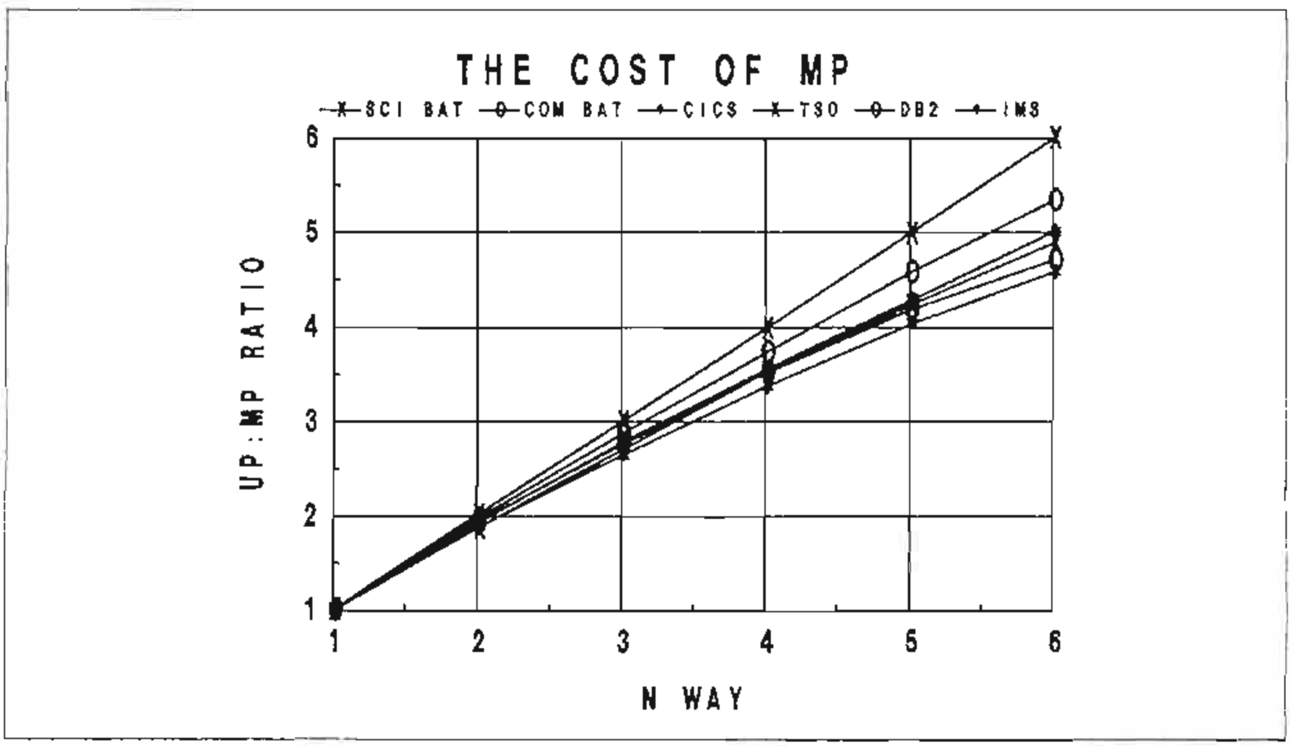


Figure 12. MP Overhead

Before looking at Example 3, let's look at the data for MP overhead. Will a 2 CP configuration be twice as fast as a 1 CP configuration? It will be close, but not exactly twice. As the number of CPs in a machine increase, the overhead increases. The CPs not only interfere with each other (accessing data) but require points of synchronization or points of serialization and a certain amount of cross talk to keep everything straight. Put another worker in my office and some amount of *interference* (albeit pleasant) is generated. Put two more workers there, and the cross talk increases.

I may be *busy* but not productive. My *ITR* falls. (ITR, Internal Throughput Rate or number of transactions per CPU second is a measure of processor power. More on this later.) Figure 12 on page 13 shows that the decline in productivity is workload dependent. Scientific batch has the least problem, IMS the most. The advantage in scientific batch is in its data and instruction stream characteristics. It does not use a lot of system services (common data and queues), it does not share data (common area or between memories), and is a single memory application. The other workload types do these things to a greater or lesser degree.

An advantage of a single system image, usually not discovered in a benchmark, is the ability of a large single system image to deliver more power to a high priority application over periods of varying demand. Excess power from one system image is usually not available to other images (outside of shared CPs in LPAR). When the high priority tasks are idle, the power can be used elsewhere.

A single system image also saves processor storage. The single image requires only one copy of the system (Nucleus, PLPA, SQA, etc). This is particularly important in smaller systems where the system itself represents a large proportion of the total storage.

This discussion does not consider any management advantages of single system image over multiple system images. One advantage usually worth the cost is the ability to maintain one application image instead of splitting the application over multiple images.

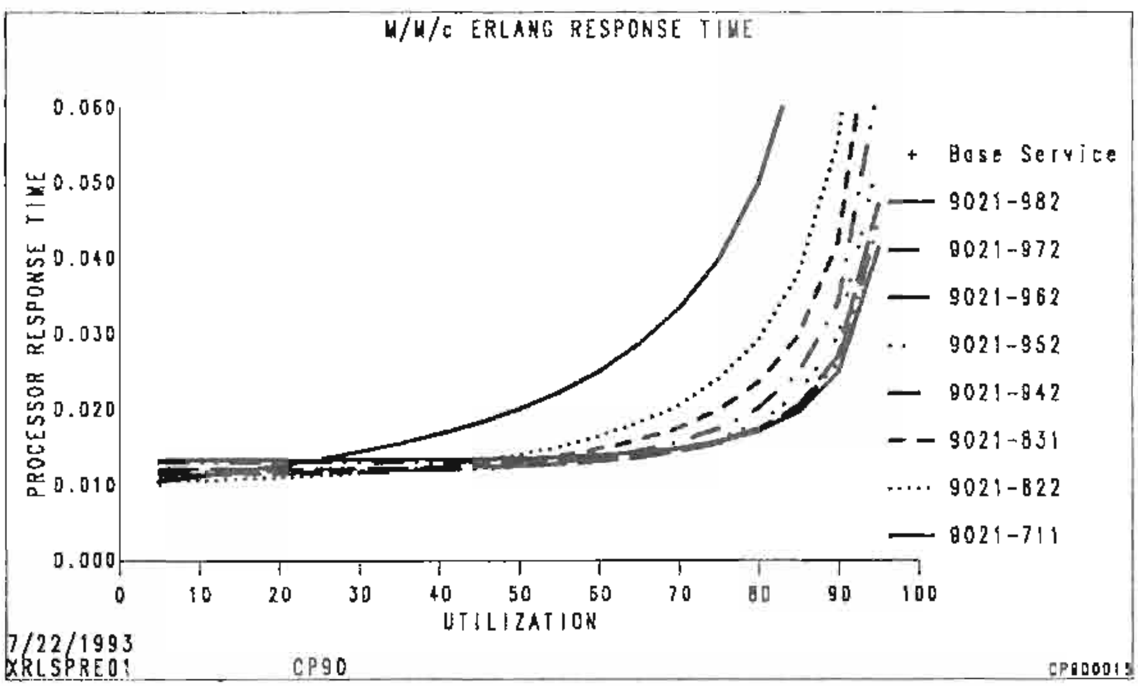


Figure 13. Example 3

Now let's continue with the examples of applying the Erlang-C formula. **Example 3** is in Figure 13. Using the 9021 family of processors which have models from one to eight CPs, one can see the impact of adding CPs. The curve flattens, as expected, as the number of servers increase.

From the discussion above, we know that the configurations with a smaller number of CPs (servers), each CP is slightly faster than the configuration of a larger number of CPs. But this slight advantage is minor compared to the advantage of having a larger number of servers. However, after about 4, the advantage appears to be one of throughput since the configuration with a larger number of CPs has more power.

That's the theory. What about assessing the amount of queueing that occurs in an actual system?

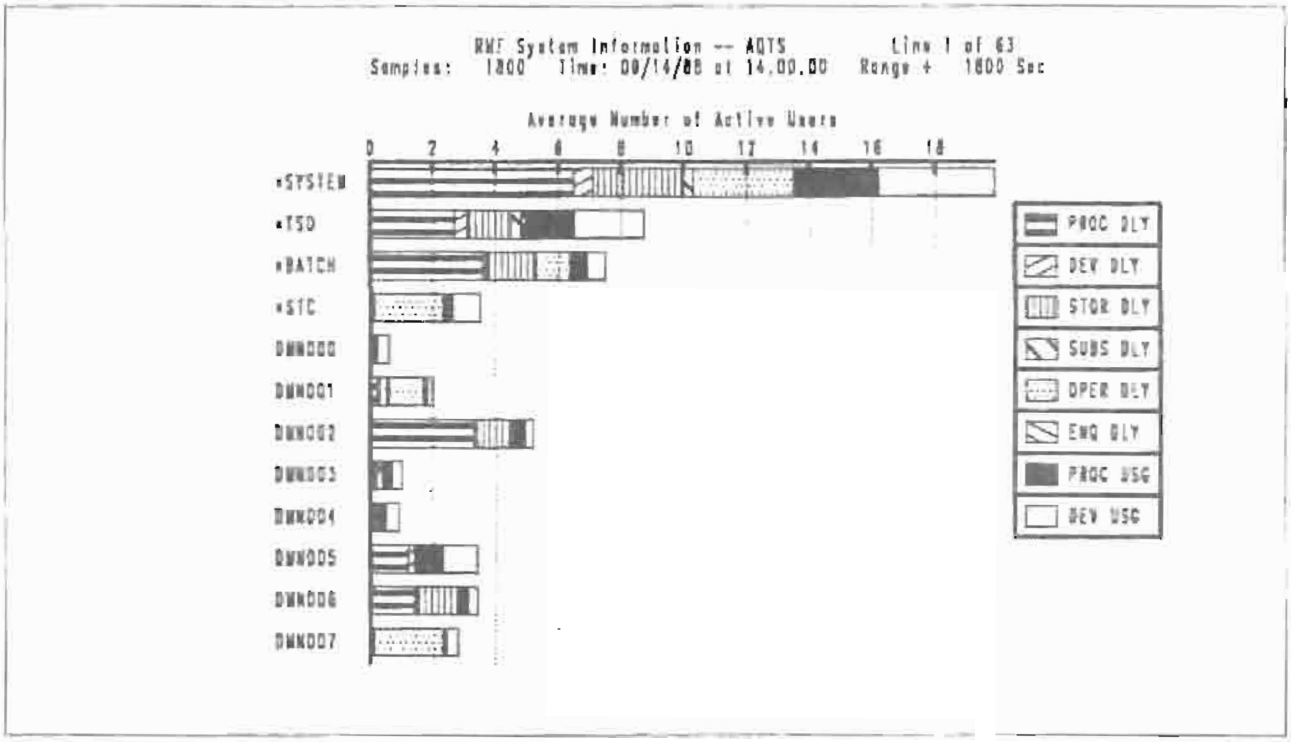


Figure 14. RMF Monitor III - System Information

What data is available to monitor processor response delay? RMF Monitor I provides the distribution of queue lengths in the CPU Activity Report. This value appears as the variable OCPU in the SLR (Service Level Reporter, IBM program no. 5665-397) Performance Management Reports. Figure 14 illustrates the RMF Monitor III System Information Report. Along with the system summary, the state of each work type (TSO, Batch, and Started Task), Domain, and Performance Group (not shown) is given. If the domains and performance groups are set up by business unit, it is possible to discover the problem of that business unit and assess the impact of a solution. Wouldn't it be nice to assess the impact of a migration from a 3090-300S to a 400S? Would the work in Domain 7 be interested in a system with more CPs? How about Domain 2? Domain 2 averages 3 users waiting for processor service. Additional engines have the potential of solving a problem for the work in this domain. Domain 7 shows little or no work being delayed by the processor. Hopefully this reflects the correct priority. The Delay report in RMF Monitor III will describe the processor delay in terms of percent of time waiting for the processor.

PROCESSOR SUMMARY

DELIVERY : CP SPEED, ARCHITECTURE,
SW & HW IMPLEMENTATION,
NUMBER OF CPs

RESPONSE: THRESHOLDS, CP SPEED,
NUMBER OF CPs

CONTROL : PRIORITY
TIME SLICE

Figure 15. CP Summary

The summary for the processor resource is shown in Figure 15. One of the key items to remember is that threshold utilizations apply to both DASD and processors. As the utilization of a server gets higher, the waiting line gets longer. It is a fact of life.

High utilization is not necessarily a problem as long as the low priority work is just that - work at the bottom of the queue and the priority is really low. High utilization becomes a problem when the work at the bottom of the queue is not meeting its performance objectives.

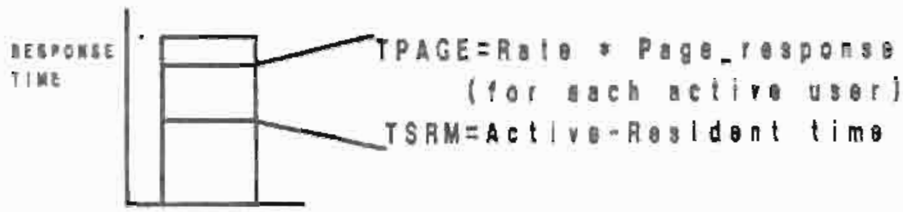
Additionally, work at the bottom of the dispatching queue, although it may be meeting its SLO, can cause storage problems. What is the job doing while it waits to be dispatched? It is holding central storage frames. This processor delay can exacerbate a storage problem.

The expectation for processor response time will depend upon the power of each CP and the number available in a single system image.[1]

[1] Although I didn't discuss it, software can have a significant effect on processor responsiveness. Is a priority scheme implemented? How about time slicing? How is it implemented? The history of system control programs illustrates the differences and advantages. Application software can have built in queueing too.

Storage.

THE PAIN OF STORAGE DELAY



- ★ MORE STORAGE
  (Central or Expanded)
- ★ CHANGE MPL
  (Paging -> Swapping)
- ★ FASTER PAGING

Figure 16. Storage Delay

The lack of processor storage causes both swap in delay (TSRM) and page delay (TPAGE). Although a page fault which is resolved by ASM can mean a delay of about 30 milliseconds, many page faults can cause an annoying delay. How much can be tolerated? That's a very individual question.

In general, delay is to be avoided. Whenever page delay can be avoided, do it. The size of page delay for any transaction is obtained by multiplying the page fault response time for a single page by the number suffered by each transaction. The delay has to be weighted by the number of users affected. If it is a TSO user, one person waits. If CICS page faults, all active users wait.

The page delay time is too high when the response time of a transaction exceeds its objective and the page delay time is a significant portion of the response time. I did not say that the page fault rate is too high. It is the page delay that counts! A page fault rate of 2/second can be worse than 5/second if the former takes 40 Ms each and the latter 10 Ms each.

Paging I/O also adds to the contention of the I/O subsystem when page data sets are not on dedicated paths.

The problem of excessive page delay can be solved in a number of ways. Reduce the page fault rate by providing the transaction with more processor storage. This is done by either setting the multiprogramming level lower (keep some users out of storage, swap instead of paging) or getting more processor storage (central or expanded). Or reduce the page fault time by providing a faster paging mechanism.[9]

The data required to determine whether you have a paging problem can be found in Figure 14 on page 15. The sample data shows about three address spaces (ASCBs) on average delayed for storage. One is enough if it is CICS or IMS. The identity of the ASCB can be found in either the domain summary or performance group summary in this report or more detailed storage reports in Monitor III. In this sample, if the work in domain 5 was the most significant in the installation, the magnitude of the storage delay would be reduced. Domain 5 isn't suffering from the pain of storage delay. However, if domain 6 was CICS, the problem takes on a different complexion. One of the CICS regions would always be in page delay. That would be serious.[10]

Other than paging and swapping, a more interesting aspect of Storage is the role of I/O elimination or Data In Memory (DIM). Of course, the elimination of paging is the elimination of I/O, but how about the elimination of useful I/O? Useful I/O is related to the movement of *MY* data, *MY* programs.

Making data available for processor operations is the name of the game in computers. After all, it is called *data processing*. Lets look at I/O, how it operates and the performance impact.

[9] See *MVS Paging Performance Considerations* by Siebo Friesenborg (GG22-9264).

[10] See *Zero Swapping/Paging I/O - A Reality* by John Ryden in 1984 CMG Conference Proceedings, Page 531.

I/O

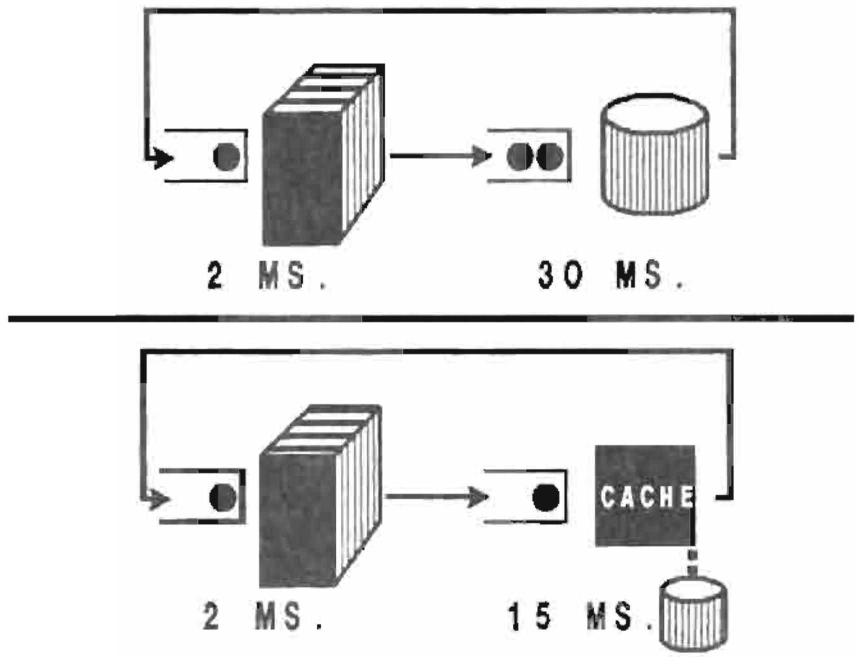


Figure 17. I/O

The I/O subsystem can be viewed as a server for I/O service requests from the processor. A system requires enough I/O to satisfy the rate of requests with a reasonable response time. From the system model illustrating the forced flow law, we can say that work flowing through the "CPU subsystem" will be at the same rate as the I/O subsystem. The flow through the CP "valves" will be the same as the I/O "valves".

From the CP point of view, you need enough I/O to keep the processor busy. Symmetrically, from the I/O point of view, you need enough processor to keep the I/O busy. It does have to be balanced to keep work flowing through the system.

On a system the size of a 3090-200S, 500 I/O operations per second would not be uncommon. That's one every 2 ms on average. An I/O operation takes about 30 Ms without cache and about 15 ms with cache.[11] With a high enough Multi-Programming Level (MPL) the processor can be kept busy. This requires about 100 uncached actuators to do this. (The capacity plan will determine an exact number.) However, each individual task, without cache, has to wait about 30 ms. Cache might reduce that significantly.

[11] See *DASD Expectations* for a complete discussion of performance expectations for DASD I/O.

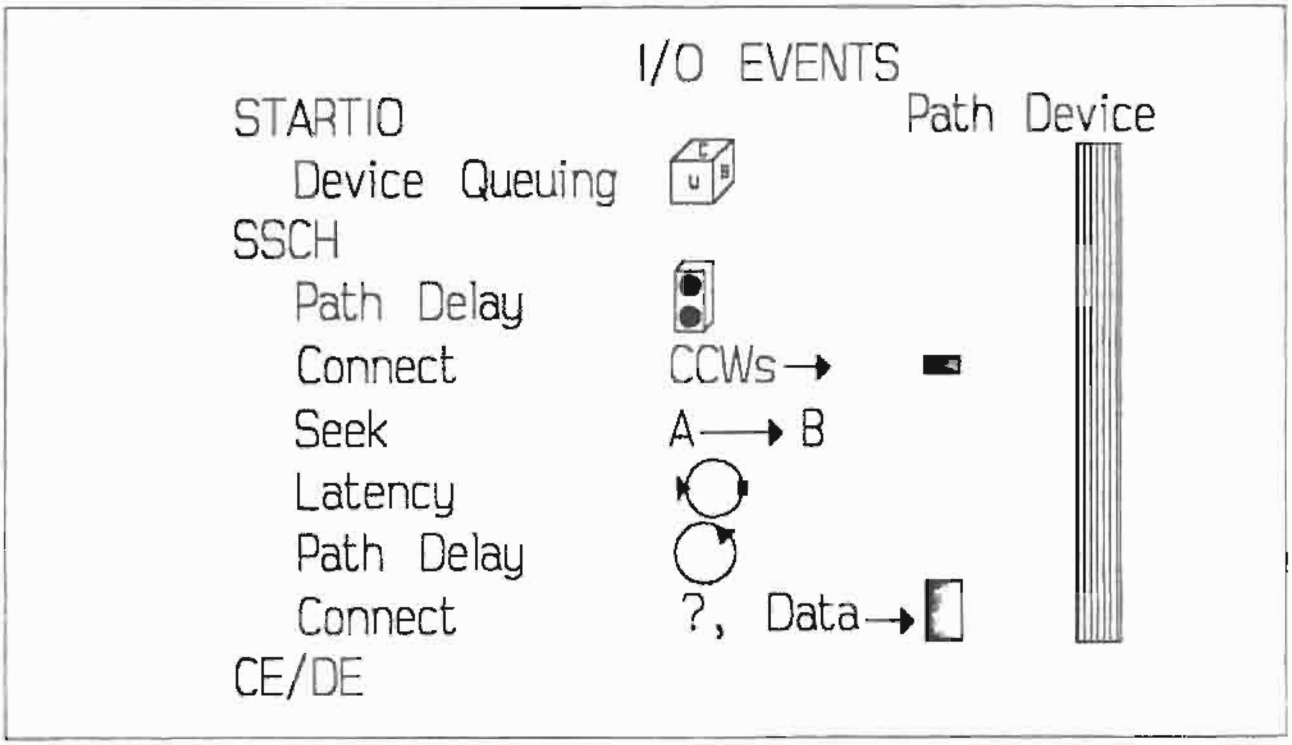


Figure 18. I/O Events

The performance of an I/O device can be more complicated than the processor. The I/O events outlined in Figure 18 begin to illustrate this complexity. Analysis of response time expectations would involve establishing expectations for the individual events. This has been done elsewhere.[11] The point of our discussion here is to note that I/O events involve mechanical and electronic events. The duration of each event will depend upon workload, I/O architecture (S/370, S/370-XA, and ESA/370), and device implementation (3380, 3390 for example).

The process or I/O analysis proceeds in a similar manner to transaction performance analysis. In Figure 4 on page 6 the transaction was decomposed to reveal "where time was spent", euphemistically, "Where is the pain?" Figure 19 on page 21 shows a similar approach for I/O. "Where is I/O time spent?"

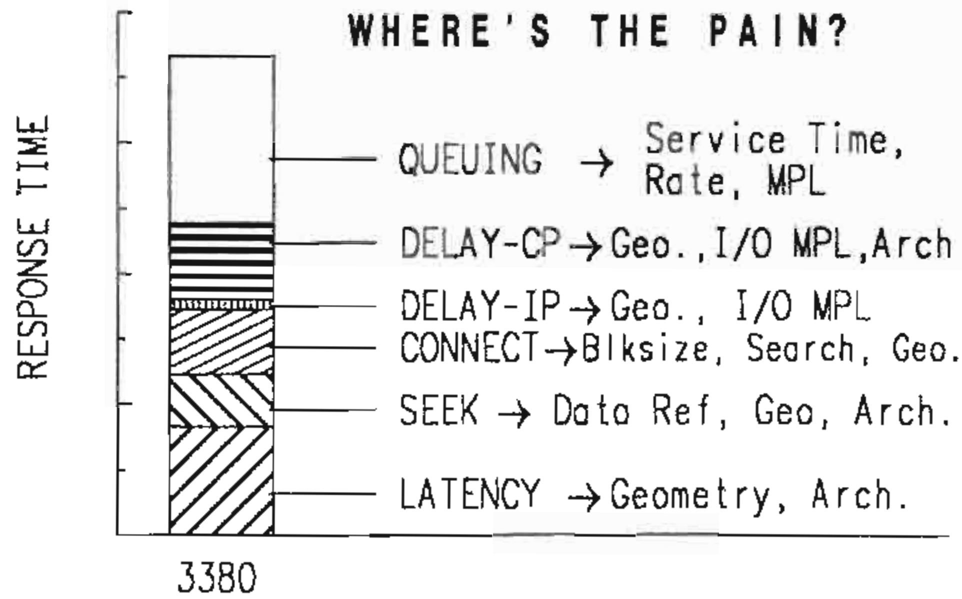


Figure 19. Where's the I/O Pain?

The analysis proceeds in identifying the size of each component and the cause of each component. For example, connect time, reported by RMF Monitors I, II, and III, is a result of the use of the interface between a device and central storage. This can occur when transferring CCWs, Data, or when the device is performing a function which requires the interface to be connected.

If the size and cause of the components can be determined, an intelligent business decision can be made about a choice of device improvements. Figure 20 on page 22 is an intelligent person's guide to DASD and controllers. It is a shopping list. The list matches specific problems with devices whose characteristics address those problems.

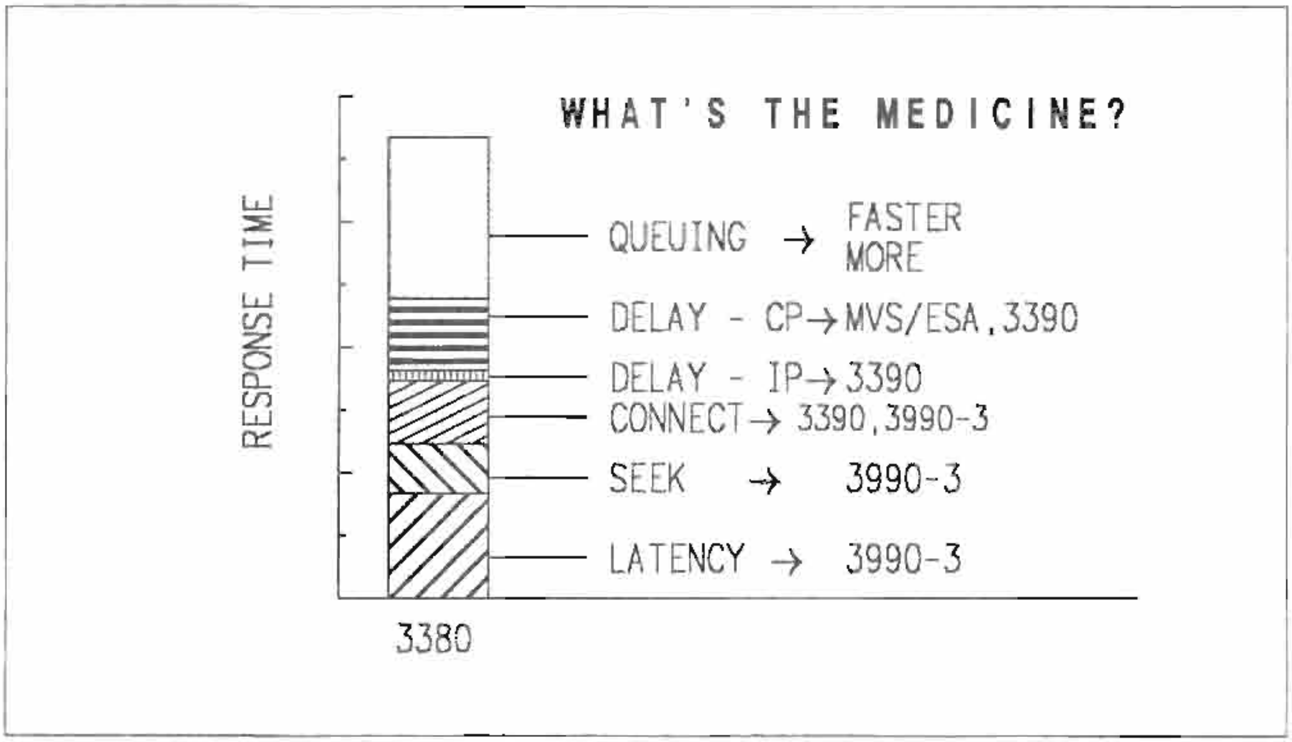


Figure 20. A Solution to Match the Pain

How is the I/O pain chart developed? First of all, is there an I/O problem? And, how can it be recognized? The process is

1. Is there a problem?
2. Who's got it?
3. Do I care?
4. What can be done about it?
5. Did I do it?

This is the general flow of *Performance Management*.

Step 1. Check Figure 14 on page 15. This RMF report indicates whether there is any device delay. It also measures the size of the problem by the number of users affected.

Step 2. Check Figure 21 on page 23. This is a sample of the delay report from another system.[12] Whose got the problem? This is more important than the number of users affected. Remember, I/O service, like the processor, is not democratic. One important job or transaction being delayed can be much more significant that a number of other transactions or jobs. The referenced report in Monitor III indicates where delays exist by specific resource by specific users. This report happens to illustrate the delays for the Batch jobs only. Here again is a useful report which aids in the assessment of performance impact. "Who will be helped?" "How much?"

Step 3. Who is being delayed often determines whether any action is necessary. Very important work requires immediate attention. Less important work? Well, maybe we'll look at it tomorrow.

[12] The data from system E90L was generated using some synthetic jobs to create pain whereas system AQTS is production data.

Step 4 To identify the solution to an I/O problem one must identify where is the I/O pain Again we appeal to RMF Monitor III The Monitor III DEV and DEVR report would give an initial assessment of the location of pain Is it device queueing, connect, disconnect or pend time? See Figure 22 on page 23

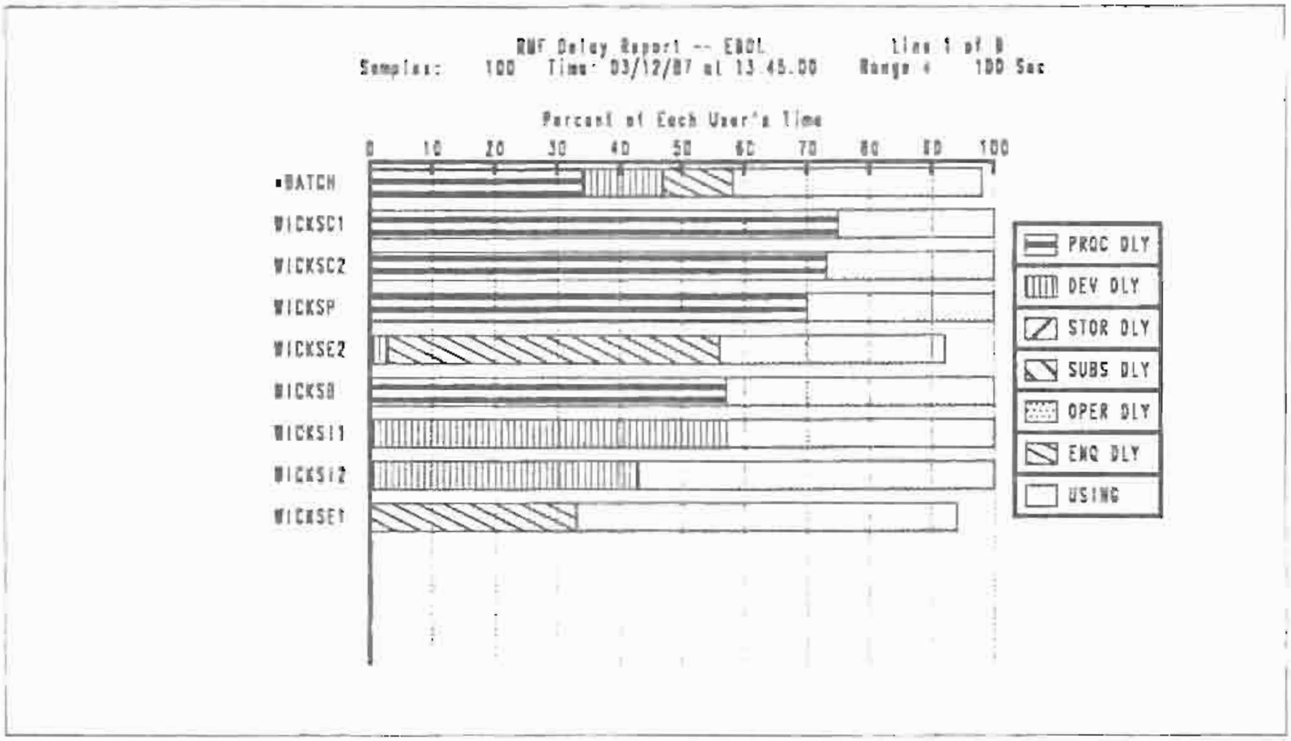


Figure 21. RMF Monitor III - DELAY Report

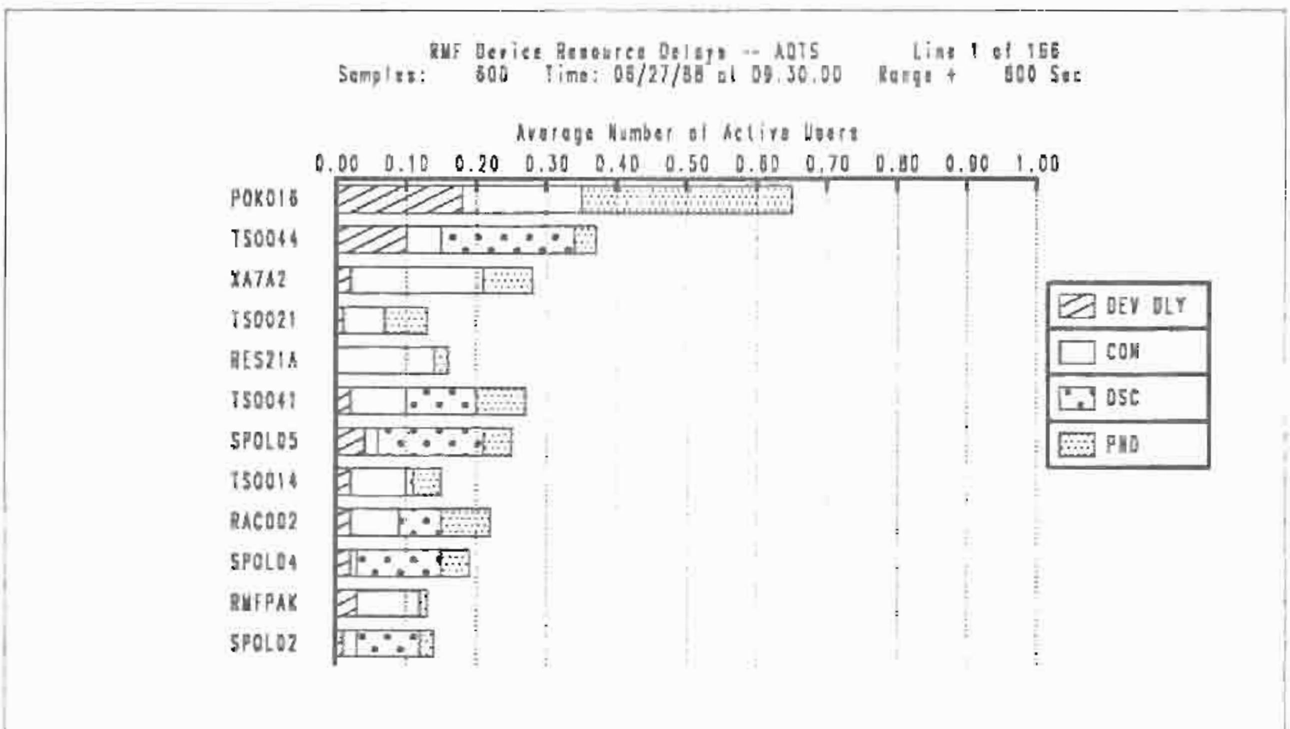


Figure 22. RMF Monitor III - DEVR Report

```
                       RMF Job Delays — EBOL                        Line 1 of 1
          Samples:     100   Time: 03/12/87 at 13.46 00     Range =    100 Sec

Job: WICKS11        Primary delay: Excessive disconnect time on volume HSM806.

Probable causes: 1) Sequential access of data with short blocksize.
                 2) Active datasets spaced widely across volume.
                 3) Overloaded channel paths causing reconnect delays.

-------------------------- Volume HSM806 Device Data ---------------------------
Number:     1C0        Active:        96%      Pending:    5%    Average Users
Device:     3380       Connect:       13%      Delay DB:   4%       Delayed
Shared:     Yes        Disconnect:    78%      Delay CU:   0%         1.0

--------------------------- Job Performance Summary ----------------------------
                  WFL USG DLY IDL UKN    ------ % Delayed for ------   Primary
C  ASID  DMN  PG   %   %   %   %   %     PROC DEV STOR SUBS OPER ENQ Reason
B  0151   11 106  43  43  57   0   0        0  57    0    0    0   0 HSM806



Command ===>                                                  Scroll ===> CSR
```

Figure 23 RMF Monitor III - Device Report

RMF Monitor III version 4.1.1 has built into it some suggested solutions to problems it identifies. Figure 23 is one of those reports. Suggested solutions are listed.

A brief set of I/O problems and solutions follow.

Device Queueing or Delay The I/O MPL for a device is too high. This could be from too many datasets being used at the same time or just dataset contention. In the former case, one could spread the datasets, run the requests sequentially, or make the I/O faster.

Pend Time This includes channel and control unit contention when starting an I/O. If it is larger than 2 milliseconds, the reason is almost exclusively shared DASD. Check the sharing systems for device contention. Large Pend time is inter system device queueing.

Connect Time This is large blocks for a sequential data set, or many blocks as in a Swap device, or search time on a PDS. In the former case, large connect time is very efficient. It usually means that lots of data is being delivered. Often it represents the same data being delivered over and over again. This problem can be reduced with caching or even I/O elimination. For example, if the device is a program library, the MVS/ESA LLA/VLF feature should dramatically reduce it.

Disconnect time If the total connect time on the paths (total rate*connect) to this device from the other devices is high, the disconnect time is probably RPS delay. Otherwise, it is seek time which is caused by many datasets in use or one very very large dataset. This is solved in a manner similar to device queueing or with a cached controller

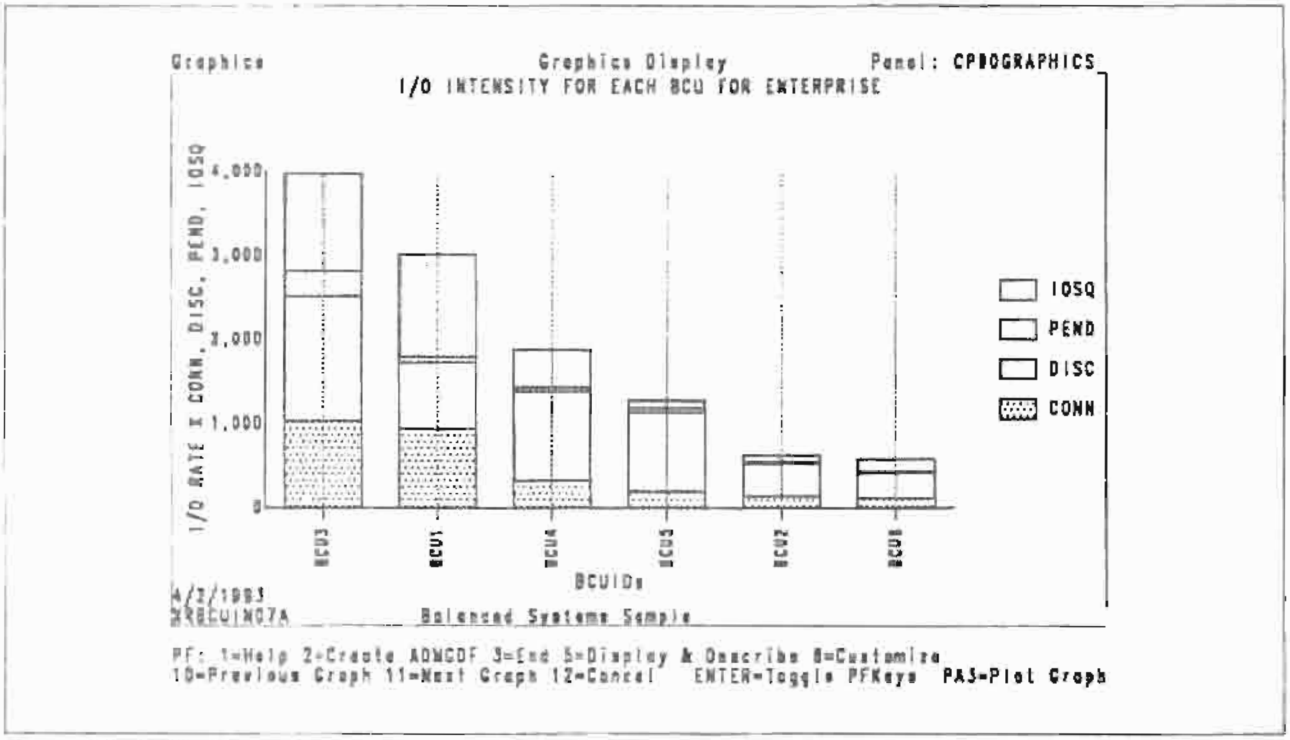


Figure 24. BCU intensity example

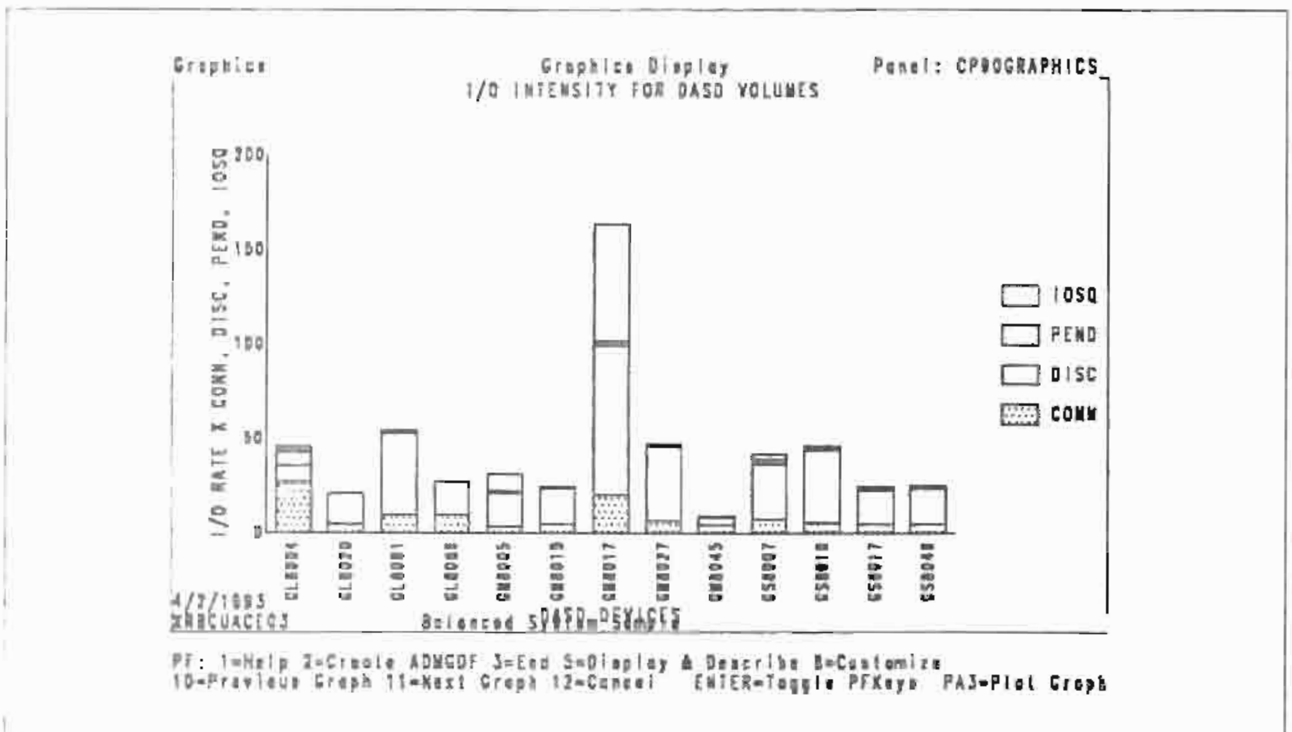


Figure 25. Actuator intensity example

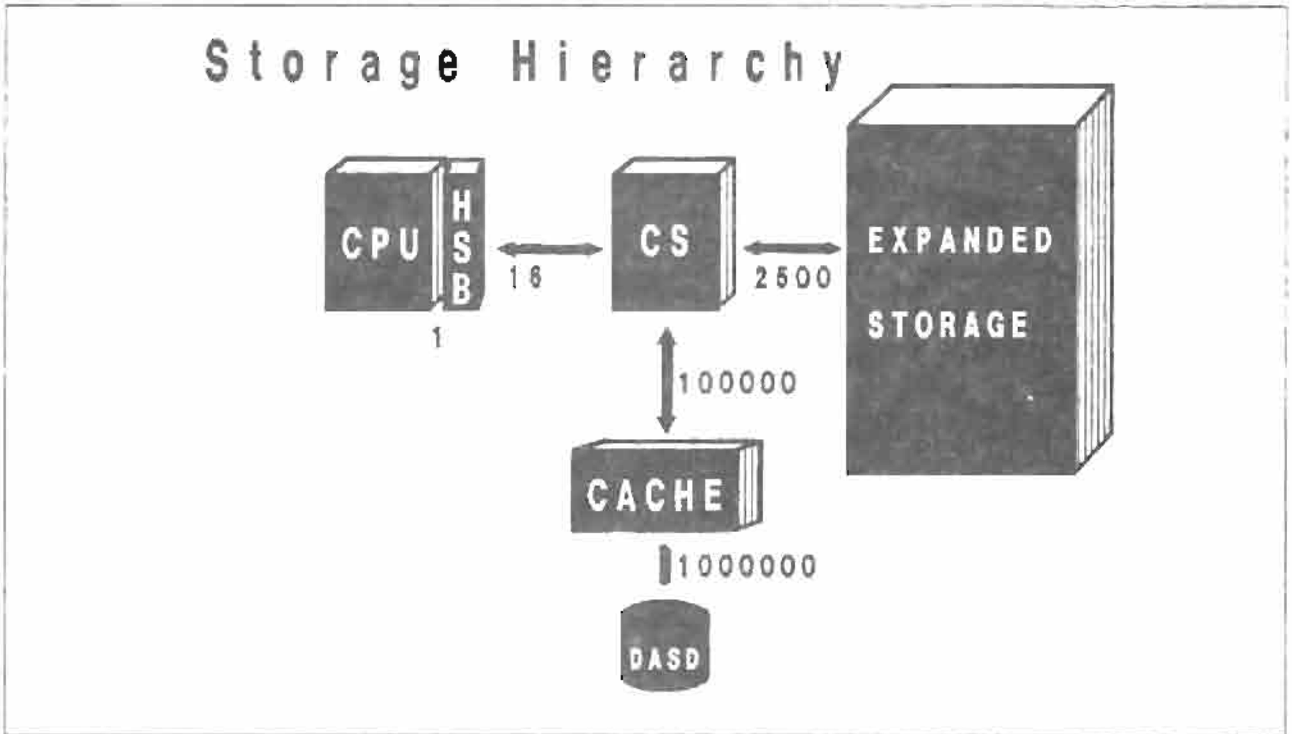


Figure 26. Data in Memory

An old joke goes something like this:

"Doctor it hurts when I do this."

"Then don't do it!"

If I/O time hurts, don't do it. In the beginning in Figure 1 on page 1, we alluded to the solution of an I/O time problem: Replace I/O operations with a memory service. Data in memory is often very easy to accomplish. The relative time it takes to get data to the processor is illustrated in Figure 26. When a task is running on a processor, the time it takes to satisfy a request for data varies considerably with the location of the data, from a relative value of 1, if it is in the high speed buffer, to around a million if the data must be retrieved from DASD.

Fortunately, the number of accesses is much higher for the faster memory. Most access by the CPU come out of the High Speed Buffer (HSB). The percent of reference *hits* in the HSB is in the high nineties. Similarly, the data reference hits in a DASD cache can be typically around 90% or better.

The placement of data in the high speed buffer is controlled by the hardware. The other locations can be controlled by software or systems programmers. The relative performance advantages can be seen in the ratios in the figure. Important data should take advantage of the performance options of the storage hierarchy. Some very general options follow.

For many sequential passes at a data set, move the data to a VIO dataset first. With MVS/SP4.2 and MVS/ESA, the VIO data set will be put in expanded storage if the space is available. After the first pass, data will be delivered extremely fast, at memory speed, not DASD speed.

If the data is referenced sequentially just once, moving the data to a faster medium is counter productive. Read it the fastest possible the first time. That's QSAM or VSAM and lots of buffers.

For Randomly accessed data, use Data in Virtual (DIV) or DIV with data spaces if the data is tabular. Otherwise use Hiperspaces, VIO or maintain a high probability that the data will be in memory (central or expanded storage) by specifying a larger number of buffers.

Data in memory in all cases assumes that memory is available - memory is not a constrained resource

Each access method has limitations.

- The number of QSAM buffers must be less than 255. But, unless the block size is very small, this limit will not be reached Any one I/O request will be limited to 31 buffers or 230-240K of virtual storage
- With MVS/DFP3 1 the number of VSAM buffers is 64k. Previously it was 64Mb in space.
- DIV window size is limited to the space available in the private area.
- Data space size is limited by 2 gigabytes. It offers data isolation in a separate address space. Currently data spaces are supported only by Assembler language.

In any case, data in memory offers a very attractive solution to the I/O time problem. In some cases it requires changes to the application. In other cases, the migration to MVS/ESA does it automatically for some system datasets

The next step in the solution is faster I/O. That's the 3990-3 and the 3390.

Summary

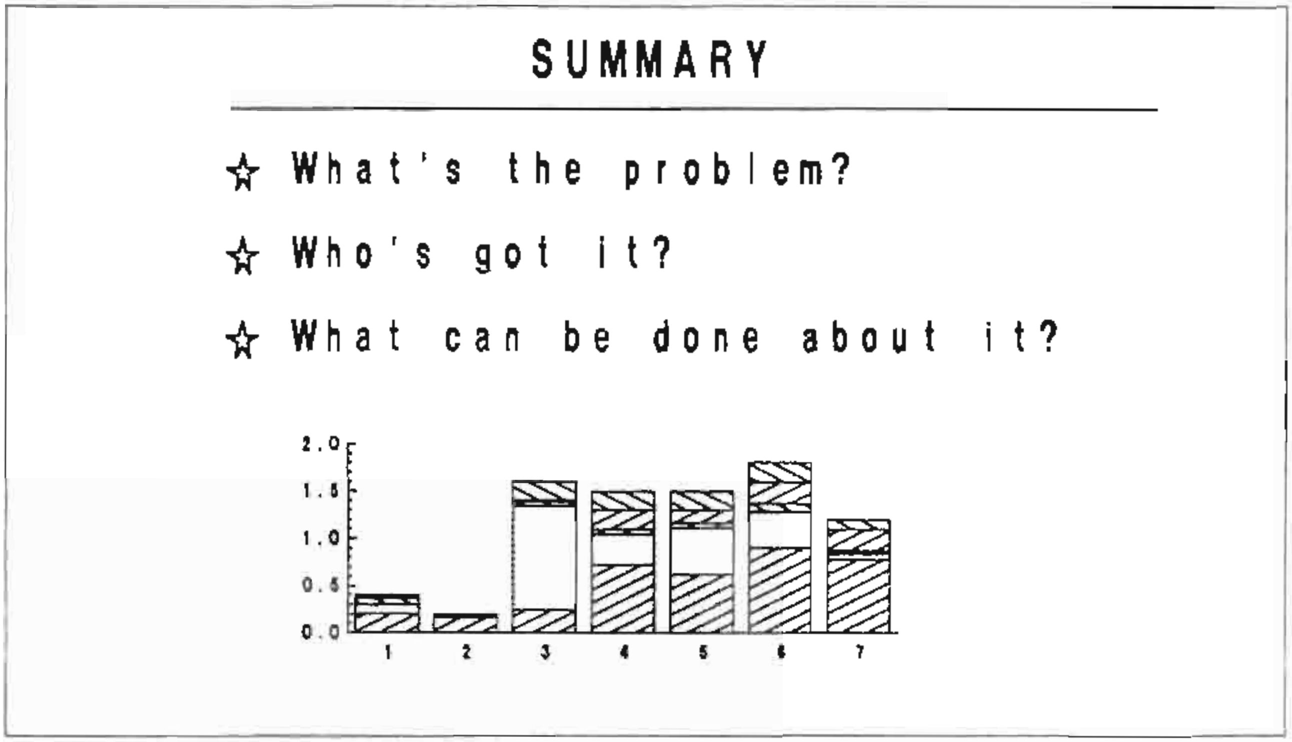


Figure 27. Performance Analysis Summary

Performance analysis and performance management steps have been outlined above. They involve establishing that there is a problem, understanding which resource is the major cause of the problem, finding out who has the problem, and finally figuring out what can be done about it?

The problem usually comes down to analysis of the behavior of the servers in the system model: processor, I/O, and memory. The analysis is greatly aided by RMF Monitor III. Resolving the problems is facilitated by the architecture in MVS/ESA which permits the trade-off of one service with another.

The next topic is capacity planning

Chapter 2. Capacity Planning Concepts

Process Flow

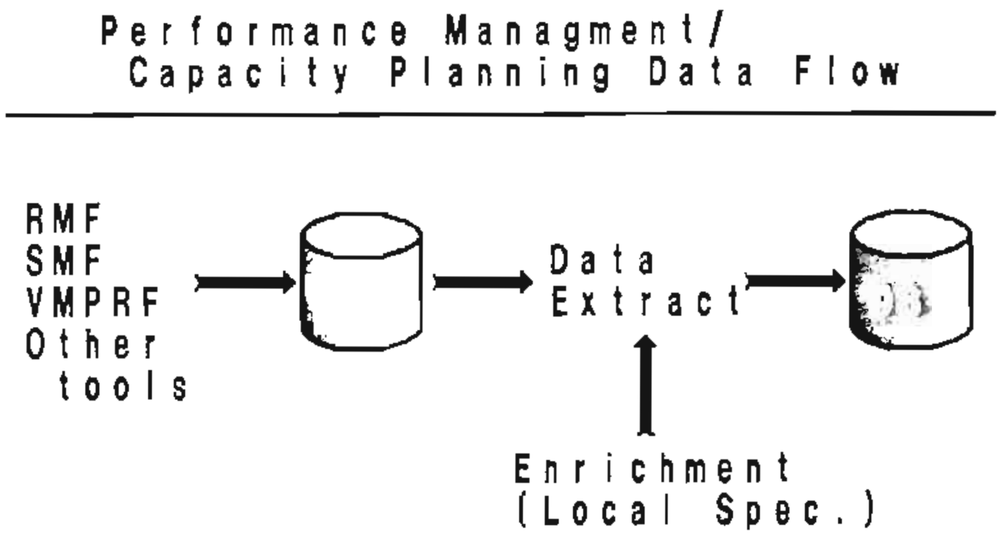


Figure 28. PM & CP Flow 1

The Capacity Planning (CP) process begins with data coming from RMF, SMF, DB2, CICS, etc. in MVS, and the many other sources in VM, and VSE. The data appears in one or more log files produced by these data collectors. (See Figure 28.)

The data contained in these logs can be voluminous. In order to use the data for capacity planning or Performance Management (PM), the data is extracted from these logs. The extraction process involves a number of operations.

- Variable reduction - all the variables in the log may not be required for capacity planning. PM usually requires a larger set of variables than capacity planning but even for PM, variables may be left behind in the log for use on an as needed basis.
- Variable merge - different logs may be combined to provide a picture of all the variables by time. Data combined by time will enable a view of one variable next to another variable at a specific time even though the variables came from different logs.
- Data enrichment - data may be added to with descriptions of variables or collections of variables which are not contained in the data provided by the data collectors. This local specification might be the relationship between performance groups and business units.

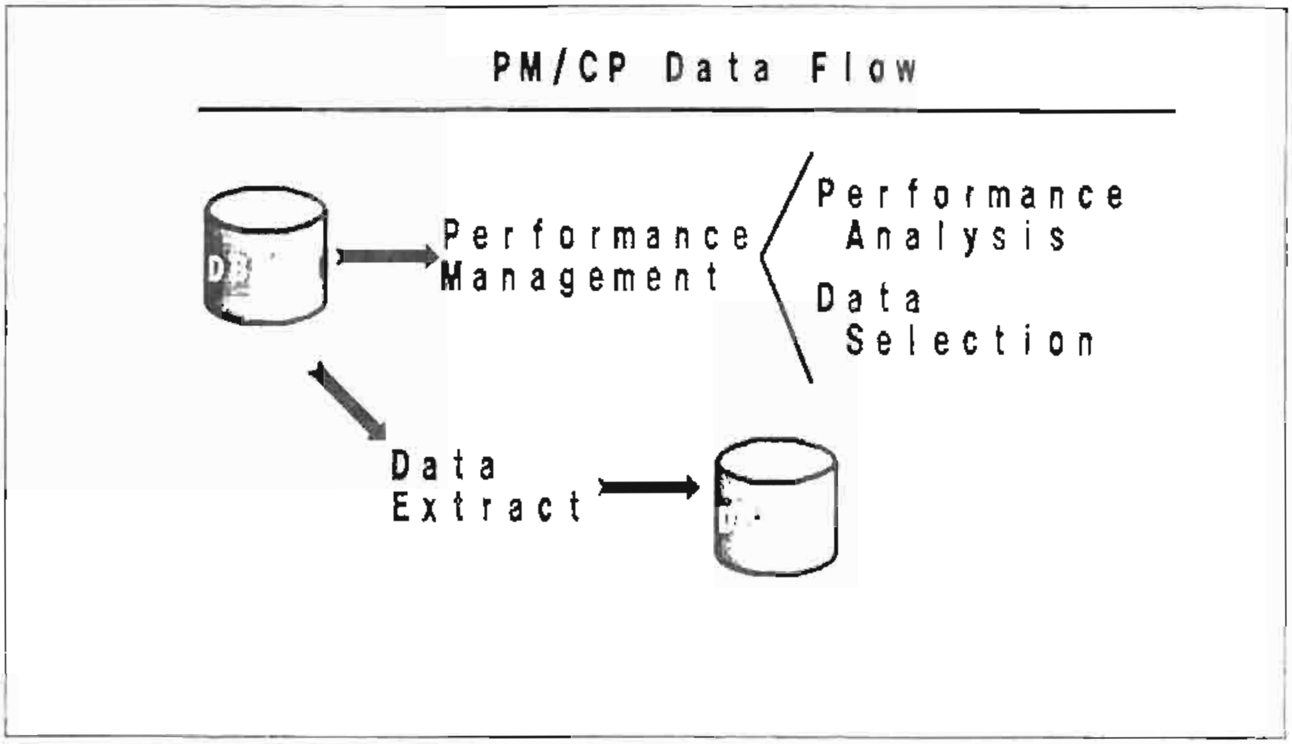


Figure 29. PM & CP Flow 2

Once the data base is formed, with a set of variables satisfactory to PM, queries can be placed against the data base. (See Figure 29.) An example is queries about the utilization of a resource at specific times. The results are returned either in tabular or graphic format.

For PM, the purpose is to do performance analysis. Am I meeting my service level objectives? Which resource is having a performance problem? What's the cause?

For capacity planning, which follows upon PM, the data result is a selection of samples which would be used for capacity planning. For example, only the hours when the DB/DC application is active may be of interest.

Upon deciding which samples to use, a further data extraction is performed. Capacity planning variables contain a subset of PM, and the interval of interest may last years, rather than a few days to a week for PM. A smaller data base is called for.

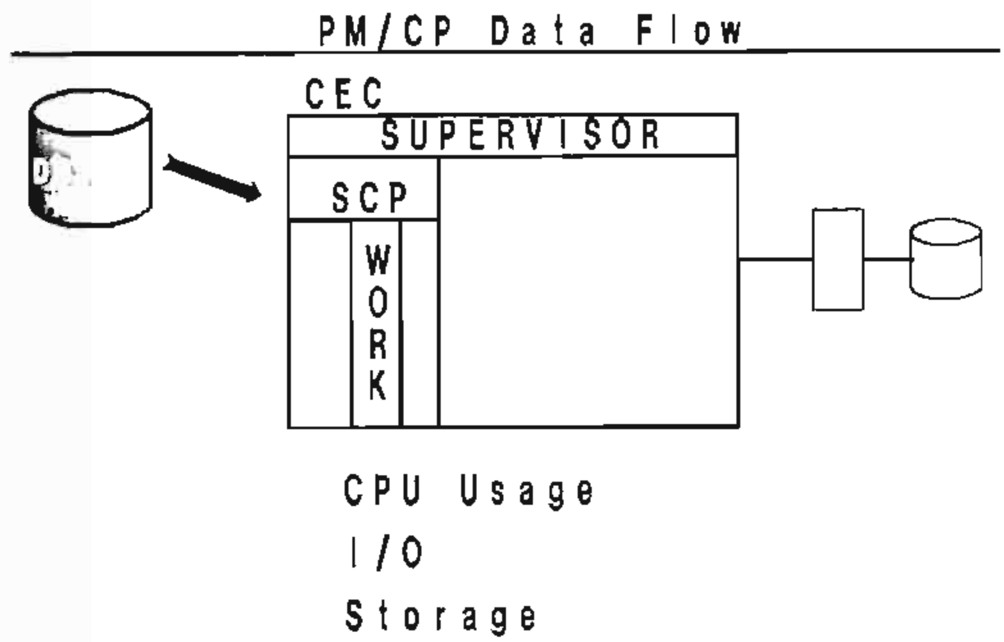


Figure 30. PM & CP Flow 3

The data in the logs may come from different systems. (See Figure 30.) The PM process may not combine data obtained from different system images although it should. Capacity planning requires that the complex of systems being considered be put into a structure which can be used for capacity planning decisions.

1. What is the Central Electronic Complex (CEC) which contains the system images?
2. If there is more than one system image running on the CEC, what is the **Supervisor** enabling this? LPAR? VM/XA?
3. What is the System Control Program? MVS/XA? MVS/ESA?
4. What are the workloads running in this system image? CICS Production? TSO Development?
5. What variables are needed to describe each workload? CPU? Storage? I/O?
6. What is the DASD configuration and usage?

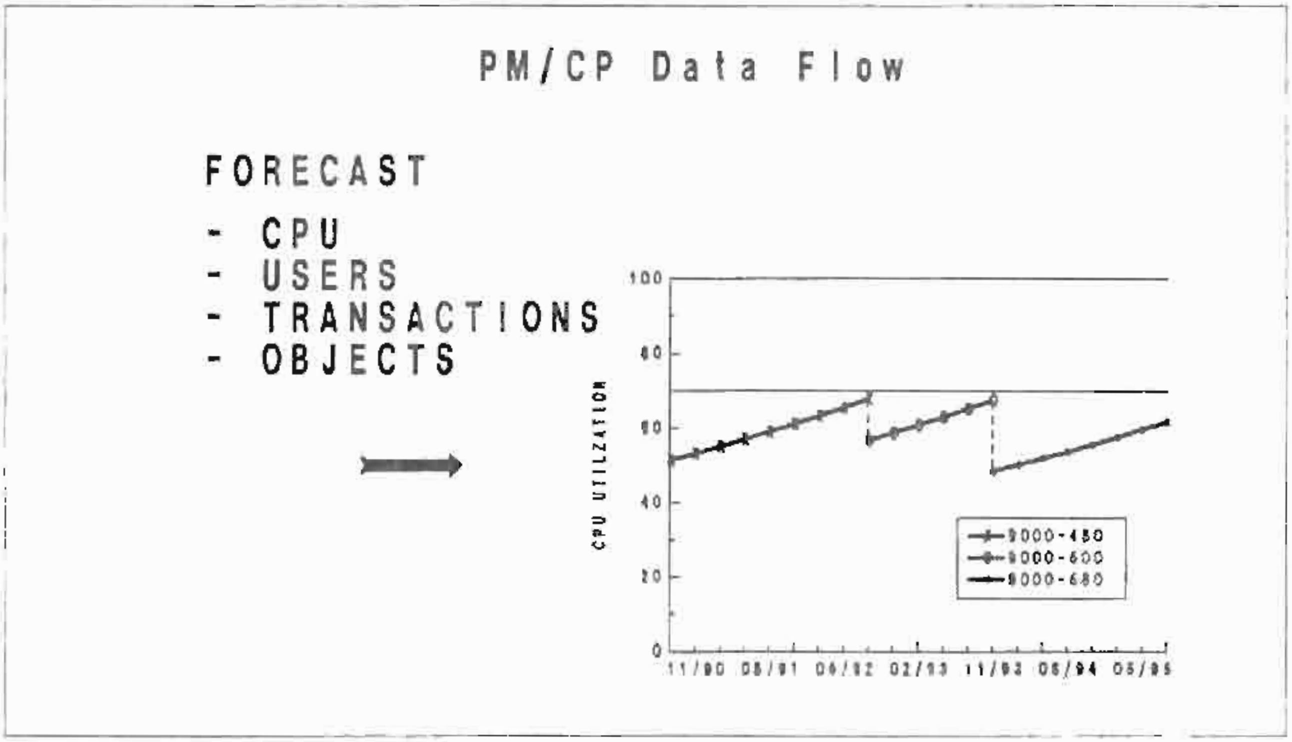


Figure 31 PM & CP Flow 4

After the step of describing **What there is** which fulfills the capacity planning structure comes the interesting part which enables capacity planning **Where is it going?** The forecast. The forecast comes in many flavors (See Figure 31.)

- **Growth by a Resource** (CPU for example.) This is the ultimate for the capacity planner Once the resource growth is known and the starting point is established the projection enables the capacity planner to decide the time of saturation - when isn't there enough resource to match the projected requirement? However this growth by resource does not usually conceptually match the business plan and often has little to do with the description of plans in the end user vocabulary **They** think in terms of users. transactions, and business objects.
- **Growth by Users** This method establishes the current number of users and a profile of the cost or resources per user With a growth by user input (next quarter there will be 10 more users) the resource requirement can be described in terms of "Users" a term understandable to the individuals making business decisions.
- **Growth by Transactions** This is similar to growth by users. The transactions are described in terms of a resource and projections are made accordingly
- **Growth by Objects** Company X makes Widgets and Gears These may have variable resource costs to produce and market The projections being considered by the company have a variety of proposals - more Widgets than Gears? Consideration is the resource cost of each and profit obtained Growth by Widgets and Gears like users and transactions, determines the resource cost of each and can handle the resource projections of a variety of alternatives

Ultimately, comes some table or graphic which embodies a recommendation showing resource growth (we hope) and points of action, such as, on which date is a larger resource obtained This is not one graph of course The business plans should have a variety of scenarios which reflect the range of economic forecasts for the company. A lean growth picture? An aggressive growth picture? And how do these different economic pictures impact resource requirements? (See also the discussion in "Growth" on page 64.)

Processor Resource and DASD I/O

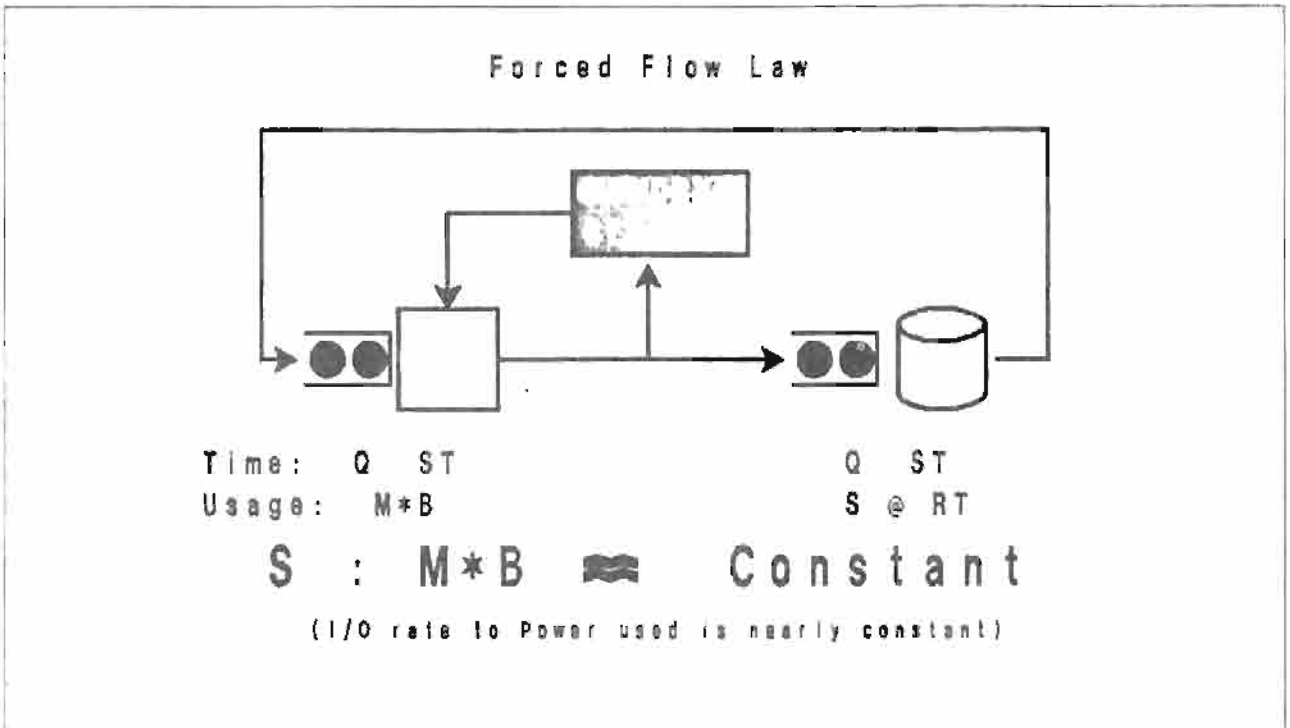


Figure 32. Forced Flow Model

Figure 32, shows the system model which will be used to **connect** the resources of Processor (CPU), DASD I/O, and memory service. The basic cycle is between processor service and I/O. The task does some CPU and then some I/O until finished. I/O can be replaced with a memory service (more buffers, Hiperspace, Data Space), but once established, the relation between CPU and I/O remains relatively constant. The amount of CPU power used (power M times proportion used B) and the amount of DASD I/O (I/O rate S at response time RT) is relatively constant.

This claim of proportion between CPU power and I/O is admittedly variable across small intervals (seconds or minutes). But capacity planning usually deals with intervals of at least an hour, and more likely, hours. This larger interval tends to smooth momentary resource bursts.

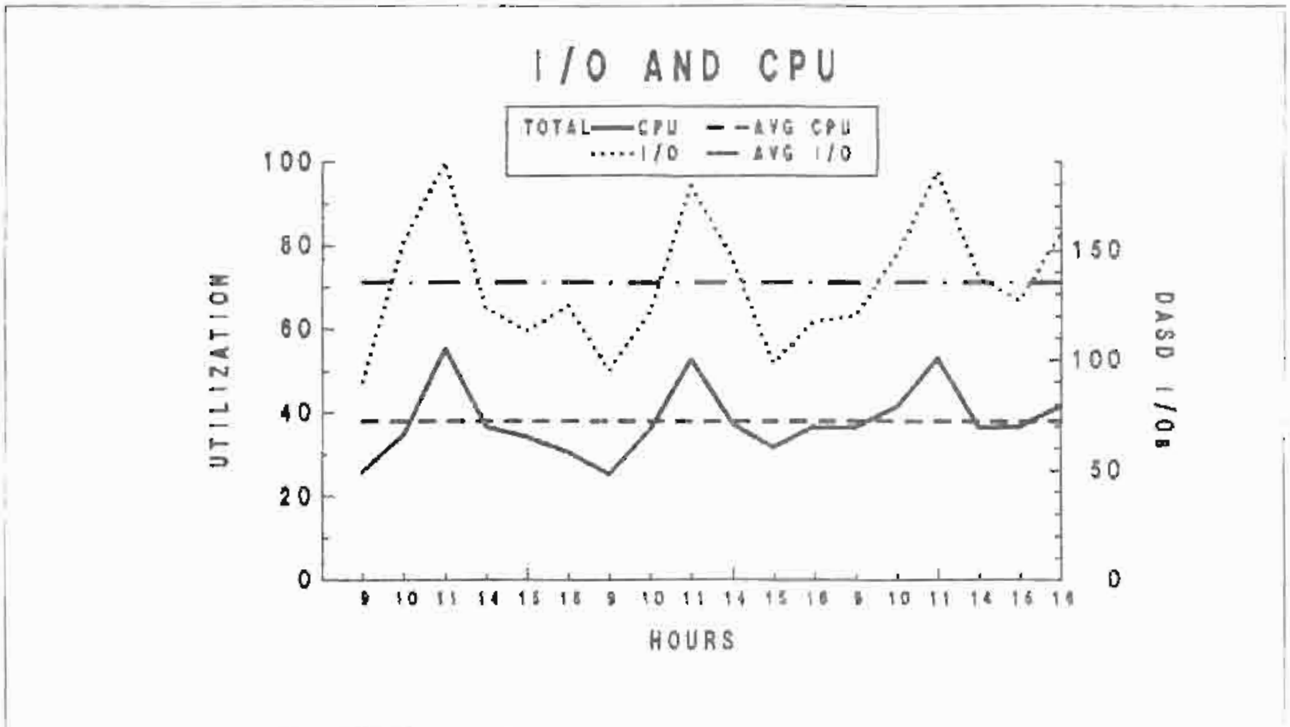


Figure 33. CPU and I/O Data

The relationship between the amount of I/O and CPU used can be seen in Figure 33. The CPU utilization and I/O rate tend to rise and fall together. This indicates that the workload is stable and the description of these workloads in terms of resource variables is acceptable. A lack of correlation between the resource consumption of CPU and I/O would discourage the application of a forecast to any of the workloads thus described. The workload grouping is not stable. If unstable, the choice is among

- Selecting different samples
- Selecting samples of larger duration (The sample duration for capacity planning usually has a minimum of 1 hour and can be an entire shift of 8 to 9 hours. Smaller intervals are subject to natural variations in workload characteristics and workload mix.)
- Splitting the plan into two plans as might occur when the samples originally came from first and second shift

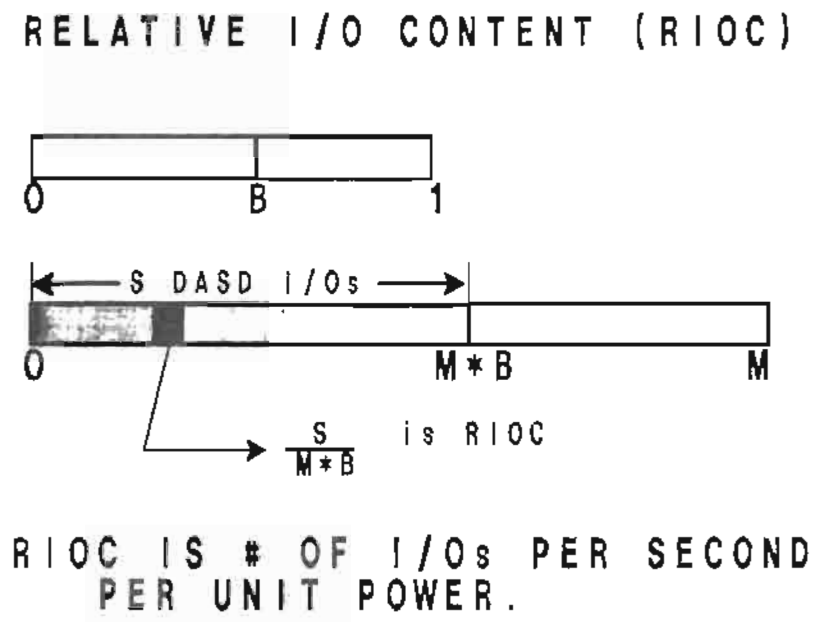


Figure 34. Relative I/O Content

The relationship between the amount of CPU and I/O is called Relative I/O Content (RIOC) (See Figure 34.) It is the ratio of the number of physical DASD I/O operations per second to the power used.

The data obtained by the data gatherers provides the CPU utilization (B, where B is in the interval (0,1)) If the specific CPU model has a power number M (the source of which is in Figure 47 on page 47), the amount of power used would be M*B. If the number of DASD physical I/Os per second was S, then the RIOC would be **S/(M*B)**. The RIOC can be viewed as the number of I/Os expected per unit of power for this workload. For example, given the following situation:

- 3090-200S
- MVS/XA,
- 67% busy
- Power number M = 1768
- DASD I/O rate of 425

Then the RIOC would be 425/(1768*0.67) or 0.36.

The RIOC is a useful concept If one knows the general range of the RIOC for an application, such as DB2, then some interesting questions can be answered. For example, a DB2 application does 25 I/Os to DASD. How much power is required?

Well, if the RIOC for DB2 is between 0.2 and 0.3, the power requirement can be computed using the equation **S = M*R*B**. We know S (25) and we know R (0.2 < R < 0.3). Can the power used (M*B) be computed?

The RIOC can change over time. Factors which affect "S" would be changes in

- Access method - change to an access method which has enhanced buffering, would reduce the physical I/Os.
- Blocking factor - more data per I/O.

- Data In Memory (DIM) - rather than a physical I/O for application data, data is retrieved from more buffers, a data space, or hiperspace.
- SCP - A new SCP could take advantage of a new architecture designed to reduce DASD I/O. For example, Expanded Storage.
- DFP and DFSMS - may change how I/O is done.

Factors which affect "M*B" would be changes in

- SCP - path length changes to support new functions
- DFP and DFSMS - more function generally means longer paths.
- Service ASCBs - changing a monitor, JES, or VTAM may add overhead
- Function - more means more.
- Complexity - As end users become more sophisticated with applications, they tend to use the full function of the application and that means more M*B.

Latent Demand

Figure 35 on page 37 shows a 52 week plot of some actual data. Each data point is a 40 hour average of prime shift CPU utilization. Notice that over time, the average and the maximum hour for the week are drifting upward.... At least until week 25

At around week 25, the maximum approaches 100%. The average's drift upward slows, but still moves upward. Did growth slow considerably after week 25? No. There was no room to grow. What happened to the growth or work that would have been there when the available resource ran out?

1. It was run later.
2. It wasn't run at all.

These options mean that new applications might not be introduced, existing work is delayed, or just that the general level of productivity decreased.

This delayed work is called latent demand. It is an amount of CPU demand that is ready to run. In Figure 35 on page 37, it appears as the area between 100% and the extension of the maximum if the maximum could continue to rise over 100%.

The measure of latent demand is found in the ratio of maximum to average processor busy called the **Peak to Average Ratio** (PAR).

If the samples are 9 Hour averages and we have 5 samples (5 samples points of prime shift averages), the **Peak** is not the maximum of the five samples. The 9 hour average is too large to be useful to compute the PAR for capacity planning. The maximum is usually the maximum value found among the RMF samples which were used to compute the 9 hour average. This samples size is usually around 15 minutes to an hour.

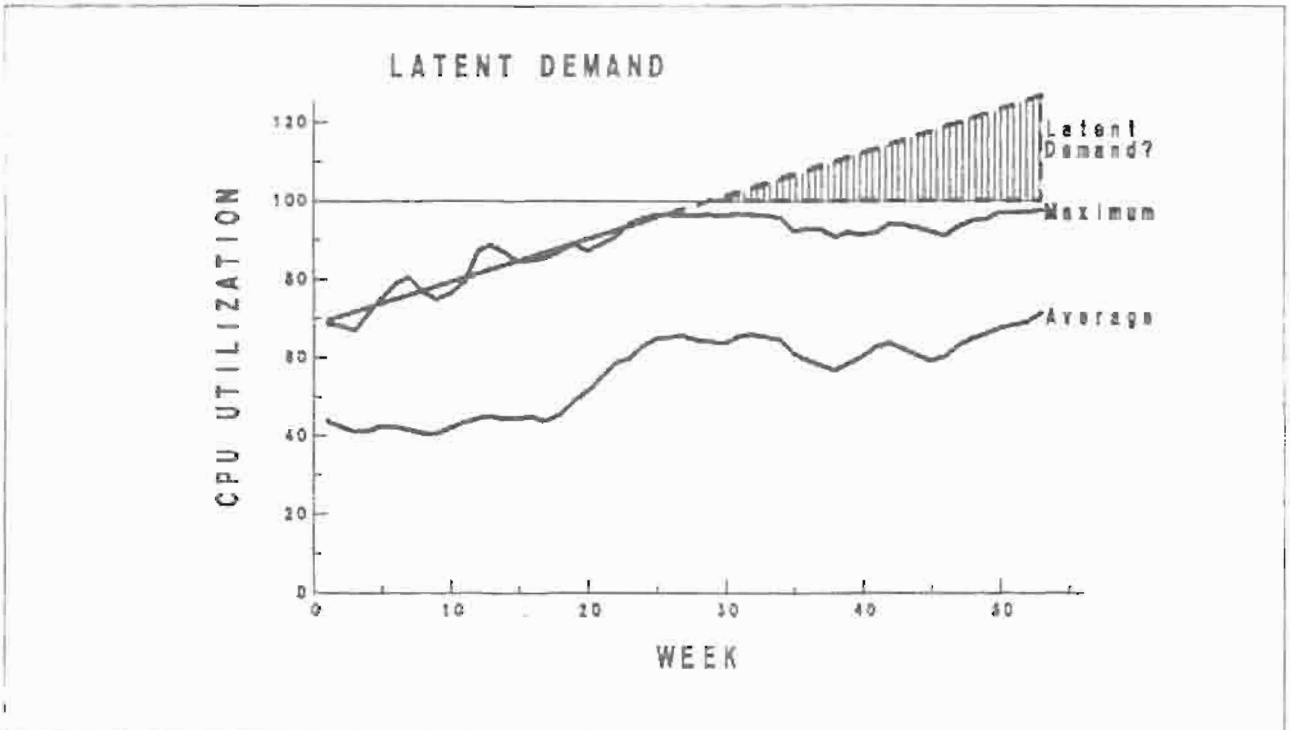


Figure 35. Latent Demand

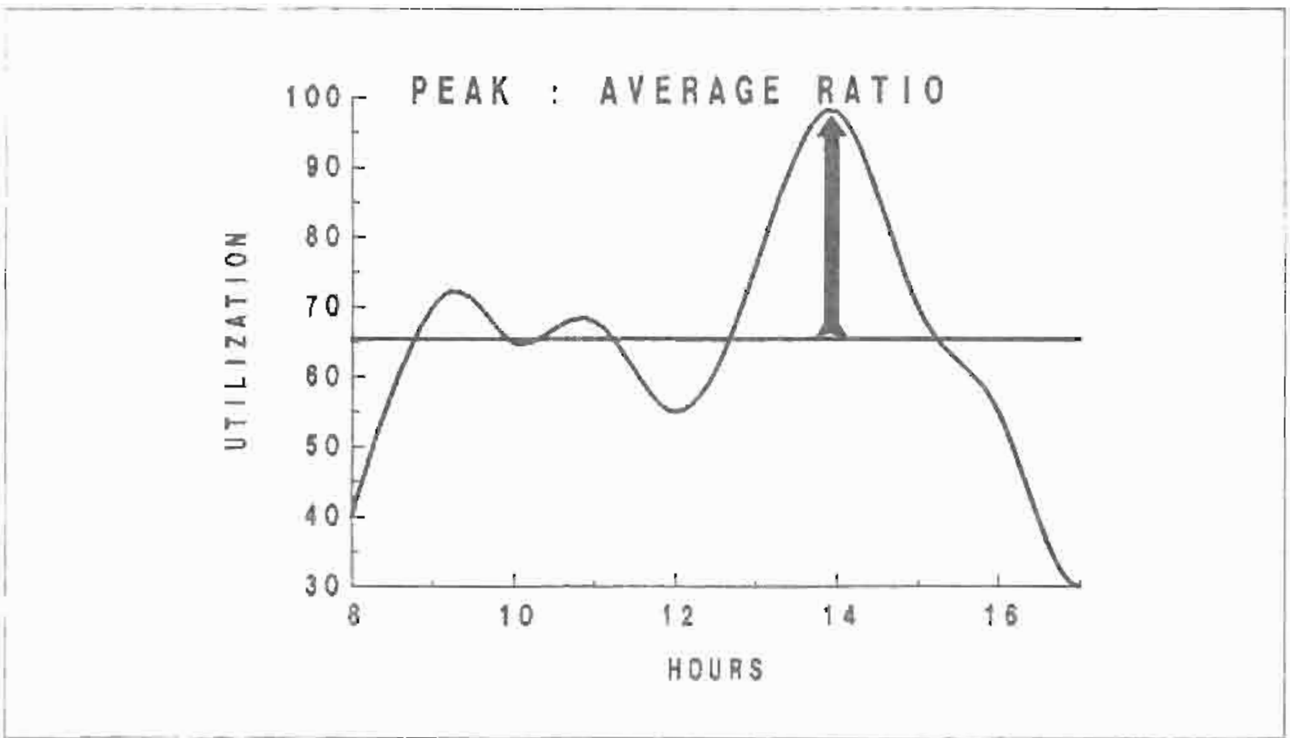


Figure 36. Peak to Average Ratio.

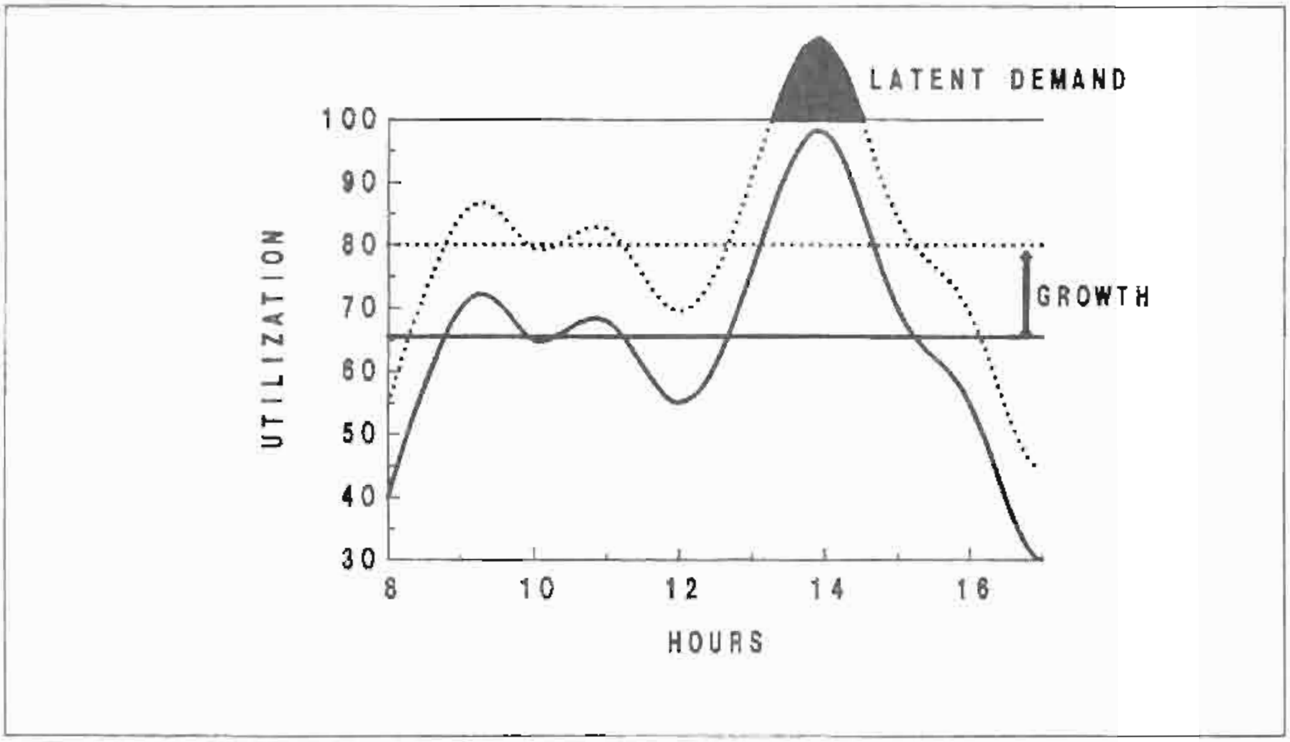


Figure 37. Latent Demand - a Measure

Figure 36 on page 37 shows a simple set of CPU utilizations taken for one day. The average is about 65% and the peak near 100%. That's a PAR of about 1.5. Since the samples are hours, the peak hour of 100% is usable. If the samples were days instead of hours, a peak day average is not sufficient. One would have to look at the samples making up the days averages to find a peak interval. For RMF, this can be found in the Summary Report developed by the RMF Post Processor.

Let's assume that the installation grew uniformly by 15%. If we could assume that the curve, with growth, is shifted upward by 15%, the curve might look at the upper curve in Figure 37. Of course, the curve cannot go above 100%. If it could, since the demand was there, it might go up to around 115%. It actually goes to 100% with a potential of 115%. This potential is called **Latent Demand**.

A latent demand of 15% says that a processor with a peak of 100%, if upgraded with another 15% more powerful, would still have a peak of 100%. Very embarrassing for a capacity planner. You will naturally be asked "Why isn't it 85%?"

To understand Latent Demand, it is imperative to track the PAR over time. The current PAR may be very misleading. In Figure 37, what happens to the PAR as the average increases? For a while it may stay the same. After the peak hits 100%, an increase in the average means a decrease in PAR

```
PAR = Peak / Average
```

The peak remains at 100%, the average increases, and the ratio decreases.

Latent demand can cause embarrassment. In Figure 38 on page 39, we have system A with application-1 and a smidgen of batch while system-B has application-2 and batch. We plan to migrate these two systems to a single system image where the processing power is approximately that of A+B.

Application-1 is the most important to the installation's business activities. What could happen?

If the figure represents the average utilization for each business unit, one can see that the peak to average ratio of Application-1 is near 1. My guess is that is was not that historically and the utilization near 100% is a red light flashing. There is latent demand here. Even if we were not tracking the latent demand by means of historical PARs, the light is still flashing red.

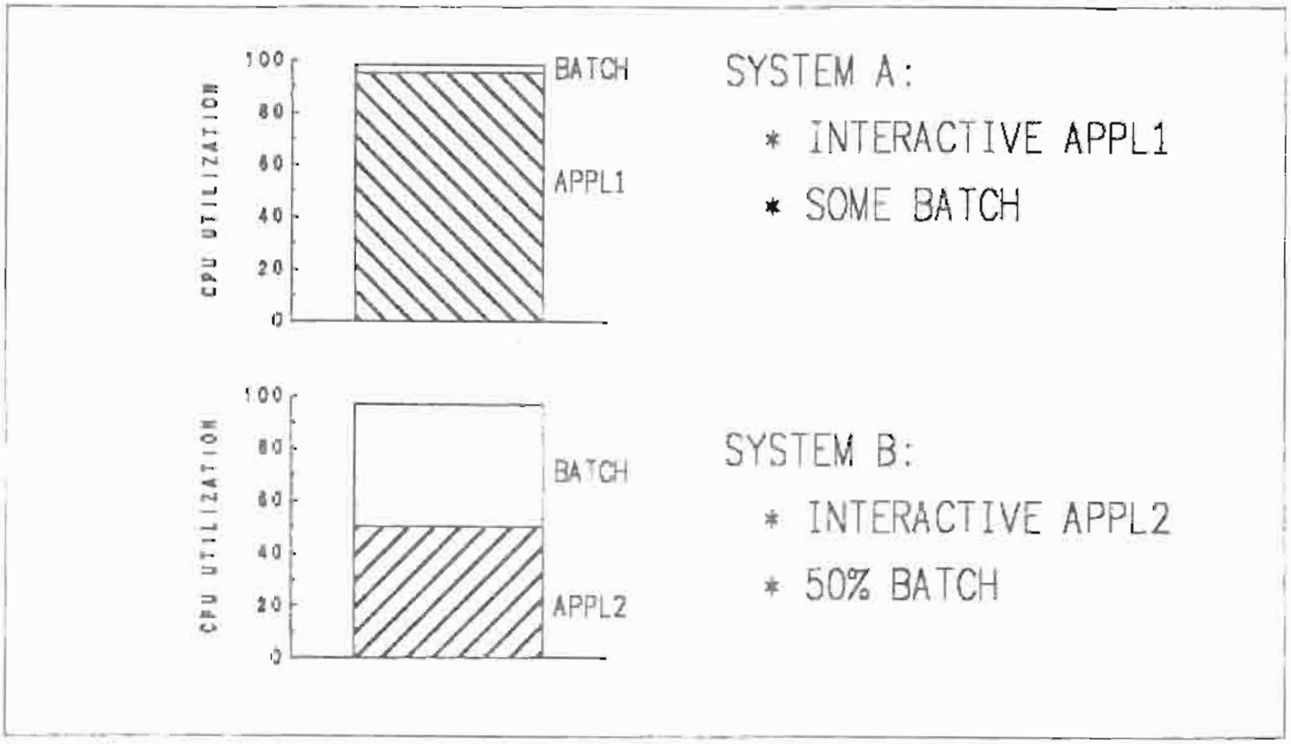


Figure 38. Merging Systems

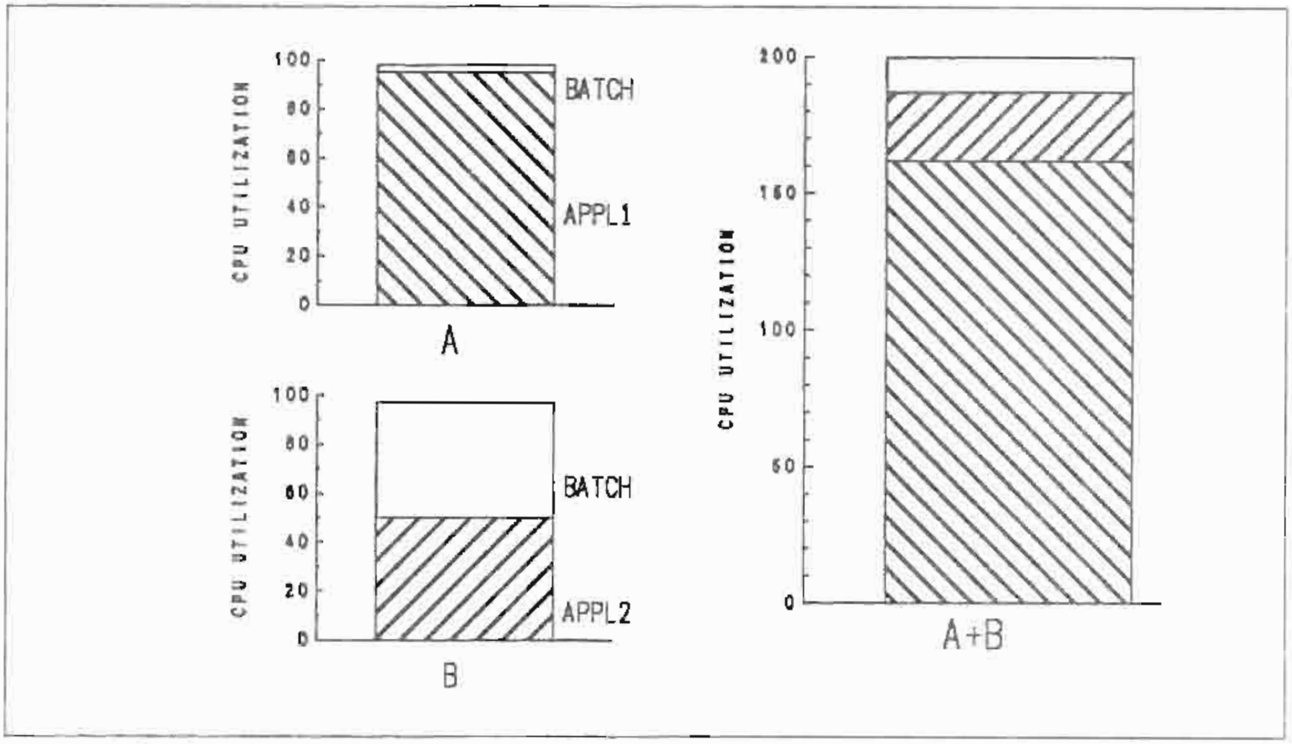


Figure 39. The Result

When the single system image is brought up, who do you think will be calling system support? Certainly not application-1 users. What happened? With application-1 at higher priority and the latent demand for CP power, application-1 began to use the increased availability of power. With a limit to the power available to lower priority work, application-2 and batch received less. This might be observed in RMF by a reduced response time or increase transaction rate for application-1 and a reverse for lower priority applications. Is this bad?

What if application-1 was THE business application? The answer might then be "This is goodness." If we wanted the same distribution of CP resources something would have to be changed to guarantee application-2 and maybe batch a larger share. This is accomplished through a combination of priority and time slicing, SRM controls. It can be adjusted. The trick is to remember that one task can use the power of only one engine at a time. On a four CP system image one task can absorb only 25% of the total system power, while on a two CP system, it is 50%.

Figure 40 on page 41 shows what might be a typical growth scenario over 30 months. Some observations:

- There was an update in months 18 and 27 or some rather large down sizing somewhere. Let's assume the former.
- The PAR in months 15, 18, 24, 27 was decreasing rapidly or negligible already.
- The growth rate was small where the PAR was decreasing. This is a warning! Ever hear the statement, "Why Upgrade? There's no growth." There's no growth because there's no place to grow into. A historical chart is very useful to put some reality into statements.

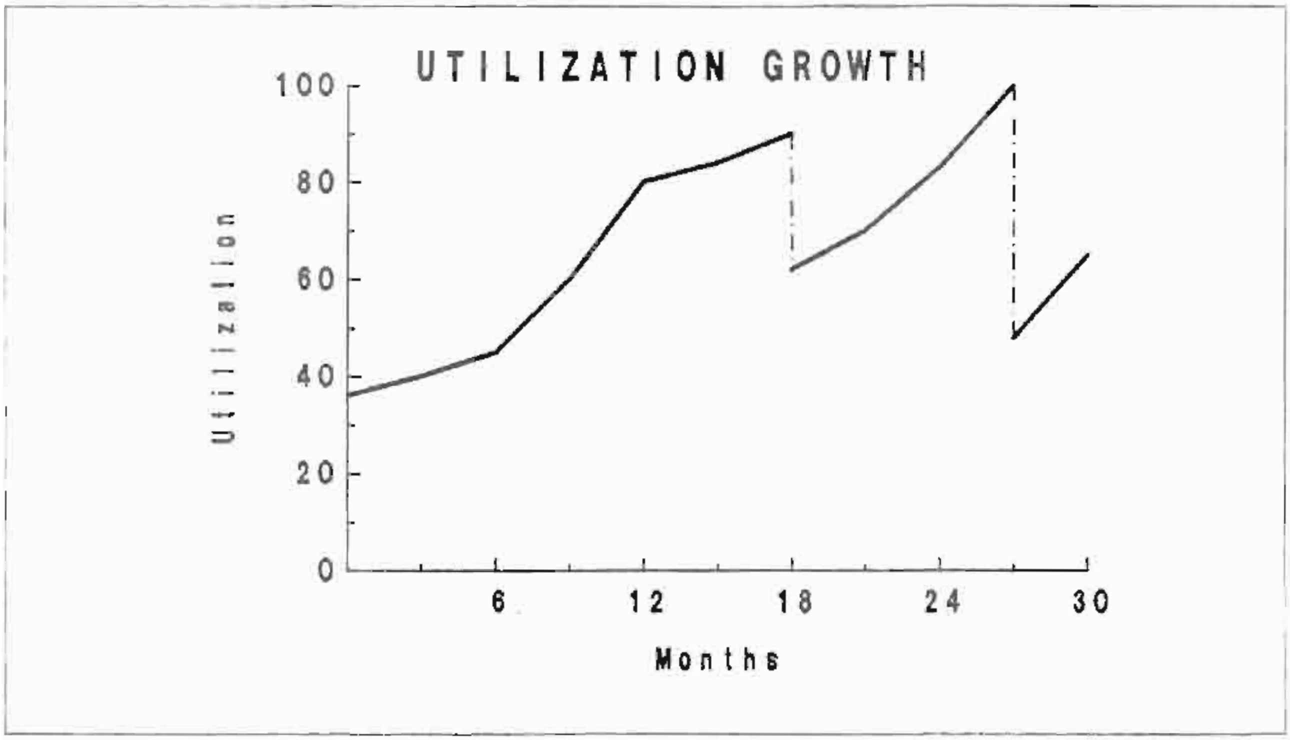


Figure 40. Some History: Growth in Utilization

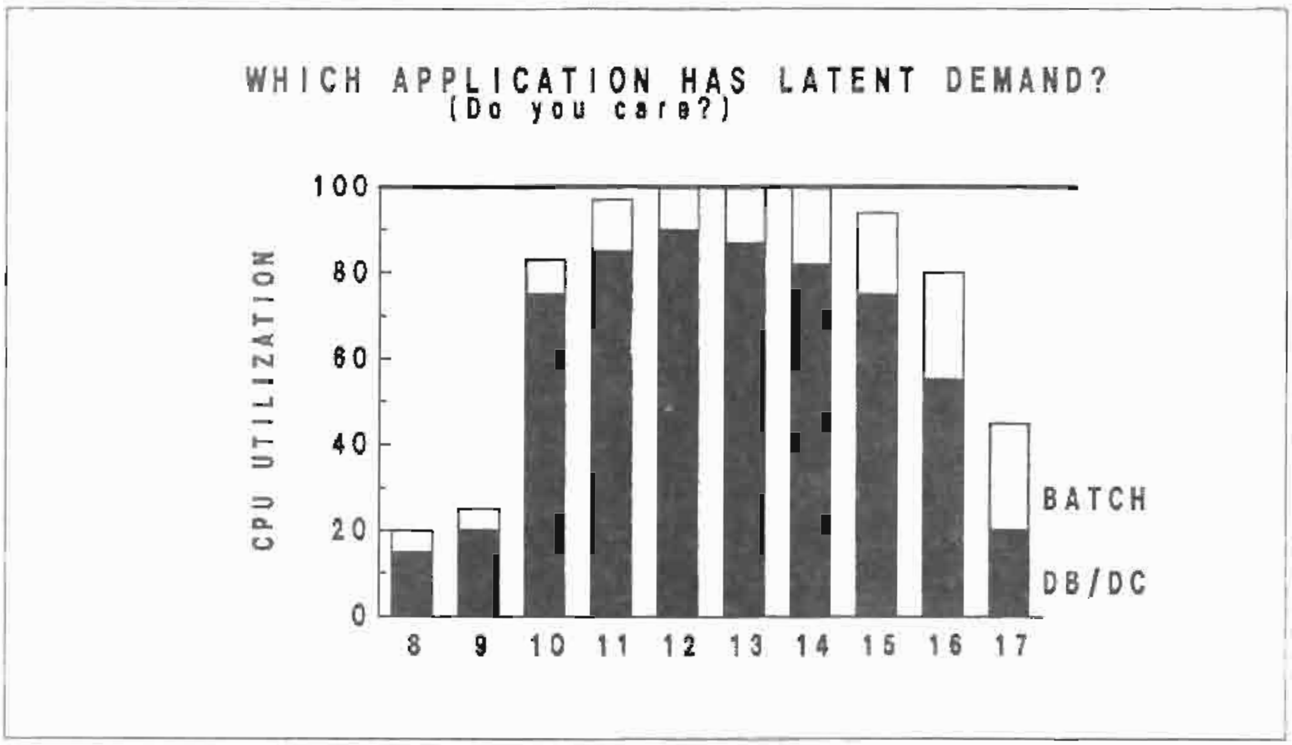


Figure 41. Latent Demand? Who Cares?

As always, you should be skeptical and ask - "When might Latent demand be of no interest?" Figure 41 show the combined utilization of two workloads, a DB/DC application and Batch. Between the hours of 11 and 14, the batch work gets squeezed out. Batch has a latent demand. Do you care? If it were **my** batch, the answer is yes. If it is yours, no.

However, one might apply the concept of PAR to just the DB/DC application

The notion of PAR and Latent demand define a concept called CPU Saturation By definition, the **Saturation Design Point (SDP)** is

```
SDP = 100 / Historical PAR
```

The historical PAR is chosen, rather than the current because the current may be zilch if the average is very high. A historical PAR of 1 4 yields a SDP of 70%

This concept of SDP is a capacity threshold. This is distinguished from a performance threshold. A performance threshold is defined as that utilization when a workload has a significant queueing problem - response time becomes large

Sample Selection

Sample selection for capacity planning differs from performance analysis in that in performance analysis there are no poor sample selections, only uninteresting ones. For capacity planning, samples are used to characterize the resource usage by business units. The samples chosen should *adequately* describe each business unit and thereby the total resource usage

Questions in sample selection

- How many samples? One is enough if it is the right one. The right one is the one which gives the picture for all the business units. In the absence of such omniscience in sample selection, a number is chosen which encompass a time interval where variation would occur
- What's the duration of each sample? Figure 42 on page 43 shows a set of data viewed for durations of 15 minutes, 1 hour, 4 hours, and 8 hours. As the duration grows, the unruly behavior diminishes. At 15 minutes the data varies considerably. Intuitively, 5 minute intervals would be even worse. It is reccommended that samples should be at least an hour in length. Four to eight hour durations are even better.
- Are all the workloads running? Are they running in the same proportion during each sample? Figure 43 on page 44 shows data for five days. It doesn't take too much to see that the workloads proportion vary considerably over the week. Friday, day 5, is primarily a batch day dominated by "P Batch." On Monday, there's none. What's to be done?
  - We could use the average but the resource description would dramatically underestimate the requirement for P Batch on Friday. The requirement could be significantly undersized if the growth rate in P Batch is large.
  - We could use just Friday. But we might then have a problem if P Batch has a small growth rate and Database W is large.
  - We could a sample at random and get a random result

  In this case, we might have to build multiple models to reflect the day dependent resource requirement. Figure 44 on page 44 offers us the opposite situation. There is great consistency across the week. Very nice data
- How many samples should be used? The answer to this comes from experience. The capacity planner should be reviewing the data at least weekly. Remember, one sample is enough if it is the right sample. I would suggest reviewing the daily averages once a week. It might not be a bad idea to even save it.

- If there are multiple images involved in the complex, you should be reviewing the data from each image. With multiple images, *the right sample* implies not only the right sample from each but the right sample for the same time! This is true because of shared DASD in a multiple image complex. Choosing different time periods makes the modeling of the DASD subsystem completely suspect. That's why using an average of a number of samples is often a superior choice.

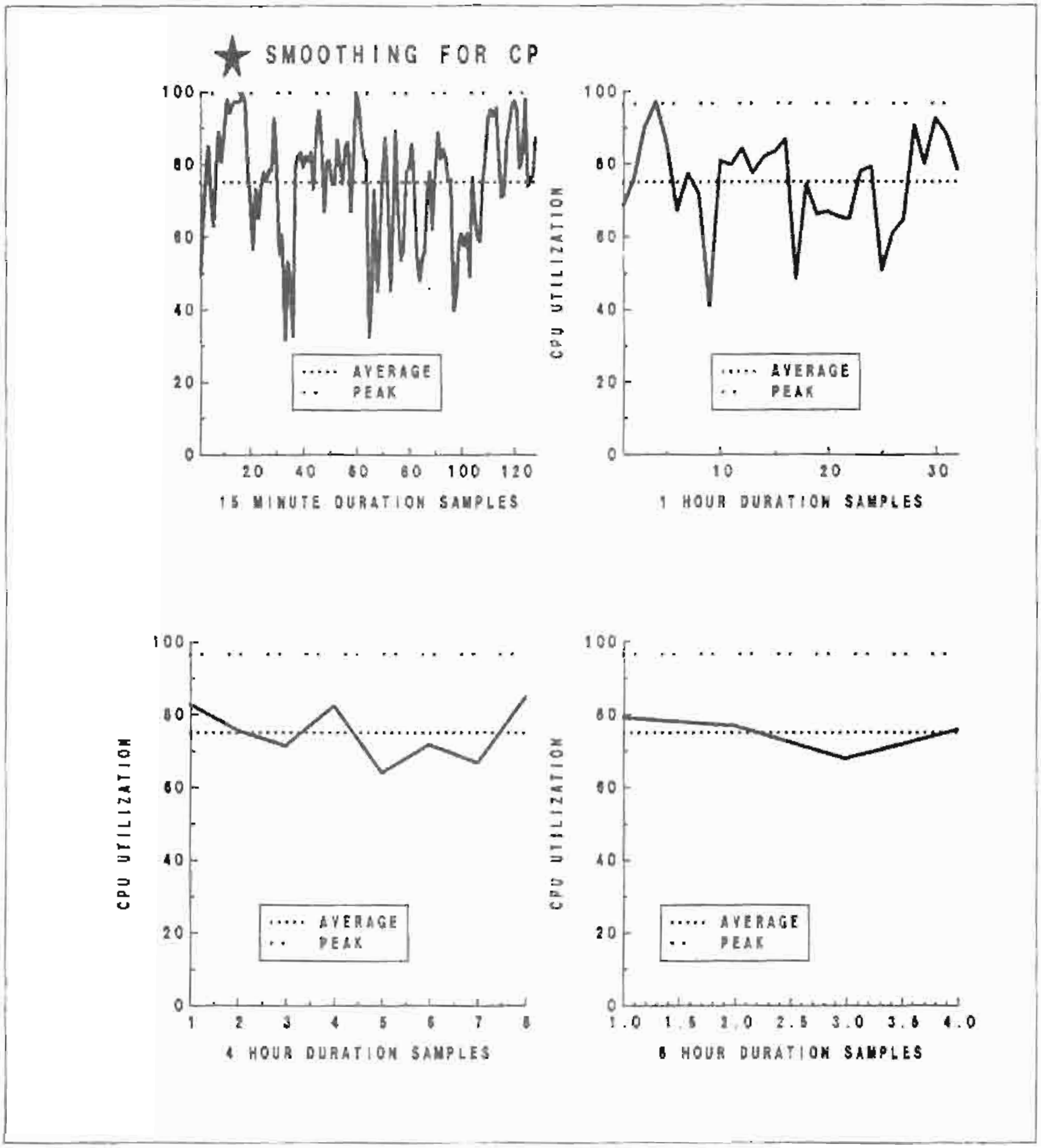
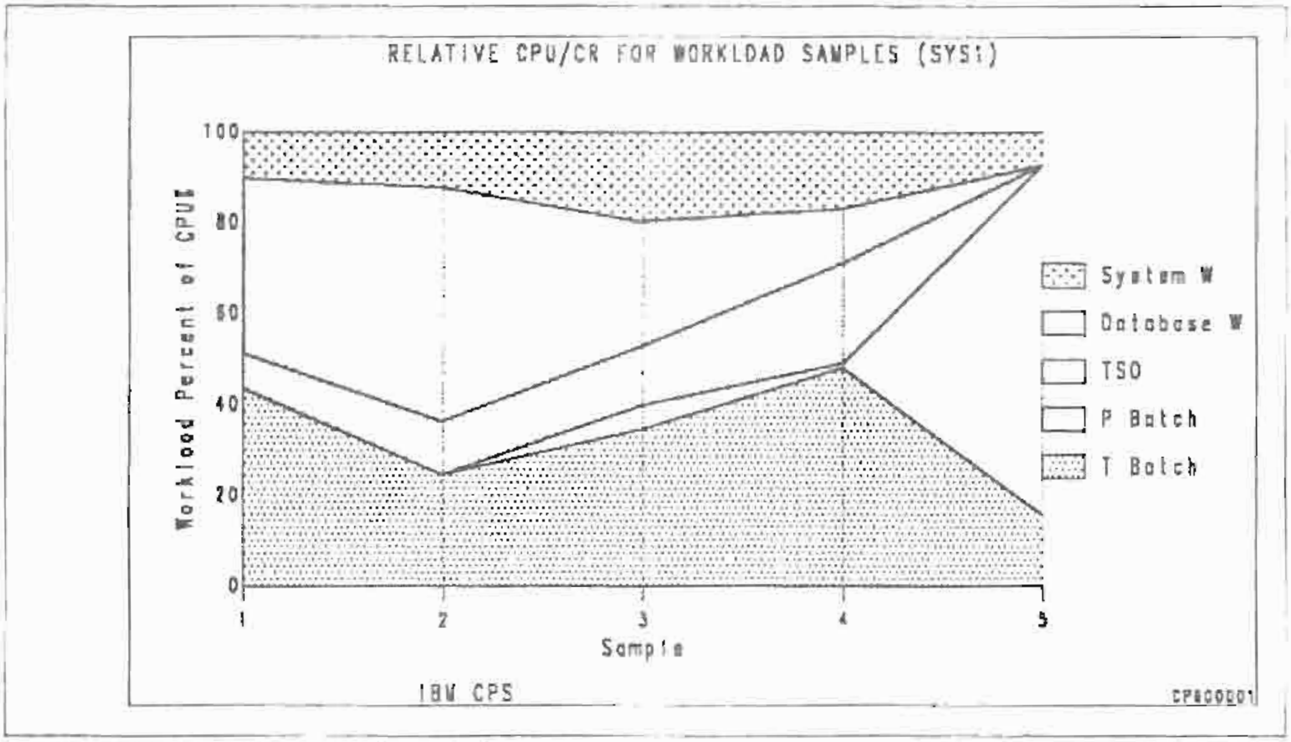
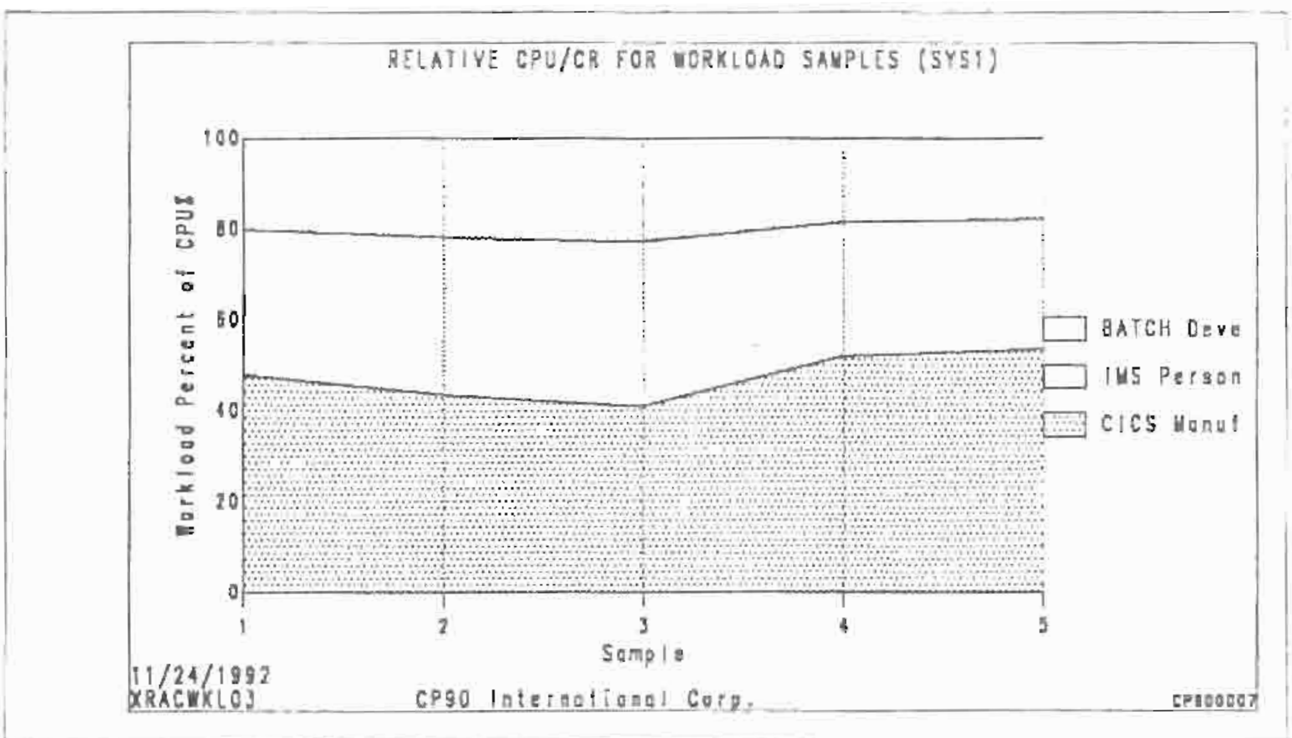


Figure 42. Sample Smoothing



| *Figure 43. Workload Proportion, Example 1*



| *Figure 44. Workload Proportion, Example 2*

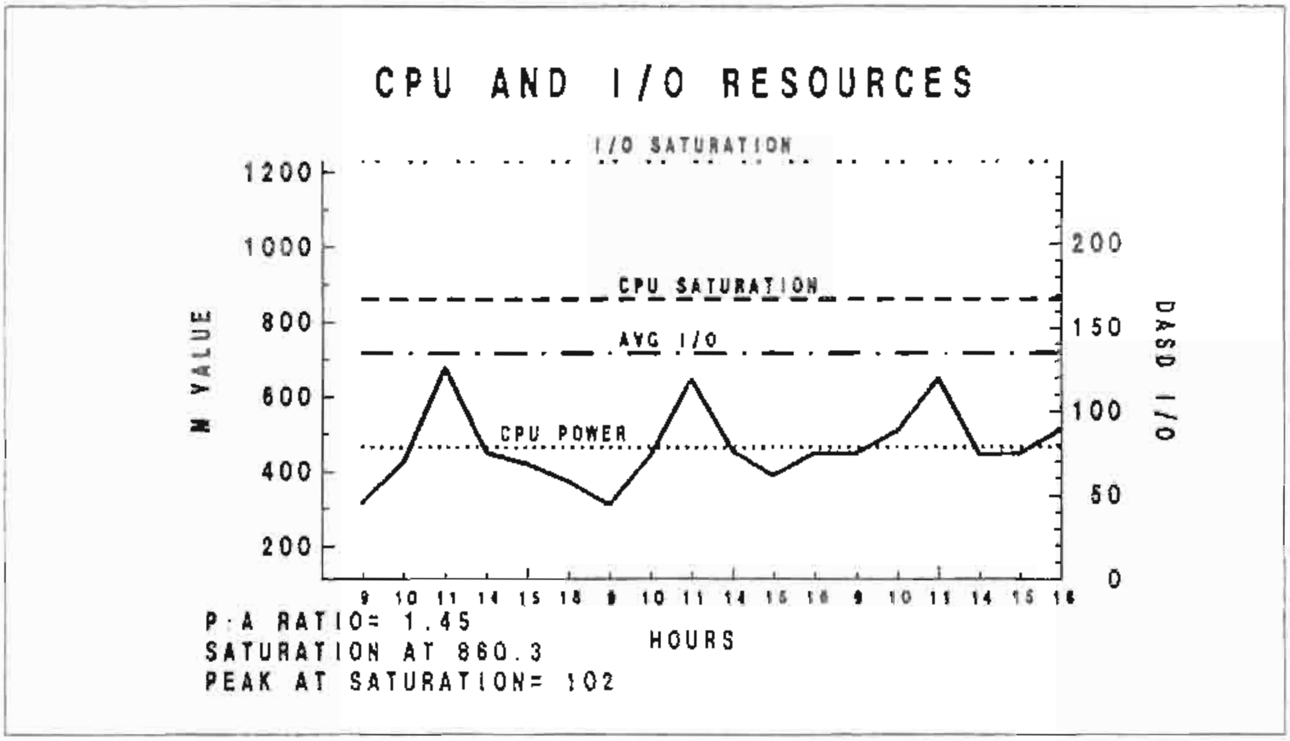


Figure 45. Saturation

Figure 45 shows some of these concepts put together on one chart, the Saturation Design Point (SDP) of 70% says that when the average reaches 70, the peaks will then be at 100. After 70% latent demand begins to build or work that could be serviced, is not. The processor is saturated.

If the ratio of I/O and power used is relatively constant (RIOC), as the CPU average rises, the I/O average should rise in proportion. If the CPU saturates at 70% in the figure, the CPU I/O rate saturates at about 250. The CPU I/O rate saturation point does **not** mean that the I/O subsystem cannot do more, it means that the DASD I/O rate coming from the CPU should average about 250. It implies that the I/O configuration should be able to handle about 250 I/Os at minimum when the SDP is reached.

Consider this example.

```
CPU Avg. 40
CPU Max. 55
I/O Avg. 135
```

The PAR is 55/40 or 1.38, the SDP is 100/1.38 or 72%. If the I/O average is 135 at 40%, what will it be at 100%? That's 135/0.4 or 338. But we saturate the CPU at 72%. So, the CPU I/O Rate Saturation is 338*0.72 or 245.

The concept of CPU and I/O saturation is interesting. You can now plot the CPU utilization and the I/O utilization on the same graph. In the example above, if one plotted the sample CPU values and the corresponding I/O samples divided by 338, a graph similar to Figure 46 on page 46 might appear. If the RIOC was relatively constant, the two lines should track with each other.

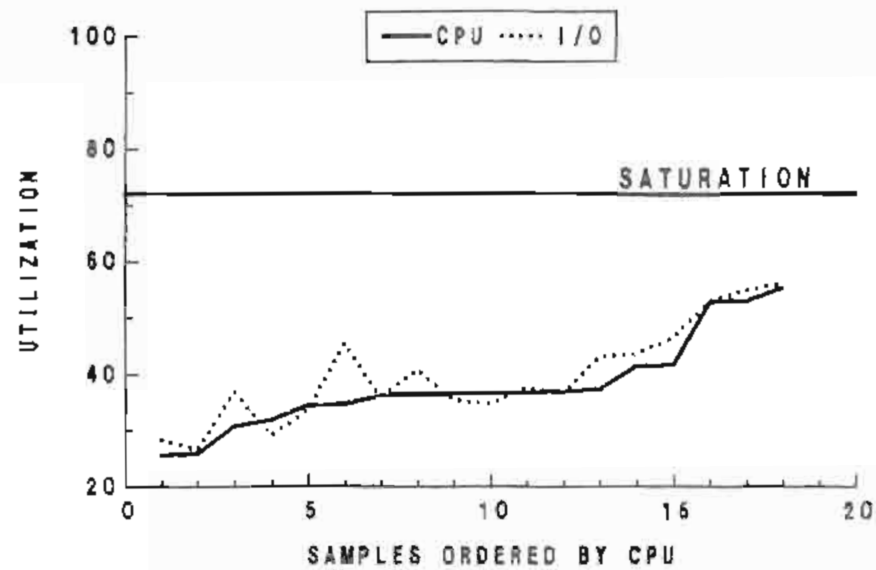


Figure 46. CPU and I/O Utilization

CPU Power

All right, where do the power numbers come from? One could use a variety of numbers.

- The SRM constant
- MIPS numbers
- ITR numbers

The **SRM** (Systems Resource Manager) constant does have some relation to CPU power. As the machine gets larger, the value goes up. How accurate is it when comparing different machines? After all, the purpose is to convert CPU seconds to service units. The answer is obtained by comparing the SRM constant with actual measurements. The answer is that for the purpose designed, the SRM constant works for the SRM. As a comparison of CPU power between models, there is often a large discrepancy between SRM constant ratios and measurement ratios. But, as always, this is better than a random guess.

The **MIPS** (Millions of 'I' per Second) numbers are probably not really measured MIPS numbers but an estimated power ratio from some base. That is, if a model 158 was really once a 1 MIPS machine, subsequent power numbers were merely an estimated ratio of power. These numbers are often based upon vendor announcements ("The new machine is between 1.8 and 2.2 times faster.").

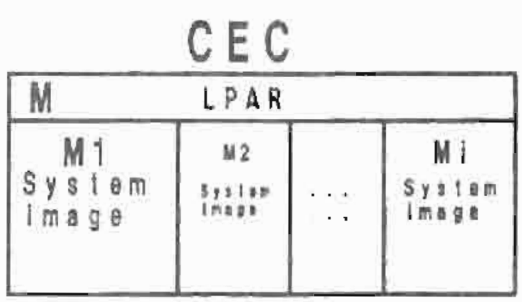
ITR (Internal Throughput Rates) numbers are measurements by workload type which determine the capability of a machine in terms of the number of transactions per CPU second. Figure 47 on page 47 shows some measured ITRs for a 3090-120E and a 3090-200 running MVS/XA. It says that the 120E can do 35.78 CICS transactions per CPU second whereas the 200 can do 135.5. The ratio of the two is the ITRR for CICS with one of the machines as a base. You can see that the ITRR varies with workload.

MVS/ESA ITRs and ITRRs

	CICS3	IMS2	TSO2	
3090-180J	84.90	53.30	19.94	ITR
3090-200J	165.36	97.66	38.41	
3090-300J	239.60	141.47	54.36	
3090-400J	308.97	176.85	71.03	
3090-500J	376.59	211.37	85.93	
3090-600J	438.31	247.44	99.25	
3090-180J	1.00	1.00	1.00	ITRR
3090-200J	1.95	1.83	1.93	
3090-300J	2.82	2.65	2.73	
3090-400J	3.64	3.32	3.56	
3090-500J	4.44	3.97	4.31	
3090-600J	5.16	4.64	4.98	

Figure 47. CPU Power computations

If a workload was proportioned as shown in Figure 47, the ITRR for that workload mix would be 3.712. The average of the ITRRs is 3.549. That means that this specific workload mix would perform about 5% better than average. The power values, M values, used later in this publication are taken as an average value of CICS, IMS, and TSO scales to a larger magnitude.



Processing Power Sensitive to:

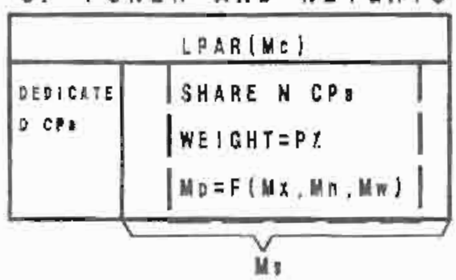
* Shared/Dedicated CPs
* Weights
* # Partitions
* Workloads

Figure 48. LPAR Requirements

LPAR, the ability to logically partition a CEC into smaller machines or system images, (see Figure 48) is a feature which makes the job of a capacity planner interesting indeed. The deliverable power to a partition depends upon the factors listed in Figure 48. Each system image **thinks** that it owns the resources. A parallel concept is found in VM.

Although the system image thinks it is running on a system with two Central Processors (CPs), for example, it may in fact be sharing the CPs with other partitions. So, the CEC probably has some power number M (although it probably varies with the LPAR configuration, SCPs, and workloads), and each partition has an apparent power number Mi

LPAR CP POWER AND WEIGHTS



Mc = Total Power of CEC as Configured
Ms = Total Power Available for Sharing
Mx = Power of Max Dispatchable CPs
Mw = Ms*P= Power Guaranteed by Weight
Mn = Power of Min Dispatchable CPs

Mp = Max(Mn,Min(Mx,Mw))
Capacity Planning Power

Figure 49. Partition Power Numbers

LPAR CP POWER AND WEIGHTS
- An Estimation

3090-400S

	LPAR	Mc=3224
D=2	S=2	S=1
MVS/XA	MVS/XA	MVS/XA
	W=40%	W=60%

Ms=1608

Mx =	1615	1588	846
Mw =	n/a	1608*.4	1608*.6
		643	965
Mn =	n/a	763	0
Mp =	1615	763	846

Figure 50. Partition Power Numbers, a Computation

The partitions which have dedicated CPs are the easy case. Those are running almost like a native system with some small overhead for being in an LPAR configuration.

The power for each shared partition will be a function of the number of shared CPs available to that partition. That would at least establish a maximum. Mx would be that power a partition might get if the other partitions sharing the same CPs fell asleep. Mx is the maximum power for any partition. (See Figure 49 on page 48.)

The Weight assigned to a partition is the amount of power the partition will get when push comes to shove at 100%. The **weight** is an attempt to guarantee some proportion P, of the available shared configuration. So, if Ms is a power number for just the shared part of the CEC, the weight says the minimum of Mx and Ms*P is the maximum delivered at 100% utilization.

Although the partition may be able to receive more power than Ms*P (it is night and the others are sleeping), the smaller of Mx and Ms*P will be used as a guide for capacity planning.

However, in the description above, it was assumed that the available resources were distributed across the sharing partitions. What if there was more power than that? For example, in Figure 50, what would happen if there was no dedicated partition and the level of sharing (S=2 and S=1) remained the same? This would mean that the shared resource could never be at 100%. One processor would always be idle. In this case, regardless of the weight, Mx would equal Mw.

To see whether the power requirements can fit into the CEC, one has to check a graph such as Figure 51 on page 50 to determine whether the total power of the CEC can accommodate the power requirements of the partitions.

The total power of the CEC can only be estimated from a CEC running without LPAR. The estimate thus built will be a combination of power numbers - a function of the SCPs running in the partitions, since the power of the CEC will depend upon the SCP even when running without LPAR.

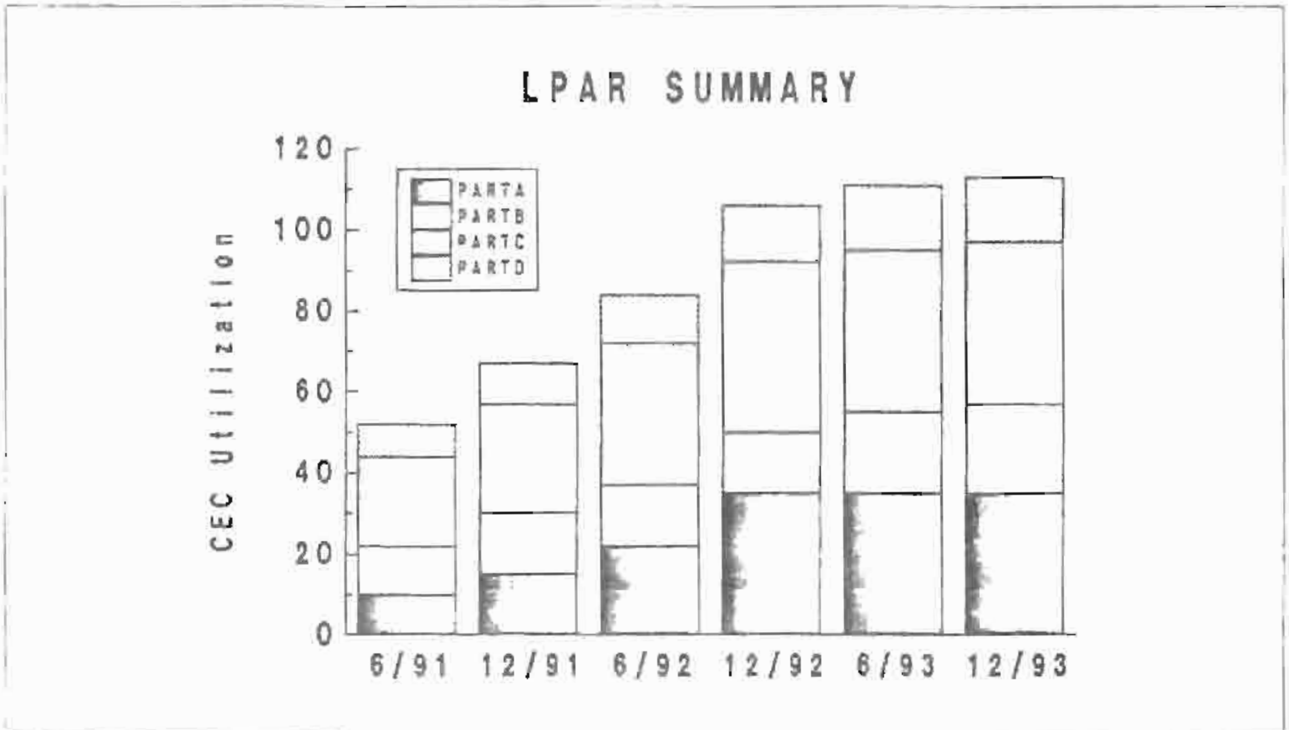


Figure 51. Partition Power Requirements

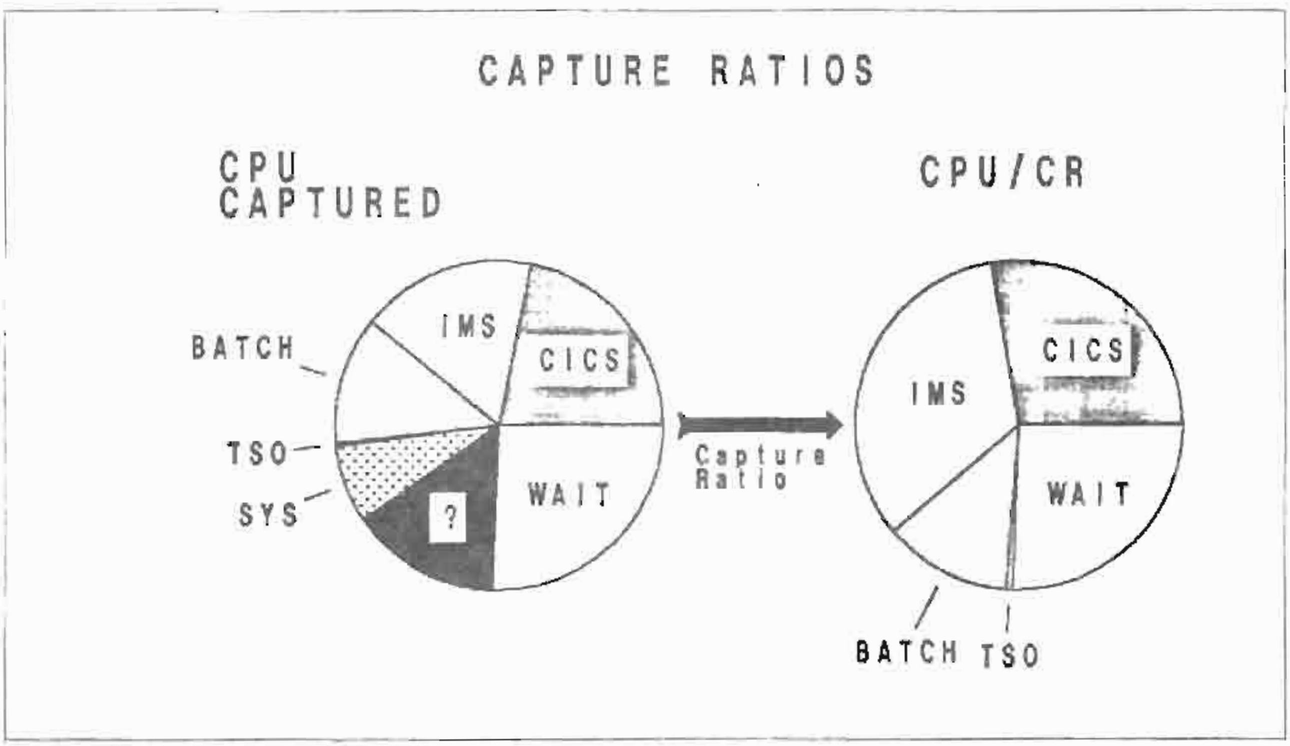


Figure 52. Business Unit Capture Ratios

One other word about CPU usage. The reported values in the data base for CPU usage by workload may not be correct. The sum of the reported values for workloads are usually less than the actual total. The actual total is available in the data base. Capture Ratios adjust the reported workload totals so that the sum of the adjusted values equals the actual. (See Figure 52 on page 50.)

In this adjustment, one usually gets rid of the system address spaces (Master, JES, VTAM, etc) too. After all, the forecast is in terms of the business units alone. That's why the graph is called *Business Unit* capture ratio rather than simply capture ratio. The traditional capture ratio would be numerically higher because it does not distribute the "System" across the business units.

See Appendix D, "Capture Ratios" on page 117 for a discussion of a variety of methodologies and tools available to compute capture ratios.

There is a consequence of applying capture ratios which is not immediately evident - what happens to the I/O rate of a workload when the CPU is adjusted? The I/O rate or the RIOC must be adjusted also. Just as the workload CPU was adjusted in order to make the total workload CPU be the same as the system CPU, the workload I/O must be adjusted to equal the system I/O.

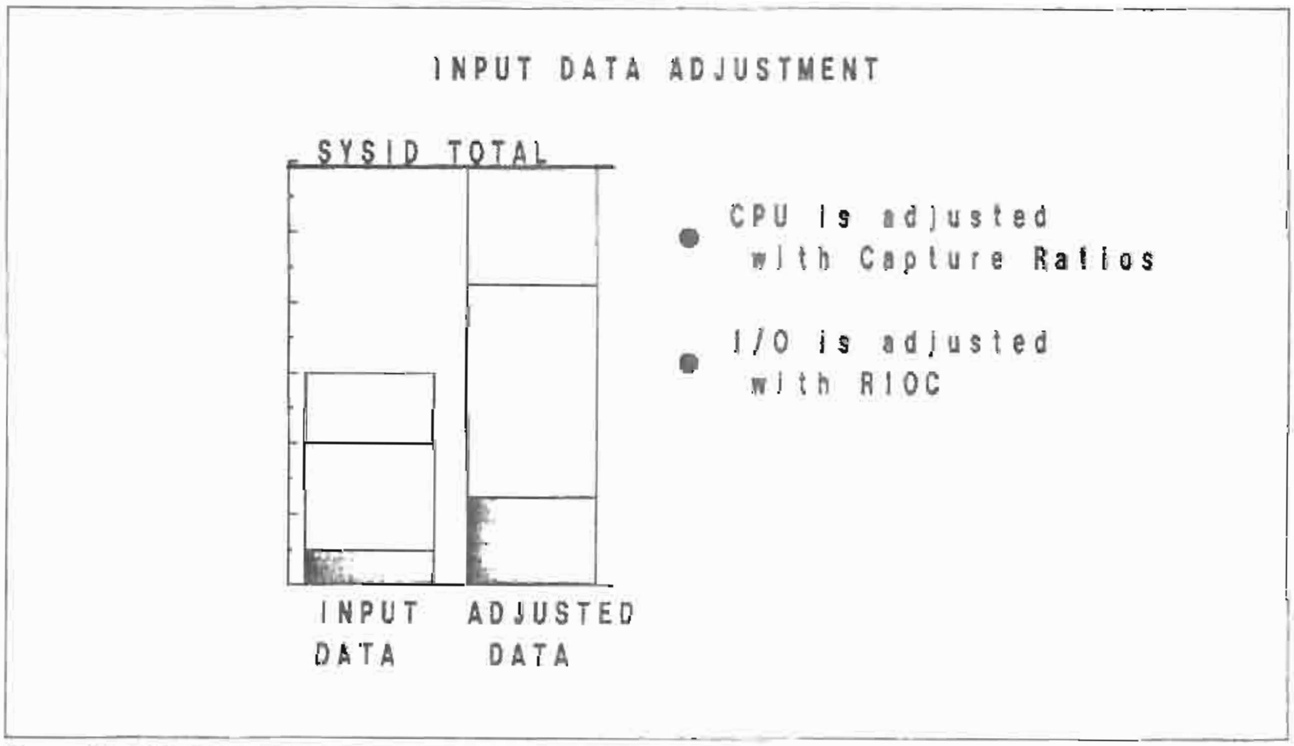


Figure 53. Data Adjustment

Figure 53 shows the capture ratio process and the subsequent RIOC adjustment. If the input data had three business units and the sum did not equal the system image (SYSID) value, the CPU data would be adjusted so that all the CPU time is accounted for by the business units. But if there is a relationship between the amount of power used (M*B) and the DASD I/O (S), adjusting M*B, may mean that the RIOC would have to be adjusted so that the totals also match for the I/O. Figure 54 on page 52 shows a numerical example.

The example starts with an average system CPU utilization of 57% and an average DASD I/O rate of 525. The two workloads add only to 50% and 510 respectively. Using capture ratios, the total is adjusted to match the system value for CPU%. However, once the CPU% is adjusted, the original RIOC values now determine an I/O rate of 579. Hence the RIOC values will have to be adjusted.

The I/O values originally shown for the workloads are in fact problematical in an MVS environment. SMF and RMF do not capture physical DASD I/Os but *block* counts. These may or may not be in one to one correspondence with physical I/Os. However, the counts could be used to proportion physical I/Os in the same proportion. Alternatively, the system RIOC could be used to distribute I/Os in the same proportion as CPU%.

```
        CPU AND I/O DATA ADJUSTMENT
    SYSID     CPU      I/O    RIOC     M=3000
              57.0%    526     .31     526/(.57*3000)

    WKLOAD1   32.0%    333     .35
    WKLOAD2   18.0%    175     .32
              50.0%    508
   Adjust CPU with Capture Ratios
    WKLOAD1   36.5%    382     .35     M*B*RIOC
    WKLOAD2   20.5%    197     .32
              57.0%    579
   Adjust RIOC to match SYSID
    WKLOAD1   36.5%    339     .31
    WKLOAD2   20.5%    185     .30
              57.0%    524
```

Figure 54. Data Adjustment Example

DASD Resource

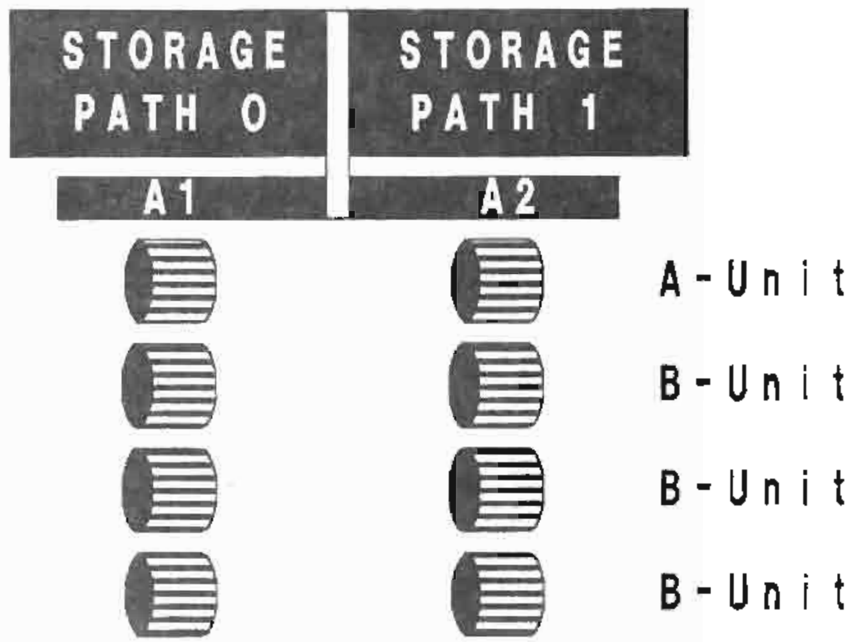


Figure 55. 2 Path BCU

There are a number of ways to approach the DASD subsystem for capacity planning. The choice can best be decided upon examining the questions to be answered. "What equipment (channels, controllers, and actuators) do I need to do the job?" is one question. The rub is in "...do the job". Does this mean have enough space to store megabytes or provide the space and response time needed?

The unit of planning that will be illustrated here is the **Basic Configurable Unit** (BCU) of which the 2 path BCU (3880 technology) is illustrated in Figure 55. The choice of the term "BCU" may be somewhat political. In most cases, as we shall see, the BCU is identical to the MVS "LCU" (Logical Control Unit). The BCU construct is used for those cases where the LCU is not identical to the BCU, for the VM family of SCPs where LCU does not appear (and VM and MVS partisans have been known to be in *friendly* competition), and for VSE SCPs where LCU does not appear.

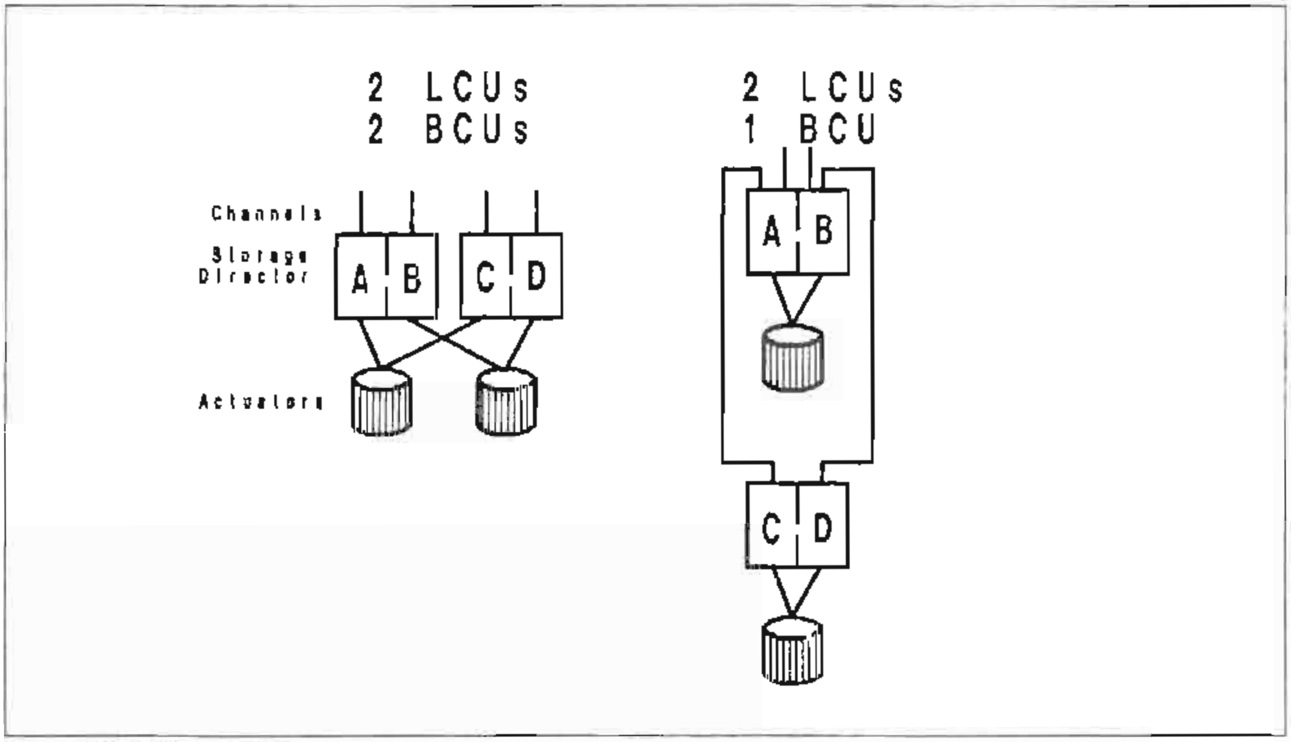


Figure 56. LCU to BCU Conversion

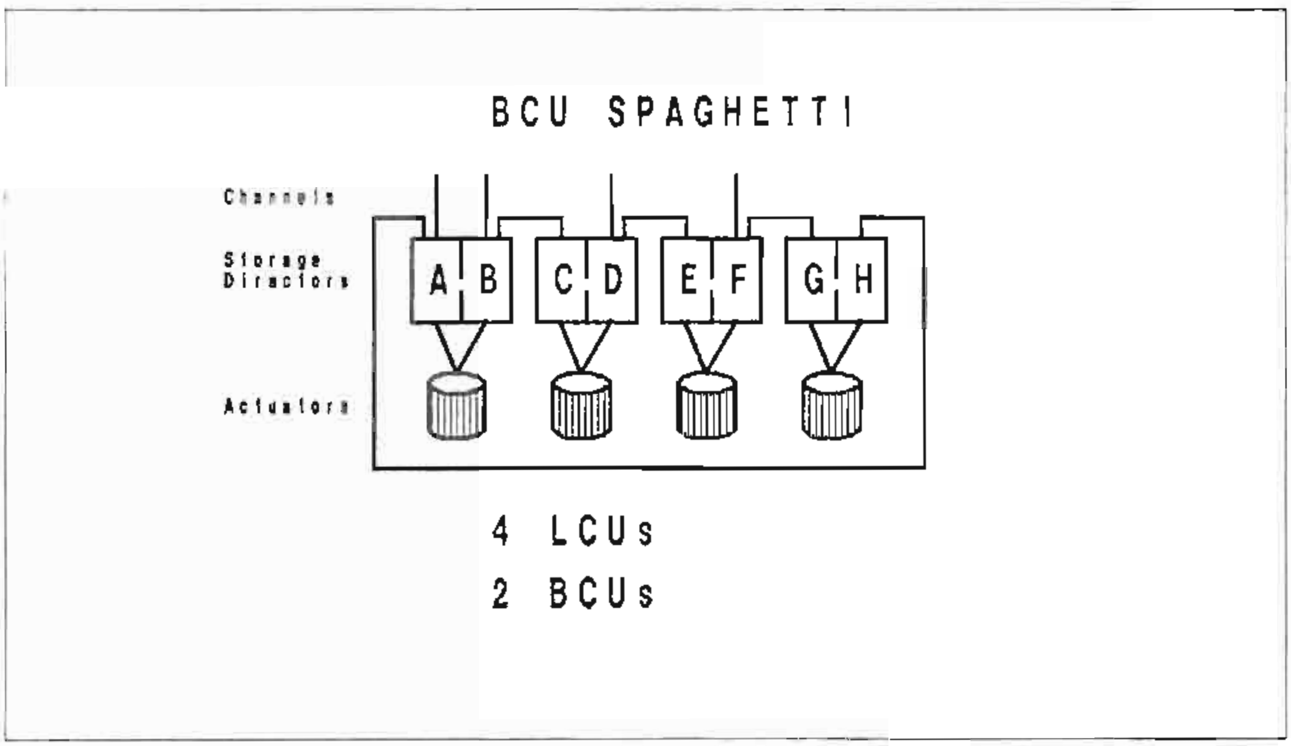


Figure 57. LCU to BCU Conversion

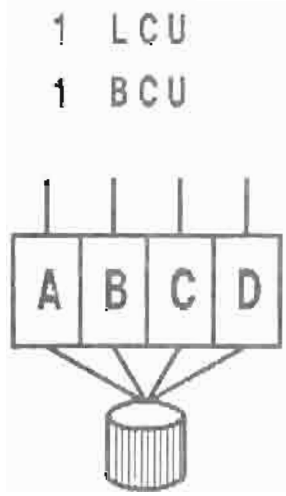


Figure 58. LCU to BCU Conversion

In Figure 56 on page 54 the configuration on the left is two BCUs. The BCU controllers, in this case, are not identical with the physical boxes. There are two controller boxes, but the mapping to BCUs is different. In fact, it is the preferred method of connection. This connection provides better availability. In this case, there are two MVS LCUs and two identical BCUs. The first BCU contains storage paths A and C, the second BCU contains storage paths B and D.

In the same figure, the configuration on the right is problematical. It is one LCU. Is it one BCU or two? There are two controllers (four storage directors) but only two paths to storage! The performance of a BCU is a function of the actuator geometry and the accessibility of paths to processor storage. The presence of the additional controller does not improve pathing. It may improve the ability of adding actuators to a storage path. For performance, this would be modelled as *one* BCU. The environmental data (floor space, heat, etc.) will be less that actual.

However, you may want to consider projecting it as two BCUs if in the future you plan to connect it in such a way as to use the full capability of the two controllers. For example, you have additional channel paths on your processor that you can use.

The BCU in Figure 57 on page 54 presents a problem to any modelling technique. Fortunately, this occurs in earlier technologies. The 3990 and cache technology prohibit this spaghetti. It is one LCU. As an approximation, I would model this as two BCUs.

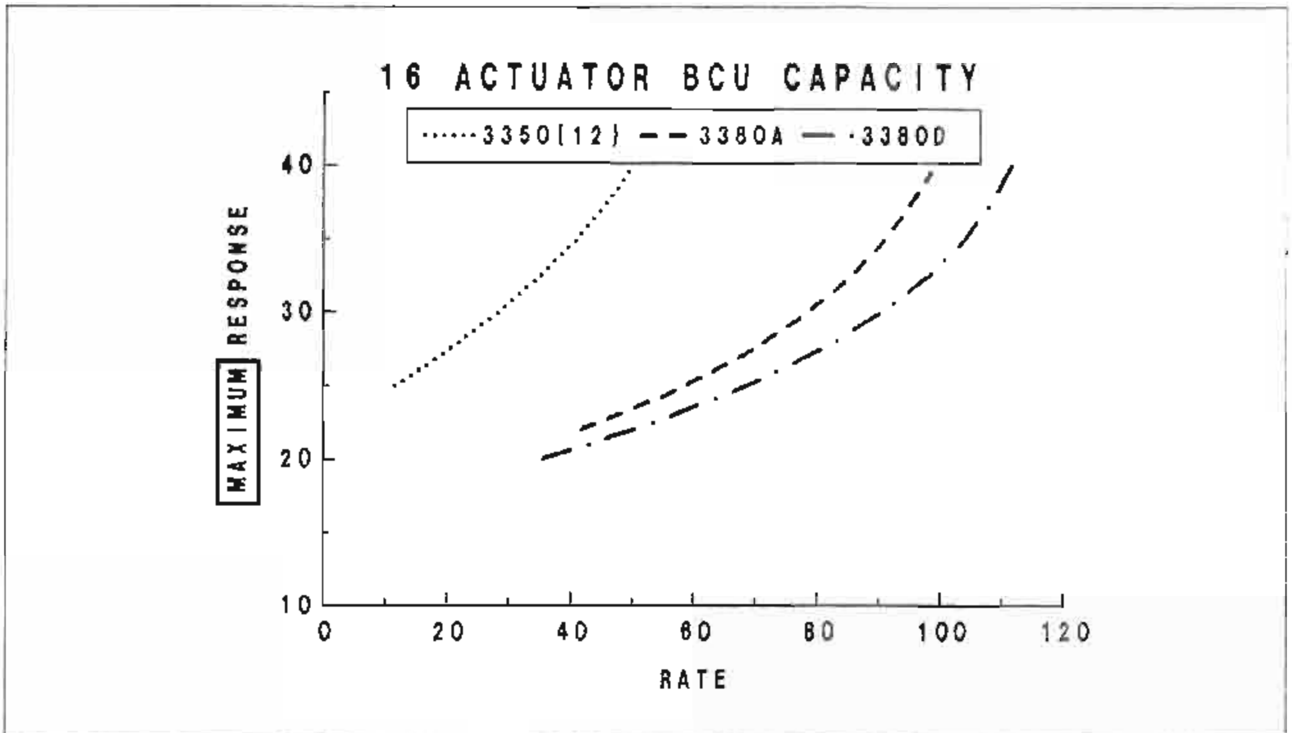


Figure 59. 2 Path BCU Capacity

In order to answer the question "How many I/O operations can a 2 path BCU perform?", reference the curves in Figure 59. Note that the 3380A/D BCUs have 16 actuators whereas the 3350 has 12. How many I/Os can a 3380 standard (3380A) BCU perform (16 3380A actuators, 3880-3 2 path controller)? 40? 60? 80? 100? That depends. It depends upon the desired response time. If the requirement is for a worst case of 20 milliseconds, then one should not plan for more than 40 I/Os per second **at that response time**. If the worst case could be 40 milliseconds, that's a capacity of about 100. A lot more.

The capacity is not only a function of the hardware, but the hardware **and** the Service Level Objective (SLO). For DASD, the maximum expected response time (for a device with an I/O rate greater than 1/second) is being used as the SLO.

Where do these curves come from? They were developed from modelling and assumed I/O characteristics. Even though the assumptions may be impeccable, the results are strictly dependent upon those assumptions. The assumptions include specification of:

- Connect time
- Disconnect time
- I/O skew across the actuators

If the assumptions claim that a 3380A BCU can do 100 I/Os at a SLO of 40 Ms, does that mean that this is the maximum rate that this hardware can attain at 40 Ms? No. Change the assumptions and the rates, and response times change accordingly. Remember, when comparing different hardware, the comparison is valid if the assumptions remain the same. Obvious, but sometimes forgotten.

Some of the other BCU configurations and curves are shown in Figure 60 on page 57 to Figure 62 on page 58.

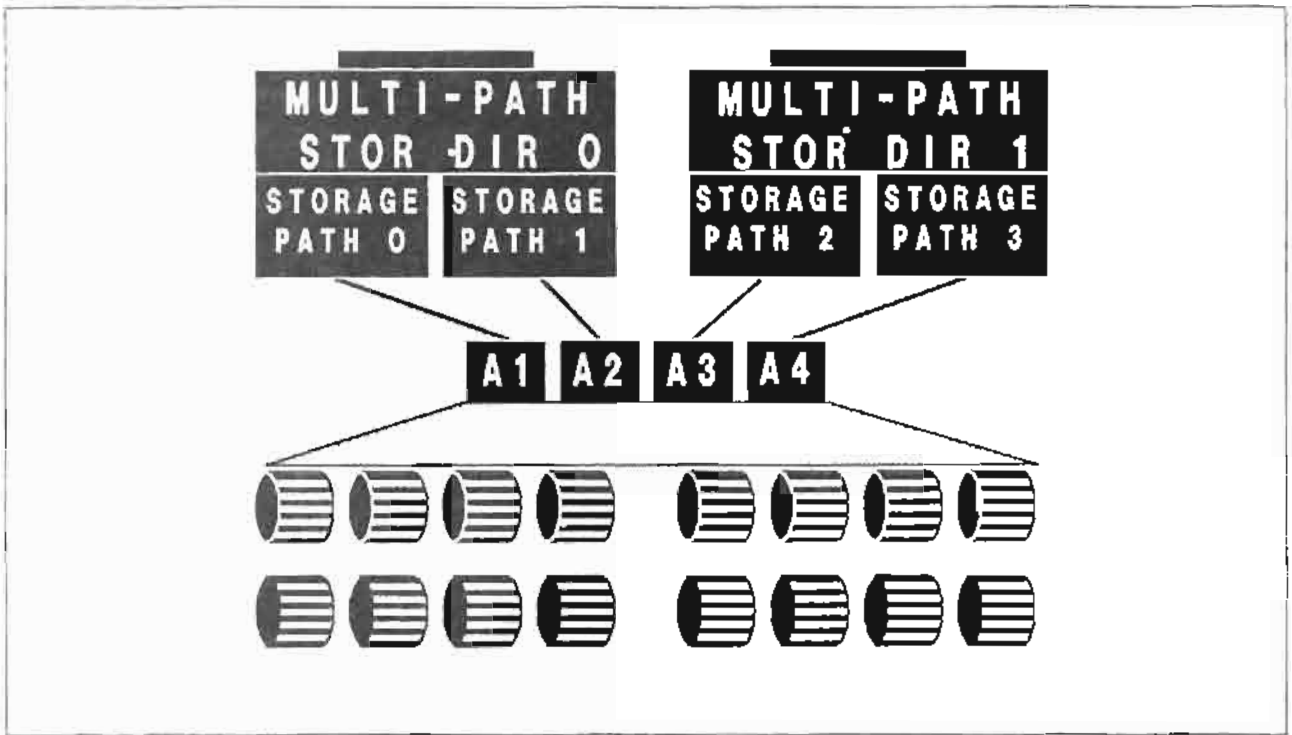


Figure 60. 4 Path BCU

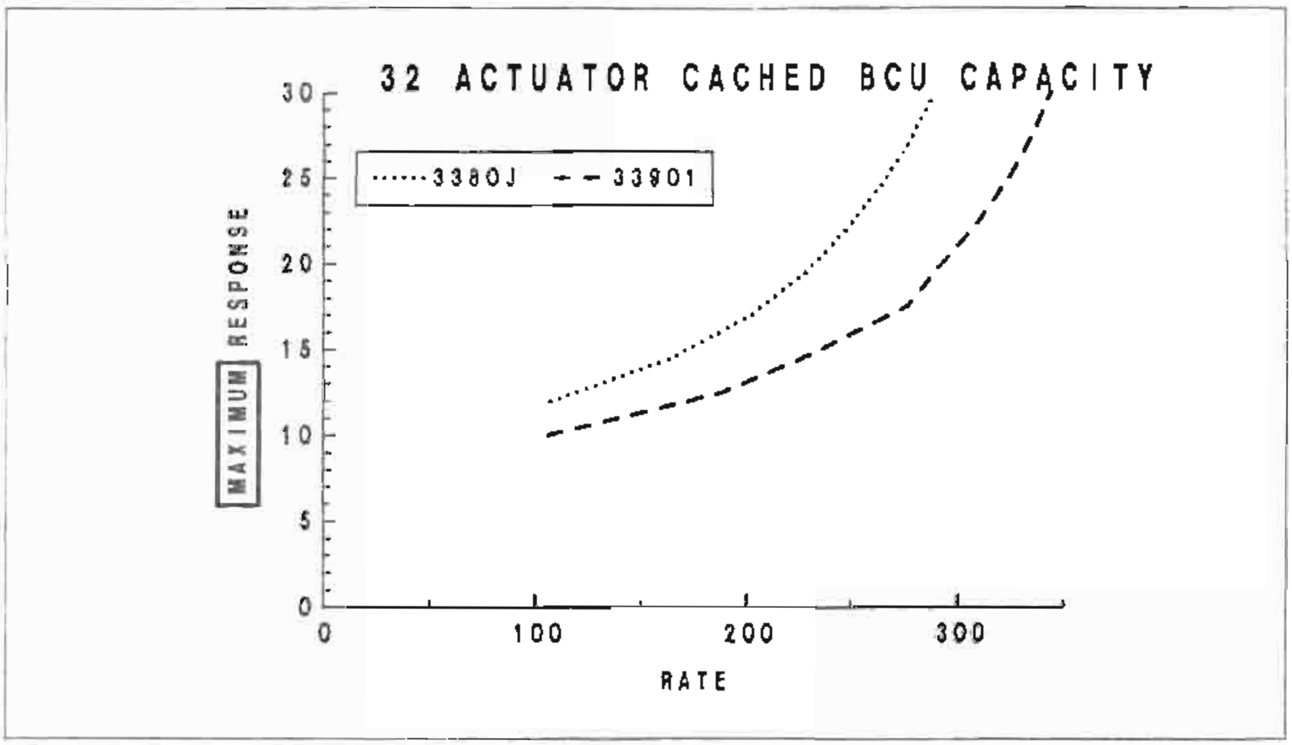


Figure 61. 4 Path BCU Capacity

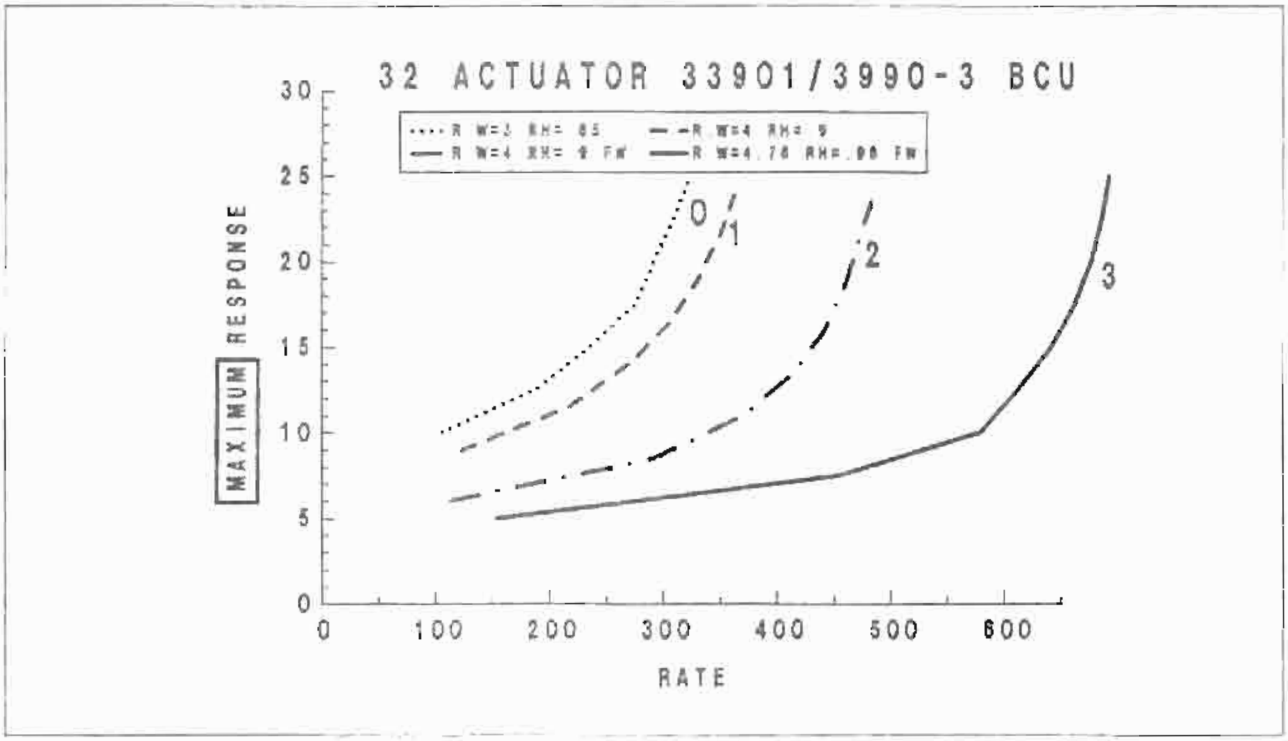


Figure 62. 4 Path BCU Capacity with Cache

Cache controllers are also subject to further assumptions of Read to Write ratios (R:W), Read Hit ratios (RH), and whether extended functions of the 3990-3, such as Fast Write (FW), are used. Figure 62 shows how the capacity curves vary with these assumptions.

When making an assessment of an I/O configurations capability, performance curves such as those illustrated will vary from actual performance achievements. Using fixed curves or tabular performance numbers for both processor and DASD will always be an approximation of reality. A better approximation of actual performance requires a detailed model such as IBMs analytic aid DCAT. There is an intermediate step between these fixed curves and a full analytic approach. That would be to construct a family of curves for all BCUs to reflect a variation of assumptions. This would be an extension of Figure 62 which offers additional performance curves for the same BCU. In that case, the family of curves is through 3 variations of Read to Write ratio, Read Hit ration, and Fast Write or not. The extension of that to all BCUs would permit a variation in connect time to reflect different application types, actuator skew modifications, and maybe even seek characteristics.

Although these curves are a rough approximation, in capacity planning the I/O configuration is changed to plan future requirements, and for most cases, this is sufficient. Change in the I/O configuration means adding and deleting boxes (actuators or controllers). Given the movement of equipment, even with a detailed model, the performance projection requires detailed knowledge of data set movement. Is the migration of whole volumes or are datasets moved to different volumes? Does the I/O characteristics of the volumes remain the same? (Modelling is usually done on a volume basis.) Unless this is carefully modelled, a lot of work, an easier method (using these curves) can often provide the level of accuracy to do the job. The easier method is illustrated here.

If BCUx can do 100 I/Os at a given SLO, will 2 of them do 200? 3 of them 300? If life was only that nice! No, we cannot expect to have a uniform distribution across controllers. Just as the devices in a BCU are not equally used (actuator skew), the controllers are not equally used (BCU skew). The BCU skew can be empirically developed from data obtained about Logical Control Units (usually equal to BCUs in MVS). Figure 63 shows that the

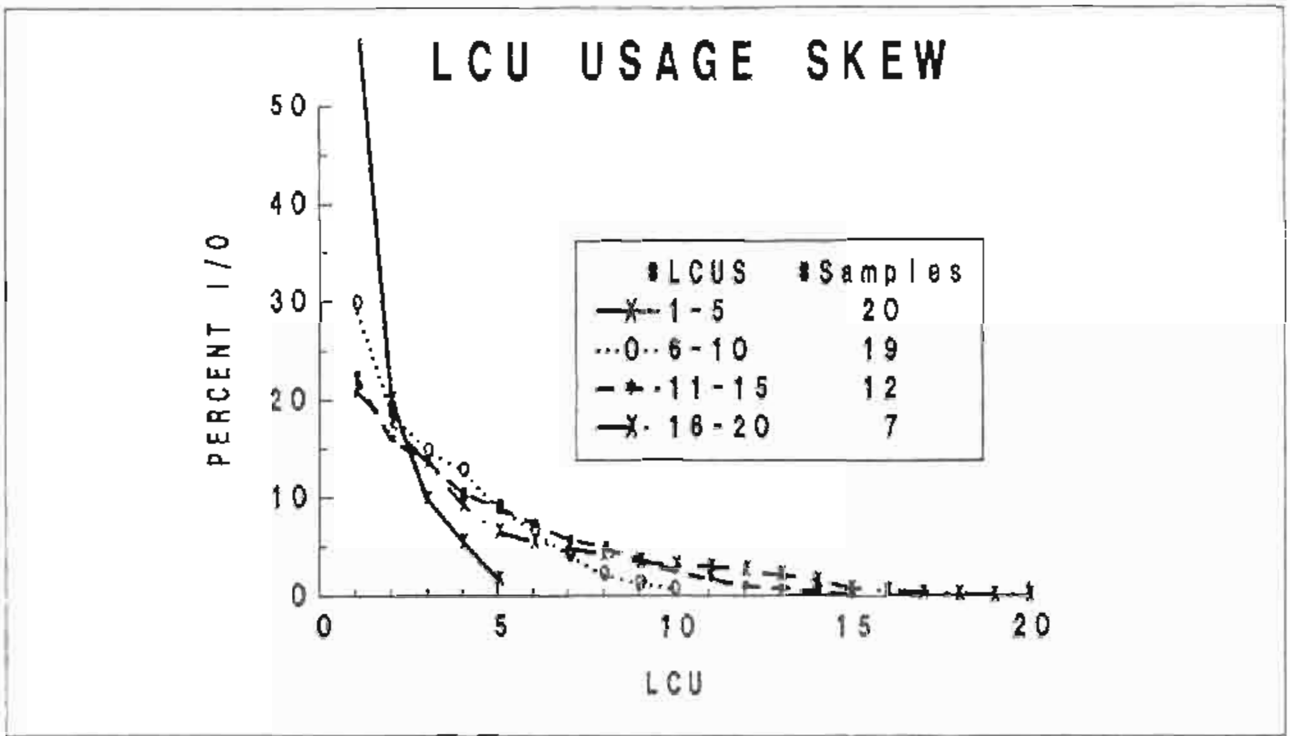


Figure 63. BCU Skew

BCU Load Skew in %

(i/K)**2/3

K	1	2	3	4	5	6	7	8	9	10
1	100									
2	63	37								
3	48	28	24							
4	40	23	20	17						
5	34	20	17	15	14					
6	30	18	15	13	12	11				
7	27	16	13	12	11	10	10			
8	25	15	12	11	10	9	9	9		
9	23	14	11	10	9	9	8	8	8	
10	22	13	11	9	9	8	8	7	7	7

Figure 64. BCU Skew Algebraic Approximation

LCU Skew, An Example

#BCUs	Distribution			
3	48%	28%	24%	
4	40%	23%	20%	17%

I/O		I/O per BCU			
100	3	48	28	24	
100	4	40	23	20	17
120	4	48	27.6	24	20.4

Figure 65. BCU Skew, an Example

skew across controllers is a function of the number of controllers. A complete discussion of this topic can be found in a paper in the CMG '90 Proceedings. (CMG reference 1 on page 123.)

The BCU configuration process is determined by the following steps.

1. Establish the I/O rate requirement by projection period
2. For each proposed BCU select an SLO
3. Determine the rate for each BCU from the performance curves as if it were the only BCU.
4. Use the BCU skew curves to determine the distribution of I/O across the given number of BCUs. Order the BCUs by capability
5. Determine whether the BCUs can service the expected total rate distributed in that manner

In Figure 65 on page 60 there is an example. At 100 I/Os per second and 3 BCUs, it is expected that the I/O will be distributed in the 48%, 28%, 24% manner. What happens immediately when a fourth BCU is added? Nothing. The distribution would be 48%, 28%, 24%, 0% since no data is on the BCU. If the total rate to the configuration remains at 100, and some I/O migrates to the fourth BCU, the rate to the other three falls. The I/O spreads out. It is not until the rate is 120 that the four BCU distribution of 40%, 23%, 20%, 17% that the first BCU is back to 48

If one knew the actual 3 BCU distribution in an installation, could you predict what the final distribution would be when the fourth BCU was added? Without a detailed migration description? This is why the theoretical distribution is used here. It is expected that boxes will be added, deleted, moved around and the detailed data movement scenario will not be available

Storage

For I/O planning the process was to establish some link between the processing power used and the I/O requirement. This relationship is embodied in the RIOC formula S=M*R*B. As (M*B) increases, S increases in the ratio determined by R. Wouldn't it be nice if there were a similar relationship between power used and processor storage required. Well, although there doesn't appear to be any a *priori* relationship, there does appear to be an empirical relationship. In another CMG paper, also by Joe Major (CMG Reference 2 on page 123) we find such an empirical relation. This is shown in Figure 66 on page 62. The amount of Processor Storage (PS) is a function of the base system (20*1.25**N) and the part for the workload (0.01*(M*B)**D)

Proc Stor = Base + F(Power used)

$$PS = 20 * 1.25^{N} + 0.01 * (M * B)^{D}$$

Base value	Workload Contribution
Software Level Adjustment	Calibration Factor

(See Joe Major, CMG '90)

Figure 66. Storage-CPU equation

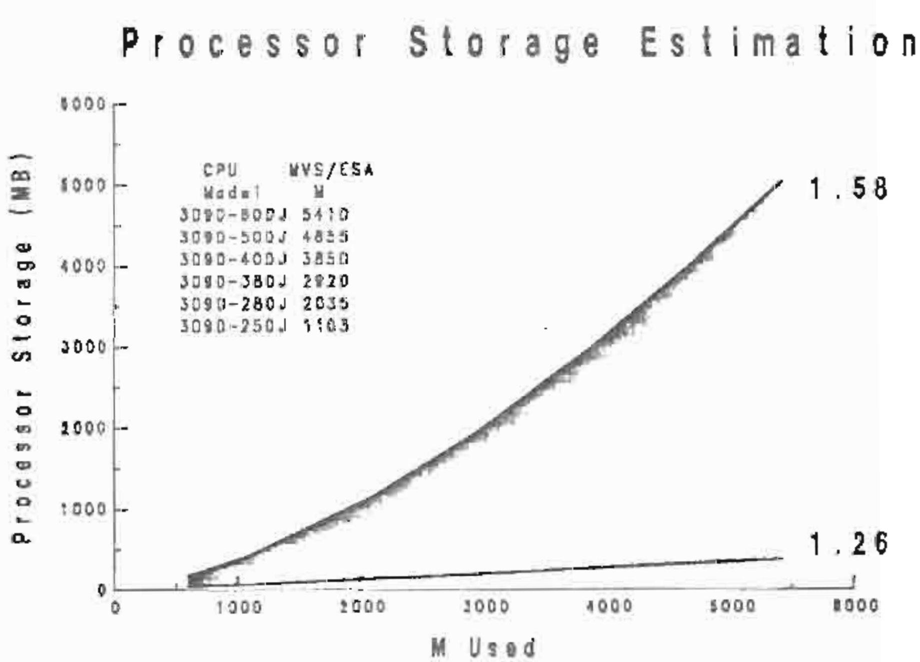


Figure 67. Storage Estimation

Using this formula, processor storage requirements can be estimated based on software currency and the amount of storage currently being used. The process would proceed as follows. (See Figure 67.)

1. Chose a software level index (N) to establish the system part of the equation. How much PS is needed for MVS to just say "Hello"?

N	System	MVS
2	25	4.2.0
1	25	4.1.0
0	20	3.1.3
-1	16	3.1.0
-2	13	2.2.0
-3	10	2.1.7
-4	8	2.1.3
-5	7	2.1.2
-6	5	2.1.1
-7	4	2.1.0

2. Calculate M*B
3. Compute PS storage currently used. That's the sum of
   - Expanded storage online less ES unused (available)
   - Central storage online less CS unused (available).
4. Given N, PS, and M*B, solve for D
5. If the actual values for the calibration part D fall within the range 1.26 to 1.58, a projection for future PS requirements can be made as a function of future M*B requirements.

The process does not say that if the calibration factor is outside of the range, that projection is impossible. It says that of all the data used, there wasn't enough in that range to support the projection.

Experience has shown that systems rich in storage tend to use the existing storage whether needed or not. The only indicators of a rich system available from RMF would be large UIC and migration age values. Experimentation can verify this. Simply vary amounts of storage offline in a staged approach to determine whether pain (storage delays) begins to appear.

The other end of the spectrum, storage paupers, are easy to detect. There's lots of variable documenting storage delays. There's hardly any instances of storage paupers that are not experiencing storage delays.

It is also recommended that the period of projection not be extended beyond 2 years. After two years, enough has changed to call into question the calibration value.

This process assumes that the page and swap I/O together constitute less than 8% of the DASD I/O rate and the workloads used in the calibration remain the same for the projection periods. What if the current paging and swapping I/O rate is greater than 8%?

If the paging I/O rate is greater than 8%, the initial storage must be increased to drop to rate below 8%. A few formula have been found which relate the change in processor storage to the change in paging rate. Intuitively, with a constant workload, an increase in processor storage reduces paging.

Formula One[13]....

[13] See Reference 2 on page 123.

```
   D = 1 - e^(-1.5*X)
where X is the proportion increase in processor storage
      D is proportion of paging I/Os eliminated
```

Roughly, if the storage is increased by 50%, the paging I/O drops by 53%. If one doubles the storage, the paging I/O decreases by 78%. This formula has not been calibrated for a while and is probably still a reasonable estimator for small PS sizes, small being around 64mb.

Formula Two[14]....

```
   New_PR = Old_PR * ((Old_PS)/(New_PS))**2
where PR is Paging Rate
      PS is Used Processor Storage
```

Formula Three[15]....

```
       L = (Y / (y + dy)**(1+Y/20)
where Y = current pageable storage in MB.
     dy = storage added
      L = Proportion of paging I/O still remaining after
          storage increase (1-D) in formula 1.
```

Growth

The best is saved for last....

After all the work is done processing the input, and that's a lot of work, capacity planning begins. It begins by gathering information about the future. The future is described in terms of growth rates applied against the business unit resource description. Given a current resource requirement of a business unit at P_0 and growth rates over some duration, say 3 months, of 1.05, 1.15, 1.17, and 1.20 the resource requirement for the next year is developed as follows.

P_0 = Current Resource Used
$P_1 = P_0 * 1.05$
$P_2 = P_1 * 1.15$
$P_3 = P_2 * 1.17$
$P_4 = P_3 * 1.20$

That's simple enough, but since we worked so hard to get accurate data for P_0, we are interested in the accuracy of P_i. Any guesses?

The answer causes the faint of heart to look for a different job. The accuracy of a forecast is as good as the growth numbers. What if the growth rate for the first period was 40% (1.40) rather than 5%? It would mean that the *real* resource requirement for this business unit is significantly larger than expected. That's a correspondingly significant problem - the result of the capacity plan is a capital acquisition which represents the business's readiness for data processing services and a poor capacity plan my understate the requirement. It could also overstate the requirement. That's a waste of capital resources.

What do you do with a forecast which is suspect? You do what a statistician does with varying data. You bound it. Figure 68 on page 65 shows a bounding

[14] Formula from Shelly Weinberg, IBM I/S Management Institute
[15] See Reference 3 on page 123.

graphic. Replacing the expected growth rates with values somewhat larger and small, an upper and lower bound can be drawn. These values are obtained by experience and knowledge of the business.

- What variation was there in the past?
- Is the business climate possibly more pessimistic than forecast?
- More optimistic?

Given these bounds, the resource projection does not yield a single point, but an interval with some point within the interval representing *best guess*.

One other reason for applying bounds to a forcast is that the *real* forecast may not be known. It is quite often that the forecast or business plan is quite confidential and the capacity planner is asked to evaluate a number of scenarios which represent a range of growth possibilities.

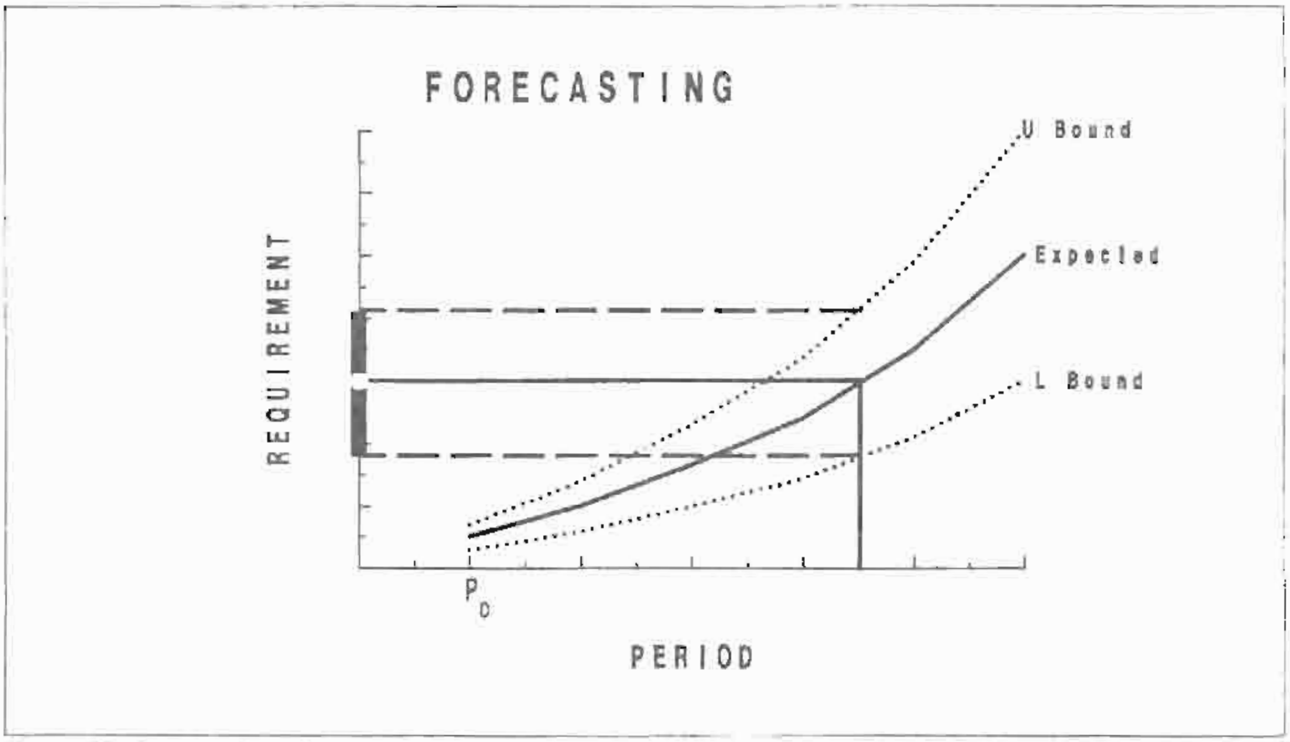


Figure 68. Forecast Bounding

The capacity planner should also be very aware of most activities related to data processing. Here's a partial list of things that can dramatically impact resource requirements.

- Latent demand as already mentioned may be *out there* ready to eat you alive.
- More work stations may be planned for an existing department without any external growth specifications. This increased access to the system can result in an impressive increase in resource usage.
- A control program application upgrade can change the resource requirement.
- Technology upgrades can translate into more work. An *operational improvement*, say DFHSM, might collapse the time waiting for data. That's more transactions albeit more productivity too.

There are some general observations about looking into the future.

- Growth requires room. If the resources are saturated, growth may appear to stop or slow significantly. Watch out for latent demand.
- A resource buffer may be required for the unexpected. If the business environment is very dynamic (acquisitions and mergers), a large amount of resource may have to be reserved to be ready for a business opportunity
- Certain software may enable business opportunities. Watch for software (or hardware) which could make the business grow.
- Is the current architecture implementation limiting growth? Would moving from MVS/XA to MVS/ESA enable applications?

Above all, keep history data. It serves as a test of truth. At least it can serve as a warning about the differences that can arise between past expectations and what really happened.

In general, most people under estimate resource requirements.

Chapter 3. Capacity Planning - A sample.

CP input Process

1. Establish the structure. Drawing a picture of the hardware and software, such as found in Figure 69, pulls together the relationships of the items drawn and the data involved. On the left is a single system image. On the right, an LPAR. The DASD and the shared BCU is shown. The picture also shows the level of detail for the data: workload data will be needed.

2. Get data for each system image (Processor, storage, I/O). The data required will depend upon the structure depicted below. For example, if the BCUs are included, you will require DASD I/O statistics for all the system images.

3. Check sample selection. Are all the workloads active? Is the data chosen taken from the same time frame? Is the data complete (all the objects described)?

4. Adjust data (Capture ratios, I/O totals). For a structure with workload (or business unit) descriptions, the data should be adjusted so that the sum of the CPU, I/O, and storage equal the total for the system. The adjustment of the data for the workloads creates a set of objects which describe the system image (see Figure 70 on page 68). For this example, the object has three dimensions - CPU, I/O, and storage. The sum of the objects fill the system image. The system image may be an object itself as part of an LPAR. Regardless, the workload is the basic object.

5. The projection process provides a description of how the workload objects grow in the three dimensions. The task is then to provide a fit container for the objects with an appropriate amount of space left for future growth. Containers are defined (new CPUs, single system image of LPAR) and the objects are moved around to get a good fit.

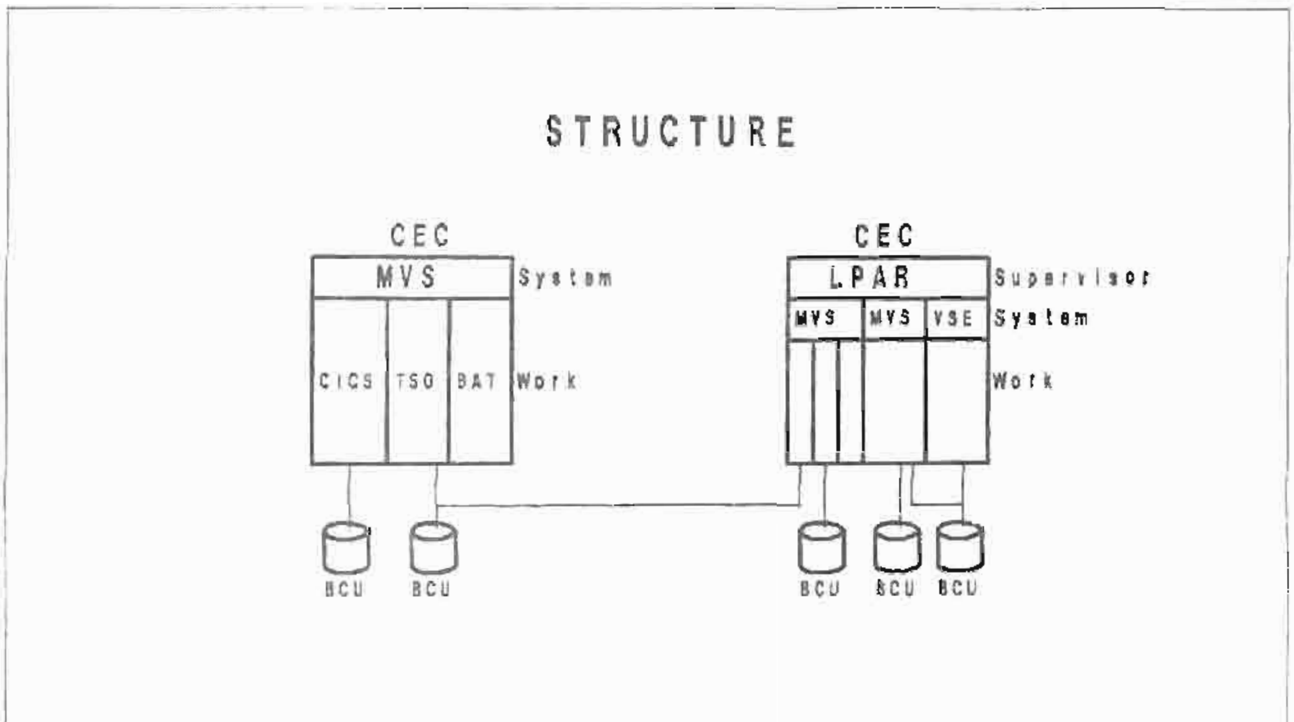


Figure 69. Overall Structure

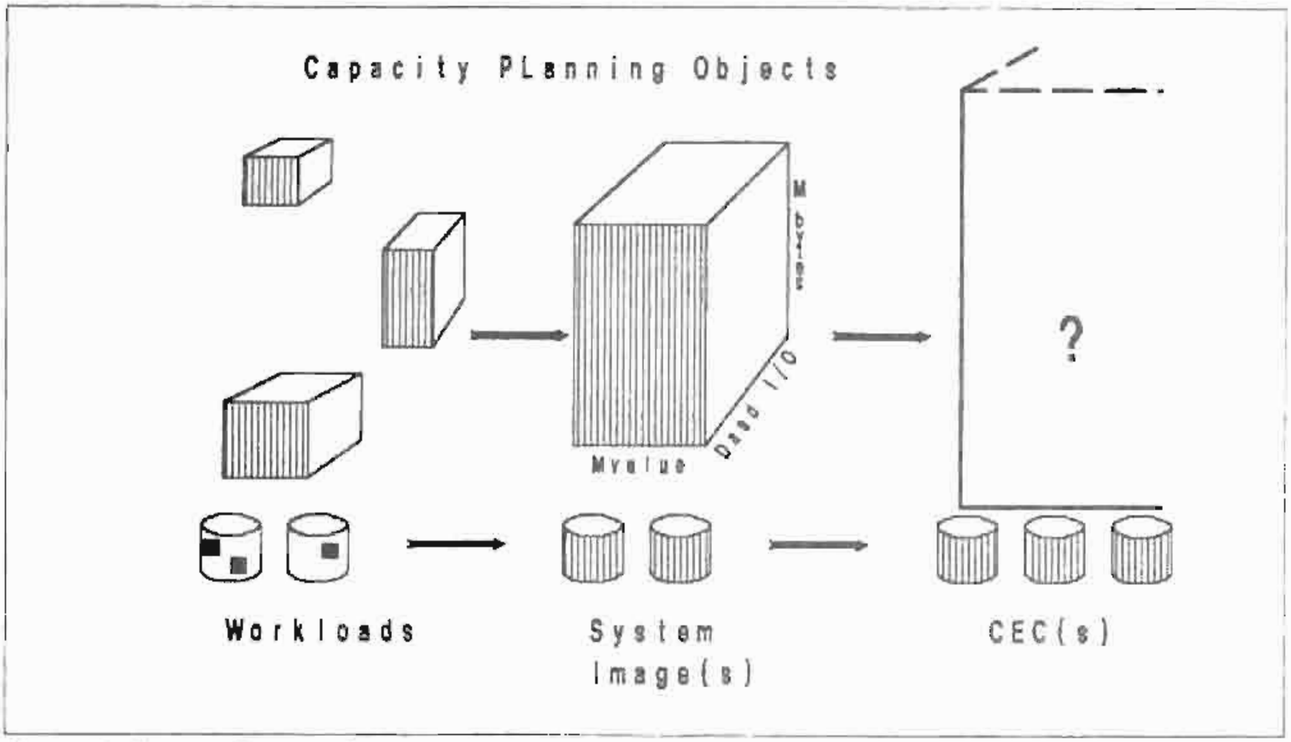


Figure 70. Capacity Planning Objects

The rest of this chapter demonstrates a capacity planning process illustrating topics discussed in the previous section. The process to be demonstrated will not be complete to avoid repetitive discussion. Here's what we shall do.

1. Input data for current systems
   - System (SYSID) data for each system which is CPU%, DASD I/O rate, and Page and Swap I/O rate
   - Storage configuration, an optional step in capacity planning
   - Workload Data, an optional step, which consists of either CPU% and I/O data or just CPU%.
   - BCU data, which is also optional
   - Adjust CPU% for workload with capture ratios if needed.
   - Adjust RIOC for workload if needed.
2. Projections and Requirements for the Future
   - Enter Growth scenarios for each workload
   - Storage Projections
   - Processor Projections
     - Migrate to larger processors
     - Combine processors onto a single system image
     - Move system images to an LPAR
   - DASD Projections
     - Delete and Add to configuration
     - Examine I/O rate requirements
     - Examine environmental requirements

CP90 - a Capacity Planning Tool

CP90 is the tool used to illustrate this capacity planning process, and to demonstrate the effective and credible methodology discussed in previous chapters. CP90 is a capacity planning tool available to IBM system engineers on their HONE (Hands On Network Environment) system. Customers can also use CP90 by subscribing to Capacity Planning Service on IBMLINK.

Capacity Planning Sample Data

The illustration that follows is a two CEC complex. The data is **constructed** to feature the topics discussed in the concepts chapter. System A, the development system, is running TSO and Batch. System B, the production system, is running CICS and Batch.

Development System

SYSID. SYSA
CPU Model: 4381-24
SCP: MVS/XA
CS Online: 32
ES Online: 0
CS Available: 1
ES Available: 0

Table 1. SYSID input data

DATE mm/dd/yy	TIME hh:mm	DURATION hh:mm	CPU%	DASD I/O Rate	Page/ Swap I/O Rate
3/19/91	8:00	8:00	62%	98	14
3/20/91	8:00	8:00	65%	97	12
3/21/91	8:00	8:00	72%	108	15
3/22/91	8:00	8:00	68%	99	14
3/23/91	8:00	8:00	75%	100	17

Development System Workloads

Workload Description: TSO Development
Growth Rate: 10% per annum

Table 2. TSO Development Sample Data

DATE mm/dd/yy	TIME hh:mm	DURATION hh:mm	CPU%
3/19/91	8:00	8:00	25%
3/20/91	8:00	8:00	27%
3/21/91	8:00	8:00	24%
3/22/91	8:00	8:00	24%
3/23/91	8:00	8:00	23%

Workload Description. Batch Development
Growth Rate: 20% per annum next year and then 15%.

Table 3. Development Batch

DATE mm/dd/yy	TIME hh:mm	DURATION hh:mm	CPU%
3/19/91	8:00	8:00	27%
3/20/91	8:00	8:00	29%
3/21/91	8:00	8:00	30%
3/22/91	8:00	8:00	30%
3/23/91	8:00	8:00	28%

Development System BCUs

Table 4. BCU Data

BCU ID	Controller Type	Actuator Type	# Actuators	I/O Rate	Maximum Response
BCU1	3880-3	3380A	12	87	42
BCU2	3880-3	3380D	16	10.51	62

Production System

SYSID: SYSB
CPU Model. 3090-200
SCP: MVS/XA
CS Online: 64
ES Online: 64
CS Available: 2
ES Available: 10

Table 5. SYSID input data

DATE mm/dd/yy	TIME hh:mm	DURATION hh:mm	CPU%	DASD I/O Rate	Page/ Swap I/O Rate
3/19/91	8:00	8:00	83%	420	3
3/20/91	8:00	8:00	79%	400	5
3/21/91	8:00	8:00	84%	440	4
3/22/91	8:00	8:00	80%	398	3
3/23/91	8:00	8:00	82%	410	5

Production System Workloads

Workload Description. CICS Production
Number ofUsers: 100
Trans/Minute: 10
Growth Rate Adding more users every 6 months. the users are getting more sophisticated, and productivity is increasing.

Table 6. Production CICS Sample Data

DATE mm/dd/yy	TIME hh:mm	DURATION hh:mm	CPU%	I/O
3/19/91	8:00	8:00	50%	250
3/20/91	8:00	8:00	52%	275
3/21/91	8:00	8:00	55%	275
3/22/91	8:00	8:00	46%	267
3/23/91	8:00	8:00	48%	282

Workload Description: Production Batch
Growth Rate. 10% per annum

Table 7. Development Batch

DATE mm/dd/yy	TIME hh:mm	DURATION hh:mm	CPU%	I/O
3/19/91	8:00	8:00	25%	80
3/20/91	8:00	8:00	27%	120
3/21/91	8:00	8:00	30%	98
3/22/91	8:00	8:00	31%	110
3/23/91	8:00	8:00	31%	109

Production System BCUs

Table 8. BCU Data

BCU ID	Controller Type	Actuator Type	# Actuators	I/O Rate	Maximum Response
BCU2	3880-3	3380D	16	17.91	83.6
BCU3	3880-23	3380E	16	189	24
BCU4	3880-3	3380E	16	87	36
BCU5	3880-3	3380E	16	66	37
BCU6	3880-3	3380E	16	27	38

Current System Definition

The execution of the capacity plan will use CP90 to illustrate the steps required

Processor data Input

Enterprise definition begins with a processor configuration specification. The CEC Specification is a combination of Supervisor and CPU model. If the Supervisor was LPAR, further definition would be required (Partition data and SCP). In this case, the Supervisor is an SCP (MVS/XA) so no further definition is required. The SCP and CPU Model will determine not only the processor capabilities (Is LPAR possible?) but also the power number (M Value in CP90 terms[16]). Each CEC is identified by a CECID (Figure 71.)

Enter the **SYSID Data** for SYSA as an example (Figure 72 on page 74.) The date, time and duration should be kept somewhere as a reminder of the data source. The time period of the data may have an impact on the results. Are any of these days a weekend? The duration is a reminder of sample size. Here we have 8 hour samples. With 8 hour samples we should expect a steady RIOC. If the samples were 15 minutes, the variation could be far greater.

Notice that the Peak to Average ratio (PAR) is only 1.1. The computed saturation design point (SDP) would be 100/1.1 or 91%. This is based upon the samples entered. But the samples entered are averages for 8 hours. Peak to average ratios should be based upon much smaller samples. So in this case we shall use 70% instead of the computed value.

```
Processor Input             Enterprise Specification            Panel: ACENT
                                   CECID: All

Enter or review CEC data.
Enter 'S' in Sel field for each CEC to be processed and press Enter.
SCP and CPU Model affect the M value. Review SYSID data when changed.


   Enterprise Name : Balanced Systems Sample                 Use PF7/8 to scroll:
   Number of CECIDs: 2                                        Index CPU Models
                                                                 1 9021-720
                    Index, SCP/ Index or           Valid SCPs/   2 9021-620
  Sel    CECID      Supervisor  CPU Model    Index Supervisors   3 9021-580
                                                                 4 9021-500
         SYSA       MVS/XA      4381-24          1 LPAR          5 9121-480
         SYSB       MVS/XA      3090-200         2 VM/XA         6 9121-440
                                                 3 MVS/ESA       7 9021-340
                                                 4 MVS/XA        8 9021-330
                                                 5 MVS/SP        9 9121-320
                                                 6 VSE          10 9121-260
                                                 7 VM/HPO       11 9121-210
                                                 8 VM/SP        12 9121-190
                                                 9 VS1          13 9221-170
                                                                14 9221-150
                                                                15 9221-130
                                                                16 9221-120
                                                                17 3090-28T

PF: 1=Help 2=Save 3=End 4=Comment 6=Select All 7=Backward 8=Forward
 9=Auto Input 10=Erase All data 11=Erase Projection data 12=Cancel
```

Figure 71. Enterprise Definition

[16] The M Value in CP90 is developed from a scaled ITRR.

```
Processor Input          CPU and I/O Specification            Panel: ACSIO
                        CECID: SYSA      SYSID:  SYSA

Edit or review CPU, I/O and Paging Data.
Caution: See help panel for unique functions of this editor (PF6).

                   Number of Samples: 5
     File Id       Date        Time     Dur      CPU      I/O      Paging
   or Samp. #      MM/DD/YY    HH:MM    HH:MM    %        Rate     SIOs      RIOC

        1          05/20/91    08:00    08:00    62.0     98.0     14        .41
        2          05/21/91    08:00    08:00    65.0     97.0     12        .39
        3          05/22/91    08:00    08:00    72.0    108.0     15        .39
        4          05/23/91    08:00    08:00    68.0     99.0     14        .38
        5          05/24/91    08:00    08:00    75.0    100.0     17        .35










                                       Average:  68.4    100.4     14        .39
  Computed PAR:    1.1
  Computed SOP:  91    SOP to be Used:  70

PF: 1=Help 2=Graphics 3=End 5=Continue 6=Edit 7=Backward 8=Forward 9=Top/Bottom
 12=Cancel
```

Figure 72. SYSA SYSID Data

Workload Data Input is similar to the SYSID input (See Figure 73 on page 75). For this case, we don't have any I/O data for the TSO Development workload. Input consists of the CPU% data. Without any I/O data, the estimate of I/O will be obtained by using the system RIOC value of 0.385. Since we don't know the capture ratios yet, we will use a CR of 1. Both the RIOC and CR may have to be adjusted later when all the workload data is entered and we can compare the workload totals to the SYSID totals and adjust the workload values accordingly.

For the second workload, we enter the data in a similar manner (Figure 74 on page 75). In both cases, the assumed method of growth will be growth by CPU%. Otherwise, the data would have to be gathered for the number of users and number of transactions and entered here with the workload.

```
Processor Input                 Workload Data                      Panel: ACSWP
                           CECID: SYSA      SYSID: SYSA
                           Workload: TSO Development

Enter or review workload sample data.
Use PF6 to select mode of entry.
  - Enter or change the CPU, CR, or RIOC to compute I/O.
  - Enter or change the CPU, CR, or I/O to compute RIOC.

     SYSID CR:  .78       System RIOC:  .385     Number of Users:       0 (Opt)
  Workload CR: 1.00     Workload RIOC:  .385    Transactions/Min:       0 (Opt)

  Date     Dur.  CPU    CPU/    I/O   Computed
 MM/DD/YY  Min. Input    CR    Input     I/O

 05/20/91  480  25.0   25.0      .0     36.7
 05/21/91  480  27.0   27.0      .0     39.6
 05/22/91  480  24.0   24.0      .0     35.2
 05/23/91  480  24.0   24.0      .0     35.2
 05/24/91  480  23.0   23.0      .0     33.7




       Average  24.6   24.6      .0     36.1

PF: 1=Help 2=Graphics 3=End 6=Enter CPU-I/O 7=Backward 8=Forward 9=Top/Bottom
 10=Previous Workload 11=Next Workload 12=Cancel
```

Figure 73. Workload Input

```
Processor Input                 Workload Data                      Panel: ACSWP
                           CECID: SYSA      SYSID: SYSA
                           Workload: Batch Development

Enter or review workload sample data.
Use PF6 to select mode of entry.
  - Enter or change the CPU, CR, or RIOC to compute I/O.
  - Enter or change the CPU, CR, or I/O to compute RIOC.

     SYSID CR:  .78       System RIOC:  .385     Number of Users:       0 (Opt)
  Workload CR:  .78     Workload RIOC:  .385    Transactions/Min:       0 (Opt)

  Date     Dur.  CPU    CPU/    I/O   Computed
 MM/DD/YY  Min. Input    CR    Input     I/O

 05/20/91  480  27.0   34.6      .0     50.8
 05/21/91  480  29.0   37.2      .0     54.5
 05/22/91  480  30.0   38.5      .0     56.4
 05/23/91  480  30.0   38.5      .0     56.4
 05/24/91  480  28.0   35.9      .0     52.7




       Average  28.8   36.9      .0     54.2

PF: 1=Help 2=Graphics 3=End 6=Enter CPU-I/O 7=Backward 8=Forward 9=Top/Bottom
 10=Previous Workload 11=Next Workload 12=Cancel
```

Figure 74. Workload Input

Once the data for all the workloads is entered. We have to compare the workload totals to the SYSID total and adjust the data. Figure 75 on page 76 shows that the CPU totals and the I/O totals are off sufficiently to warrant adjustment. The data displayed already is adjusted with the workload capture ratios provided earlier, but with a place-holder of 1. A capture ratio of 1 does nothing.

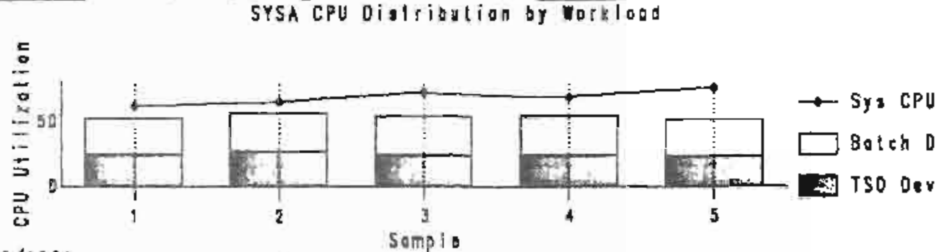
Capture ratios can be computed in a number of ways. The graphic in Figure 76 on page 77 shows the difference between the workload data (the bars) and the system total (the line). The capture ratios attempt to make them as close as possible. In this case, we shall use the system capture ratio of 0.78. Figure 77 on page 77 shows the data adjusted to the new value.

This adjusts the CPU totals to match. For this case, the I/O is close enough to the SYSID total. We shall see an RIOC adjustment later.

```
Processor Input               Workload Summary                  Panel: ACWKL
                          CECID: SYSA      SYSID: SYSA
                          Workload: All

Enter or review workload descriptions.
Enter 'S' in Sel field for each workload to be processed and press Enter.


              Unique Workload                         Adjusted
          Sel  Description                            CPU       I/O

           •   TSO Development                        24.6       36
           •   Batch Development                      28.8       42










                                    Workload Total    53.4       79
                                       SYSID TOTAL    68.4      100
Workload CPU total differs from System CPU Utilization (53%). Adjust CR.
PF: 1=Help 2=Graphics 3=End 5=Continue 6=Select All 7=Backward 8=Forward
 9=Top/Bottom 10=Adjust CR 11=Adjust RIOC 12=Cancel
```

Figure 75. SYSA Workload Summary

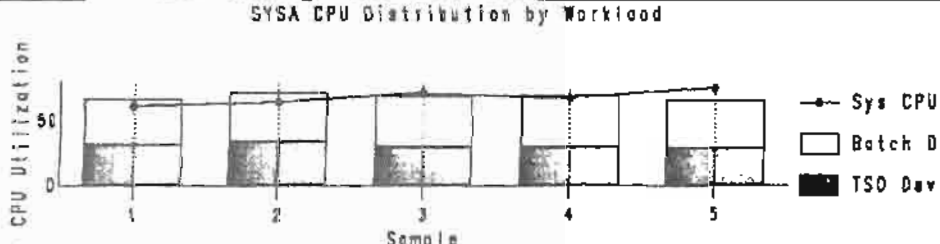
```
Processor Input          Capture Ratio (CR) Adjustment          Panel: ACCR
Workload: All           CECID: SYSA       SYSID: SYSA

Adjust CR Manually, or use function keys for assistance in adjusting CR.

              Workload              Input Calc.   Workload
                  #      I/O    CR    CPU   CPU    Description
                  1      36   1.00   24.6  24.6 TSO Development
                  2      42   1.00   28.8  28.8 Batch Development


           Total Workloads:         53.4  53.4 (-21.9%)
              Total SYSID:_    .78  68.4  68.4
```




```
5/30/1991
XRACCR                  Balanced Systems Sample

PF: 1=Help 2=Zoom 3=End 4=Reset CR 7=Backward 8=Forward 9=Regression
 10=Use Sys CR 11=Prop. to I/O 12=Cancel
```

Figure 76. Capture Ratio - Before

```
Processor Input          Capture Ratio (CR) Adjustment          Panel: ACCR
Workload: All           CECID: SYSA       SYSID: SYSA

Adjust CR Manually, or use function keys for assistance in adjusting CR.

              Workload              Input Calc.   Workload
                  #      I/O    CR    CPU   CPU    Description
                  1      36    .78   24.6  31.5 TSO Development
                  2      42    .78   28.8  36.9 Batch Development


           Total Workloads:         53.4  68.5 (    .1%)
              Total SYSID:_    .78  68.4  68.4
```




```
5/30/1991
XRACCR                  Balanced Systems Sample

PF: 1=Help 2=Zoom 3=End 4=Reset CR 7=Backward 8=Forward 9=Regression
 10=Use Sys CR 11=Prop. to I/O 12=Cancel
```

Figure 77. Capture Ratio - After

Storage data input (Figure 78 on page 78) is rather straight forward. However, we are appropriately warned that the paging I/O rate is over 8% and will have to be adjusted prior to any sucessful storage projection. The *available* values are for those unused frames. That would be the average available values on the storage report in MVS.

```
Processor Input            Storage Specification             Panel: ACSTO
                          CECID: SYSA       SYSID: SYSA

Enter or review storage data.
Select "Set SSE" function to adjust SCP level, if necessary.


                                   SCP: MVS/XA
                             CPU Model: 4381-24
                              CPU Text: Development System
         System Software Environment: SP2.2.0

                            CS Online:    32 MB.
                            ES Online:     0 MB.
                         CS Available:     1 MB.
                         ES Available:     0 MB.

                   Page/Swap I/O Rate:    14 Page I/O rate is excessive. It
                        DASD I/O Rate:   100 should be < 8% of DASD I/O
             Paging Percent of DASD I/O:  14 % rate. At Processor Projection
                                             time, it will have to be
                                             adjusted.


PF: 1=Help 3=End 5=Continue 9=Set SSE 12=Cancel
```

Figure 78. SYSA Storage Input

Figure 79 shows the summary of data put into the model for SYSA. Notice that the Historical Peak to Average Ratio is 1. If we replaced that with 1.6, the latent demand would indicate 100-1.6*68.4 or 9%. The DASD input for SYSA will not be shown here, but will be shown shortly for SYSB in this chapter.

```
Processor Input          System Image Specification          Panel: ACSIM
                         CECID: SYSA       SYSID: SYSA

Enter 'S' in Sel field for each function to be processed and press enter.
Enter or review system data on this panel.

                  +--- Input Data-Function Selection -----+
                  | Sel Data Type             Data Exists |
                  |                                       |
                  |  •  CPU      (Required)        Y      |
                  |  •  Workload (Optional)        Y      |
                  |  •  Storage  (Optional)        Y      |
                  |  •  DASD     (Optional)        Y      |
                  +---------------------------------------+

                                   SCP: MVS/XA
                             CPU Model: 4381-24
                               M-Value: 381
                             CPU Text : Development System
                 Number of CPs Online: 2
                     CPs Shared (Y/N): N
                     Average CPU Busy:  68.4 %
               Saturation Design Point:  70.0 %
                    Peak:Average Ratio: 1.1
         Historical Peak:Average Ratio: 1          Latent Demand: 0    %
                        Total DASD I/O:  100.4
                                  RIOC:   .385
                           Paging Rate: 14

Input added or updated.
PF: 1=Help 3=End 4=Comment 6=Select All 10=Previous SYSID 11=Next SYSID
   12=Cancel
```

Figure 79. SYSA SYSID Summary

The input process for SYSB follows a similar process - CPU data input, workload data, storage data, and then DASD data. Figure 80 on page 79 shows the workload summary for SYSBs workloads after the capture ratios have adjusted the CPU data. The workload total for I/O is different enough from the SYSID total to warrant RIOC adjustment.

Figure 81 on page 80 shows the original input. The difference between workload and SYSID I/O is -9.8%. Although the input workload I/O came from RMF data which reports only "block" counts, the RIOC adjustment could be made using these somewhat fictitious I/O counts. One could distribute the physical I/O count in the same proportion as the block counts. Figure 82 on page 80 shows the results after this adjustment. The other alternative is to use the system RIOC for each workload. This would assume that each workload produces the same I/O rate for the same amount of CPU.

The SYSID summary for SYSB is shown in Figure 93 on page 88.

```
Processor Input              Workload Summary                     Panel: ACWKL
                        CECID: SYSB      SYSID: SYSB
                        Workload: All

Enter or review workload descriptions.
Enter 'S' in Sel field for each workload to be processed and press Enter.


                  Unique Workload                   Adjusted
             Sel  Description                      CPU      I/O

              *   CICS Production                 50.2      270
                  Batch Production                29.2      103










                             Workload Total   79.4      373
                                SYSID TOTAL   81.6      414
Workload I/O total differs from System I/O (5%). Adjust RIOC.
PF: 1=Help 2=Graphics 3=End 5=Continue 6=Select All 7=Backward 8=Forward
 9=Top/Bottom 10=Adjust CR 11=Adjust RIOC 12=Cancel
```

Figure 80. Workload Summary

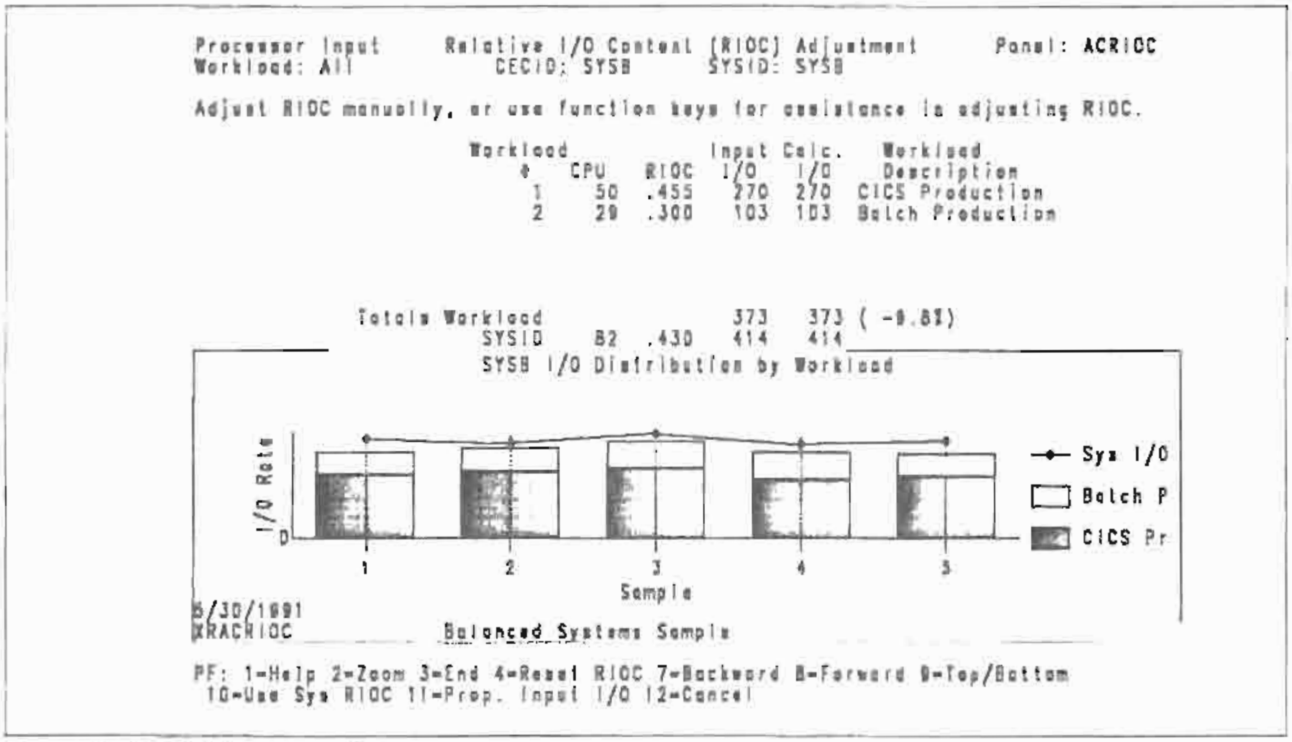


Figure 81. RIOC Adjustment - Before

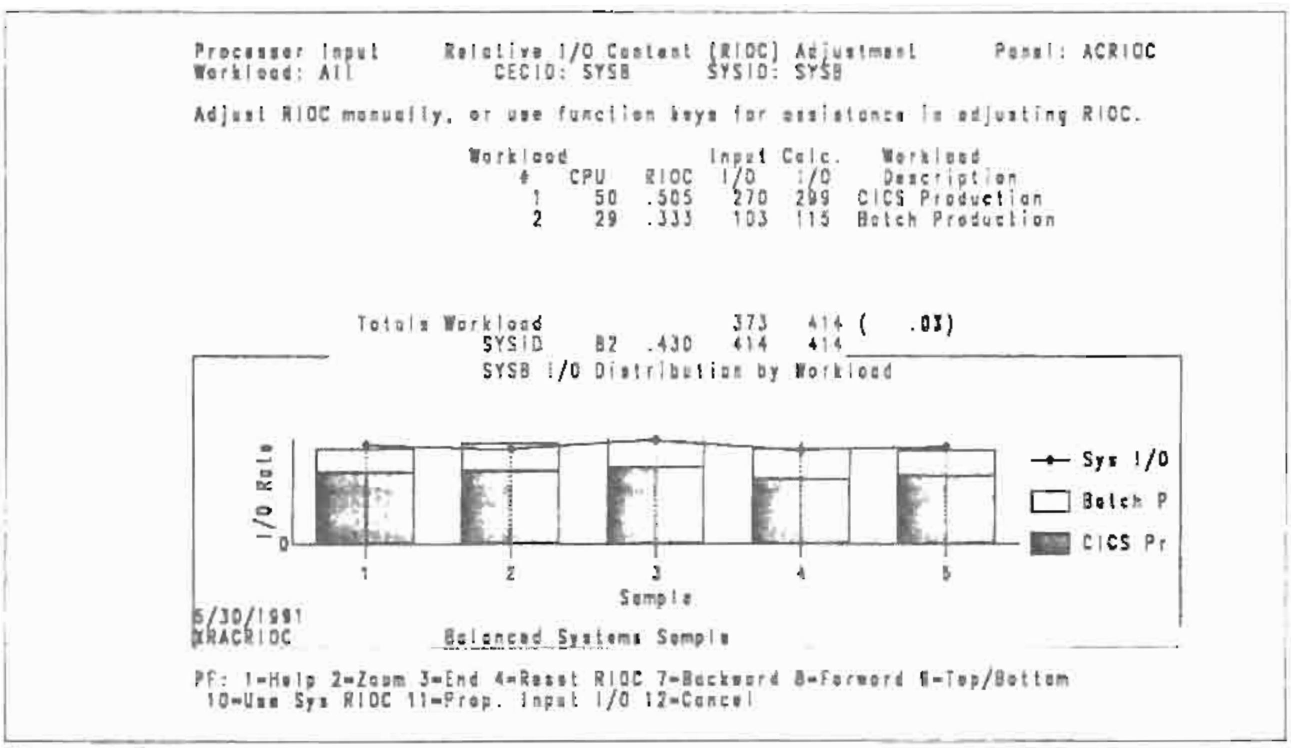


Figure 82. RIOC Adjustment - After

DASD data Input

The DASD data for SYSB is entered in two phases. The overall performance picture and then the environmental information. The performance data may be entered in multiple levels. The first level is entered in Figure 83 on page 82 in terms of the controller type, DASD type, I/O rate, and **maximum** I/O Response time.[17]

Response time data on the BCU level may be optionally entered as shown in Figure 84 on page 82. This data can be used to review the components of response time as shown in Figure 18 on page 20.

In the event that we need to review further the performance data of actuators on each BCU, this is shown in Figure 85 on page 83. This data is not entered by manual input, but is included by automated input tools to CP90.[18] In this way, DASD performance data may be used as an overview as well as for detail performance management. Figure 85 on page 83 shows the review of actuator data for BCU2 which has a maximum response time of 83.8 Ms. This will be discussed further.

On panel displayed in Figure 83 on page 82, response time entered is that of the actuator with the largest average response time observed on the BCU. Why choose the maximum rather than the average? The maximum was chosen on the assumption that the performance evaluation of the I/O subsystem should be centered around the worst case which can often be the busiest and most important device. The average could be used if accompanied by a response time skew value such as Maximum/Average. Graph in Figure 86 on page 83 generated by CP90 shows the relationship between maximum and average response time can be used for this purpose. In this case there is no great variation between the maximum and average response time for the BCUs except for BCU2. Therefore it is appropriate to use maximum response time as the service level objective of all BCUs except for BCU2. For BCU2, maximum response time far exceeds average response, then you probably don't want to use that as the service level objective. We will show how this is changed in the projection phase.

The BCUID is used to identify which BCUs have DASD sharing. BCU2 was identified in both SYSA and SYSB. The data for this BCU will have to be combined to obtain the complete picture of the activity on BCU2. Figure 87 on page 84 shows the I/O distribution across BCUs. BCU2, as indicated, has I/O from both SYSA and SYSB.

[17] The maximum response time is the actuator with the largest average response time observed on the BCU. The I/O rate to that actuator should have a non trivial I/O rate of something greater than 1.

[18] These tools include CPAIP (ARIAS), CP90EXTR, and the bridge from MXG. ARIAS is part of CP90. CP90EXTR and the MXG bridge are available from the authors of this bulletin.

```
DASD Input                      BCU Data (1 of 3)                  Panel: BCUINP
                           CECID: SYSB      SYSID: SYSB

Enter or review BCU data.
Pop Window (PF4) requires the cursor be on a field in the window to be popped.
Enter 'S' in Sel field to select BCU actuator data.

  +-------------------DASD Basic Configuration Unit Types-------------------+
  | Index Type      Index Type      Index Type      Index Type      Index Type |
  |   1   39901        9  399D1C      17  3880DJ      29  9345T1      41  9332  |
  |   2   39902       10  399D2C      18  3880K       30  9345T2      42  9335  |
  |   3   39903       11  399D3C      19  3880D       31  9345F1      43  9336  |
  |   4   3990J       12  399DJC      20  3880E       32  9345F2      44  23051 |
  |   5   3990K       13  399DKC      21  3880A       33  93CC21      45  23052 |
+------------------------------- BCU Data ------------------------------------+
|                  BCU       Storage             I/O        Max
| Sel  BCUID       Type      Director    DASD    Rate       Response
       BCU2        3880DJ    3880-3      3380    17.91      83.6
       BCU3        388DEC    3880-23     3380   189.00      24.0
       BCU4        388DE     3880-3      3380    87.00      36.0
       BCU5        388DE     3880-3      3380    66.00      37.0
       BCU6        388DE     3880-3      3380    27.00      38.0
|
+-----------------------------------------------------------------------------+
Total Number of Actuators:     80   Total I/O:      387
Enter or review BCU response time data: N

PF: 1=Help 2=Graphics 3=End 4=Pop Window 5=Continue 6=Select all 7=Backward
 8=Forward 9=Top/Bottom 12=Cancel
```

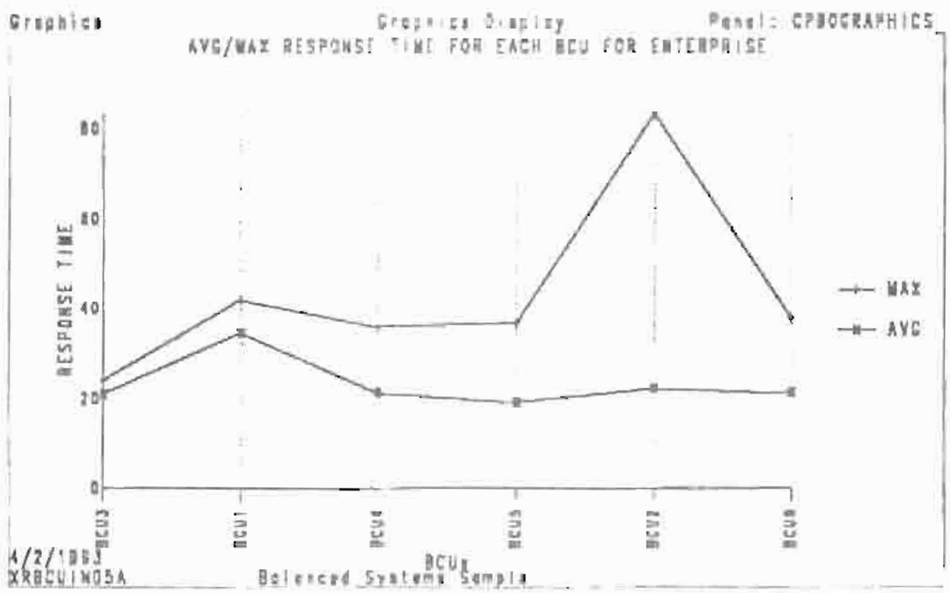
Figure 83. BCU Input

```
DASD Input                      BCU Data                           Panel: BCURT
                           CECID: SYSB      SYSID: SYSB

Enter or review BCU data.
Press PF4 to sort numeric field on which the cursor is positioned.
Enter 'S' in Sel field to select BCU actuator data.
+------------------------------- BCU Data ------------------------------------+
|                BCU      I/O     Max       Average
     Sel BCUID   Type     Rate   Response  Response  CONN    DISC    PEND    IOSQ
         BCU2    3880DD   17.91   83.6      17.8     3.8     13.3    .5       .20
         BCU3    388DEC  189.00   24.0      21.0     5.4     7.9     1.6     6.10
         BCU4    3880E    87.00   36.0      21.5     3.7     12.1    .5      5.20
         BCU5    3880E    66.00   37.0      19.2     3.1     14.0    .8      1.30
         BCU6    3880E    27.00   38.0      21.5     3.8     11.6    .4      5.70
|
+-----------------------------------------------------------------------------+
Total Number of Actuators:     80   Total I/O:      387

PF  1=Help 2=Graphics 3=End 4=Sort 6=Select all 7=Backward 8=Forward
 9=Top/Bottom 12=Cancel
```

Figure 84. BCU Response Time Components

```
DASD Input                        BCU Actuator Data                     Panel: BCUACT
                             CECID: SYSB       SYSID: SYSB

Review BCU actuator data.
Press PF4 to sort field on which the cursor is positioned.

+-----------------------------BCU-Actuator Data -----------------------------+
                                      DASD    I/O  Average
 XCLUD BCUID   Volume Address Type    Rate Response CONN    DISC   PEND   IOSQ
       BCU2     GL8004   0180 33800     1.5    83.8  58.0   14.3   13.4     .0
       BCU2     GS8049   0186 33800      .5    20.5   4.1   14.8    1.4     .3
       BCU2     GL8059   0191 33800     1.2    23.0   8.5   14.3     .2     .0
       BCU2     GM8027   0195 33800     2.8    18.1   2.1   13.7     .3     .1
       BCU2     GS8017   0196 33800      .8    18.4   3.5   13.8     .9    1.2
       BCU2     GL8051   01AD 33800     3.4    15.9   2.9   12.8     .3     .0
       BCU2     GL6027   01A3 33800      .2    16.8   3.4   13.0     .4     .0
       BCU2     GL8086   01A5 33800      .2    17.4   3.1   14.0     .3     .0
       BCU2     GS8010   01AB 33800     2.1    17.8   1.8   15.0     .8     .2
       BCU2     GL8104   01B1 33800      .2    21.8   8.7   12.8     .3     .0
       BCU2     GS8007   01B3 33800     1.2    19.6   2.7   14.7     .7    1.5
       BCU2     GL8019   01B8 33800      .2    18.4   3.1   13.0     .3     .0
       BCU2     GM8015   01B9 33800     1.4    18.5   3.1   12.8     .8     .0
       BCU2     GL8020   01BB 33800     1.5    14.1   3.1   10.8     .3     .0
       BCU2     GM8017   01BE 33800      .1    15.8   2.8   12.4     .4     .0
       BCU2     GL8021   01BF 33800      .5    19.8   2.8   14.8    2.2     .0

+--------------------------------------- <-------------Average------------> +
No. of Actuators:     16 Total I/O     17.9    17.7   3.5   13.4      8     .2

PF: 1=Help 2=Graphics 3=End 4=Sort 5=Restore Original Selection 7=Backward
 8=Forward 9=Top/Bottom 12=Cancel
```

Figure 85. Actuator Response Time Components




```
PF: 1=Help 2=Create ADMCDF 3=End 5=Display & Describe 8=Customize
10=Previous Graph 11=Next Graph 12=Cancel   ENTER=Toggle PFKeys  PA3=Plot Graph
```

Figure 86. Graph showing maximum versus average response time

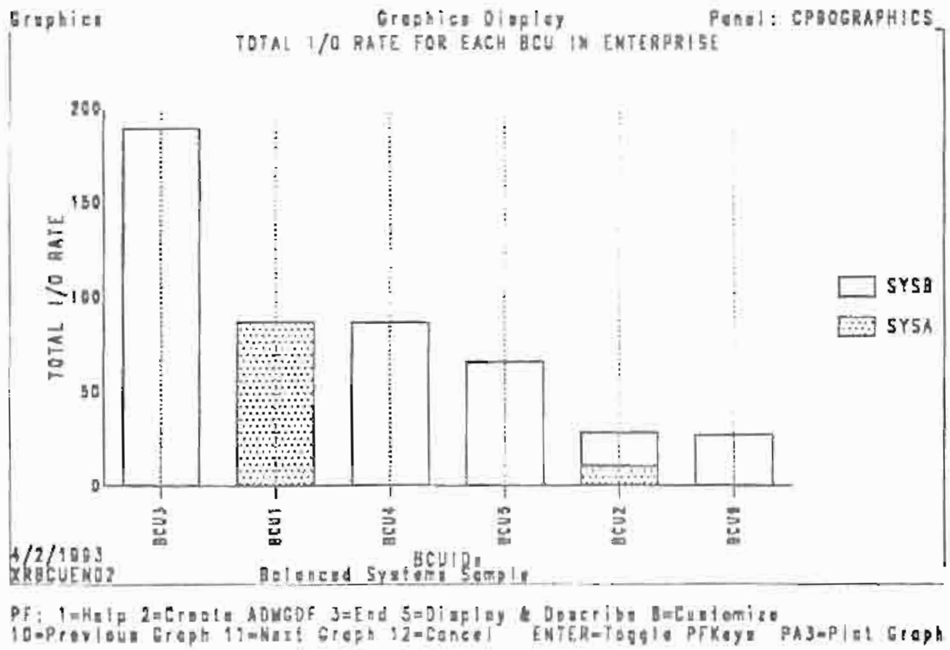


Figure 87. BCU Graphic Summary

The number and type of actuators on each BCU are shown in Figure 88. This environmental information will be used to generate data for megabytes, heat, power, and floor space.

```
DASD Input                    BCU Data (2 of 3)                 Panel: BCUIN
                         CECID: SYSB       SYSID: SYSB

Enter or review actuator data.

+------------------------------ Actuator Data -------------------------------+
           BCU    Storage        DASD   DASD   DASD   DASD   DASD   DASD   DASD   DASD
  BCUID    Type   Director DASD  Type-#Type-#Type-#Type-#Type-#Type-#Type-#Type-#
  BCU2     3880D  3880-3   3380  D 16   J  0   K  0   E  0   A  0      0      0      0
  BCU3     3880EC 3880-23  3380  E 16   J  0   K  0   D  0   A  0      0      0      0
  BCU4     3880E  3880-3   3380  E 16   J  0   K  0   D  0   A  0      0      0      0
  BCU5     3880E  3880-3   3380  E 16   J  0   K  0   D  0   A  0      0      0      0
  BCU6     3880E  3880-3   3380  E 16   J  0   K  0   D  0   A  0      0      0      0











+----------------------------------------------------------------------------+
  Total Number of Actuators:    80

PF: 1=Help 3=End 5=Continue 7=Backward 8=Forward 9=Top/Bottom 12=Cancel
```

Figure 88. Actuator Input

Finally the data from all the SYSIDs can be combined into an enterprise picture of the DASD (Figure 89 on page 85).

```
DASD Input              Enterprise Performance Summary          Panel: BCUENT
                       CECID: All        SYSID: All

Review enterprise BCU performance data.
DASD shared by multiple systems has been combined.
+------------------------ Enterprise BCU Performance -------------------------+
                                Actual       Actual       Expected   Variation
                    BCU         I/O          Max          Max        from
 Sel BCUID          Type        Rate         Response     Response   Expected
     BCU1           3880A          87           42.00      32             31%
     BCU2           3880D          28           83.80      30            179%
     BCU3           3880EC        189           24.00      25             -4%
     BCU4           3880E          87           36.00      30             20%
     BCU5           3880E          66           37.00      30             23%
     BCU6           3880E          27           38.00      30             27%

+-----------------------------------------------------------------------------+
                    Total: 484       Avg. 35.4
Review BCU response time data: N

PF: 1=Help 2=Graphics 3=End 4=Sort by MAXREP 5=Sort by I/O Rate 6=Select All
 7=Backward 8=Forward 9=Top/Bottom 12=Cancel
```

Figure 89. BCU Summary

Since the DASD projection method is based upon a fixed relationship between maximum response time and I/O rate, this figure also shows the difference between the reported maximum response time and the expected maximum response time as obtained from the curves. The curves, such as Figure 59 on page 56, can be looked upon as functions relating either I/O rate to maximum response time or the other way around. As indicated in the previous chapter, the tabular or curve method of projection will not adjust the curves to actual data. Figure 89 is essential to determine the difference between the input data and what the model considers a reasonable maximum. This assessment should act as a warning if the input data indicates poor response times. The projection process will enable the users to correct this.

Figure 89, shows that BCU2 varies greatly from the acceptable maximum response time. We can choose to look further into the DASD performance data to see if this is a problem: we can look at response time data on the BCU level as well as the actuator level. Response time data on the BCU level is shown in Figure 84 on page 82. Simply looking at response time for a BCU will not help you to decipher problem areas in your DASD subsystem. A BCU with long response time and low I/O rate is usually not a problem. To review the combination of response time and I/O rate, we can look at the I/O intensity of a BCU as depicted in Figure 90 on page 86. I/O intensity is calculated by mutiplying I/O rate with response time or the components of response time. Sorting the BCUs by their respective I/O intensity gives you a good picture of the relative importance of each BCU to the DASD subsystem. The channel load generated by each BCU can be calculated by multiplying the I/O rate and connect time of each BCU. By reviewing the data, it is shown that BCU1 and BCU3 carry high channel loads whereas BCU2 does not present itself to have great impact on the DASD subsystem.

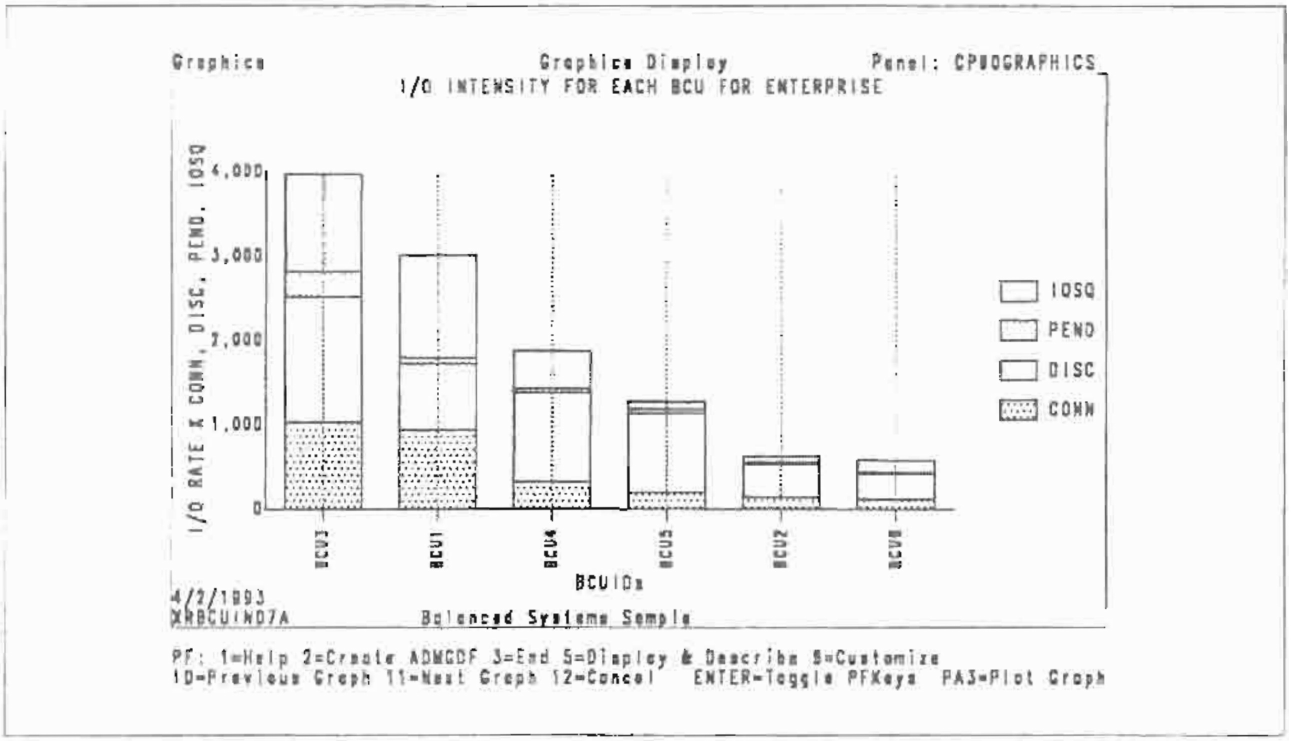
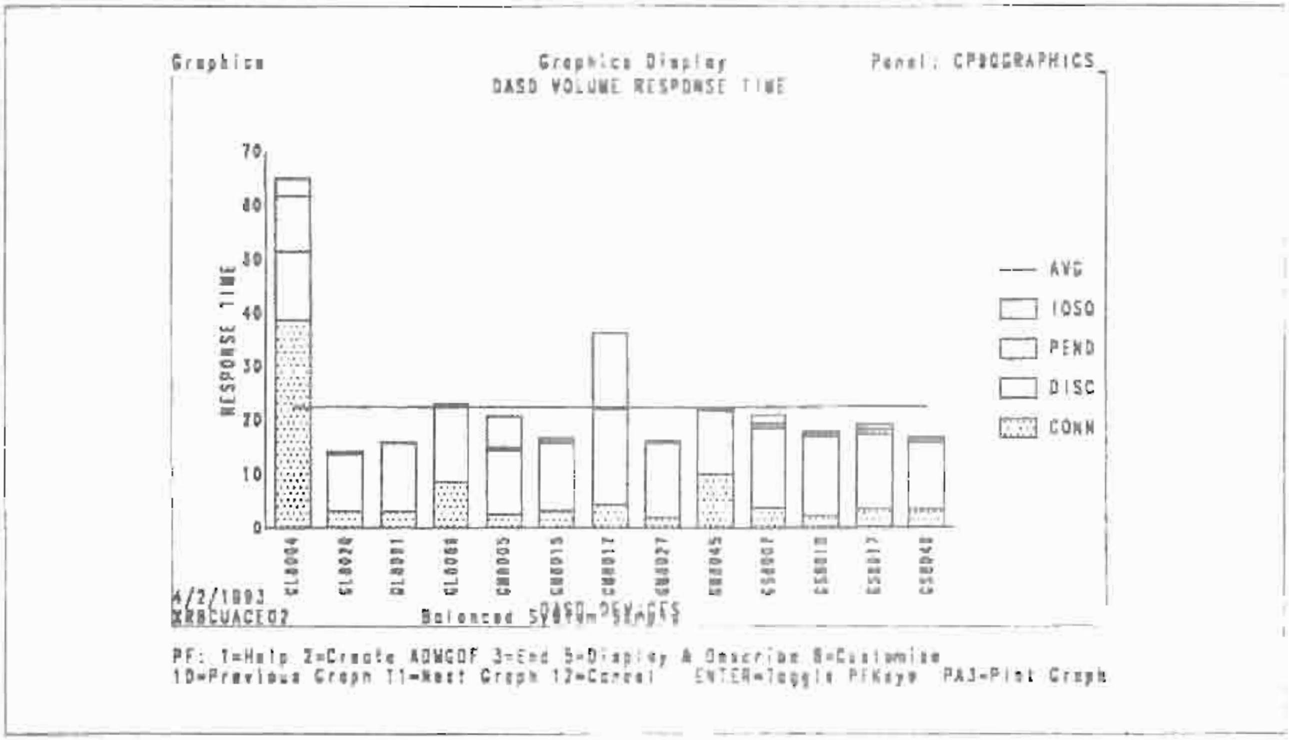


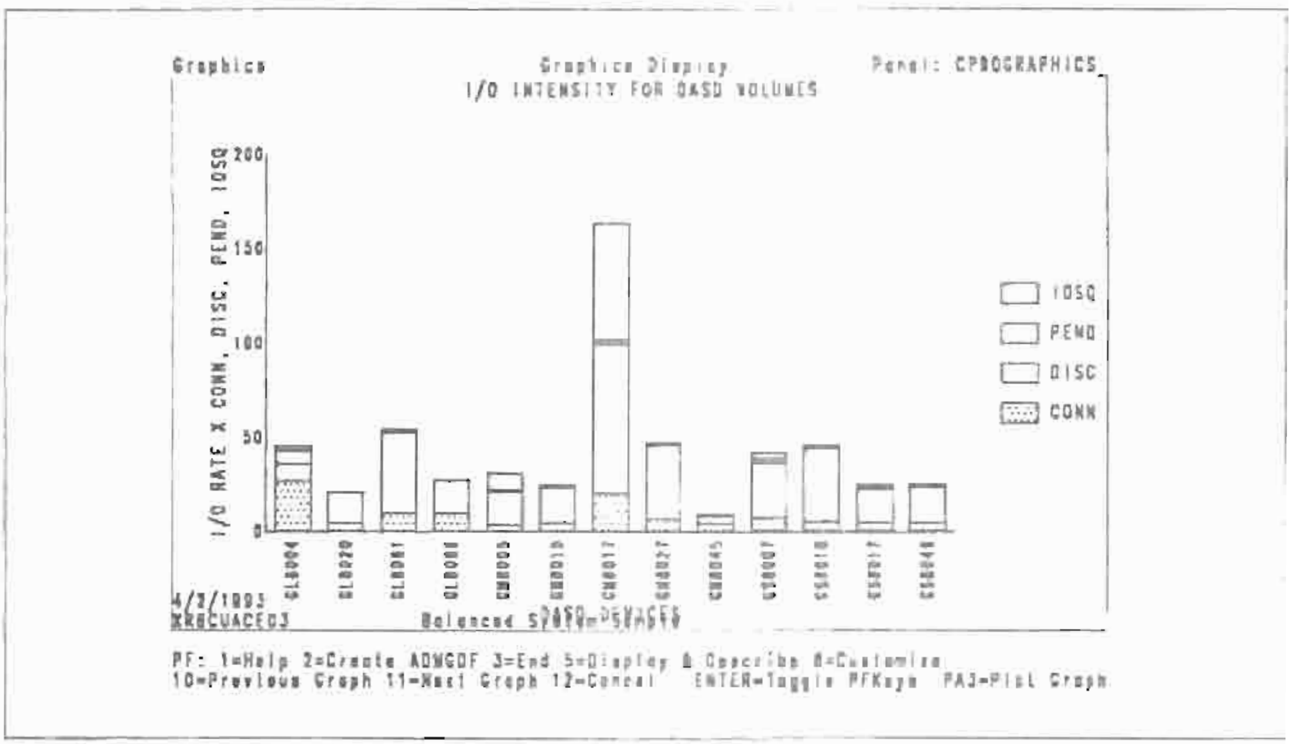
Figure 90. BCU Response Time Components

Of course, in this relatively simple case all we need to do is to see that the average response time is only 17.8 Ms for BCU2 in Figure 84 on page 82, and conclude that BCU2 is not a performance problem. However, we can look at this further for illustration purposes by reviewing the actuator response time components of each BCU, either collectively or individually. This actuator performance data is shown in Figure 85 on page 83. In this example, we will look further into BCU2 by reviewing its actuator data. It is shown that the the high response time is mainly from one actuator GL8004. This is shown in Figure 91 on page 87. Again, we will look at the I/O intensity as shown in Figure 92 on page 87. This picture again shows the volume GL8004 does not do much I/O, and therefore I/O tunning for this volume is probably not necessary. However, if this example turns out to be truly a performance problem, we can do some tuning by concentrating on this volume after we have gone through steps to review DASD data on a BCU level as well as actuator level.

We can go through the same steps shown to determine the channel load problems on BCU1 and BCU3. A performance tool like RMF Monitor III may be used to determine if this is caused by multiple applications. An example of this is shown in Figure 22 on page 23. The data can either be split onto multiple volumes on less busy controllers to solve channel load problem, or performance can be improved by replacing some DASD configuration with improved technology, such as 3990 and 3390s.



| *Figure 91. BCU2 Actuator Response Time Components*



| *Figure 92. BCU2 Actuator I/O Intensity*

The summary of SYSB input is shown in Figure 93 on page 88.

```
Processor Input            System Image Specification           Panel: ACSIM
                         CECID: SYSB         SYSID: SYSB

Enter 'S' in Sel field for each function to be processed and press enter.
Enter or review system data on this panel.

                    +--- Input Data-Function Selection ----+
                    | Sel Data Type             Data Exists |
                    |                                       |
                    |  s  CPU      (Required)        Y      |
                    |  s  Workload (Optional)        Y      |
                    |  s  Storage  (Optional)        Y      |
                    |  s  DASD     (Optional)        Y      |
                    +---------------------------------------+

                                    SCP: MVS/XA
                              CPU Model: 3090-200
                                M-Value: 1180
                              CPU Text : Production System
                    Number of CPs Online: 2
                        CPs Shared (Y/N): N
                        Average CPU Busy:  81.6 %
               Saturation Design Point:  70.0 %
                      Peak:Average Ratio: 1.03
          Historical Peak:Average Ratio: 1.42       Latent Demand: 15.8 %
                         Total DASD I/O:  413.6
                                   RIOC:   .430
                            Paging Rate: 4

Input added or updated.
PF: 1=Help 3=End 4=Comment 5=Select All 10=Previous SYSID 11=Next SYSID
    12=Cancel
```

Figure 93. SYSB SYSID Summary

Processor Projection

Once the basis of our capacity plan, the current state, is established, projection can begin. Projection requires a minimum of a specification of growth over some period. Figure 94 on page 89 indicates that for the sample here, we shall use a horizon of 3 years - 3 month periods and 12 periods.

Growth can be specified in a number of ways. Here we shall use growth by CPU and growth by users. Figure 95 on page 90 shows three scenarios for growth by CPU for the TSO Development workload. The initial problem had a specified growth of 10%. As usual, one should bound this 10% value in some way. "10%" is not an absolute number but one reflecting the business planners estimate. Economic opportunities or problems could mean a more pessimistic or optimistic growth. With this in mind the 10%, and the other specifications were placed in scenario 2 with a lower estimate in scenario 1 and a higher estimate in scenario 3.

Growth by users, in Figure 96 on page 90, is used for the CICS Production workload. The problem stated that more users would be added every 6 months, the users would be getting more sophisticated, and their productivity is increasing. Let's translate that.

Given the number of users, transaction rate, and current CPU load for this workload, each user can be assigned a CPU load. Increasing the number of users simply translates into increased CPU.

Productivity means that the number of transactions per user will increase. That also translates into CPU.

Sophistication means more CPU time for each transaction. As the user learns the application, the complexity of transactions increases. The cost of each transaction goes up. In Figure 96, this is represented as a Complexity Factor (CF) or multiplier on the CPU per transaction.

CF = 1.05 (a 5% increase in CPU cost per period) is shown as an example. Figure 97 on page 91 shows the effect of changes in users and transaction rate per user.

```
Processor Projection        Enterprise Specification             Panel: AFENT
                                      CECID: All

Enter or review Starting Date, Period, and Number of Periods.
Enter or review CEC data.
Enter 'S' in Sel field for each CEC to be processed and press Enter.
Use PF5 (All Workloads) to add, change, and move workloads, and enter Growth.

   Enterprise Name : Balanced Systems Sample

   Workload Projection Parameters:
     Starting Date: 06/91       Period (months): 3        No. of Periods: 12

                     Index, SCP/ Index or          Valid SCPs/         Valid
  Sel   CECID        Supervisor  CPU Model   Index Supervisors  Index CPU Models

        SYSA         MVS/XA      4381-24         1 LPAR              1 9021-720
        SYSB         MVS/XA      3090-200        2 VM/XA             2 9021-620
                                                 3 MVS/ESA           3 9021-580
                                                 4 MVS/XA            4 9021-500
                                                 5 MVS/SP            5 9121-480
                                                 6 VSE               6 9121-440
                                                 7 VM/HPO            7 9021-340
                                                 8 VM/SP             8 9021-330
                                                 9 VSI               9 9121-320
                                                                    10 9121-260
                                                                    11 9121-210
                                                                    12 9121-190

PF: 1=Help 2=Save 3=End 4=Graphics 5=All Workloads 6=Select All 7=Backward
 8=Forward 9=Top/Bottom 11=Erase Projection 12=Cancel
```

Figure 94. Inital Processor Configuration

Processor Projection — CPU Growth - CPU Utilization — Panel: AFCROC

CECID: SYSA   SYSID: SYSA

Workload: TSO Development

Enter or review Growth data.

Start Date (MM/YY): 06/91
Length of Period (Months): 3
Number of Periods: 12
Base CPU Mode : 4381-24
Base SCP: MVS/XA
Base CPU%: 31.5 %

Growth Scenarios   Growth% is Annual Compound Growth.

	1		2		3	
Periods	Growth%	CPU%	Growth%	CPU%	Growth%	CPU%
09/91	8.00	32.15	10.00	32.30	12.00	32.44
12/91	8.00	32.78	10.00	33.08	12.00	33.38
03/92	8.00	33.41	10.00	33.88	12.00	34.34
06/92	8.00	34.06	10.00	34.69	12.00	35.32
09/92	8.00	34.72	10.00	35.53	12.00	36.34
12/92	8.00	35.40	10.00	36.39	12.00	37.38
03/93	8.00	36.09	10.00	37.26	12.00	38.46
06/93	8.00	36.79	10.00	38.10	12.00	39.56
09/93	8.00	37.50	10.00	39.08	12.00	40.70
12/93	8.00	38.23	10.00	40.02	12.00	41.87
03/94	8.00	38.97	10.00	40.99	12.00	43.07
06/94	8.00	39.73	10.00	41.98	12.00	44.31

PF: 1=Help 2=Graphics 3=End 5=Use Periodic Growth 6=Enter CPU% 7=Backward
8=Forward 9=Top/Bottom 10=Propagate Growth 12=Cancel

Figure 95. Growth by CPU

Processor Projection — User Growth — Panel: AFCROU

CECID: SYSB   SYSID: SYSB

Workload: CICS Production

Enter or review User Growth data.

Start Date (MM/YY): 06/91   Scenario 1
Length of Period (Months): 3
Number of Periods: 12

Base Data
Users: 100   CPU/User: .5   CPU Model: 3090-200
Tran Rate: 10   Tran/User: .10
CPU Util: 50.2

Period	# of Users	Tran/ User	Complexity Factor	CPU/ User	CPU/ Util %	Annual CGR %	I/O Rate
09/91	100	.100	1.05	.5	52.7	22	305
12/91	100	.100	1.05	.6	55.3	22	320
03/92	100	.100	1.05	.6	58.1	22	336
06/92	100	.100	1.05	.6	61.0	22	353
09/92	100	.100	1.05	.6	64.1	22	370
12/92	100	.100	1.05	.7	67.3	22	389
03/93	100	.100	1.05	.7	70.6	22	408
06/93	100	.100	1.05	.7	74.2	22	429
09/93	100	.100	1.05	.8	77.9	22	450
12/93	100	.100	1.05	.8	81.8	22	473

PF: 1=Help 2=Graphics 3=End 7=Backward 8=Forward 9=Top/Bottom
10=Previous Scenario 11=Next Scenario 12=Cancel

Figure 96. Growth by Users - Default

```
Processor Projection              User Growth                    Panel: AFGROU
                          CECID: SYSB       SYSID: SYSB
                          Workload: CICS Production

Enter or review User Growth data.

      Start Date (MM/YY): 06/91         Scenario  1
Length of Period (Months): 3
       Number of Periods: 12

      Base Data
  # Users:     100         CPU/User:      .5        CPU Model: 3090-200
 Tran Rate:     10        Tran/User:     .10
  CPU Util:   50.2
```

Period	# of Users	Tran/ User	Complexity Factor	CPU/ User	CPU/ Util %	Annual CGR %	I/O Rate
09/91	100	.100	1.05	.5	52.7	22	305
12/91	100	.100	1.05	.6	55.3	22	320
03/92	110	.100	1.05	.6	63.9	78	370
06/92	110	.100	1.05	.6	67.1	22	388
09/92	120	.110	1.05	.7	84.6	152	485
12/92	120	.110	1.05	.7	88.5	22	51[illegible]
03/93	130	.110	1.05	.8	101.0	87	584
06/93	130	.110	1.05	.8	106.1	22	613
09/93	140	.120	1.05	.9	130.8	132	756
12/93	140	.120	1.05	1.0	137.4	22	794

```
PF: 1=Help 2=Graphics 3=End 7=Backward 8=Forward 9=Top/Bottom
 10=Previous Scenario 11=Next Scenario 12=Cancel
```

Figure 97. Growth by Users Specified

Figure 98 on page 92 summarizes the growth for all the workloads. Batch Development has a constant annual growth of 15, 20, and 25 percent for the three scenarios and the growth of Batch Production is the same as Batch Development. The CICS production is not shown since this is a complex specification by user and transaction.

With the growth specified, our first look into the future is in Figure 99 on page 92. SYSA utilization begins near the saturation design point of 70% (SDP). If we take the SDP seriously, the 4381 should be upgraded in 9/91 according to scenario 1. Depending upon the size of the upgrade, we may have to upgrade again later in the planning period.

The 4381-24 has a power number (M Value) of 381 for MVS/XA. Through some iteration, a 9121-210 whose M value is 491 will replace the 4381 in 9/91. The 9121-260 whose M value is 690 will be installed in 12/92. The *saw tooth* curve for this proposal is shown in Figure 100 on page 93.

Remember that this is only one scenario and the proposal has to be bounded with alternatives. Even for the same hardware upgrades, the upgrade dates may move in and out with changes in the business climate. That's what the different scenarios are about.

```
Processor Projection          Workload Summary                     Panel: AFWKL
                           CECID: All        SYSID: All

Move or delete workloads by changing CECID/SYSID; use PF4 to Add Workloads.
Enter constant '% Annual CPU Growth' directly on this panel.
To enter other workload growth types, select the workloads and press Enter.

   Workload Projection Parameters:
     Starting Date: 06/91       Period (months): 3         No. of Periods: 12

                                                                  % Annual
Sel CECID     SYSID            Workload              M Used   I/O    CPU Growth
                               Description                            1   2   3

 * SYSA       SYSA        TSO Development               120      47    8  10  12
 * SYSA       SYSA        Batch Development             141      55   15  20  25
 * SYSB       SYSB        CICS Production               592     290    *   *   *
 * SYSB       SYSB        Batch Production              344     103    8  10  12




                                           Totals      1197     495

PF: 1=Help 3=End 4=Add Workload 5=Select All 7=Backward 8=Forward 9=Top/Bottom
 12=Cancel
```

Figure 98. Workload Summary

```
Processor Projection       System Image Specification            Panel: AFSIM
                           CECID: SYSA       SYSID: SYSA

Enter or review system data on this panel.
Use PF6 to perform Storage (and Paging) Analysis.
                                                          Start
                                  Model        M-Value    Period   Display
                   CPU Number 1: 4381-24       381        06/91
                   CPU Number 2: 9121-210      491        09/91
                   CPU Number 3: 9121-260      690        12/92

                           SCP: MVS/XA
                      CPU Text: Development System
         Number of CPs Online: 2
             CPs SHARED (Y/N): N                                   Valid
        Latent Demand in CPU%:      0 %                 Index   CPU Models   M Value
 Saturation Design Point(SDP):   70.0 %                     1   9021-900        9400
               DASD I/O Rate:   101.7                       2   9021-860        8320
            RIOC % Reduction: 0                             3   9021-820        7000
                                                            4   9021-740        5380
           Projected % Utilization (4381-24)                5   9021-720        4700
                      Scenarios                             6   9021-660        4030
 Periods        1           2           3                   7   9021-640        3880
   06/91      68.5        68.5        68.5                  8   9021-620        3450
   09/91      70.4        70.9        71.5                  9   9021-580        2750
   12/91      72.4        73.5        74.7                 10   9021-520        2110
   03/92      74.4        76.2        78.0                 11   9021-500        1970
   06/92      76.5        79.0        81.5                 12   9021-340        1060
   09/92      78.7        81.9        85.1                 13   9021-330         955

PF: 1=Help 2=Graphics 3=End 4=Comment 5=All Workloads 6=Storage Analysis
 7=Backward 8=Forward 9=Top/Bottom 10=Previous SYSID 11=Next SYSID 12=Cancel
```

Figure 99. SYSA

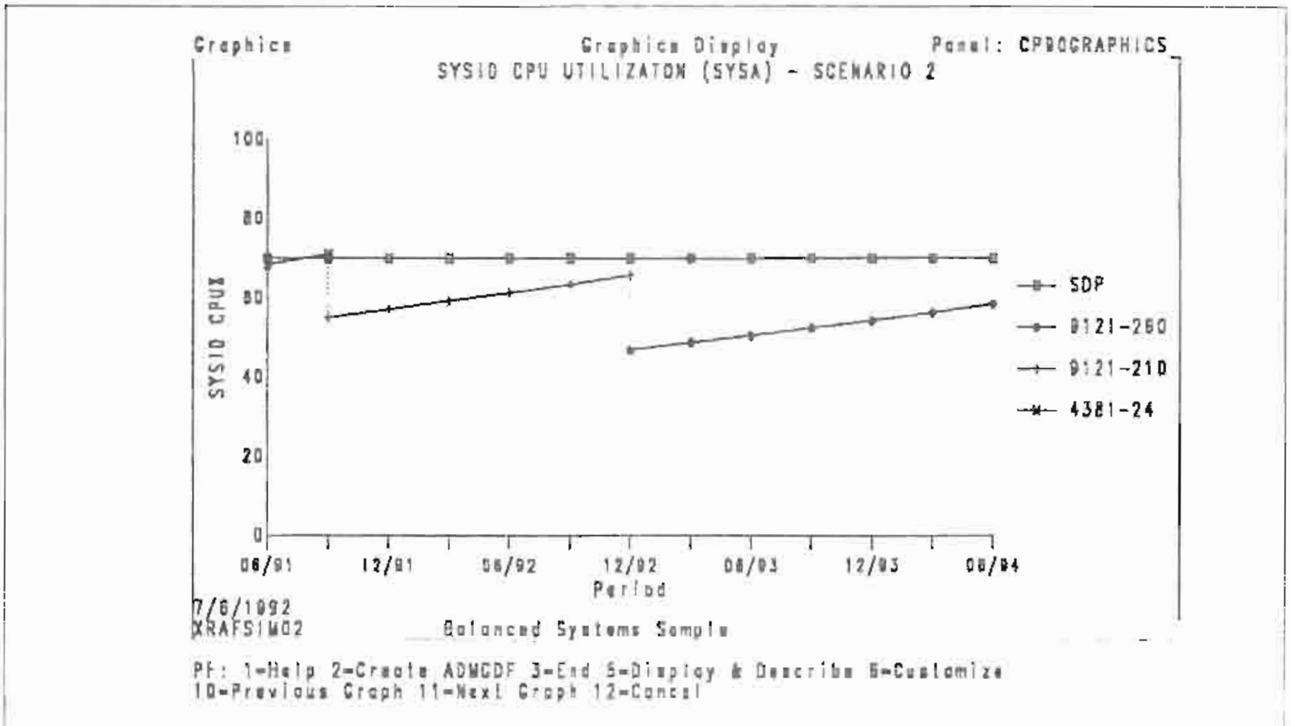


Figure 100. SYSA Projection

The storage projection begins in Figure 101 on page 94 with a reminder that the amount of paging is too great. Before storage projection, we will have to improve storage to drop paging below 8%.

Using the algorithm of the previous chapter (see section "Storage" on page 61), paging can be reduced to 5% of the DASD I/O total with a minimum upgrade of 19 Mb. (See Figure 102 on page 94) To satisfy the algorithm, 19 Mb. is enough even though one cannot order only 19 Mb. The 19 Mb. can be either Central storage or Expanded Storage. The theory does not distinguish.

After the storage has been adjusted to reduce paging, Figure 103 on page 95 indicates that the values of storage used, software level, and power used will not calibrate. This means that for this amount of storage, the calibration point on Figure 67 on page 62 falls outside of the shaded area. Figure 104 on page 95 indicates that the current storage line (64 Mb.) should be between the minimum and maximum line to fit the theory. This does not indicate any error but that this application combination requires more storage than the samples used in the theory. (Remember also that this is a constructed set of data to show these things.) Ahead we shall see that SYSB calibrates.

```
Processor Projection          Storage Analysis                 Panel: AFSTO
                      CECID: SYSA      SYSID: SYSA

Enter or review storage data on this panel.
Use PF9 (Set SSE) to adjust SCP level, if necessary.
Do not use this panel (Storage Analysis) if you have moved workloads to or
  from this SYSID.

                                     SCP: MVS/XA
                               CPU Model: 4381-24
                                CPU Text: Development System

           System Software Environment: SP2.2.0
                             CS Online:    32 MB
                             ES Online:     0 MB
                          CS Available:     1 MB
                          ES Available:     0 MB
   Storage projection calibration factor:
                         Page/Swap I/O rate:    14 Page I/O rate is excessive. It
                              DASD I/O Rate:   100 should be < 8% of DASD I/O
                 Paging Percent of DASD I/O:  14 % rate. Storage must be adjusted
                                                  before storage projection
                                                  calibration. Use PFkey.




PF: 1=Help 2=Graphics 3=End 9=Set SSE 10=Adjust Storage 12=Cancel
```

Figure 101. SYSA Storage

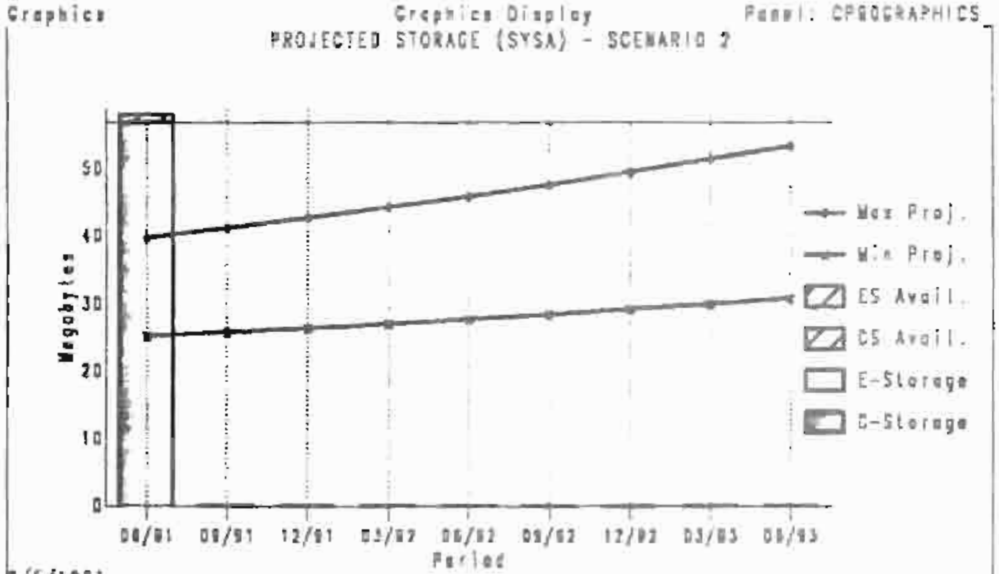
```
Processor Projection         Storage Adjustment                Panel AFSTOC
                      CECID: SYSA      SYSID: SYSA

Enter the amount of Central and Expanded Storage to provide at least the
  indicated increment of Processor Storage.
The indicated amount should reduce the Page/Swap I/O to the new target.

                                     SCP: MVS/XA
                               CPU Model: 4381-24
                                CPU Text: Development System

                           Input     Increment    Total
   Central Storage:       32 Mb       26 Mb        58
  Expanded Storage:        0 Mb        0 Mb         0 Mb
 Processor Storage:       32 Mb       26 Mb        58 Mb

Minimum PS Increment required      26 Mb

                          Input       Target
Page/Swap I/O Rate:        14           5
     DASD I/O Rate:       100          91
   % Page/Swap I/O:        14           5




PF: 1=Help 3=End 12=Cancel
```

Figure 102. SYSA Storage Adjustment

```
Processor Projection          Storage Analysis                 Panel: AFSTO
                     CECID: SYSA       SYSID: SYSA

Enter or review storage data on this panel.
Use PF9 (Set SSE) to adjust SCP level, if necessary.
Do not use this panel (Storage Analysis) if you have moved workloads to or
  from this SYSID.

                                  SCP: MVS/XA
                            CPU Model: 4381-24
                             CPU Text: Development System

            System Software Environment: SP2.2.0
                              CS Online:    58 MB
                              ES Online:     0 MB.
                           CS Available:     1 MB.
                           ES Available:     0 MB.
   Storage projection calibration factor:  1.51
                      Page/Swap I/O rate:    5 Storage projection factor does
                           DASD I/O Rate:   92 not calibrate. Acceptable
                Paging Percent of DASD I/O:   5 % range is (1.28,1.42). Computed
                                               value was  1.51.




PF: 1=Help 2=Graphics 3=End 9=Set SSE 10=Adjust Storage 12=Cancel
```

Figure 103. SYSA Storage Calibration

```
Graphics                    Graphics Display              Panel: CPGGRAPHICS
```




```
8/5/1991
XRAFST002            Balanced Systems Sample

PF: 1=Help 2=Create ADMGDF 3=End 5=Customize 10=Previous Graph 11=Next Graph
12=Cancel
```

Figure 104. SYSA Storage Projection

The initial look at SYSB in Figure 105 on page 96 shows a rather busy situation for the future. Just looking at the upgrade option after some alternatives were looked at, we see a 9021-500 upgrade in 9/91 and another upgrade in 9/92 to a 9021-620. Figure 106 on page 97 shows a graphic summarizing all the scenarios.

The 70% saturation line is clearly overrun for the current 3090-200. The 9021-500 in 9/92 and the 9021-620 in the 9/93 to 6/94 time frame.

Notice that the installation time for the 620 is a function of the scenario. If the aggressive growth scenario 3 is to be followed, the 620 is required in 9/93. If business growth slows as reflected in scenario 1, the 500 is fine for all periods projected. With this graphic, the business can make decisions on capital plan and business plan side by side.

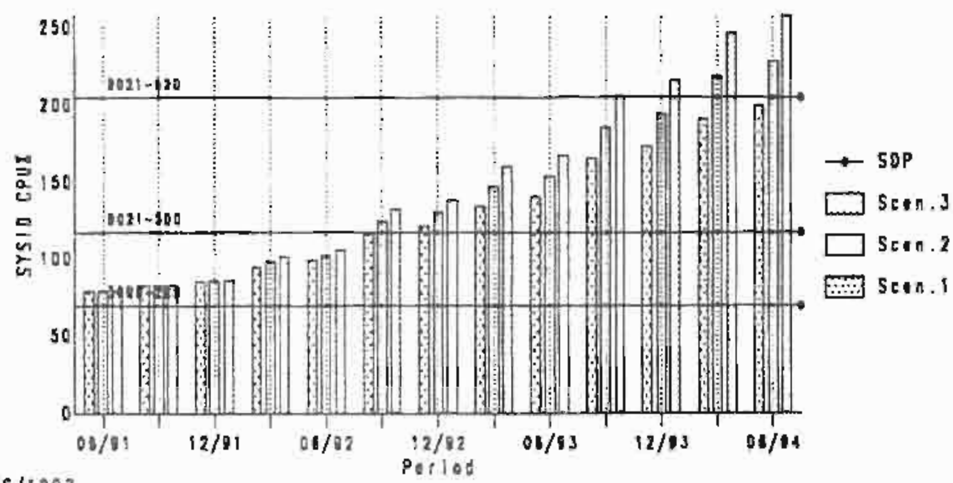
Here is a time when one considers the flexibility of an architecture or processor which enables easy upgrade paths. One could opt for the more conservative scenario 1 or 2 and still be able to upgrade quickly if the business climate improves.

Figure 107 on page 97 shows the storage picture for SYSB. Unlike SYSA, the paging I/O percentage is below 8% and the calibration factor is within range. The graphic showing the projected storage required is in Figure 108 on page 98. The graphic shows the projection for only 2 years. The theory would say that too much changes, for example workload change or movement, to project beyond that horizon.

```
Processor Projection      System Image Specification            Panel: AFSIM
                        CECID: SYSB       SYSID: SYSB

Enter or review system data on this panel.
Use PF6 to perform Storage (and Paging) Analysis.
                                                         Start
                                   Model       M-Value   Period   Display
                  CPU Number 1: 3090-200      1180      06/91
                  CPU Number 2: 9021-500      1970      09/91
                  CPU Number 3: 9021-620      3450      09/92

                           SCP: MVS/XA
                      CPU Text: Production System
          Number of CPs Online: 2
              CPs SHARED (Y/N): N                                 Valid
          Latent Demand in CPU%:        0 %             Index   CPU Models    M Value
   Saturation Design Point(SDP):     70.0 %                 1   9021-900        9400
                 DASD I/O Rate:    393.5                    2   9021-860        8320
             RIOC % Reduction: 0                           3   9021-820        7000
                                                            4   9021-740        5360
          Projected % Utilization (3090-200)                5   9021-720        4700
                      Scenarios                             6   9021-660        4030
   Periods       1           2           3                  7   9021-640        3880
     06/91     79.4        79.4        79.4                 8   9021-620        3450
     09/91     82.4        82.6        82.7                 9   9021-580        2750
     12/91     85.7        85.9        86.2                10   9021-520        2110
     03/92     94.8        98.2       101.5                11   9021-500        1970
     06/92     98.6       102.3       105.9                12   9021-340        1060
     09/92    116.7       124.5       132.3                13   9021-330         955

PF: 1=Help 2=Graphics 3=End 4=Comment 5=All Workloads 6=Storage Analysis
 7=Backward 8=Forward 9=Top/Bottom 10=Previous SYSID 11=Next SYSID 12=Cancel
```

Figure 105. SYSB

```
Graphics                    Graphics Display          Panel: CPBOGRAPHICS
              SYSID (SYSB) PROJECTED CPU% (3090-200) - SUMMARY
```




```
7/6/1992
XRAFSIM07             Balanced Systems_Sample

PF: 1=Help 2=Create ADMGDF 3=End 5=Display & Describe 6=Customize
10=Previous Graph 11=Next Graph 12=Cancel
```

Figure 106. SYSB Projection

```
Processor Projection            Storage Analysis                 Panel: AFSTO
                         CECID: SYSB       SYSID: SYSB

Enter or review storage data on this panel.
Use PF9 (Set SSE) to adjust SCP level, if necessary.
Do not use this panel (Storage Analysis) if you have moved workloads to or
  from this SYSID.

                                  SCP: MVS/XA
                            CPU Model: 3090-200
                             CPU Text: Production System

              System Software Environment: SP2.2.0
                                CS Online:    64 MB.
                                ES Online:    64 MB.
                             CS Available:     2 MB.
                             ES Available:    10 MB.
    Storage projection calibration factor:   1.35
                        Page/Swap I/O rate:      4
                             DASD I/O Rate:    395
             Paging Percent of DASD I/O:       1 %

Storage calibration complete. Graphics indicate future projections.
PF: 1=Help 2=Graphics 3=End 9=Set SSE 10=Adjust Storage 12=Cancel
```

Figure 107. SYSB Storage

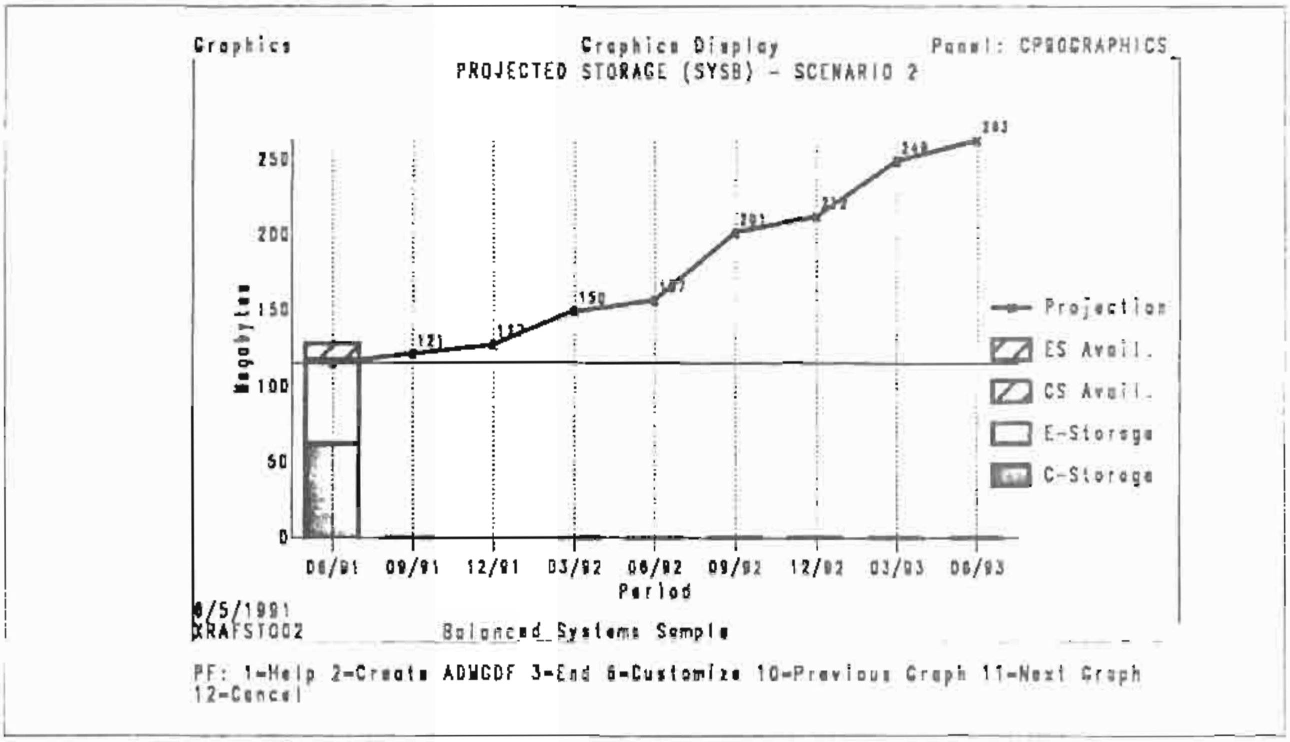


Figure 108. SYSB Storage

As an alternative proposal, let's consider combining the 2 CECs onto one CEC using LPAR on a 9021-580. The steps begin in Figure 109 on page 99 with the definition of a new CEC. At this point the CEC is *empty*. No workloads are defined or partitions made available. Figure 110 on page 99 illustrates the proposed partition structure. The 9021-580 has 3 CPs. Both partitions will be shared with the number of CPs indicated. There are 2 M-values for each partition. The Maximum is the potential power deliverable if the partition got all it wanted. The Weight value reflects what happens at 100% CEC utilization under the weight distribution mechanism.

The workloads will be moved to the new system by system image. The new SYSAX will receive SYSA and SYSBX will get SYSB. This movement is accomplished by setting the new target CECID (NEWSYS) and SYSID (SYSAX or SYSBX). For LPAR, there are distinguishing identifiers for the CEC and System images running in LPAR. (See Figure 30 on page 31.) Figure 112 on page 100 and Figure 113 on page 101 show the appropriate system image summaries for each partition. The graphic is Figure 114 on page 101 shows the partition summary from the partition point of view. The graphic appears as a non LPAR display for the most part. 100% is the potential maximum value. The saturation point of 70% is also shown. The only difference is the presence of a Weight line to indicate the impact of reaching 100%.

Keep in mind that the view from SYSAX is the view that it is running on a single CP configuration. It is a picture relative to the LPAR configuration. A change of CPs to Share or Dedicated will affect the power delivered and the utilization levels illustrated. To view the CEC utilization with the absolute SYSID contribution, look at Figure 116 on page 102.

```
Processor Projection        Enterprise Specification          Panel: AFENT
                                 CECID: All

Enter or review Starting Date, Period, and Number of Periods.
Enter or review CEC data.
Enter 'S' in Sel field for each CEC to be processed and press Enter.
Use PF5 (All Workloads) to add, change, and move workloads, and enter Growth.

   Enterprise Name : Balanced Systems Sample

   Workload Projection Parameters:
     Starting Date: 06/91      Period (months): 3       No. of Periods: 12

                 Index, SCP/  Index or          Valid SCPs/        Valid
   Sel  CECID    Supervisor   CPU Model   Index Supervisors  Index CPU Models

        SYSA     MVS/XA       4381-24        1 LPAR             1 9021-720
        SYSB     MVS/XA       3090-200       2 VM/XA            2 9021-820
        NEWSYS   LPAR         9021-580       3 MVS/ESA          3 9021-580
                                             4 MVS/XA           4 9021-500
                                             5 MVS/SP           5 9121-480
                                             6 VSE              6 9121-440
                                             7 VM/HPO           7 9021-340
                                             8 VM/SP            8 9021-330
                                             9 VS1              9 9121-320
                                                               10 9121-260
                                                               11 9121-210
                                                               12 9121-190

PF: 1=Help 2=Save 3=End 4=Graphics 5=All Workloads 6=Select All 7=Backward
 8=Forward 9=Top/Bottom 11=Erase Projection 12=Cancel
```

Figure 109. Add A New CEC

```
Processor Projection             CEC Specification            Panel: AFCEC
                                   CECID: NEWSYS

Enter or review SYSID (partition) data.
Enter 'S' in Sel field for each SYSID to be processed and press Enter.


                                                              CEC M-Value:   2855
   CPU Model: 9021-580      Supervisor: LPAR              Computed CEC SDP:   70 %
 # CPs Online: 3                 # SCPs: 2                    Used CEC SDP:   70 %

                   Index or      #  Share        Maximum Planning
 Sel SYSID           SCP       CPs  Y/N Weight  M-Val    M-Val      Index  Valid SCPs

   SYSAX         MVS/ESA        1    Y    20     1021     571          1  VM/ESA
   SYSBX         MVS/ESA        3    Y    80     2785    2284          2  VM/XA
                                                                       3  MVS/ESA
                                                                       4  MVS/XA
                                                                       5  MVS/SP
                                                                       6  VSE/ESA
                                                                       7  VSE/ESA/VR
                                                                       8  VSE
                                                                       9  VM/HPO
                                                                      10  VM/SP

PF: 1=Help 2=Graphics 3=End 4=Save 5=All Workloads 6=Select All 7=Backward
 8=Forward 9=Comment 10=Previous CEC 11=Next CEC 12=Cancel
```

Figure 110. CEC Definition

```
Processor Projection         Workload Summary                     Panel: AFWKL
                        CECID: All        SYSID: All

Move or delete workloads by changing CECID/SYSID; use PF4 to Add Workloads.
Enter constant '% Annual CPU Growth' directly on this panel.
To enter other workload growth types, select the workloads and press Enter.

   Workload Projection Parameters:
     Starting Date: 06/91      Period (months): 3        No. of Periods: 12

                                                                  % Annual
Sel CECID      SYSID          Workload              M Used   I/O    CPU Growth
                              Description                           1   2   3

   NEWSYS    SYSAX    TSO Development                 123      47   8  10  12
   NEWSYS    SYSAX    Batch Development               144      55  15  20  25
   NEWSYS    SYSBX    CICS Production                 625     290   *   *   *
   NEWSYS    SYSBX    Batch Production                383     103   8  10  12




                                        Totals       1255     495

PF: 1=Help 3=End 4=Add Workload 6=Select All 7=Backward 8=Forward 9=Top/Bottom
 12=Cancel
```

Figure 111. Workload Movement

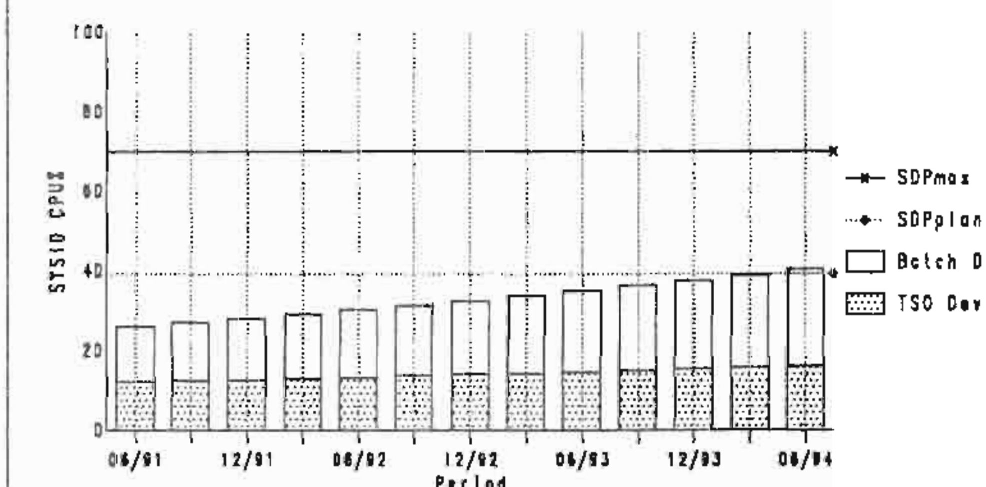
```
Processor Projection    System Image Specification (LPAR)        Panel: AFSIML
                         CECID: NEWSYS     SYSID: SYSAX

Enter or review partition data on this panel.
Use PF5 to perform Storage (and Paging) Analysis.


                                SCP: MVS/ESA
                          CPU Model: 9021-580
                            M-Value: 1013
                       Start Period: 06/91
                           CPU Text: Development
             Number of CPs Online: 1
                   CPs SHARED (Y/N): Y
              Latent Demand in CPU%:     0 %
       Saturation Design Point(SDP):  70.0 %
                      DASD I/O Rate:  101.7
                    RIOC % Reduction: 0

                Projected % Utilization (9021-580)
                         Scenarios
      Periods      1            2           3
        06/91    26.4         26.4        26.4
        09/91    27.1         27.3        27.5
        12/91    27.9         28.3        28.7
        03/92    28.7         29.3        30.0
        06/92    29.5         30.4        31.4
        09/92    30.3         31.5        32.8

PF: 1=Help 2=Graphics 3=End 4=Comment 6=Storage Analysis 7=Backward 8=Forward
 9=Top/Bottom 10=Previous SYSID 11=Next SYSID 12=Cancel
```

Figure 112. New SYSA

```
Processor Projection    System Image Specification (LPAR)       Panel: AFSIML
                        CECID: NEWSYS    SYSID: SYSBX

Enter or review partition data on this panel.
Use PF6 to perform Storage (and Paging) Analysis.

                          SCP: MVS/ESA
                    CPU Model: 9021-580
                      M-Value: 2763
                 Start Period: 06/91
                     CPU Text: Production
         Number of CPs Online: 3
            CPs SHARED (Y/N): Y
        Latent Demand in CPU%:       0 %
   Saturation Design Point(SDP):  70.0 %
                DASD I/O Rate:  383.5
             RIOC % Reduction: 0

              Projected % Utilization (9021-580)
                          Scenarios
   Periods      1           2           3
    06/91      35.8        35.8        35.8
    09/91      37.1        37.2        37.3
    12/91      38.8        38.7        38.8
    03/92      42.7        44.2        45.7
    06/92      44.4        46.1        47.7
    09/92      52.6        56.1        59.6

PF: 1=Help 2=Graphics 3=End 4=Comment 6=Storage Analysis 7=Backward 8=Forward
 9=Top/Bottom 10=Previous SYSID 11=Next SYSID 12=Cancel
```

Figure 113. New SYSB

```
Graphics                    Graphics Display              Panel: CP9OGRAPHICS
            SYSID (SYSAX) PROJECTED CPU% (LPAR 9021-580) - SCENARIO 2
```




```
7/6/1992
XRAFSIM05            Balanced Systems Sample

PF: 1=Help 2=Create ADMCDF 3=End 5=Display & Describe 6=Customize
10=Previous Graph 11=Next Graph 12=Cancel
```

Figure 114. SYSAX Graphic

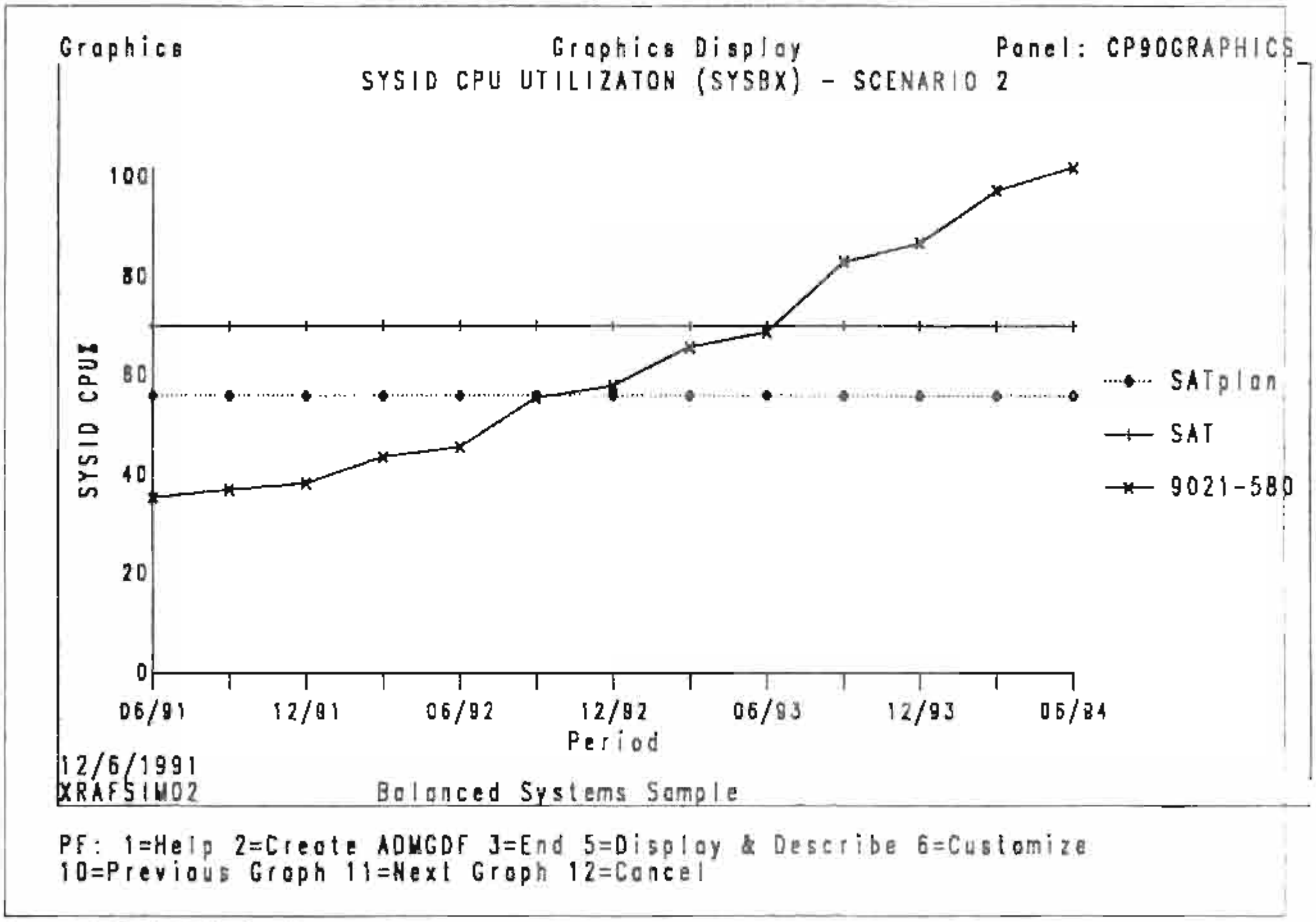


Figure 115. SYSAY Graphic

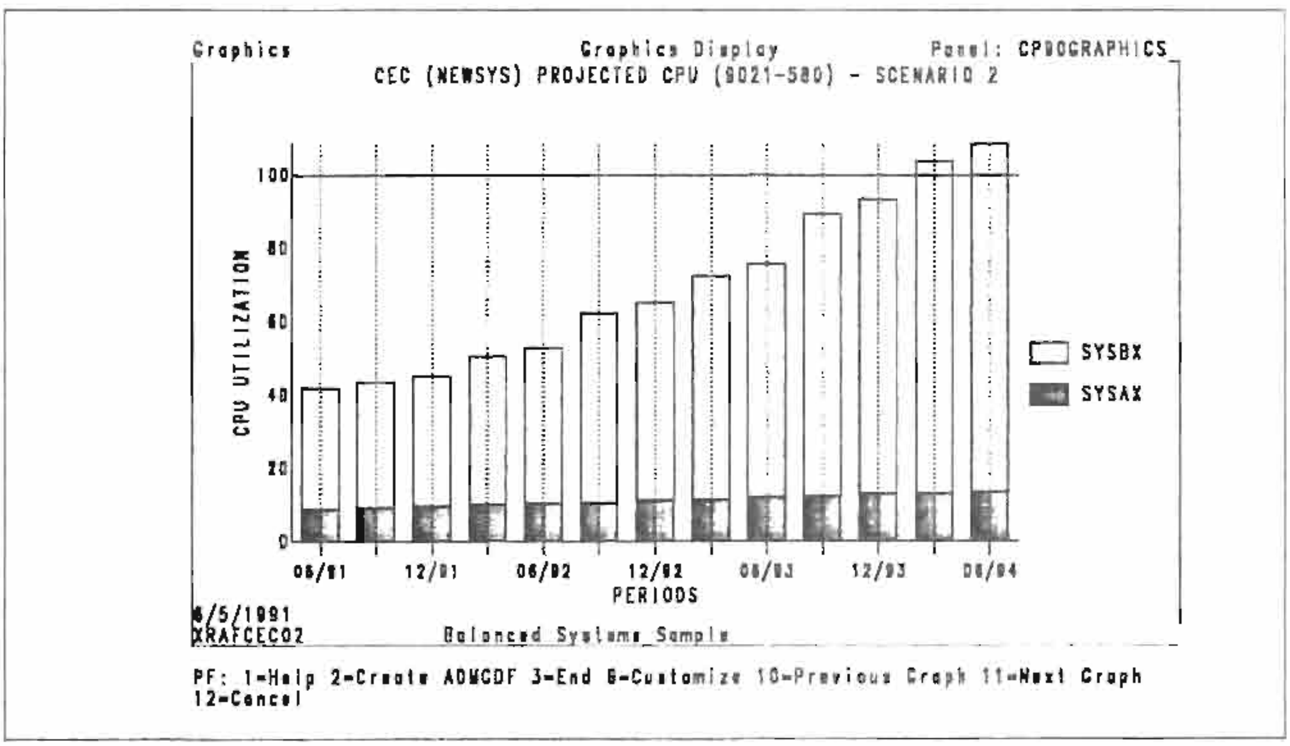


Figure 116. NEW System

DASD Projection

The input BCU data combined from all the SYSIDs is shown in Figure 117 on page 104. The Service Levels are from the input data illustrated earlier. The service level for BCU2 has been changed from 83 Ms (maximum response time on input) to a more reasonable 40 Ms. Those BCUs that were shared, are consolidated. Let's assume that we need improvement in the I/O subsystem in response time while planning for future I/O requirements. Currently, the performance curves indicate that a maximum response time of 24 Ms for a 3880EC BCU can be attained when the I/O rate is 166 or less.

Remember that the projection of 166 comes from all the assumptions mentioned earlier in the concepts chapter. To examine the capacity of the existing configuration at an improved SLO, Figure 118 on page 104 shows the I/O rate expected at the improved SLO. Expecting faster from the same equipment means expecting less. Before the expected total was 695. After improving the SLOs, the expected total is 583.

If you compare the maximum I/O Rate at the service level in Figure 118 with the Projected I/O rate, you will see that there isn't too much room for growth anywhere and especially for BCU3 and BCU2.[19] This is illustrated in Figure 119 on page 105. The different levels in the bar charts indicate the increasing I/O rate generated by processor growth. As the CPU requirement increases, the I/O requirement increases in the RIOC ratio.

To satisfy the I/O rate requirement and response time SLOs, the following changes were made: BCU1 was changed to a 39902C (3990-3 with 3390 mode 2), at performance level 2, and an aggressive SLO of 15 Ms **maximum**. Figure 121 on page 106 shows the result of the change. The new BCU1 has a maximum rate of 430, well above the immediate requirement of 149. Remember, the number 430 does not mean that 430 is the maximum capacity of that BCU, but the expected capacity with the performance curve assumptions at that SLO.

Figure 121 on page 106 shows a dramatic change in scale. In fact, all the BCU bars are below the maximum rate (the line).

Figure 122 on page 106 shows the resulting environmental changes resulting from replacing a 3880A BCU (3880-3, 12 x 3380A) with a 39902C BCU with 24 actuators, a net increase of 12 actuators. The gigabytes increased by about 40%, power up a couple percent, and floor space decline a couple percent.

[19] The projected I/O rate is obtained by taking the current total I/O rate of 405 and dispersing it by means of the I/O distribution shown. Remember that the distribution is an algebraically calculated one which is a function of the number of BCUs.

```
DASD Projection          Performance Projection Summary          Panel: AFBCU3
                        CECID: All        SYSID: All

Caution: See help panel for unique functions of this editor (PF6).

+------------------------ Performance Projection Data ------------------------+
| Start Date: 06/91             Period displayed: 06/91  Scenario displayed: 1 |
|          BCU    Storage    DASD   Perf  Service    BCU Max  I/O  Projected   |
|  BCUID   Type   Director   Type   Level Level      I/O Rate Dist I/O Rate Path|
| BCU3     3880EC 3880-23    3380E   1      24          166  30%      149   2  |
| BCU4     3880E  3880-3     3380E   0      35          103  18%       89   2  |
| BCU5     3880E  3880-3     3380E   0      37          105  15%       74   2  |
| BCU6     3880E  3880-3     3380E   0      38          107  13%       64   2  |
| BCU2     3880D  3880-3     3380D   0      40          112  12%       59   2  |
| BCU1     3880A  3880-3     3380A   0      42          102  11%       54   2  |
|                                                                              |
+------------------------------------------------------------------------------+
Restore "Input" Configuration: N          Totals:     695  100%      489   12
Allow I/O Distribution Update: N                       BCUs:     6 Acts:   92
Press PF5 to continue
PF: 1=Help 2=Graphics 3=End 4=Pop Window 5=Continue 6=Edit 7=Backward 8=Forward
 9=Top/Bottom 10=Environment 11=Sort by Max. I/O Rate 12=Cancel
```

Figure 117. Original BCU Configuration

```
DASD Projection          Performance Projection Summary          Panel: AFBCU3
                        CECID: All        SYSID: All

Caution: See help panel for unique functions of this editor (PF6).

+------------------------ Performance Projection Data ------------------------+
| Start Date: 06/91             Period displayed: 06/91  Scenario displayed: 1 |
|          BCU    Storage    DASD   Perf  Service    BCU Max  I/O  Projected   |
|  BCUID   Type   Director   Type   Level Level      I/O Rate Dist I/O Rate Path|
| BCU3     3880EC 3880-23    3380E   1      20          152  30%      149   2  |
| BCU2     3880D  3880-3     3380D   0      30           91  18%       89   2  |
| BCU4     3880E  3880-3     3380E   0      30           87  15%       74   2  |
| BCU5     3880E  3880-3     3380E   0      30           87  13%       64   2  |
| BCU6     3880E  3880-3     3380E   0      30           87  12%       59   2  |
| BCU1     3880A  3880-3     3380A   0      30           79  11%       54   2  |
|                                                                              |
+------------------------------------------------------------------------------+
Restore "Input" Configuration: N          Totals:     583  100%      489   12
Allow I/O Distribution Update: N                       BCUs:     6 Acts:   92
Press PF5 to continue
PF: 1=Help 2=Graphics 3=End 4=Pop Window 5=Continue 6=Edit 7=Backward 8=Forward
 9=Top/Bottom 10=Environment 11=Sort by Max. I/O Rate 12=Cancel
```

Figure 118. Change SLOs

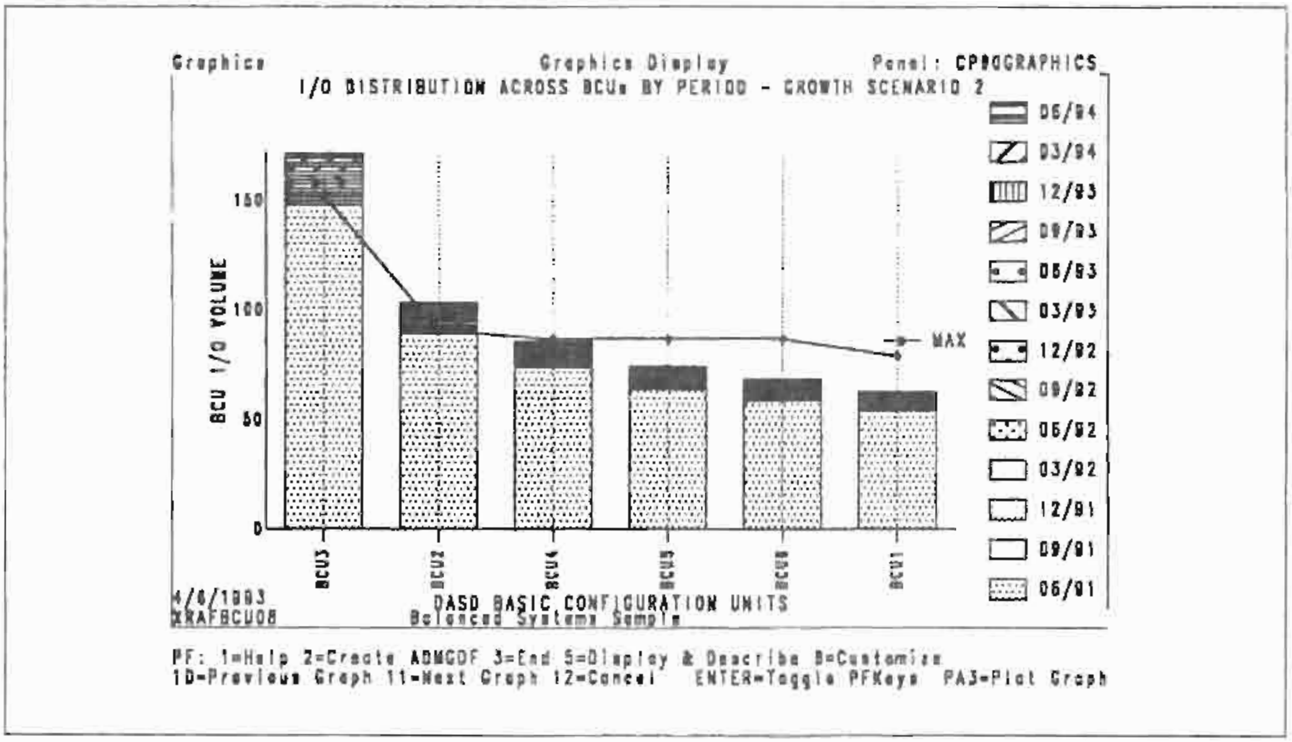


Figure 119. Projection

```
DASD Projection          Performance Projection Summary          Panel: AF8CU3
                         CECID: All         SYSID: All

Caution: See help panel for unique functions of this editor [PF6].

+------------------------ Performance Projection Data -------------------------+
| Start Date: 06/91               Period displayed: 06/91  Scenario displayed: 1
|           BCU    Storage   DASD   Perf  Service   BCU Max   I/O  Projected
|  BCUID    Type   Director  Type   Level Level     I/O Rate  Dist I/O Rate Path
|  BCU1    39902C  3990-3    3390-2  2     15          430    30%       149   4
|  BCU3    3880EC  3880-23   3380E   1     20          152    18%        89   2
|  BCU2    3880D   3880-3    3380D   0     30           91    15%        74   2
|  BCU4    3880E   3880-3    3380E   0     30           87    13%        64   2
|  BCU5    3880E   3880-3    3380E   0     30           87    12%        59   2
|  BCU6    3880E   3880-3    3380E   0     30           87    11%        54   2
|
+------------------------------------------------------------------------------+
Restore "Input" Configuration: N           Totals:     934  100%       489   14
Allow I/O Distribution Update: N                           BCUs:    6 Acts:  104
Press PF5 to continue
PF: 1=Help 2=Graphics 3=End 4=Pop Window 5=Continue 6=Edit 7=Backward 8=Forward
 9=Top/Bottom 10=Environment 11=Sort by Max. I/O Rate 12=Cancel
```

Figure 120. New BCU Configuration

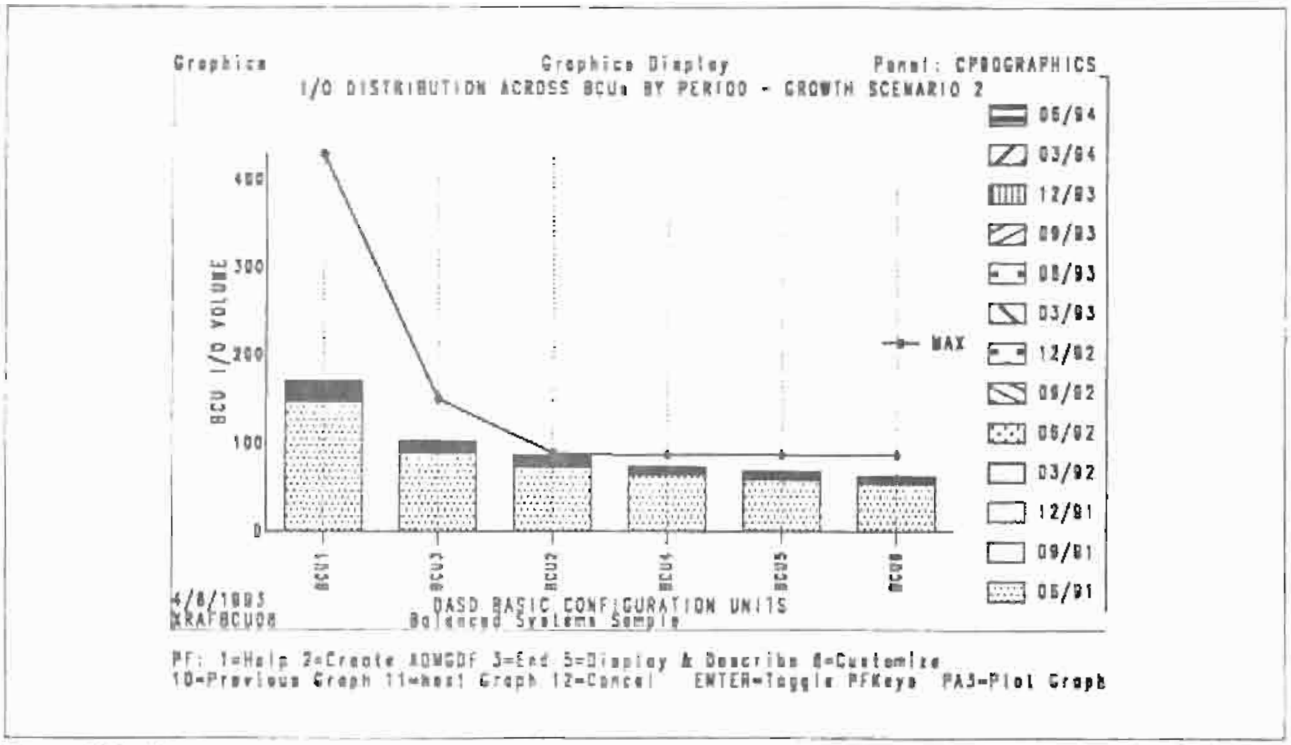


Figure 121. Projection

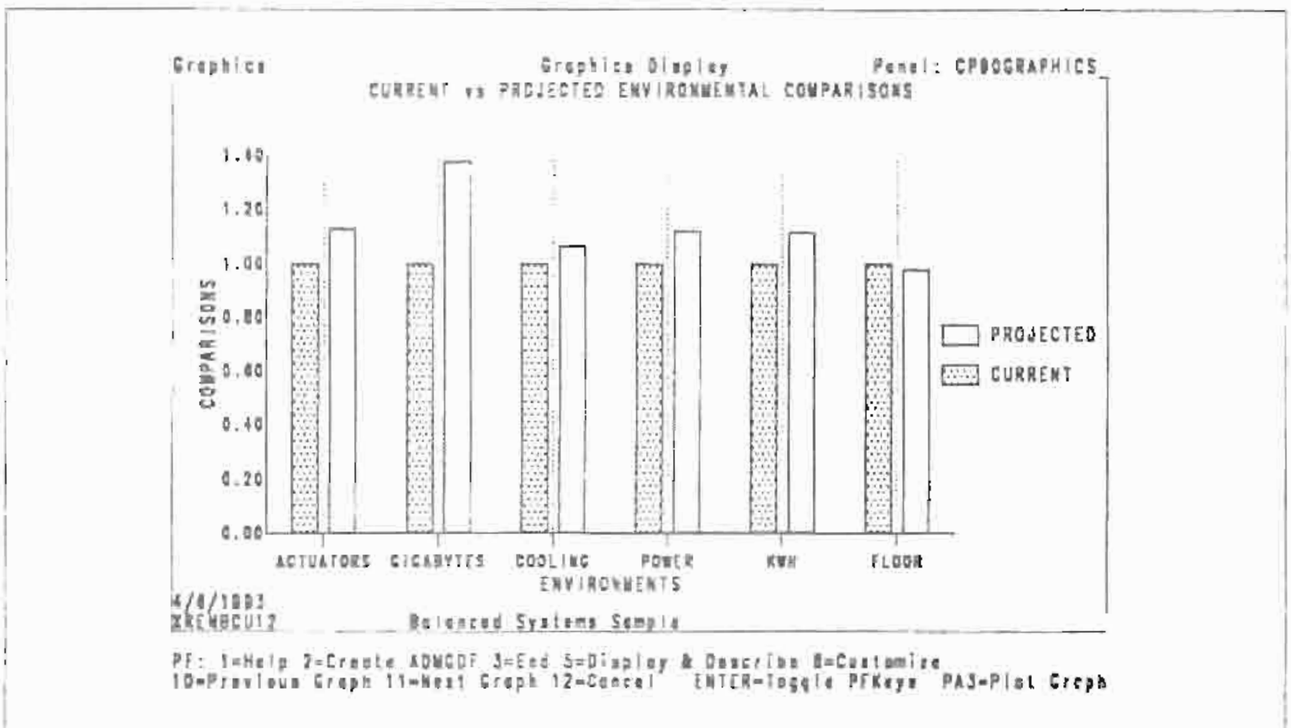


Figure 122. Environmentals

Appendix A. RIOC Conversation

Joe Major who has been working with RIOC values for years, gave me his impression of RIOC values

During the last two years the following pattern seems to have emerged: The greater the capacity of a processor, the lower is the RIOC of its workload mix. The 3090 line shows 0.35 in the 200E to 400E range, 0.30 above that and 0.40 for the 180's. Below that it is 0.45 to 0.50. Low RIOC components are the 4th generation language loads (0.1 to 0.2), CICS/IDMS loads and workloads that capitalize on data in memory capabilities. TSO is 0.25 to 0.30 when there is no page/swap to AUX storage. IMS full function is 0.30 to 0.40 The smaller systems that I see tend to have small memories and 5 - 10% of their I/O's are page/swap I/O's. These systems have not as yet discovered the advantages of DIM. RIOC reduces when CICS/MRO is implemented. It reduces as well when RACF or similar products are implemented and used at low level of access control. The annual reduction is 10 - 15%. MVS/ESA Data in Memory may accelerate that rate of reduction. IMS Fast Path in banking shows 0.55

What is Relative I/O Content?

Relative I/O Content is the ratio of the total number of I/Os to the total power consumed for the same period of time. It is a way of describing a workload in terms of I/O content, where a higher number indicates a workload of higher I/O content. This is shown by turning the I/O-CPU equation around to calculate the RIOC:

$$RIOC = \frac{S}{M * B}$$

Relative I/O Content varies by software, so typical measured systems show different values for RIOC. Thus, TSO will be different from CICS, CICS/VSAM from CICS DL/1, etc.

Some observations concerning RIOC[20] are:

- If the workload is always unchanged, then the RIOC will remain constant.
- If the workload is changed, then the RIOC will change.
- Since 1980, RIOC has been observed to be steadily decreasing year after year (at a rate of 10 - 15%) as software and applications have exploited larger central and expanded storage.

The assumption is made that the RIOC of any given workload will remain constant over time. As the workload grows in the future, the processing power required is projected based on the I/O-CPU equation. Since the RIOC is constant, the power required (M) will be determined by the increase in the processor busy percentage (B) and the I/O rate (S). However, the RIOC remains

[20] For further discussion of RIOC, you should refer to the *IBM Systems Journal, Vol. 20, No. 1*, "Processor, I/O Path and DASD Configuration Capacity."

constant only if the software and application design are unchanged. Some factors which may change, and therefore affect RIOC values, are the system control program, microcode, and application characteristics. For example, when a workload is transferred from one processor to another, the *relative* power of the processors must be taken into consideration. This means that as the relative processor power (M) increases, the utilization (B) will decrease.

It is important to review the RIOC of both the current and future workloads because the DASD I/O content of the future workload is a key factor in deciding both processor and DASD requirements. If the workload you are studying is one with which you are not familiar, you may not be able to assess the validity of the RIOC. In this case, you should accept the system RIOC.

Table 9 which is shown below, shows some typical RIOCs based on impressions of experienced capacity planners in 1990. Table 10 shows a reduction factor for every year after 1990. These reduction factors can be applied to the workload RIOC for 1990 to calculate the RIOC for any year up to the year 2000. For example, you can calculate the RIOC for TSO in 1995 by multiplying .3 by .50.

Table 9. Some RIOCs, Vintage 1990

Workload Type	Likely RIOC or Range
Commercial Batch	0.3 - 0.6
Commercial MVS Batch	0.29
Commercial OS/VS1 Batch	0.49
Commercial DOS/VSE Batch	0.75
Engineering Batch	0.12
TSO	0.2 - 0.4
Prime Shift, Application Development	0.20
Off-shift, Systems Programming	0.38
IMS CICS	0.2 - 0.45
with DL/1	0.30
Fast Path/High Performance	0.46
IDMS	0.17
VM	0.3 - 0.6
General Workload Mix	0.35
CMS Time Sharing	0.58
CMS, Minidisk cache	0.25
VM Guest Operating System	0.1 - 0.4
MVS Production Workload	0.13
MVS, Preferred	0.30
DOS/VSE Production Workload	0.35
Fourth Generation Language	0.15 - .25
PROFS	0.8
DISSOS	1.6

Table 10. RIOC Reduction Factors

Year	Reduction Factor
1990	1.00
1991	0.87
1992	0.76
1993	0.66
1994	0.57
1995	0.50
1996	0.43
1997	0.38
1998	0.33
1999	0.29
2000	0.25

Please keep in mind when using these tables that the RIOCs were *not* based on a systematic analysis of data. Rather, they were either

- based on impressions of typical systems in 1989 and 1990, or
- derived from RIOC measurements taken in 1980 and 1985 (with reduction factors applied).

When preparing an accurate capacity plan, you must adjust the Relative I/O Content to ensure that the variance between the workload total and the SYSID total is as small as possible.

Appendix B. A Word on Distributions

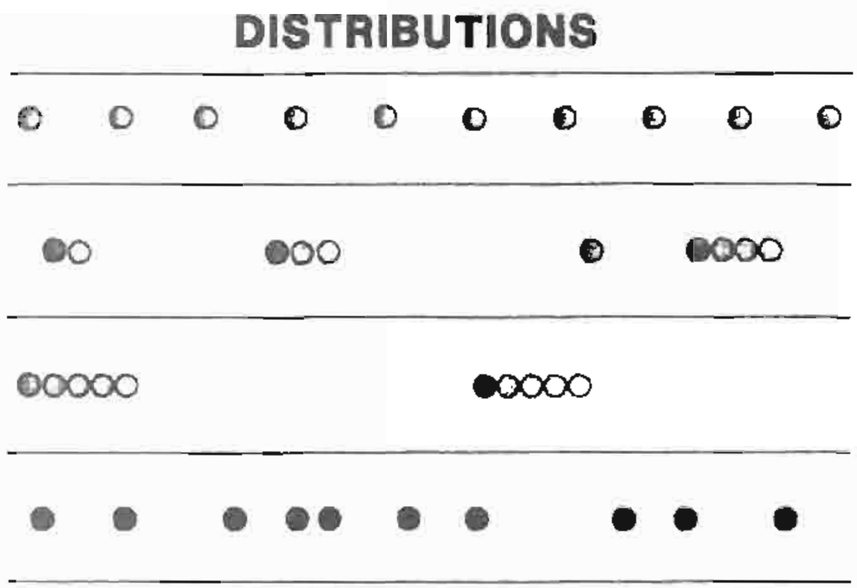


Figure 123. Various distributions.

Figure 123 shows a variety of distributions of events. If the interval was the same for all, the average time between events would be the same even though they arrive in different patterns. The effect of the pattern of arrival is significant in that our analysis is interested in service time, queueing time, and server utilization.

The top distribution is the most well behaved. The time between events is constant. If in this well behaved situation, the service of the events was just less than the time between arrivals, the server would be nearly 100% utilized, AND no one would wait! If the formula in Figure 7 on page 9 was used to estimate queueing time, the formula would greatly exaggerate waiting. But for this distribution service time is high, server utilization is 100%, and there's no queueing.

The next distribution I call the "Cafeteria" distribution. It represents the observed behavior of people getting rid of their trays (service time). People eat in groups and leave in groups. Their arrival for service is clustered. How about queueing time? The first in the group doesn't wait, while the rest wait for the one server to finish with the person in front. Here service time is short, queueing is proportionately high, and server utilization is low.

The next distribution (third from the top) is what I call the "London Bus" distribution. If you have been in London you know it's true. There, buses arrive in an extreme form of the cafeteria distribution.

The last distribution is one similar to the assumed distribution in most queueing theory formula. Although the average is the same as the distributions above, the distance between events cannot be predicted very well. This might be an illustration of the Poisson arrival or exponential distribution. Knowing what just happen does help with the next event. The effects of this distribution can be predicted knowing the average service time and arrival rate as seen earlier in the body of this publication.

You can begin to see that analytic assumptions say something about the universe of discussion. Fortunately for us, a lot of the computer system behavior is Poisson. Are there exceptions? Yes. You are editing a dataset using ISPF. The entire dataset is read, you spend time editing, and then you write it back out. The I/O requests in this instance are similar to the *London Bus* distribution.

Appendix C. SCP Migration with ITRRs in CP90

Measurements for ITRs (Internal Throughput Rates are the number of transactions per CPU second) are made with the same SCP and various CPU Models and workloads With the change of SCPs the workloads can change also This is evident with a change from VM/XA to MVS/XA However, this is also true for a less radical change such as MVS/XA to MVS/ESA

This means that if one looked at the ITR data for a 3090-300J running MVS/XA and compared it to a 3090-300J running MVS/ESA, the ITRs would not be comparable since the workloads were changed significantly for MVS/ESA The MVS/XA ITR is 49.07 whereas the MVS/ESA is 54.36 This difference is not due to the difference in SCP but largely due to the difference in workload.

For each SCP, the ITRRs are different. In other words, each SCP power range has a different scale Not only will the power values be different between SCPs for the same model but the ratio of models for the same SCP will be different For example, the ITRR for a 3090-180J to 3090-600J comparison for MVS/XA running TSO is 4.53 whereas the same comparison for VM/XA is 5.88. The difference is not only a comparison of SCPs and workloads but a relationship between SCP and CPU model

Power values M values, are developed from ITRRs by multiplying the ITRRs by a factor Let's look at three migration cases

The simplest SCP migration is shown in Figure 124 on page 114 In this instance the CPU model does not change only the SCP The base migration assumption is that the Utilization does not change when the migration takes place This is accomplished by adjusting the M used in such a manner to have to CPU utilization unchanged This effectively changes scales from the old SCP scale to the new.

After that, an optional adjustment could be made where it is known that, for example, MVS/ESA reduces CPU utilization by 8% on the specific model when migrating from MVS/XA

The second case is a two step The migration is to a new SCP and CPU model but the old CPU model runs both new and old SCP. This is shown in Figure 125 on page 114 The migration process would entail a first step of changing the SCP on the existing model, a power scale change followed by a power change to the new CPU model within the new power scale

The third case (Figure 126 on page 115) is a bit more complicated. The new SCP does not run on the old CPU model For example, one may be running MVS/XA on a 3090-200 and wish to run MVS/ESA. This SCP migration will assume a CPU model change too since MVS/ESA does not run on a 3090-200 In order to minimize the impact of scale change the process attempt to find a reference CPU which does run both SCPs This reference CPU is searched for as close as possible to the old CPU After that reference model is located the process is similar to the two step above

In the migration between SCPs, the model takes care of scale changes as outlined. The user must supply an optional changes in utilization expected from the SCP migration Input into the model is shown in Figure 127 on page 115.

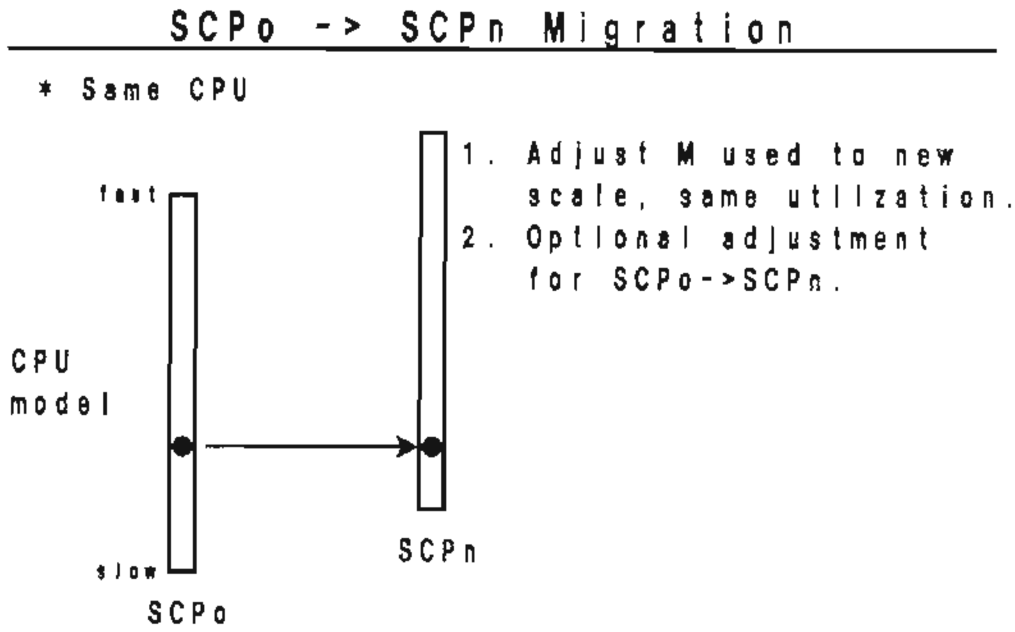


Figure 124. Case 1

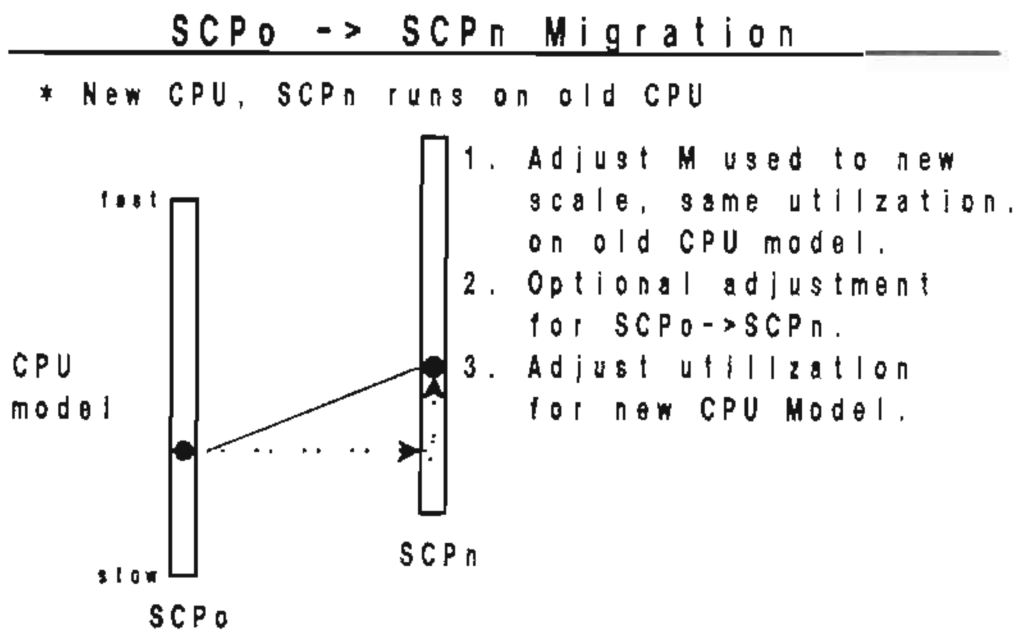


Figure 125. Case 2

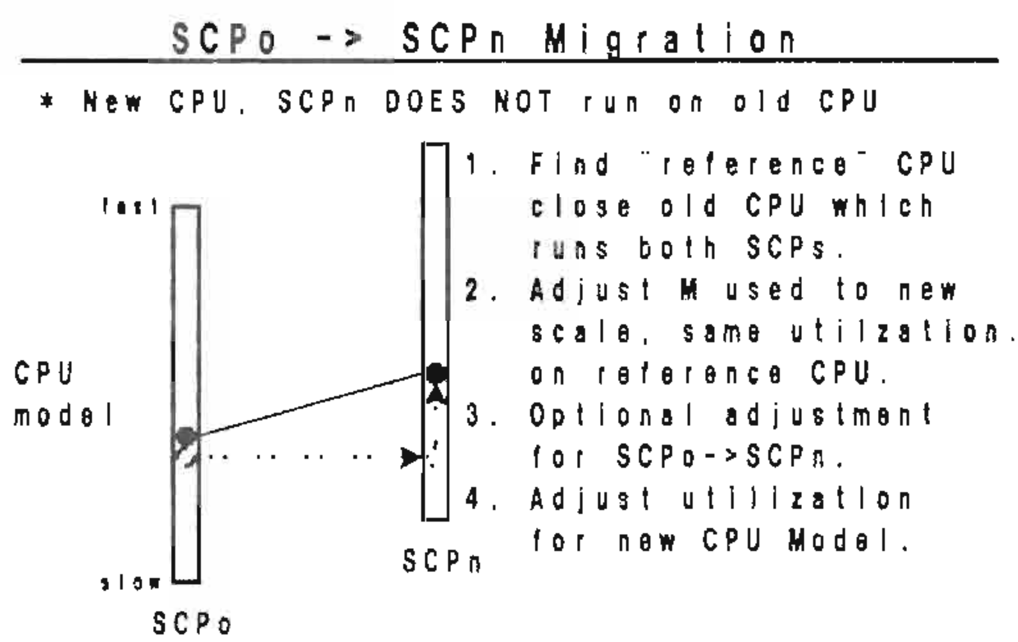


Figure 126. Case 3

```
Processor Projection          Workload Migration              Panel: AFMIGR#
                          CECID:  NEWSYS   SYSID:  SYSAX
                          Workload: TSO Development

The workload has been moved to a SYSID with a different SCP.
The M used may be adjusted for this new SCP.
Enter or review migration factors.

The reference model used for this scaling is: 3090-120J

                     Old SCP: MVS/XA
                     New SCP: MVS/ESA

In moving to a new SCP, it will be assumed that if the CPU Model remains the
same, the CPU% and the I/O rate will be the same. This assumption can be
adjusted below.

Reasonable ranges might be between 0.9 and 1.1.
When in doubt, leave the default of 1.0.
Values between 0 and 2 will be accepted.

  Enter CPU% Migration Factor: 1.00
    I/O Rate Migration factor: 1.00

                             For Reference Model
                                  Old       New
                       CPU%:      28.7      28.7
                   I/O Rate:      46.9      46.9

Warning: The CPU model used for SCP migration scaling is 3090-120J
PF: 1-Help 3-End 12-Cancel
```

Figure 127. CP90 Migration Panel

Appendix D. Capture Ratios

This section describes the technique to calculate capture ratios for business units. The objective is to describe the total system resource consumption in terms of the resource consumption of just the business units. If I have n business units, BU_1, ... ,BU_n, and a total resource consumption of *T*, a capture ratio for each business unit would accomplish the following:

$$T = BU_1/CR_1 + BU_2/CR_2 + \ldots + BU_2/CR_n$$

where $0 < CR_i <= 1$ for each business unit. CR *adjusts* each business unit resource so that the total approximates the system total as best as possible.

This is done with a recursive application of regression analysis. *Although this sounds complicated, it need not be hard.*

The Regresssion Technique

The first step is to collect a number of samples. Each sample consists of the resource consumption of the business units. The business units may be groups of TSO users, batch types, IMS applications, etc. These may be distinct performance groups in an installation with foresight to map performance groups to departments or work types. Figure 128 on page 119 shows a sample of data which we shall use to illustrate the technique.

The data used here is arbitrary. Data used in capacity planning would represent a long term sample over large enough intervals. The data would represent the resource usage for *stable* components. Stability is defined as a workload having a consistent CPU to I/O ratio.

Let's use the data in Figure 128 on page 119 to illustrate the technique. In regression analysis terms, the matrix of samples will be the *X* matrix which has three business units (columns) and 15 samples (rows). The totals will be the 15 by 1 *Y* matrix. Regression analysis techniques will provide coefficients b_j, such that

$$Y_i = \text{Sum from } \langle j=1 \text{ to } 3\rangle \text{ of } b_j X_{ij}$$

Note that this form of regression has no constant term b_0. Many packages and techniques can be used. These will be illustrated in "Illustrations." on page 120. The problem is that the solutions may not be *suitable*. We require the values of b_i to be such that $1 <= b_i$. This has to be in order that $0 < CR_i <= 1$ since $CR_i = 1/b_i$.

What if b_i is not greater than 1? The best solution would be to make $b_i = 1$, for those cases where regression analysis says that the solution is some value less than 1. For our sample in Figure 128 on page 119, the first regression analysis yielded the following solution: B=2.441 2.588 0.805. The value for b_3 is not acceptable. We shall then set b_3 1. Once we have *accounted* for b_3 we will go back and solve again for b_1 and b_2. However, since we already have b_3, we will have to subtract that resource from the total. Figure 129 on page 119 shows that and Figure 130 on page 120 illustrates the procedure.

Regression analysis again would yield B=1.780 2.579 for the remaining columns. These values are acceptable. Since we now have all three, the inverse of each yields the capture ratios. CR=0.562 0.388 1.0.

In the absence of the above procedure there are some quicker alternatives

1. Simply distribute the CPU time for the service address spaces (Master, JES, VTAM....) and the uncaptured CPU time in proportion to the captured CPU time or the I/O rates of the business units.
2. Before proceeding with the previous step, distribute the CPU time of the service address spaces which use the service. For example, distribute VTAM time to terminal users.

All processes should be tested. This is accomplished by dividing the samples in two parts. Compute the CRs with one part and test them with the other.

Y	X		
400.0	51.5	61.9	172.0
336.0	36.2	57.4	115.2
316.8	37.3	52.3	85.2
398.4	48.2	45.4	203.3
395.6	58.1	51.5	169.0
354.8	38.8	28.4	185.6
400.0	41.9	52.3	209.3
354.4	42.5	47.1	144.9
323.6	41.0	64.8	68.4
338.0	33.7	52.9	129.6
353.6	44.3	45.6	129.8
398.8	36.6	59.6	200.7
333.6	35.5	57.0	103.0
399.6	40.4	56.0	219.5
370.0	57.0	60.8	124.7

Sample 6

Total Resource

Resource reported by business unit

Figure 128. Sample data

Subtract

Y	X		
400.0	51.5	61.9	172.0
336.0	36.2	57.4	115.2
316.8	37.3	52.3	85.2
398.4	48.2	45.4	203.3
395.6	58.1	51.5	169.0
354.8	38.8	28.4	185.6
400.0	41.9	52.3	209.3
354.4	42.5	47.1	144.9
323.6	41.0	64.8	68.4
338.0	33.7	52.9	129.6
353.6	44.3	45.6	129.8
398.8	36.6	59.6	200.7
333.6	35.5	57.0	103.0
399.6	40.4	56.0	219.5
370.0	57.0	60.8	124.7

Figure 129. Recursion step 1

Y	X	
228.0	51.5	61.9
220.8	36.2	57.4
231.6	37.3	52.3
195.1	48.2	45.4
226.6	58.1	51.5
169.2	39.9	28.4
190.7	41.9	52.3
209.5	42.5	47.1
254.2	41.0	64.8
208.4	33.7	52.9
223.8	44.3	45.6
198.1	36.6	59.6
230.6	35.5	57.0
180.1	40.4	56.0
245.3	57.0	60.8

Sample 6

Total Resource — Resource reported by business unit

Figure 130. Recursion step 2

Illustrations.

APL

One statement in APL will perform a least squares regression.

```
B←(⌹(⍉X)+.×X)+.×(⍉X)+.×Y
```

The iteration steps above would look like the following.

```
      CR←3⍴0
      ⎕←6 3⍕B←(⌹(⍉X)+.×X)+.×(⍉X)+.×Y
 2.441  2.588  .805
      CR[3]←1
      Y←Y-X[;3]
      X←X[;1 2]
      ⎕←6 3⍕B←(⌹(⍉X)+.×X)+.×(⍉X)+.×Y
 1.780  2.579
      CR[1 2]←B
      6 3⍕÷CR
 .562  .388  1.000
```

Lotus

A similar procedure can be followed using 1-2-3® Version 2. With V2, a regression capability is included.

SAS

The following SAS® procedure (Statistical Analysis System is a product offering of SAS Institute) will yield the first estimate of the coefficients (parameter estimate in SAS terms) of B=2.441 2.588 .805

```
//WICKS  JOB (????,????),WICKS,MSGLEVEL=1,MSGCLASS=0,NOTIFY=WICKS
//SAS EXEC SAS
//SYSIN DD *
DATA CAPTURE;
 INPUT Y X1-X3;
 CARDS;
   400.0    51.5    61.9    172.0
   336.0    36.2    57.4    115.2
   316.8    37.3    52.3     85.2
   398.4    48.2    45.4    203.3
   395.6    58.1    51.5    169.0
   354.8    39.9    28.4    185.6
   400.0    41.9    52.3    209.3
   354.4    42.5    47.1    144.9
   323.6    41.0    64.8     69.4
   336.0    33.7    52.9    129.6
   353.6    44.3    45.6    129.8
   396.8    36.6    59.6    200.7
   333.6    35.5    57.8    103.0
   399.6    40.4    56.0    219.5
   370.0    57.0    60.8    124.7
PROC REG;
 MODEL Y = X1-X3/NOINT;
/*
```

After getting the fist estimate. The process of eliminating X3 can be accomplished by setting Y1=Y-X3. and reducing the model as follows.

```
//WICKS  JOB (????,????),WICKS,MSGLEVEL=1,MSGCLASS=0,NOTIFY=WICKS
//SAS EXEC SAS
//SYSIN DD *
DATA CAPTURE;
 INPUT Y X1-X3;
 Y1=Y-X3;
 CARDS;
   400.0    51.5    61.9    172.0
   ....
   370.0    57.0    60.8    124.7
PROC REG;
 MODEL Y1 = X1-X2/NOINT;
/*
```

The second iteration will yield the parameter estimate of B=1.780 2.579. SAS also provides an estimate of goodness of fit.

Other Techniques

Other easier methods do exist which will distribute the difference between the total resource used T, and the sum of the business units B in an equitable manner. The simplest would be to distribute the difference T-B in proportion to the size of the business units. If there were two business units B1 and B2 reporting equal amounts of CPU usage, divide T-B equally between B1 and B2. If B1 and B2 were in 80 20 proportion, split T-B so that B1 received 80% of T-B and B2 received 20%. This method distributes T-B in porportion to the business units CPU time.

Alternatively, distribute T-B in proportion to another resource such as the I/O count. if the I/O counts for B1 and B2 are in 80 20 proportion, distribute T-B in a similar manner. This technique is simple in concept but difficult in implementation. The mapping of physical I/Os to business unit is not simple. RMF does provide block counts or "EXCP"s, but these do not map one-to-one to DASD I/Os. It remains for each instance to establish a level of confidence that the block count proportion is a good basis for distributing DASD I/Os.

Also, one could distribute in a manner which reflects a specific knowledge of the workloads. For example, if it is known that a service address space relates only to specific business units, one could distribute that before any generic distibution.

Appendix E. Bibliography

Probability, Statistics, and Queuing Theory, by Arnold O. Allen Academic Press Inc.

The Art of Computer Systems Performance Analysis, by Raj Jain, Wiley

Quantitative System Performance, by E. Lazowska, J Zahorjan, G. Scott, and K. Sevcik Prentice Hall

MVS I/O Subsystems, by Gilbert E. Houtekamer and H. Pat Artis, McGraw-Hill

MVS Performance Management. by Steve Samson McGraw-Hill

IBM 3090 performance. A balanced system approach by Y. Singh, G.M King, and J.W Anderson. IBM Systems Journal, Vol. 25, No. 1, 1986.

Processor, I/O Path, and DASD Configuration Capacity by Joe Major, IBM Systems Journal Vol 20, No. 1, 1981.

Almost any volume of the Computer Measurement Group (CMG) Proceedings is worth looking at for performance and capacity planning articles. In particular...

1. *DASD Control Unit Planning and Technology Usage Trends* by Joe Major (CMG '90 Proceedings)
2. *The CPU - Memory Equation* by Joe Major (CMG '90 Proceedings)
3. *MVS/ESA Paging and Main Storage* by Joe Major (CMG '91 Proceedings)

The following are some IBM documents which provide information on the products and methodologies included in this document.

GG22-9264	*MVS Paging Performance Considerations* by Siebo Friesenborg
GG22-9351	*MVS Performance Management* by Siebo Friesenborg and Gary Hall
GG22-9363-03	*DASD Expectations* by Siebo Friesenborg and Ray Wicks
GG66-0232-00	*Introduction to Large Processor Capacity and Performance Evaluation* by Ken Radecki
GG66-0254-00	*Capacity Planning Overview* by Dick Armstrong
GG66-3135-00	*Performance Management 101: A Notebook for SLR V3 Performance Managment* by Linda August and Gary Hall
LY28-1007	*MVS/ESA Analyzing RMF Version 4 Monitor I and II Reports*
LY28-1008	*MVS/ESA Analyzing RMF Version 4 Monitor III Reports*
LY28-1042	*RMF Support for LPAR Management Time*
SG22-1051-00	*Large Systems Performance Reference* by John Fitch
SH19-6442-0	*SLR Version 3 Users Guide: Performance Management*
TR00.3629	*Evaluating Processor Upgrade Performance Benefits*, IBM Technical report by Linda Mier

Index

IBM®

Printed in U.S.A.

GG22-9299-04