DRAFT (Revised)
13 July 1966
RGM

MANAGEMENT PROBLEMS OF MULTI-ACCESS COMPUTER SYSTEMS

Introduction

The next three or four years will see the installation of a significant number of major computer facilities—perhaps as many as fifty—whose implementation will draw heavily upon the new computer system technology now under development at MIT Project MAC and elsewhere. These new multi-access systems, which will represent an investment in computer hardware alone of more than three hundred million dollars ($300,000,000), constitute the basis of an important new class of tools for attacking business, commercial, industrial, scientific and engineering problems of many types. As must be expected with the evolution of any new technology, new techniques for the employment of these tools must be and are being developed. Correspondingly, there also arise new issues with respect to the management of these computer systems or—more to the point—of the resource which they represent.

This document is a hasty attempt to point out some of the administrative and accounting problems relating to the operation of multi-access systems. It is addressed to a proposed working group at MIT whose purpose would be to deal with these new management issues in a way which can have long-term benefit, not only for the Institute and other universities, but for all activities,

commercial, military, and government, which may operate multi-access computer systems in the future.

The decisions made at MIT with respect to the management of the Project MAC system may establish tenacious precedents. We must, as we make these decisions, recognize our responsibility to develop an approach which is complete, general, and which meets the needs of users, managers, accountants and auditors.

Objectives of the Working Group

The immediate objective of the proposed working group would be to establish a basis for charging for the services of the Project MAC computer facilities. The system-development effort is expected to reach a stage at which "public" use is possible in the second quarter of 1967, and at that time MAC expects to go on a charging basis for all use of the computer.

Because of the management and accounting questions raised by the innovations in the MAC system concept, this first objective cannot be met without giving overall consideration to its implications for internal accounting practice and interaction with external--particularly Federal Government--purchasing regulations and audit procedures. A second objective, then, would be to establish policies, procedures and guidelines for accounting and charging which constitute a carefully-considered, general

framework which satisfies the immediate needs and which also meets broader and more general needs.

A third objective of the working group would be to communicate to appropriate authorities, both internal to MIT and external (again, notably the Federal Government), the technical aspects of the new form of computing service of which the Multics system represents an instance and to work closely with appropriate authorithies to determine if modifications to existing policies, regulations and practices are called for. If so, the group would work together to formulate such revisions and new practices as may be required, having in mind the problem in its full generality as well as the specific instance at MIT.

Assumptions and Problem Areas

The basic goal in implementing a computer service facility must be to meet the community's demand for computation by supplying the best possible service at the lowest possible price. A basic assumption here and throughout this paper is that system users will pay for all computer service and that the total revenue received from these payments will approximately equal the total cost of providing the service. It is also assumed that if the system capacity is insufficient to meet the demand--real demand, backed by willingness to pay--of the community, then the system

capacity will be increased until that demand is met. This may seem elementary, but remember that in the past, overload situations have had to be met by such unsatisfactory schemes as rationing of resources and installation of "private" computers by some users. The natural expansibility of multi-access systems permits this more desirable approach.

It is clear that under the assumptions stated above the aggregate amount of computer service provided the community will be determined by the cost of "units of service" and by the total funds available to buy them. Again, this is hardly a startling observation, but it emphasizes that computer service is "for sale" in the same "market" as other pertinent commodities. In an academic community, for instance, computer service competes with personnel, floor space, research equipment, etc., for the available funds.

Against this background one may say that the value of the multi-access systems lies in the extent to which they promise to improve both the quality and the quantity of service available per dollar.

Some of the specific problems of meeting the "best service at lowest price" goal with a multi-access computer system are discussed below. This is not intended to be an exhaustive treatment, and in such abbreviated form it cannot be expected to convey the full implications and limitations which may be involved.

In general it should be said that multi-access computer systems represent a totally different class of tool from any that has existed heretofore; one should expect from the outset that the management and accounting practices associated with them must be correspondingly different. It is possible that thoughtful and open-minded consideration of the issues will lead to radical departures from long-established business tradition.

A New Charging Basis

The classical charging unit for the use of computer systems has been the "system hour." That is, the charge for computer service was determined only by the number of hours, measured to the nearest hundredth, used by a given job. Frequently, a different (lower) price per system hour was charged for use during times other than the normal 40-hour work week. In unusual instances the charging rate per system hour might vary slightly depending upon the system configuration actually used; that is, a job which used only two of the dozen or so tape drives connected to a computer might be charged at a slightly lower rate than one which used eight. Even in this case, however, the basic unit of charging has typically been the system hour—in effect, a charge for the use of the entire "roomful" of computer.

Such a charging basis obviously discriminates against the users whose programs involve mainly processing and favors those who fully utilize all of the system components. This inequity is significant in any system in which the cost of peripheral equipment is a large fraction of total system cost (as it is in large, general-use installations, both conventional and multi-access). It may actually prove economic for a heavy "processing-only" user to set up his own stripped-down computer in preference to paying for underutilized peripherals at a central installation. The Laboratory for Nuclear Science is an instance at

MIT of this situation. The unfavorable implications for a central computer-service facility are clear.

Well-designed multi-access systems inherently include means for accounting for the use of the various system components (processors, input/output controllers, etc.) in rather microscopic detail. Instead of a gross charge for use of an entire computer, charges can be associated with the particular devices used, and device usage can be measured with extreme precision (possibly down to the microsecond). The reason for this level of detail is that it is needed by the supervisory programs, quite independent of accounting considerations, to properly schedule and control the various devices which make up the computer system. Roughly speaking, the accounting programs can just "tap off" information which must be present for operational reasons.

As a consequence of the availability of extremely detailed "metering" of the use of each device (e.g., processor, core memory unit, disk-storage controller, etc.), almost all charges for system use, broken down by device, become directly assignable to a user. The proportion of system costs which have to be lumped together and allocated on some necessarily artificial basis becomes very small. The overhead component of system charges virtually disappears, and the man who uses only a processor and a core memory may no longer find it economic to install his own machine.

Other aspects of improved system management also result from precise metering. For example, if a high percentage of idle time is observed on some device—say a drum storage unit—one might consider terminating rental of that device. In general, loads on each system component are individually measurable, so that the system configuration can be altered and its capacity adjusted to meet variations and trends in the aggregate demand imposed by the using community.

Problems of the New Charging Basis

Following is a discussion of some of the problems relative to use measurement that might be considered by the proposed working group.

The first issue that comes to mind is that of pricing—what is a microsecond of processor time worth? A "word-microsecond" of core memory? A full set of prices must be established.

More involved, and more important, how may these prices be permitted to vary? Here it should be pointed out that if only a single fixed charging rate is established for each computer device regardless of system load or time of day then the system will experience extremely heavy loads during the 9-to-5 work day. Remembering that we have for argument sake taken the position that our response to any adequately-funded demand for service is to

expand the system if necessary to meet it, we are soon operating an enormous computer facility to meet the peak demand. During off-peak hours, however, such a large system would be seriously underutilized, and as a result, a unit of computer service ~~has~~ would have to be priced very high in order to recover the total system cost. This peak-loading effect on the system can be alleviated by the usual approach of offering bargain rates for off-peak use. We propose such a rate structure for MAC, possibly involving several steps throughout the 24-hour day. The rate structure may also recognize weekend days as deserving of special treatment.

It is even conceivable that in some later stage of development we might permit the prices to vary from minute to minute as determined by actual system load measurements, or we might permit prospective users "bid" for system services, with high bidders paying more but receiving preferential treatment.

All of these possibilities and a multitude of others are technically feasible, but how would they affect the users, system managers, accountants and auditors? Permitting prices to fluctuate in a manner not completely known (or understood) by a user may cause "customer-relations" problems. Even reasonably conservative "stepwise" reductions in off-peak prices introduce complications. For instance, 1) the cost of running exactly the same job can vary,

depending on when it is run, 2) auditing "proof" of detailed charges becomes messy, and 3) a priori cost estimating becomes more difficult.

There is a possibility, if certain pricing strategies are adopted, that auditors may discover that they must prove the procedure--possibly the actual computer program--used to compute charges, rather than the charges themselves. Is this a realistic possibility?

Multiprogramming and Multiprocessing

In a slightly different but related area lies another set of problems which are raised by the techniques of multiprogramming and multiprocessing.

In the conventional use of computer, jobs are "queued up" and run sequentially; that is, one user's program is run to completion before the next user is started. There is no way in which one user could, by accident or design, affect the execution time (and, therefore, the cost) of another user's job.

Multi-access systems operate in a fundamentally different manner. One can visualize the difference by imagining the entire queue of users turning 90 degrees and approaching the computer broadside. By using techniques of "multiprogramming" and "multiprocessing" (which will not be precisely defined here) this

broadside demand can be met. However, the new situation has implications for both the system and for the users.

From the system point of view, there is now a much wider selection of tasks awaiting the attention of one or another of the devices that make up the computer. A clever supervisory program can therefore juggle the scheduling of this multitude of tasks so as to keep most of the computer devices busy most of the time. The vastly-improved hardware utilization achievable in systems of this kind is one of the primary benefits of the new technology.

From the user's viewpoint there are also great benefits. For one thing, he can operate "on line" directly to the computer without worrying about, or--more important--being charged for, computer time during periods when he is "connected" to the system but not actually receiving service. The unexpectedly great value of this new kind of access to computers is the greatest single motivation for the work of Project MAC.

However, there are some potential problems in providing this access. One which may be troublesome from a management point of view is that the service received by a user may be strongly affected by the presence of other users on the system. Make no mistake; this does not mean that the service seen by active users of the system at peak-load periods will be sluggish and unsatisfactory; long experience has made clear that it is better to provide

adequate service to a restricted set of users than unsatisfactory service to all customers. Thus an important measure of the extent of system overload is the number of users who are denied access, and system capacity must be increased when this measure ~~is~~ indicates the need.

Another problem, although a minor one, is the fact that, completely aside from any variable-pricing schemes, the cost to run a given job will vary from time to time in a manner which is both not predictable and not easily understandable by the average user. The range of variation will probably be small, and it will diminish as the system capacity grows and the population from which the statistics are drawn becomes large.

Problems of Multiprocessing and Multiprogramming

Problems for the working group relate particularly to this system-related variability in service and cost. Will users accept it? Will auditors, sponsors, and proposal evaluators? Will there be a need for one or more of these kinds of people to develop test jobs to be run occasionally for the purpose of measuring variance? These difficulties will affect primarily so-called production jobs; that is, relatively long-running jobs involving little or no interaction and using programs which are thoroughly debugged. For this kind of work if the cost variability concerns the user, then he may be offered a "flat rate" for his job based on past experience. If the range of variability actually proves to be very

limit flops

small, he may prefer to ignore it and take his chances by running on a normal "metered" basis.

Other Problem Areas

So far we have discussed some of the management problems of a basic and long-term nature. Here we will just mention ~~two~~ three others and then move on to state some of the more immediate problems relative to the MAC system and its interaction with the rest of the Institute and with the MAC sponsors.

Aside from system use and storage occupancy, there may be three other valuable resources for which users can be charged. The first of these is access. There is a distinct value associated with the ability to approach the system at will. An obvious part of the charge for access is made up of at least some of the cost of the terminals and communication circuits used. There may be other parts not so clearly discernible.

The concept of demand as a multi-access system resource must be considered. The term is borrowed from the "demand charge" component of commercial electric-power rates, and the analogy is good, if not perfect. The power company imposes a demand charge to recover some portion of the cost of the physical plant it has had to install in order to meet the heaviest ~~single~~ load a user may impose on the system. On a multi-access computer system a

demand charge is imposed if the user requires some guaranteed reservation of system resources. Such a guarantee removes those resources from the common pool which the supervisory program would otherwise be able to share among the community of users; therefore the users who make the reservation must pay for it.

A third possible candidate for inclusion as a chargeable resource is responsiveness. Speed of response is certainly related to system capacity, and it should be considered in setting up a charging structure.

In a more practical vein, here are several of the issues which will directly confront MAC when its new system becomes operational:

How can "surplus" MAC system capacity be "sold" to non-MAC users? The proceeds should reduce the cost to MAC of its computer facility. What is an accounting mechanism which permits the sale and satisfies MIT interests, the MAC-contract monitors and the monitoring agency for the purchaser's contract? What account should the funds be paid into? Can the funds be used for purposes other than paying direct costs of operating the system?

At least for a year or two, the MAC system will be a unique technical facility with capabilities of great worldwide interest. Suppose the system is made accessible to non-MIT agencies; how should external use be managed? Are there different considerations

for different users? Government? Not-for-profit research? Professional groups? Other universities? Industry? Others?

Conclusion

We view adequate consideration of the relevant management issues as an inseparable part of the introduction of the new technology of multi-access systems. It is for this reason that these issues, beyond their immediate implications for MIT, are of long-term research interest to Project MAC. We may discover a need to overturn well-established business-management traditions, or to alter widespread accounting practices.

There is immediate and urgent need for a local determination of policy. It seems both reasonable and necessary to keep in mind the wider implications while meeting the local need.

A specific proposal for a charging structure and accounting procedure is in preparation and will be distributed soon.