This paper extends a particular capacity planning approach to include usage accounting by business element of Input/Output path and Direct Access Storage Device resources. A simple non-linear procedure is outlined to size host configurations that can process workloads at specified rates. An important feature of the procedure is to take account of a law of diminishing returns, which is that doubling the number of components does not double the amount of work done. Discussed are configuration relationships involving TSO and DB/DC subsystem sizing, tuning, workload variability, data considerations, and hardware and software considerations. Typical, but hypothetical, examples are presented.

Processor, I/O path, and DASD configuration capacity

by J. B. Major

Processor capacity planning (Reference 1) decomposes an installation workload into business elements, identifies current processor usage associated with these elements, projects business element growth (and processor usage), accumulates business element processor usage requirements, and translates the overall requirement into a processor configuration. Projection by business element is key: growth could be disproportional, and thus affect the overall processor requirements severely in an otherwise unanticipated fashion. Furthermore, current processor usage of some workload components (e.g., testing and development) may have a direct bearing on future usage of others (e.g., production), thus giving additional forecasting insight. Finally, responsiveness requirements of on-line services limit the average realizable processor time allocated to these services in proportion to the average-to-peak on-line workload—hence the need to track separately the on-line processor usage.

A decomposition of business into its elements and usage accounting by business element presupposes processor usage data by business element. The usually available performance monitors and accounting packages, discussed for example in References 2 to 4, do not produce the information in the required form. Reinterpretation of the data is necessary. Basically, we have to address an estimation problem as follows.

Let u_1, u_2, $\cdots$, u_n represent processor usage as shown by the System Management Facilities (SMF) of the Multiple Virtual Storage (MVS) operating system for each of the n business elements. A discussion of SMF is given in Reference 5. (These and all subsequent acronyms are listed in the Appendix.) Note that usage here is incomplete; the individual numbers do not add up to the total usage U. However, accurate information on U is available. Knowing the nature of the software used to process each business element, rational statements can be made about the fraction of actual usage that SMF accounts for. Thus capture ratios can be stipulated for the real processor usage of each business element for a variety of operating system environments.[1] Let f_i denote the reciprocal of the ith capture ratio. Then $U = f_1u_1 + f_2u_2 + \cdots + f_nu_n$. This statement is true if the capture ratio estimates are correct. Otherwise, adjustment of some or all of the ratios is necessary. Guidelines derived from experience usually do the job quite acceptably. Some installations opt for a more methodical factor estimation process that utilizes a fair amount of historical data and (constrained) linear regression.

Business-element-based accounting assumes additivity of resource usage requirements. If each element is processed in isolation, processor times add up to the composite time with 97–98 percent accuracy, and I/O resource usage additivity is about 85 percent accurate. The greatest difference is observed in systems consisting of very different subsystems (Data Base/Data Communication (DB/DC) with time-sharing and batch).

This paper extends the capacity planning approach of Reference 1 to include usage accounting by business element of Input/Output (I/O) path and Direct Access Storage Device (DASD) resources. I/O path includes block multiplexor channel (BMPX) and I/O device control unit combination. Disk I/O handling extends the I/O path definition to include a head-of-string controller as well. I/O usage accounting introduces additional adjustment factors. Guidelines are given for the initial choice of the relevant factors. When the accumulated usage matches the systemwide total (accurately reported by the available monitors), business-element-based host resource usage projection becomes possible.

The estimation factors introduced here are software and hardware environment-dependent. When practical, they are converted to dimensionless quantities to minimize the software/hardware dependence. Data analysis of many installations indicates a good measure of stability of the factors presented here. As our approach is put to test in more and more situations, the estimation factors become more and more refined.

Capacity planning is basically the following three-step procedure: (1) account for current usage, (2) forecast future usage, and (3) match host resource requirements to a future configuration.

The methodology outlined in the paper has been put to practical use by the author and several of his colleagues. It emerged from the author's previous work that utilized queuing models to study configuration relationships. These studies were time-consuming and hard to communicate. The problems addressed then and many others now can be put into a simple context. Data reduction and coefficient estimation turn out to be a relatively quick (1–2-day) process. As expected, installations with I/O path or DASD usage above the recommended guidelines usually experience responsiveness problems. Configuration balancing for both the current and projected workload processing levels is a straightforward, systematic, and credible effort.

The methodology is illustrated through several examples; no attempt is made to depict any one specific installation. The numbers chosen, however, do tend to describe likely actual situations.

Workload characterization

relative I/O content

Workload is viewed here as a collection of processor and I/O resource requests during a particular period of interest. Workloads are compared through a measure called *relative I/O content.* The concept is explained for the methodology discussed here as follows.

Consider an MVS installation that processes a particular general-purpose commercial batch workload during a logical shift, i.e., a period of interest. Processor, I/O path, and DASD resource consumption, I/O activity, and other workload processing parameters can be obtained from resource usage monitors.[2-4] A measure of *I/O intensity*, i.e., the I/O content of workload, can be defined as the BMPX channel activity count per second divided by the processor time (ms/s) used. This measure is related to the concept of instructions per byte transferred that is sometimes used to characterize workloads. Furthermore, this measure is processor-dependent, i.e., the same amount of code execution requires different amounts of BMPX channel activity per second on different-speed processors.

I/O content varies as the workload varies: Commercial batch differs from engineering batch, and batch (in general) differs from time-sharing. DB/DC gives rise to yet another I/O content, and time-sharing systems differ from one another. It also varies as the workload system usage parameters (block size, buffering, and storage allocation) change. One might suspect that I/O content is an unstable characterization of workloads. Experience shows the contrary.

Table 1 Relative I/O content and kiloinstructions (ki) per Start I/O by workload type

Workload type	*Relative I/O content*	*ki/SIO*
Commercial MVS batch	1.00	18.0
Commercial OS/VS1 batch	1.70	10.6
Commercial DOS/VSE batch	2.60	6.9
Engineering batch	0.40	45.0
TSO, prime shift, application development	0.60	30.0
TSO, off-shift, systems programming	1.30	13.8
IMS, CICS with DL/1	1.20	15.0
IMS Fast Path, CICS high performance	1.60	11.2
VM, general workload mix	1.20	15.0
VM, CMS time-sharing	2.00	9.0
VM, guest operating system MVS production workload	0.45	40.0
VM, guest operating system DOS/VSE production workload	1.20	15.0

Note: SIO is the physical I/O as represented by the channel activity count in monitor reports.[2–4]

Suppose we define the I/O content for the average commercial batch MVS installation (workload) as a base unit of 1.0 when processed on a given processor. A processor-independent measure of I/O content can be defined to relate workloads to one another. *Relative I/O content* of a workload with respect to the base workload is the ratio of its I/O content to the I/O content of the base when processed on the same processor. Relative I/O content and kiloinstructions (ki) per START I/O (SIO) from a large number of installations are given in Table 1.

The relative I/O content and the associated ki/SIO data are averages observed by the author. Processor independence of the relative I/O content is incomplete: Two processors may have different internal processing power ratios for different software environments (e.g., MVS commercial batch and DB/DC). Fortunately, the variations are not too great except in the IBM 3033 N group of processors, which show a higher relative I/O content for engineering batch. This group of processors is implemented with relatively less execution parallelism and with smaller cache memory relative to its processing speeds. Possible reasons for data deviations are given later in this paper under the heading "A perspective on data."

The higher the relative I/O content of a workload component, the more SIOs are processed for each unit of processor time, hence the need to plan I/O configuration requirements of these workloads. Although relative I/O content is very important, it is not the only descriptor. Workload processing implies a given number of I/Os per unit time. Thus, workloads can be described in terms of their main hardware component requirements. We are led to the following parameters:

- Relative I/O content (processor usage per I/O).
- I/O path usage per I/O.
- DASD usage per I/O.
- I/Os per second to process the workload in time.

In summary, host configuration is based on establishing workload processing and resource usage requirements. These, in turn, are related to resource usability experience to estimate the number and size of resources.

application of I/O content

Conventional system planning approaches estimate future processor requirements from workload breakdown into business elements and individual growth projections and processor usage accounting for these elements. For moderate to low relative I/O content, the most important task is the choice of processor. Within the System/370 product line almost any processor can be configured to handle the I/O requirements. The number of DASD devices depends on data space and availability. The number of I/O paths (channel, control unit, head-of-string) depends on availability and recovery considerations.

Workloads of high relative I/O content require special attention when based on processor requirements alone, particularly when the processor choice is a channel-constrained attached processor (AP) or powerful uniprocessor (UP). It might not be possible to realize the high SIO rates that such workloads call for on the selected processor. Faster DASD (or tapes) may be needed than are installed or available in the current product line, more channels and control units may be needed than the configuration provides for. Shorter strings of DASD may be required than the ones intuitively felt.

Another problem is presented by the fact that any increase in the I/O processing requirements cannot usually be met by a proportional increase in the I/O configuration. We need to size system requirements rather than processor requirements. A simple extension to handle these requirements is now presented.

case study of configuration sizing

A 3031 MVS installation undertakes a capacity planning study based on USAGE.[1] Current and anticipated requirements 18 months hence are given in the first three columns of Table 2.

The prime shift usage of 110 hours (68% processor busy during 160 hours of 20 business days, 9:00 a.m. to 5:00 p.m.) is expected to grow to 233 hours, necessitating a faster processor. The on-line

Table 2 Summary data for an installation's current and anticipated 18-month requirements for CPU and I/O usage; 160 hours of operation, 20 prime shifts

Business elements	*CPU usage hours*		*Relative I/O content*	*I/O usage*			
				now		*future*	
	now	*future*		*Relative I/O usage hours*	*SIO/s*	*Relative I/O usage hours*	*SIO/s*
Batch production	30	30	1.0	30	11.7	30	11.7
Test and develop							
time shared, TSO	5	15	0.6	3	1.2	9	3.5
batch	5	8	1.0	5	2.0	8	3.1
Support							
time shared, TSO	15	35	1.3	20	7.8	46	18.0
batch	5	10	1.0	5	2.0	10	3.9
Maintenance, TSO	10	15	0.6	6	2.3	9	3.5
Order processing high performance DB/DC, CICS	40	120	1.6	64	25.0	192	75.2
Total CPU hours	110	233					
Total SIO/s					52		119
Total relative I/O usage				133		304	
Average relative I/O content				1.21		1.30	

portion of the processing load is 185 hours of 3031 during 160 hours of operation. According to USAGE guidelines, 64 hours is the recommended maximum. A processor of nearly three times the power of the 3031 is needed. Thus a 3033 N is considered for installation because it has sufficient processing power. There are also questions about the I/O subsystem. Will one channel director be sufficient? Can five I/O paths handle the I/O load? To answer, configuration rules are needed.

The relative I/O content factors of Table 2 are taken from Table 1 as approximations for the workload represented by the business elements. I/O usage figures are derived by multiplying the Central Processing Unit (CPU) usage hours (now and future, respectively) by the relative I/O content factors. As an example in Support, time shared, TSO, 35 future CPU hours times the 1.3 factor yield 46 relative I/O usage hours (future). Total SIO/s data (now) are obtained from Resource Measurement Facility (RMF) channel activity reports for the type of time periods (averaged) depicted under I/O usage in Table 2. Our guidelines have been derived under the assumption that byte multiplexor channel activity is ignored. For our example, the SIO/s activity rate is 52. The total relative I/O usage (133, now) divided by the CPU usage hours (110, now) defines the average relative I/O content of the total workload for the selected time period. The result is 133/110 = 1.21. Similarly, the future workload has a relative I/O content of 304/233 = 1.30, an increase over the current level.

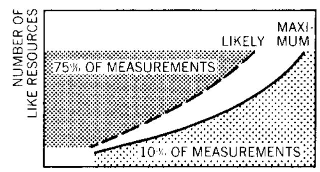
Each business element can be assumed to have SIO/s rates in proportion to its relative I/O usage. Thus, for example, Support, time shared, TSO (now) has an SIO/s rate of $52 \times (20/133) = 7.8$. And again, the SIO/s requirement of the future workload is obtained by multiplying the current rate (52) by the ratio of relative I/O usage, i.e., future over current. Thus, $52 \times (304/133)$ yields 119 SIO/s future requirement. The business-element-based SIO/s requirements for the future workload are calculated as in the previous paragraph. Thus, Support, time shared, TSO (future) has an SIO/s rate of $119 \times (46/304) = 18$.

Law of diminishing returns

concepts

System tuning aims at balanced (approximately equal) utilization of like resources. The greater the number of like resources, the greater the difficulty of making the balance. Furthermore, utilization balancing becomes more difficult with shrinkage of the period during which equal utilizations are desired.

Figure 1 Law of diminishing returns



The author's study of several installations has resulted in the observed overall utilizations shown in Table 3 and illustrated in Figure 1. Increased overall utilization, as previously discussed, may require an inordinate amount of ongoing tuning effort. Further, as hardware costs become lower, the average realized utilization of like components is expected to become lower as well to compensate for cost savings from reduced tuning effort.

Host system behavior obeys a law of diminishing returns in that a system with $2N$ like components cannot do twice the work of one with N like components of the same type. Table 3 summarizes our empirical findings for I/O path and DASD resources for IBM 3330, 3350-type DASD, 3420-type tapes, 3830, and 3880 DASD controllers, and also for 3803 tape controllers. A theoretical explanation is discussed later in this paper under the heading "Justification of guidelines." Systems consisting of isolated subsystems (and isolated workloads) may exhibit higher overall utilization. Likewise, data base systems with randomized record positions over the DASD actuators and random requests for records may exhibit DASD utilizations substantially above the overall level shown here for that many actuators, but such systems have probably undergone substantial tuning.

application

Projections of workload growth should take into account a nonproportional expansion requirement. Plots of traffic intensity against the number of like resources (I/O paths or DASD) in Figure 1 show traffic intensity maxima that can be realized for given numbers of like resources. Here, traffic intensity is the sum of utilizations of like components of a subsystem, and utilization is measured as the fraction of time a component is busy. Figure 1 is

Table 3 Empirical findings for I/O path and DASD resources

Block channels	*Traffic* (ms) *maximum*	*Traffic* (ms) *likely*	*DASD*	*Traffic* (ms) *maximum*	*Traffic* (ms) *likely*
2	520	300	8	1460	1100
3	750	500	10	1800	1350
4	980	750	12	2130	1640
5	1200	900	14	2450	1900
6	1400	1050	16	2770	2130
7	1600	1220	18	3070	2380
8	1780	1400	20	3360	2600
9	1940	1550	24	3920	3070
10	2100	1680	28	4440	3500
11	2240	1810	32	4920	3920
12	2380	1940	40	5760	4700
14	2600	2180	48	6450	5360
16	2780	2380	56	6990	5940
18	2910	2560	64	7370	6450
20	3000	2700	72	7600	6870
24	3050	2910	80	7700	7200
28	3080	3020	88	7800	7440
			96	7850	7600

not drawn to scale. It illustrates the author's finding that 75 per cent of the measurements fall below a likely overall utilization (DASD and I/O path), and 10 per cent show utilization in excess of the maximum. Responsiveness problems are manifest above the likely level. Specific behavior is given in Table 3.

procedures

Let R be the relative I/O content of a workload. Unconstrained I/O processing capacity S of various processors can be defined as SIO/s processed when the processor is 100% busy, as follows:

$$S = RM \tag{1}$$

where M is a processor multiplier and is given for the IBM processor types shown in Table 4. These M values are the author's best estimates, based on processor comparisons within his experience. We see that workloads of low relative I/O content (small R) result in low SIO rates.

Table 4 Processor multipliers (M) for a range of processor types

IBM processor type	*M*
4341 MG1	44
158-1	47
158-3	55
3031	64
158-3 AP, MP	88
3031 AP	106
168-1	122
168-3, 3032	138
3033 N	193
168-3 AP, MP	221
3033	256
3033 AP, MP	422

Equation (1) together with Table 3 is used to compute the size of the host. The stepwise procedure is given as follows:

1. Choose a host processor from Table 4. The unconstrained I/O processing capacity must not be less than the workload projection. If none qualifies, consider loosely coupled processors. If, for example, S and R are obtained from Table 2, then $S = 119$, $R = 1.3$, and $M \geq 92$. Thus a 3031 AP or larger processor is required.

2. Apply the workload projections for SIO/s and ms/SIO to obtain the required number of channels and DASD. If Table 2 is used again, we are considering a CICS-TSO-batch mixture with likely average service times of 9–10 ms on the I/O path and 32–36 ms on DASD. The result is a best-case (9,32) and a worst-case (10,36) configuration as follows:

I/O paths	Best case		Worst case	
Traffic intensity	$119 \times 9 = 1071$ ms		$119 \times 10 = 1190$ ms	
I/O paths	minimum	reasonable	minimum	reasonable
From Table 3	5	6 or 7	5	7

DASD	Best case		Worst case	
Traffic intensity	$119 \times 32 = 3808$ ms		$119 \times 36 = 4284$ ms	
DASD	minimum	reasonable	minimum	reasonable
From Table 3	24	32	27	36

We deduce that a 3031 AP as found in step 1 is sufficient, but with the following further understanding:

- It would be $100 \times 92/106 = 87\%$ busy, which is somewhat high for planning purposes.
- All of its I/O paths (maximum five) are required, which requires significant tuning to operate at high I/O path performance.
- There is no difficulty in configuring the DASD.
- A 3033 N processor with one channel director is a more attractive system. There is field upgradability if availability justifies another director, or if performance requires more than five I/O paths.

3. In general, a configuration based on maximum performance measures is usually only realizable through substantial tuning. Realistic planning should increase the calculated values by one-third. The traffic intensity data under "likely" in Table 3 have been derived with that safety consideration in mind. Systems vary around that level of performance. Thus we desire the following configuration:

- 7 I/O paths (2 channel directors, 7 DASD control units)
- 32–36 DASD
- $M > 122$ (3032 processor or up)

4. Estimate the required number of DASD on-line to the host from space requirement considerations. Use the larger number obtained from steps 3 and 4.

Denote B as the maximum CPU busy fraction that is acceptable for peak workload within the installation responsiveness specifications. From Equation (1), the new, constrained I/O processing capacity S is expressed as follows:

$$S = RMB. \quad (2)$$

Figure 2 Configuration relationships

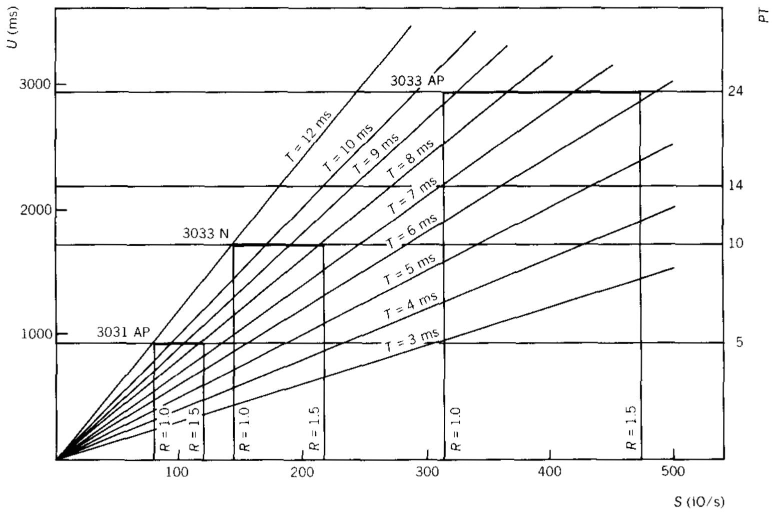


Table 5 Selected characteristic SIO times

Workload type	*I/O path* (ms/SIO)	*DASD* (ms/SIO)
Commercial batch, MVS	9–11	30–35
General purpose TSO	10–14	30–40
DB/DC, CMS	6–8	25–35

I/O service times

Current I/O service times are available for MVS systems from RMF reports (Channel and DASD activity). Similar information is available from Virtual Machine Analysis Program (VMAP) reports for VM/370 starting with Release 6. The Virtual System Extended/Performance Tool (VSE/PT) provides equivalent information for Disk Operating System/Virtual Storage Extended (DOS/VSE) systems.

Future I/O service times depend on future workload composition and physical characteristics of future configurations. Thus if device speeds and channel usage modes change, service time estimation has to be performed in terms of the new specifications.

Characteristic I/O times for current products are obtained from experience. When specific measurement information is not available, one may use the information given in Table 5. The lower end of the channel service time range tends to correspond to 3350-type DASD and/or fast tapes, whereas the higher end describes 3330-type DASD and/or slower tapes. Extensive use of fixed-head storage devices also reduces DASD and channel service times.

Table 6 Estimated I/O path (channel activity) service times

Business elements	*SIO/s*		*ms/SIO* (I/O path time)		*Total ms* (total for I/O paths)	
	Now	*Future*	*Now*	*Future*	*Now*	*Future*
Batch production	11.7	11.7	9	9	105.3	105.3
Test and develop						
time shared, TSO	1.2	3.5	12	12	14.4	42.0
batch	2.0	3.1	9	9	18.0	27.9
Support						
time shared, TSO	7.8	18.0	14	14	109.2	252.0
batch	2.0	3.9	9	9	18.0	35.1
Maintenance, TSO	2.3	3.5	12	12	27.6	42.0
Order processing high performance DB/DC, CICS	25.0	75.2	8	8	200.0	601.6
Total	52	119			492	1106
Average					9.5	9.3

When measurements, extrapolations of same, or estimates of channel and DASD service times are available by business element, extension of the USAGE table is possible. Table 2 shows estimated SIO/s information by business element for current and future workloads. If the I/O service time estimates exist for the current workload, the corresponding estimates of I/O path times are given in Table 6. The ms/SIO entries for I/O paths might be different in the future if new peripherals (for example, IBM 3380-type DASD) and/or new accessing techniques are used. Table 6 can be extended or analogously produced for DASD SIO estimates of current service times and SIO rates (DASD only) plus new service times for future configurations to yield total ms/DASD information for use with Table 3.

Configuration relationships

Equation (2) is used to generate the relationships of Figure 2. Let $B = 0.75$, $R = 1$ and 1.5. Consider a 3031 AP with its maximum five I/O paths, a 3033 N with its maximum ten I/O paths and a 3033 AP with its announced maximum 24 I/O paths. Each processor is illustrated by a rectangle, the left vertical side at the maximum I/O rate for $R = 1$ and $B = 0.75$, the right vertical side at $R = 1.5$ and $B = 0.75$. The horizontal lines correspond to the likely I/O path traffic (Table 3) for the configured number of paths PT.

I/O path usage U is proportional to the number S of I/Os and to the I/O path time T of the average I/O. Path times of 3–12 ms are chosen and path usage is depicted by rays. Workloads of high (R = 1.5) I/O content experience I/O path problems on the 3031 AP and on the 3033 N when $T \geq 8$ ms. To fully utilize ($B = 0.75$) the 3033 AP with $R = 1.5$, a workload with $T \leq 6$ ms is needed.

Configuration limitations of the 303X processors point to the need for reduced I/O path times through high-transfer-rate DASD and tape devices. Reduction of R is another possibility. High incidence of logical swapping—a feature of MVS System Extensions (MVS/SE) when adequate main storage is configured—also helps overcome configuration limitations.

The case study of configuration sizing poses the question about the sufficiency of a one-director (maximum five I/O paths) 3033 N. Figure 2 shows such a system to be tight at the $T = 5$ ms level. In the case study, $T = 9.3$ ms for the assumed 3350 DASD. With 3380-type DASD used in the future, T could reduce to about 5 ms. However, one director allows for only two high-speed (data stream mode) DASD paths, which is insufficient in this case. On the other hand, two directors with four high-speed and four regular DASD paths result in $T = 7$ ms for this mixture of DASD, and provide an I/O configuration sufficient for this workload.

TSO subsystem sizing

A system with P active programmers at terminals is to be sized. When in session, each programmer generates three transactions per minute. It has been observed that if a system has an average response time of ≤ 1.5 s, the programmers tend to generate many transactions in a short time; otherwise, they generate fewer.

A typical TSO transaction induces 26–36 SIOs. At the above programmer activity level, this results in 78–108 SIOs per minute or 1.3–1.8 SIOs per second per active programmer.

SIO service times are approximately 12 ms on the I/O path and 36 ms on DASD.

Relative I/O content of the given TSO workload is estimated as 0.4 when foreground compiles are practiced and 0.5 otherwise. Table 1 shows a relative I/O content of 0.6 for TSO application development. Variations around that number are also possible. We assume significant logical swapping activity under MVS/SE.

From Equation (2),

$M = S/(RB)$.

For simplicity, assume 1.5 SIOs per second per programmer. For $B = 0.75$,

$M = 5P$ with foreground compiles,
$M = 4P$ without foreground compiles.

Table 7 TSO system sizing with and without foreground compile

Active programmers		Host type	Number of I/O paths		Number of DASD	
with foreground compile	without foreground compile		with foreground compile	without foreground compile	with foreground compile	without foreground compile
9	12	158-1	2	2	3	5
11	14	158-3	2	2	4	6
13	16	3031	2	2	5	6
17	22	158-3 AP, MP	2	3	6	9
21	26	3031 AP	3	3	8	10
24	30	168-1	3	4	10	12
27	34	168-3, 3032	3	4	11	14
39	48	3033 N	4	5	16	20
44	55	168-3 AP, MP	5	6	20	25
51	64	3033	6	7	22	28
84	105	3033 AP, MP	9	12	37	50

With the above information, we can use Tables 3 and 4 to size the required configurations, with results summarized in Table 7.

System sizing is not proportional to the number of active programmers due to diminishing returns. To determine the specific numbers we follow through the sizing procedure, using the example of the 3033 N processor as follows:

1. Table 4 has $M = 193$ as the maximum $S/(RB)$. Since each programmer generates 1.5 SIO/s, the maximum $P = 193RB/1.5$. For $B = 0.75$ and $R = 0.4$, $P = 39$, but $P = 48$ when $R = 0.5$.
2. I/O path traffic intensity is $12S$ or $18P$ and DASD traffic intensity is $36S$ or $54P$. Referring to Table 3, 39 active users generate $39 \times 18 = 702$ ms of I/O path traffic (4 I/O paths under the likely heading) and also $39 \times 54 = 2106$ ms of DASD traffic (16 DASD under the likely heading). Therefore, we obtain the entries under the foreground compile heading previously discussed.

DB/DC subsystem sizing

A system with P active users is to be sized in which each user has a terminal and generates one transaction per minute. Consider the 303X systems only.

A typical DB/DC transaction induces 8 SIOs (light, interactive transactions representing simple inquiry/update-type work). At the given user activity level, that results in 0.13 SIO per second per active user. SIO service times are approximately 8 ms on the channel and 32 ms on DASD.

Relative I/O content of this DB/DC workload is 1.2, which is consistent with Table 1. Given 100 active users, the calculations are 100

Table 8 Summary of system sizing for DB/DC

Users	*SIO/s*	*Traffic ms/s*		*Configuration required*			*Minimum M*
		I/O paths	*DASD*	*CPU type*	*I/O paths*	*DASD*	
100	13	104	416	3031	1	4	13.5
200	26	208	832	3031	2	6	27.1
300	39	312	1248	3031	3	10	40.6
400	52	416	1664	3031	3	14	54.2
500	65	520	2080	3031AP	4	16	67.7
600	78	624	2496	3031AP	4	20	81.3
700	91	728	2912	3031AP	4	24	94.8
800	104	832	3328	3032	5	28	108.3
900	117	936	3744	3032	6	32	121.9
1000	130	1040	4160	3033N	6	36	135.4
1100	143	1144	4576	3033N	7	40	149.0
1200	156	1248	4992	3033N	8	44	162.5
1300	169	1352	5408	3033N	8	50	176.0
1400	182	1456	5824	3033N	9	56	189.6

$\times$ 0.13 = 13 SIO/s, 13 $\times$ 8 = 104 ms/s I/O path traffic, and 13 $\times$ 32 = 416 ms/s of DASD traffic. The DB/DC system is likely to be responsive at processor utilizations of up to 0.8.

We have $R = 1.2$, $B = 0.8$, and $S = 0.13P$. Using Equation (2), we calculate $M = 0.135P$. Let T_P and T_D denote I/O path and DASD traffic (ms/s) respectively. Then $T_P = 0.104P$ and $T_D = 0.416P$. Using Table 4 for processor sizing and Table 3 (likely values) for I/O sizing, we configure for a variable number of users as shown in Table 8.

Here none of the processors is configuration-limited, implying that it is possible to configure sufficient I/O to handle the specified DB/DC workload.

tuning

The objective of tuning is to improve the throughput and/or responsiveness of a system within the context of a particular configuration. Sometimes minor configuration changes are considered part of the tuning. These objectives are usually achieved through two kinds of improvements:

- Elimination of unnecessary work.
- Balanced utilization of like components.

In the first case, a given throughput is obtained with less resource utilization than otherwise (e.g., more main storage is used to do fewer SIOs through larger block sizes, and lower I/O path, DASD, and CPU utilizations are needed for a given quantity of work).

In the second case, the diminishing returns effect is tempered. Through subsystem isolation and balancing, the overall system behaves as though it were processing independent, isolated workloads in parallel, each using its own I/O subsystem, while sharing the host processor(s) and main storage. Such subsystems would have to be sized in their own right, e.g., a TSO subsystem could have its own I/O paths and DASD, and a DB/DC subsystem would have its own I/O paths and DASD, while still sharing processor(s), main storage, and, to some extent, I/O (system data set volumes and I/O paths). Sizing the two subsystems could be almost independent from subsystem to subsystem, with the total requiring somewhat more resources (due to sharing) than the sum of the two. On the other hand, combined sizing could require more resources than the sum of the two components.

Consider the example of two TSO subsystems that are independent of one another. One subsystem supports 16 programmers, and the other supports 48. There is no foreground compile/run in either system. According to the results in the previous section, the first system requires six DASD, the second requires 20 DASD, and 26 DASD are required in all. A single, nonisolated system for 64 programmers needs 28 DASD. Thus, isolation, if achieved, is good.

Some systems are readily isolated, and DB/DC systems are good examples of this. Another class of examples is given by tape jobs and DASD jobs. Tape channels and I/O paths are separate from the DASD I/O paths. Thus, the subsystem requirements can be sized separately. Some sizing gains could be realized. Likewise, paging workload of TSO systems can be isolated from nonpaging I/O work with improved configuration economy and performance.

As a final note, tuning tends to be costly in skilled people resources as well as in hardware. Tuning is an ongoing effort, and the results are often unstable. Human resources are growing more costly, whereas hardware is decreasing in cost. Therefore, tuning should not be an objective; rather, it should merely be an unavoidable temporary measure. The scope of tuning is usually within the maximum resource traffic numbers given by Table 3. Thus an installation may find temporary savings of less than 25 percent on the host hardware cost through sustained tuning—to be balanced against the cost of hardware resources, tuning monitors/aids, and people.

workload variability with time

Workload composition and relative I/O content change from shift to shift as well as within shifts. Therefore, plans should be made for all relevant workload profiles. Configuration sizing should accommodate them all. If, for example, N_1 units of a component are required to process type 1 workload and N_2 units are needed to process type 2 workload, the configuration should have max (N_1,N_2) components. Here max (N_1,N_2) means the greater of the quantities N_1 and N_2. If some of these units are to be exclusively

Table 9 Validation data

Software environment	*Source systems*	*Measurements*		*Relative I/O content*	
		number	*length*	*mean*	*standard deviation*
VM/370 mixed workload	IBM	82	3 months prime shift	1.2	0.4
Commercial MVS batch	large commercial systems, Canada	29 13	≥1 shift and ≤1 hour	1.0	0.2
TSO	large commercial systems, Canada	23 12	1 shift and ≤1 hour	0.9*	0.5
IMS DB/DC, CICS DL/1	railroad, utilities, government in Canada, SHARE 49 IBM tests[7]	6 6 4	1 shift and ≤1 hour and ≤1 hour	1.2	0.2
High performance DB/DC	banking industry, Canada	3 6	1 shift and ≤1 hour	1.6	0.3
DOS/VSE	consulting companies, Canada	14	1 hour	2.7	1.2
VM/370 with MVS guest production system	IBM	10**	15 hours	0.4	0.3
VM/370 with DOS/VSE guest production system	IBM	10**	15 hours	1.2	0.3
VM/370 CMS	IBM, utilities, Canada	10** 10 1	15 hours 1 shift 2 hours	2.0	0.7

* Extrapolations based on batch and TSO mix considerations.
**Extrapolations based on user resource utilization reports.

used by the respective workload (e.g., dedicated, on-line DASD), the configuration should contain N units as follows:

$$\max (N_1, N_2 \leq N \leq N_1 + N_2.$$

The law of diminishing returns suggests that multiple host systems may be configured to process a specified amount of work with less I/O resource than a single combined system. The gain is illusory in that when workload is split into several parts and each is directed to a specific host, each host becomes loaded as its workload varies with time. Subsystem capacity should then accommodate subsystem peaks as they occur. Since two subsystems hardly ever peak at exactly the same time, the consolidated system has an average load equal to the sum of the subsystem averages, but with a peak that is less than the sum of the subsystem peaks. Thus the consolidated system need not be as large, compared to its average load, as some consolidated subsystems should be, compared to their respective average loads. There is no general conclusion other than that consolidation may or may not be economical, depending on the degree of subsystem synchronization.

One needs also to consider the consequences of multiple system data sets (or data-set sharing to avoid multiplicity) and operational costs of the isolated subsystems as well as possible gains in subsystem integrity, availability, and operability.

Multiple isolated or even loosely coupled systems also obey the law of diminishing returns. Thus Table 3 can be extended to show for N like processors the amount of realizable processor usage (maximum and likely). Multiple host systems of an installation can then be configured.

The author has not yet collected sufficient data to propose multisystem performance guidelines under diminishing returns conditions. Based on experience, however, the following opinion is put forward. Given a multiple processor installation with a combined workload and N (≤ 5) processors with M factors of M_1, M_2, M_3, M_4, and M_5 in decreasing order, the overall workload processing capacity is the inner product of the **K** and **M** vectors (for their first N components) as follows:

$N =$	1	2	3	4	5
$\mathbf{K} =$	1	0.75	0.65	0.5	0.4
$\mathbf{M} =$	M_1	M_2	M_3	M_4	M_5

These numbers, however, are quite tentative.

Justification of guidelines

During 1978–1980 the author collected measurement data from several Canadian railroad, utility, banking, education, and government installations. In addition, measurements of IBM's own systems were reviewed. A tabulation of findings is shown in Table 9. An installation is represented by 1–5 sets of measurements, although several installations provided many more measurement reports. Processor utilization averages 72 percent for the MVS installations in Table 9 and 60 percent for the other systems.

The results as shown in Table 9 match closely but not completely the relative I/O content factors as proposed in Table 1. To isolate TSO workload resource usage, we applied the approach in Reference 1 for the processor resource. I/O resource usage was obtained as follows:

- Estimate batch workload parameters from stand-alone runs.
- Apply batch workload parameters to mixed runs to estimate TSO or CICS workload parameters.
- Estimate CICS or TSO I/O resource usage from knowledge of dedicated I/O resources and their usage (as reported by RMF) of these systems, when sufficient information is available.

When both of the last two steps could be performed, the results were comparable.

a perspective on data

Consider an installation with two modes of operation, A and B. Let the same workload be processed under both modes and introduce the following notation for mode A:

X_A number of SIOs to process the workload
P_A ki/SIO (thousand instructions per SIO)
R_A total ki required to process the workload
C_A average channel service time (ms) on I/O path

Likewise, for mode B we define X_B, P_B, R_B, and C_B.

If C_A is greater than C_B, the block sizes used in mode A are probably greater than those in mode B, in proportion C_A/C_B. Workload processing in mode B requires roughly C_A/C_B times X_A SIOs, and R_B exceeds R_A by the amount it takes to process the extra SIOs with, say, an I/O path length of S ki/SIO. Thus, we derive the following relationships for R_B and R_A:

$$R_B = R_A + SX_A(C_A - C_B)/C_B.$$

Also,

$$R_B = P_BX_B = P_BX_AC_A/C_B$$

and

$$R_A = P_AX_A.$$

Thus,

$$P_BX_AC_A/C_B = P_AX_A + SX_A(C_A - C_B)/C_B$$

or

$$P_B = P_AC_B/C_A + S(C_A - C_B)/C_A.$$

Some examples show the following results:

P_A	18	18	18	18
C_A	10	9	8	10
C_B	8	7	6	12
S	5	5	5	5
P_B	15.6	15.1	14.8	20.6

In other words, installations that tend to show 15 ki/SIO and also tend to have 7-8 ms average I/O path service time are consistent with our sampled installations having $P_A = 18$ and $C_A = 9$ to 10 ms.

hardware and software notes

The above consideration can be extended to explain the high relative I/O content of DOS/VSE commercial batch. (See Table 1.) DOS/VSE installations of today tend to have evolved from DOS installations with very little main storage (in today's terms) on their System/360 or System/370, and data set block sizes are quite small (50-500 bytes). Sequential data set buffering is limited to 1 or 2, and little if any chaining takes place. More I/Os are needed to process a given amount of user data than in MVS systems. Page size is 2K bytes as opposed to the 4K-byte page size in MVS. Once again, more I/Os are needed to do the same work.

When fewer bytes are transferred per I/O, less I/O path time should be required for the same tape and DASD speeds. DOS/VSE installation hardware can, however, differ, having slower tapes and different DASD. The DOS/VSE DASD Rotational Position Sensing (RPS) is not necessarily used. As a result, 8-10 ms are added to the I/O path time for each I/O. DOS/VSE measurement interpretation and projection should take this into account.

VMAP reports resource usage by virtual machine (user). When business elements can be related to users or groups of users, the configuration procedure is performed as for other environments. Processor usage reporting by user is incomplete, and multiplication factors are necessary, as in SMF-based systems. Observations show that the ratio between the total user processor resource usage, as shown by the User Resource Utilization Summary report of VMAP,[3] and total processor usage, as shown by its Monitor Statistical Summary report, is usually 0.95 for uniprocessor configurations and about 0.8 for attached processors. A simple approach is to multiply user processor resource usage by the reciprocals of these numbers, as applicable. The incompleteness is probably due to spooling and not allocating the processor resource used in paging I/Os and physical printing. Adjustment of processor usage may be based on user page and print (spool) I/O usage, compared to the systemwide usage of same (in terms of the quantities reported by VMAP). Another approximation prorates the unallocated processor time by taking one third

Table 10 Diminishing returns illustration for a single-processor, N-I/O-server system

Multitasking level	*Number of DASD*	*Percent CPU busy*
4	4	27
4	8	32
4	16	34
8	4	36
8	8	48
8	16	57

of it and prorating it in proportion to the page I/O (read plus write) usage, while the remaining two thirds are allocated in proportion to the user print (spool) I/O usage.

Workloads under the Operating System/Virtual Storage 1 (OS/VS1) are similar to the workloads of MVS systems, with some qualifications. Page size is 2K bytes and blocking factors are probably lower than in MVS environments. With limited OS/VS1 performance data analyzed to date, a relative I/O content of 1.7 is estimated for OS/VS1 commercial batch.

There is no particular reason to change DB/DC relative I/O content estimates for various operating systems because the essential processing requirements and data transfer characteristics do not change, as long as the DB access methods are the same.

Indexed Sequential Access Method (ISAM) data set I/Os are peculiar: data record transfer requires searching on the track from home address to the record. During the search, the I/O path is busy, i.e., the path time per I/O is 8–10 ms longer than that for a comparable Virtual Storage Access Method (VSAM) data transfer.

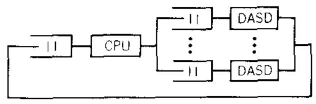
Other software may affect the value of the proposed (default) workload characterization parameters. Deviations should be explained. Defaults can then be customized for an installation, and projection and configuration become more realistic.

law of diminishing returns

The same measurement sample that yielded the relative I/O content characterization findings is used to show DASD and channel utilization behavior. The specific choices for maximum and likely traffic intensities when compared to actual observations are found to be low in about 10 percent of the (maximum) cases and 25 percent of those that are likely.

Another view of diminishing returns is illustrated in Table 10 and Figure 3 by a central server queuing example of a single-processor, N-I/O-server system. If the service times have coefficients

Figure 3 Two-stage central server model, with a single processor and N I/Os



of variation of 5 and 0.5 for the CPU and DASD, respectively, and are gamma-distributed, approximate throughputs can be evaluated under the assumption that wait plus service time of any one server is exponentially distributed. Assume that the total task processing time under conditions of no contention is split into ten units, of which one unit is CPU time and nine units are DASD time. The percentages of CPU busy occurrences for a specified number of DASD and multitasking levels are given in Table 10.

Analytical solution of the queuing model is based on Reference 6, which treats the M/G/1/k queue. In our example, queue times of a server, for example the CPU, represent holding times for the other server, the DASD, and vice versa. This two-stage system can be solved iteratively. If queue times of the servers are nonexponential, the results are inaccurate: Simulation studies show that the CPU busy level is overestimated by 1-10 percent. The estimation error decreases as the multitasking level and/or the number of DASD increases.

We see that doubling the number of like components under otherwise equal conditions does not double the CPU utilization. Like components here means the DASD resource on one hand, and the multitasking resource on the other. In both cases, the returns diminish. No consideration is given to I/O paths or to realistic CPU-I/O behavior. The purpose of the model is simply to illustrate the point of diminishing returns.

tape considerations

It is generally believed that tape channel utilization may exceed DASD channel utilizations, while still giving good responsiveness. The author agrees with that view where the host has unlimited or generous main storage. Tape channel contention in the busy tape channel environment slows down the tasks using the tape channels, increases main storage residency, and results in paging that otherwise need not happen. Thus responsiveness may be affected, unless paging is insignificant.

Tape channels behave differently from DASD channels. The effects of contention are different. However, the channel busy indication given by RMF is consistently low compared to hardware monitor measurements of tape channels that have come to the author's attention. In fact, RMF may show tape channel utilizations in the range of the DASD channel utilizations when tape channels are 1.5 times as busy. Thus, for our purposes, no practical difference exists between tape and DASD channels when RMF-based channel busy data are used.

No information is available to the writer at this time on the accuracy of tape channel usage data reported by VMAP or VSE/PT. In the limited number of investigations to date, the information has been treated the same way as RMF-derived data.

Analysis of a sample of small systems shows relatively large I/O path service times due to the lack of Rotational Position Sensing (RPS). Such systems exhibit DASD channel traffic ms data above the maximum guideline of Table 3. Compensation comes from the presence of a dedicated and inactive tape channel to bring the overall channel traffic (of 2–4 channels) close to the guideline. DASD configuration of these systems is invariably governed by space requirements, as long as 3340-type (low space capacity) disks are used.

Concluding remarks

A simple nonlinear procedure is outlined to size host configurations that can process workload at specified rates. When workload processing involves multiple components of a given kind, doubling the processing rate requires more than doubling the resources. This law of diminishing returns is invariably overlooked even by experienced systems specialists. This paper attempts to shed light on this situation. Workloads are characterized by three parameters: relative I/O content, I/O path, and DASD service time requirements. Workload processing implies a fourth parameter: I/Os processed within unit time. Several workloads have been analyzed to obtain representative estimates of these parameters when specific information is not otherwise available. This characterization is an acceptable first estimate and one that has been found usable for host system capacity planning.

Appendix: acronym descriptions

Acronym	*Description*
AP	Attached processor
BMPX	Block multiplexor
CICS	Customer Information Control System
CMS	Conversational Monitor System
CPU	Central Processing Unit
DASD	Direct Access Storage Device
DB/DC	Data Base/Data Communication
DOS/VSE	Disk Operating System/Virtual Storage Extended
IMS	Information Management System
ISAM	Indexed Sequential Access Method
I/O	Input/Output
MG1	Model Group 1
MP	Multiprocessor
MVS	Multiple Virtual Storages
MVS/SE	MVS System Extensions
OS/VS1	Operating System/Virtual Storage 1
RMF	Resource Measurement Facility
RPS	Rotational Position Sensing

SIO	Start I/O
SMF	System Management Facilities
TSO	Time Sharing Option
UP	Uniprocessor
VMAP	Virtual Machine Analysis Program
VM/370	Virtual Machine /370
VSE/PT	VSE Performance Tool

ACKNOWLEDGMENTS

The author thanks G. Galambos and C. Rolfson for their comments while he was developing the concept; J. Cooper for his suggestions on presenting the ideas; L. Beauchemin for his help in clarifying inconsistencies in earlier drafts; and D. Oldfield for preparing material for the several classes on the subject presented in Quebec and the Maritimes by the author and L. Beauchemin. He also thanks the anonymous referees and G. C. Stierhoff for their critical review and comments.

CITED REFERENCES

1. J. C. Cooper, "A capacity planning methodology," *IBM Systems Journal* **19**, No. 1, 28–45 (1980).
2. *OS/VS2 MVS Resource Management Facility (RMF)*, SC28-0922 (1977); available through IBM branch offices.
3. *VMAP—VM/370 Performance Monitor Analysis*, SB21-2101 and SN20-6537 (1978); available through IBM branch offices.
4. *VSE Performance Tool (VSE/PT)*, SH20-2171 (1980); available through IBM branch offices.
5. *System Management Facilities (SMF)*, SC28-0706 (1976); available through IBM branch offices.
6. L. Takacs, "On a stochastic process concerning some waiting line problems," *Theory of Probability and its Applications* **2**, No. 1, 90–103 (1957).
7. R. G. Duncan and R. A. Brouse, *SHARE/IMS User Profile*, IBM Palo Alto Systems Center Bulletin, Issue PG-RBD04 (March 1975); available through the IBM Marketing Support Systems Center, 1501 California Avenue, Palo Alto, CA 94304.

The author is located at the Eastern Region Field Support Centre, IBM Canada Ltd., 5 Place Ville-Marie, Montreal, Quebec, Canada H3B-2G3.

Reprint Order No. G321-5141.