FERMILAB CENTRAL COMPUTING STRATEGY FOR THE MID-80's

Al Brenner

Introduction

The intent of this paper is to outline in broad terms the overall directions and plans which are being laid for the evolution of the Central Computing Facility at Fermilab. No detailed time schedule is given because of the number of factors which make that very difficult. The computing industry and the availability of computing components and the styles and packages in which they come are becoming more and more diverse at a very rapid rate. Also, the integration of these rapidly changing technologies and approaches into an ongoing dynamic operational situation is very complex. Finally, the whole activity is paced almost entirely by the level of funding and staffing that one can apply to this problem.

During the last decade, corresponding to the operational life of the Laboratory, the computing industry has moved along briskly, but in a relatively monolithic direction. Now, however, the level and style of interaction possible, and also the mechanisms of accomplishing number crunching procedures, are in the process of a major diversification. This gives rise to enormous opportunities for improving our future computing capabilities, but carries with it the heavy burden of bringing all this together into a unified system. The Report of the **ad hoc** Committee on Future Computing Needs for Fermilab (the Ballam Committee Report) summarized in an accompanying article, has recognized these opportunities and burdens and has recommended that the Laboratory assign more resources to support these computing activities then it has heretofore. The rate at which the program outlined here will come to pass is strongly dependent upon our ability to carry out the recommendations of that Report.

The Current State of Affairs

The Fermilab Central Computing Facility currently consists of a complex of three Cyber 175 computers which were installed at the end of 1978. They are currently quite saturated, although normal priority production jobs typically do turn around overnight. Interactive (terminal - CRT) jobs are responsive, but are limited in both memory allocation and in CP time for each user session. The latter is hardly a constraint, but the former is becoming a more and more serious limitation to the user community.

About 90% of the CP time is utilized by the experimental physics community setting up for and analyzing data. The remaining 10% is mostly utilized for accelerator design, structural analysis, magnet design, and other engineering functions. A small fraction currently is utilized by the theoretical physics community. A substantial number of non-physics users also utilize the facility and constitute a third or more of the interactive population. Their use of CP time, however, is negligible. These include most of the secretaries at the Laboratory and a small number of people doing computer-aided design (CAD) work.

All users on the site with direct hard-wired connections to the computing facility, as well as those off-site accessing the facilities via common carrier lines, have access to the computers under management control of the Computing Department, as well as a few other computers. This is accomplished with a very modern MICOM port-selector "cross-bar" switch which gives good flexibility to the user community. Bandwidths up to 9600 baud are possible. Thus, from any given terminal, access is possible to any of the connected computers, or with a dial-out capability over common carrier lines, to any of a limited number of allowed facilities outside the Laboratory.

Currently, there is no direct connection to ARPANET or to any of the value-added networks. ARPANET is accessible when required through a connection to Argonne. Currently, about a dozen user groups are connnected over dedicated common carrier lines, usually at 9.6 Kb, to the Fermilab MICOM data switch. By the end of this year, we will be connected into the BITNET network.

We have recently committed and started installing a broadband multiple T1 (1.54 MHz) system connecting Wilson Hall and the general B0/Industrial Area complex. There will also be installed in that complex a satellite MICOM port selector. That will give us broadband communications for a number of different lines between the B0/Industrial complex and Wilson Hall. Furthermore, that connection will be extended farther out to the other three experimental areas during the next year. Some of these connections will be used for machine-to-machine communications; in the first instance between the CDF data-acquisition and off-line VAX's.

Projections for the immediate future in terms of raw computing power required are as difficult as ever to make, but for the near-term early Tevatron period, the techniques we have used heretofore should continue to work. Thus, we project a need for a factor or two more computing than we have right now before the end of calendar '85. By that time, TeV I will start operating and the projections further into the future are very much more difficult to make. Without doubt, the growth slope will start increasing at that point. Thus, by the end of 1986 we will be in a very serious overload condition and will again require major additional equipment.

Our system is nicely backed-up for archival purposes with the Automatic Tape Library (ATL). This system is fully operational and is an enormous aid in terms of storing and retrieving files with a minimum of hassle for the operations group and with minimum delay for the user. The ATL in this environment is very effective. It is able to contain at current rates one year's worth of active disk files.

For interactive terminals, which are quite uniform across the Laboratory, we have a large number of dumb ADM-3A terminals. A substantial number of these have graphics-board extensions. The terminals have been very effective, inexpensive, and stable. New competition has come into play in the last year or so with somewhat more-powerful terminals. We have slowly moved in that direction as well. We have also been searching for an appropriate terminal on which to allow for output previewing. We will make a decision on that terminal in the near future.

With respect to the question of upgrading our facilities to take advantage of modern interactive graphics techniques, over the years we have developed our own internal global device-independent graphics system (DIGS) similar to what most of the laboratories have done, each in its own individual style. For interactive graphics, we mostly use 4010 protocols and we support a number of different terminals using 4010 graphics mode, including Retrographics Corporation enhancements to the dumb ADM-3A terminals that we use. These have found widespread and satisfactory use, at a very good price, across the whole Laboratory. The device independent graphics system also allows us to handle a variety of on-line graphics devices, including a number of Calcomp drum plotters and two high-quality Benson Varian medium speed electrostatic plotters. The latter, in particular, have been very successful in supplying high-quality graphics output with very good turn around.

Recently, after a search for more than two years, we have chosen a low-cost, reasonably high resolution, high quality color graphic system. A number of these Envision terminals are on order and are beginning to be allocated to the user community. As part of this process, we are setting up a do-it-yourself transparency and slide making facility for the convenience of the users.

In other dimensions, we have made available simple database-management systems, likely to be convenient for the management of scientific as well as administrative data. We also are in the process of introducing new scientific high-quality printing terminals so that papers and letters can be output directly with an extended character set including Greek and mathematical symbols. Associated with this, we are also having developed a new dumb terminal which will present both Greek and mathematical symbols directly as typed.

The Near-Term Plan

Conceptually, the plan at the time of the FY79 acquisition was to add "a fourth mainframe" during the FY83/84 time period to enhance the currently-installed system. All the necessary connections were provided at the time of that acquisition to make this possible. Since the "fourth mainframe" in the sense described did not exist as a marketplace item, earlier this year we broadly opened a competitive procurement to obtain enhancement at the same level, i.e., to at least twice that which we currently have installed to carry us through this early Tevatron period. Since a shortage of computing cycles was viewed as the greatest short-term (and indeed longer term) problem, the RFP procurement most strongly valued this aspect of the system. Bids were recieved from Amdahl, CDC, Cray, and IBM. A more detailed report on this computer acquisition appeared in the November 1983 issue of **Fermilab Report.**

CDC's bid for a dual CPU Cyber 875 system with 750K words of memory has been selected. The first CPU was delivered early in December. The remaining equipment and software is scheduled for delivery during the next 21 months. The total installed computing power in the Central Computing Facility at that time will be the equivalent of seven Cyber 175 CPU's. In addition to the CPU's and associated channels and controllers that are required, additional peripheral devices will come with the system. The installed disk capacity will double in size from about 15 gigabytes to a total of 30 gigabytes. Also, the high-speed printing capacity will be doubled. As part of this acquisition, during the next year a major upgrade of the operating system from NOS 1 to NOS 2 will also occur. With this change will come a number of added features giving rise to a more powerful operating environment for the user community.

The Cyber 875 will be connected to the existing Cyber and other local systems via a Network Systems Hyperchannel connection. We also have committed and will shortly install two quite large VAX 780 computers which will be configured in a cluster network themselves, as well as being separately connected to the Hyperchannel. This cluster will serve as an alternative front end to all systems. It is our intention to continue this expansion of options for alternate interactive mechanisms using the existing old architecture machines and the computers about to be acquired as the computing workhorse engines, behind these new interactive systems. Next year will see the addition of an IBM style alternative front end to the network.

We are planning to introduce Ethernet-type connections throughout Wilson Hall for general interconnectability via a gateway to the central local area network. The details of this connection are only conceptualized and are not yet detailed enough to proceed. It is only recently that Ethernet has stabilized and become a real product, supported by industry with real hardware and software.

With respect to graphics enhancements, we are watching very closely the status of the U.S. Standards Committee relative to both the GKS and Core (the U.S. nominal standard) protocols. We fully expect GKS or some minor variant of it to be adopted as the U.S. standard, and as soon as that happens, we will put major effort behind changing our local, unique DIGS standard to that one.

With the current explosion in the rate of introduction of new powerful desktop workstations, we are continuously tracking the availability in the marketplace in this arena. Clearly new ideas and new options are being introduced at a very rapid rate. Currently, our perception is that there is no obviously correct system to which Fermilab should commit. Furthermore, we fully expect to have our hands full making all the new items we have coming in with the current new acquisition work. It is our intention to continue to track the availablity and performance of workstations. Probably next year, when we have stablilized our new acquisition and integrated it into a total system and made it useful to the user community, we will turn our attention once again to the workstation question. We hope that new and stabilized workstation equipment will then be available, and we will evaluate seriously the choice of workstation(s) appropriate to the Fermilab needs.

As indicated earlier, we will continue to expand and increase the level of integration of a number of networking systems that we use. We will install an Ethernet system, enlarge the Hyperchannel network and connect BO/Industrial Area to the Central Laboratory via a T1 broad-band connection. The latter will then be expanded to other experimental area locations and also to the cross gallery.

We are also interested in expanding our support for engineering computing during this period. This includes an enhanced CAD capability, extended finite-element analysis tools, magnet-design programs, alternative word-processing capability, and more direct support of microcomputers.

Longer-Range Expectations

The major problem with any medium-term solution now is that none of the normal architectural enhancements, neither IBM-like architectures nor any of the other extant architectures, is likely to bring with it at a cost that is even imaginably tractable the necessary computing that we will require, once the Tevatron starts functioning and producing data at the rates expected. Even assuming that we do everything we can to preprocess in real time as much of the data as possible, we will still need a factor over our current computing level exceeding 10 and possibly approaching 100 to handle the expected data rate as we approach the end of the decade. We will start needing more than the factor of two that we are currently installing as early as

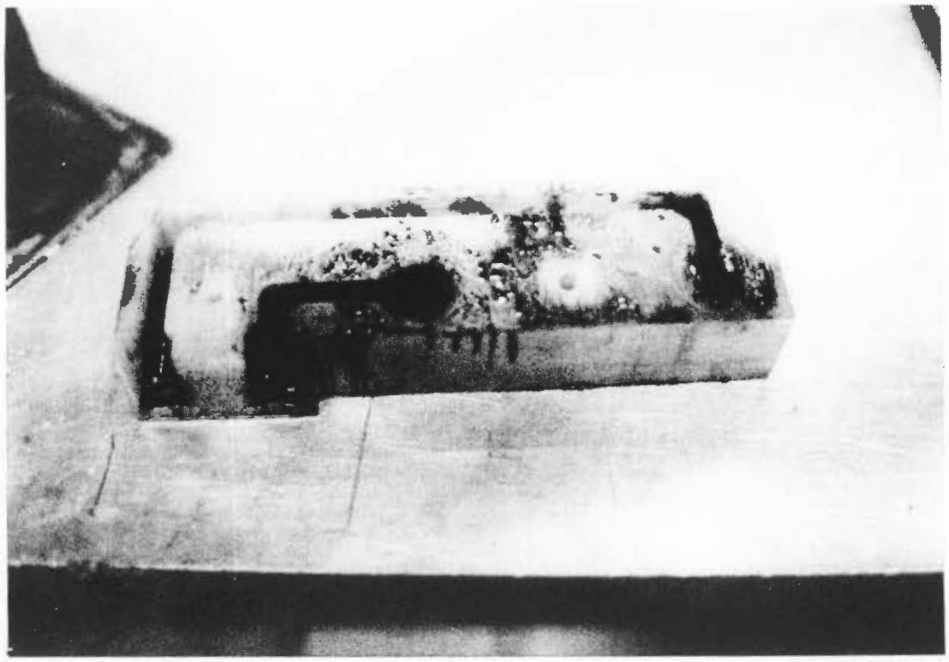
the end of 1985. No announced product is likely to be able to provide at reasonable cost the level of computing that we project as our need. On the other hand, there are a number of relevant activities currently in progress, with lots of interest on the part of government, laboratories, and universities. The message is being heard; the potential Japanese commercial encroachment also is having useful effect in inducing the various companies in this marketplace to come up with new products, some of which may be appropriate for the high-energy physics community.

Also, there are a number of activities being pursued in developing special-purpose processors, particularly cost effective for the needs of high-energy physics data-analysis computing. These include IBM emulators, the 3081/E being designed by SLAC and CERN and the 370E being designed by the Weizman Institute in collaboration with a number of European laboratories, the product of the Fermilab Advanced Computer Project (ACP) and a number of other university and industry-based projects. One or more of these will certainly play an important role in the solution to the Fermilab computing problem during the second half of the decade.

In our long-range Automatic Data Processing Plan, we have set aside $15M for major new architectural equipment in FY86/87. As indicated above, there is no collection of machines today that carries with it in a natural way the capability of computing at levels which we require. All indications are that the vector machines, even the dual-headed vector machines, i.e., the Cray X-MP, will not solve our problems. Neither do any of the IBM or IBM look-alikes, nor do any of the other available machines that function today. On the other hand, all the important manufacturers, e.g., Amdahl, CDC, Cray, Denelcor, IBM, and Trilogy as well as all three of the major Japanese companies, Fijitsu, Hitachi, and NEC, are committed to big new machines in the relevant time frame. In many cases, however, these are vector machines and what performance we would achieve with our job mix is totally unknown to us at this point. Nevertheless, we have been making known and our benchmarks have demonstrated our needs to these firms. In this way, we are trying to encourage all of the companies, especially the new ones like Denelcor, whose architecture appears to be better suited to our needs than that of any other machine, to pursue their efforts with new technology and, most importantly, with new architectural approaches.

However, if we take a route like the Denelcor route, the chances of there being a sophisticated operating system with all sorts of productivity-enhancing tools on it, is very small. The only way in which such a machine will end up being both a useful computing engine and a useful productivity tool for the user is by appropriately networking it with other machines. We are in the process of learning how to network various machines, as is the rest of the world just learning. We are sensitive to the availability of these different tools, and we believe that we will put together an appropriate mixture of productivity tools

connected to computing engines to handle the Fermilab computing needs as they grow.



Part of a burned-out spool piece (correction magnet assembly) removed from the Energy Saver.
(Photograph by Fermilab Photo Unit)