Contents

**Computing** 287

1 Introduction 287

2 Past Progress and Current Status 287

3 Mode of operation 289

4 Technology Needs 289

5 Recommendations 292

# COMPUTING

KENNETH MCFARLANE

*CEBAF, Newport News, VA 23606, USA*

## Working Group Members

S.C. Loken, *Lawrence Berkeley Laboratory*
R.P. Mount *California Institute of Technology*
C. Watson, *CEBAF*
J. Womersley *Fermilab*

## 1 Introduction

High-energy physics has made heavy use of computational technology throughout the development of both fields. Common roots include the first "AND" gate using vacuum tubes, and the development of Monte Carlo simulation techniques for neutron transport during the Manhattan Project. Now, "computing in HEP" is taken to mean application of (primarily) digital computer hardware and software in a variety of overlapping areas, including DAQ/event building, control and monitor systems, on-line and off-line filter and reconstruction systems, networks and interconnections, data handling and storage systems, user interfaces and visualization, information systems and communications, simulations, and analysis systems.

We discuss these various aspects of computing under three headings:

- software systems and tools,
- computing hardware, and
- information systems and communications (including both hardware and software).

The remarkable progress in computing, along with early HEP adoption and adaptation of hardware developments in computing and electronics, has enabled the HEP program to be successful. There are, however, issues in each of the above areas which will require significant R&D and capital investment to meet the future needs of HEP. Within the U.S., HEP laboratories have not produced a coherent body of software with scope anything close to that produced by CERN. With the cancelation of the SSC project, the U.S. has lost both a source of funding for, and a vehicle for coordination of, computing R&D. We believe that the proportion of the HEP effort devoted to computing, and the level of coordination, must be increased if HEP computing technology is to keep pace with other detector developments.

In the following, we discuss past progress and the current status, review likely technology needs, and make recommendations for the future.

## 2 Past Progress and Current Status

Advances in computing made within HEP have been critical to the success of the HEP program, and some developments have had, and are having, wide influence on computing in general, despite the fact that HEP is a now a relatively small part of the user community. There have been notable HEP-industry associations, from which both have benefited. HEP-industry projects have influenced IBM (HEPVM, REXX, storage management, tape robots, real-time AIX, WDSF), DEC (VAX/VMS to STK silo connection), STK (tests of high-capacity cartridges). An L3-IBM project has implemented a high-performance system using HiPPI. HEP representatives sit on computing industry Customer Councils and Advisory boards, as well as standards bodies. Computer "farms" (*e.g.*, ACP at FNAL, SHIFT at CERN, PDSF at SSCL) have influenced commercial products.

The advent of the SSC and LHC created new challenges to computing systems. After some study, it appeared that (Level 1) trigger systems could reduce the trigger rate for general-purpose detectors by about three orders of magnitude, from the collision rate of 60 MHz. Given expected event sizes, the task of DAQ and computing was to take a data rate of about 30 GB/s of event fragments and make that consistent with storage technology (a rate to storage of order 100 MB/s seemed possible) and analysis requirements, in a cost-effective way. Most of the resulting issues re-appear in near-future experimental environments (*e.g.*, Tevatron collider upgrades and high-rate fixed-target experiments such as KTeV) and future colliders (*e.g.*, LHC, SLAC PEP-II, and RHIC). To address the challenges, a variety of R&D projects in DAQ and computing were begun: in the SSCL Generic R&D Program, in the SDC and GEM R&D programs, within the SSCL, and in concert with the DOE part of HPCCI. (Some funding also came from the TNRLC.)

Projects carried out in the SSC context included:

- studies of massively-parallel closely-coupled computing systems for DAQ (for B physics),
- studies of CPU farms for on-line filters and production,

- trials of commercial network components for DAQ,
- design of DAQ/event builders in combination with on-line filtering,
- R&D on data-base technology for event storage (the PASS project),
- development of detailed software concepts for off-line computing, and
- donstruction and operation of heterogeneous computing systems based on UNIX (*e.g.*, PDSF at SSCL).

A notable characteristic of these efforts was the use of modeling and discrete simulation of the hardware systems, to an extent new in our field. Several vendors participated in projects, and the PASS (Petabyte Access Storage Solutions) project (in which SSCL was one participant) is continuing. Also, there was a significant effort to involve computer professionals in the SSC program. Below, we give some examples of past progress and achievements (this is not an exhaustive list and we hope that any omissions are drawn to our attention), and describe the current status of computing in HEP. Next, we comment on the mode of operation and funding approaches.

### *2.1 Software Systems and Tools*

The character of HEP has resulted in many software developments designed to enhance the productivity of physicists who are not computer specialists; these developments have been freely available and shared internationally. Notable ones include:

- standard packages, beginning with such programs as SUMX, KIOWA, and TVGP, which were crucial in the era of resonance studies,
- early use of interactive/time-share computing (Wylbur),
- HEPVM, with priority batch and tape-handling systems,
- tools built to extend the capabilities of FORTRAN (Mortran and ZEBRA),
- EGS, for electromagnetic shower simulation, and
- the CERNLIB package, with many tools, including detector simulation (GEANT), analysis (PAW), fitting (MINUIT), numerical analysis, and the ZEBRA memory management/data structure system, that were ported to many systems.

EGS has been widely adopted for simulation of medical radiation treatment systems, and CERNLIB has been used outside the HEP community. Industry software standards committees, such as those for POSIX and FORTRAN) have had HEP representatives. The current status of software technology in HEP can be summarized as follows:

- Tools such as GEANT and PAW are the HEP standard (*e.g.*, CDF's primary tool for their top analysis was PAW).
- database technology has been clearly shown to be necessary for Large experiments.
- UNIX has gained acceptance in HEP (for many experiments, the bulk of the CPU power is in UNIX systems).
- Beginning at the detector end (DAQ, slow control), C and C++ are supplanting FORTRAN as the language of choice; many projects using an object-oriented approach are underway.
- Software methodologies and tools (particularly SA/SD) have been used with some success; tools for developing object-oriented systems are coming into use.

The software situation is less than satisfactory: there have been several cases where insufficient attention to software has led to loss of results. HEP is behind the computing industry in the use of new operating systems and languages, and software engineering approaches. Whereas 10 years ago HEP graduate students received a valuable training in computing, many potential students are now repelled by the need to learn and use outdated and inefficient software techniques.

### *2.2 Computing Hardware*

The HEP program has always had a high demand for cost-effective CPU power, and has needed access to very large data sets. Data acquisition from electronic detectors is also a critical application of computing and storage hardware technologies. These needs have driven a number of innovations and adaptations:

- emulators as an inexpensive source of CPU cycles, for analysis and on-line filters;
- pioneering use of computers for real-time DAQ (PDP-1, 8, 11), in conjunction with standards such as CAMAC (CAMAC was adopted by IEEE as a U.S. standard for DAQ and control);
- early trial of videotape for storage;
- farms, adopted by industry for some applications; and
- massively-parallel computing systems (MPP) developed for lattice QCD calculations (following the demonstration that problems can be decomposed onto MPP) have been crucial to the development of industry MPP systems.

HEP representatives have been involved in developing hardware standards: SCI and CAMAC are examples.

The current technical status of computing hardware in HEP can be summarized as follows:

- Data-compression at the front end is increasingly used to reduce bandwidth requirements, at the expense of more CPU power being needed later.

- The improving cost, performance, and power consumption of VLSI CPU's is leading to their use closer and closer to the front end of the DAQ chain.
- CPU farms, using volume-produced commercial boxes, are established as a cost-effective and manageable approach to computing power.
- Modern multiprocessor "mainframes" with I/O capacity of hundreds MB/s (megabytes/second) are used for data-intensive physics analysis.
- Tape, and tape robots, can handle the present generation of experiments.
- High-performance tapes (*e.g.*, Ampex DD-2) are being used for data storage.

The availability of hardware resources to U.S. HEP is nearly adequate: fortunately the extremely rapid cost and performance improvement of industry products has more or less kept pace with (even unanticipated) HEP needs and budgets. There is presently a CPU resource shortage (of a factor of two to four) which affects the ability to run simulations of existing and proposed detectors and to carry out needed reconstruction and DST production of existing data. Up to its demise in mid-1994, the PDSF at SSCL was still heavily exploited by U.S. HEP.

### *2.3 Information and Communication Systems*

Information and communication systems have been critical to the rapid progress in HEP; HEP has been a model for "getting online." The experiments run by large collaborations could not be managed without e-mail. File and information sharing over networks is also critical to rapid analysis of data and dissemination of results.

Notable developments include:

- early extensive use of e-mail for collaborative efforts;
- the SPIRES data bases and the QSPIRES no-login access system;
- World-Wide Web; this hypertext information system, which does not require logging into the computing system carrying the information, was originated at CERN and has now spread far beyond HEP;
- networking installed in Europe for HEP has been extremely influential in developing the infrastructure there, and links to China, Russia, and Brazil have had important impacts; and
- HEP is an active participant in development of desktop video-conferencing over the network.

The current technical status of information and communication systems in HEP can be summarized as follows:

- Video-conferencing is in increasing use and was essential in the development of SSC collaborations, including links to Japan.
- World-Wide Web is being used extensively to support collaborative efforts.
- Commercial packages for documents (text and figures) are being used; there are incompatibilities which inhibit collaborative efforts.

Information and communication systems are currently adequate within the U.S. and parts of Europe. International links have been restricted by financial and political considerations, and by lack of communications infrastructure within regions that are active in HEP, such as China and the former Soviet Union.

## 3 Mode of operation

The current administrative status for computing in HEP is unsatisfactory from a number of perspectives. The demise of the SSC has removed an important source of funding. While the CERN LHC RD program continues with its support of computing projects, the U.S. has lost a central focus and vehicle for coordination. Budgets for computing organizations in HEP are flat, and many costs are hidden and dispersed, making cost/benefit optimization and coordination of programs difficult. R&D funds for computing in HEP are small compared with other R&D funds: in the SSC Generic R&D program, funding for computing was 3% of the total.

Besides funding of R&D work by DOE and other agencies, another avenue for making progress is through joint work with industry. This has been quite successful at CERN, which is seen by vendors as a display point for Europe; U.S. HEP laboratories have not been viewed as demonstration sites to the same extent.

For the U.S. high energy physicist, computing is not seen as an advantageous career path, though computing systems need to be built just as detectors need to be built. This has had disadvantages: problems are solved only in the immediate context, new languages and operating systems are adopted slowly, and software is poorly documented and maintained. Within the U.S., HEP laboratories have not cooperated to produce a coherent body of software with scope anything close to that produced by CERN.

## 4 Technology Needs

Technology needs of near-term and future experiments can be measured by the Level 1 trigger rate, event size (which is a measure of complexity), and acceptable rate to storage. Current experiments operate at a DAQ rate (product of L1 rate and event size) of about 30 MB/s with data rates to storage of order 3 MB/s. In the future, fixed-target experiments will have DAQ rates of 2 GB/s (*e.g.*, KTeV) and high-luminosity hadron collider experiments, 40 GB/s (see Table I, taken from various sources). Rates to storage will be about 3 MB/s and 100 MB/s respectively. For typical runs, this implies data samples up to 1 PB. The rates anticipated for LHC detectors very likely will be the maximum for the foreseeable future: the next generation of colliders will likely trade higher energy for lower luminosity.

The high data rates and large data samples place demands on hardware, and the large reductions from DAQ rates to storage rates, along with the need to simulate and analyze, place demands on software. In addition, large collaborations need improved information and communication systems. Both simulation and reconstruction will require the computing power available to an experiment to be increased by three orders of magnitude. We must fully exploit available technologies so that a shrinking or stable community can maximize its productivity.

Opportunities for HEP to contribute to the computing industry may include the systems that can handle DAQ and analysis of large samples of "events." The analogy of a shopping cart full of items to an HEP event has been suggested. A large chain of stores might generate up to 100 Hz of customer checkouts, with perhaps 2 kB of data per checkout, a rate quite comparable with LEP experiments, but occurring day in and day out. Each event will have correlations between purchases; these are already being exploited by retailers with point-of-sale terminals to increase the market impact of coupon campaigns, but the ability to record and analyze large data sets efficiently may give industry more sophisticated capabilities.

### *4.1 Software Systems and Tools*

Software systems must be improved so as to enhance physicist productivity. An "engineering" approach should be taken: the practices of the most successful experiments and the best tools of industry should become standard. A way must be found within this framework to overcome the inevitable conservatism of HEP software: the very success of the HEP tools in use inhibits a shift to new methods. Complex systems with many tools that depend on detailed information about the data structures become very hard to change.

Issues to be addressed are:

- software management for distributed efforts;
- certification and control of code, especially for on-line filter systems;
- flexibility, to take advantage of evolving technology and requirements without disruption of on-going work;
- the use of commercial tools (*e.g.*, software development environments, analysis packages, distributed system management);
- analysis frameworks that facilitate the posing of physics questions;
- data bases and data access tools;
- computing professionals as part of teams, or hired to produce specific software systems; and
- interoperability, to allow the use of heterogeneous systems.

### *4.2 Computing Hardware*

Computing hardware provides CPU power, interconnections in distributed systems (including DAQ systems), data storage, and data access systems. The technical needs of near-term experiments appear to be satisfiable with existing hardware, or hardware available in the near term. Future experiments are more demanding and extrapolations of computing cost and performance less certain.

The main areas of need for high-luminosity experiments are:

- specialized systems for trigger and DAQ,
- data storage and data access, and
- CPU power for physics simulations and event reconstruction.

Trigger and DAQ systems will use some combination of off-the-shelf items and special-purpose items, and will be tailored to fit available technology and the specific needs of experiments. Estimates of CPU power needed for experiments vary from 1-10 kMIPS for KTeV in the next two or three years, to 1,000 to 10,000 kMIPS for LHC in ten years. This includes CPU power needed for simulation, which is likely to equal or exceed that needed for reconstruction. It should be noted that the Physics and Simulation Facility (PDSF) at SSCL was a 4 kMIPS facility and was heavily used, after the cancelation of SSC, for about a year for HEP work (unrelated to SSC). Rates in data-storage and data-access systems can be very high; systems currently being envisaged will require many links operating at 100 MB/s. For access to data and distributed analysis, network bandwidths of order 1 GB/s will be needed. New data base technologies will impose new demands on hardware, as will desktop video-conferencing.

There will be two major needs in HEP that will likely require substantial adaptation of what we can expect from industry:

- data compression and pipelining in the front end and DAQ, which will require CPU power and high-speed communications in specially-configured systems to build events and unpack the data;
- storage of data and access for analysis. The data-access requirements will dominate the bandwidth and structure of the storage systems.

Computing systems have been evolving especially rapidly since 1980, in performance and cost, as a result of a number of factors: VLSI technology, hardware-software optimization (RISC), portable operating systems (UNIX), volume production, competition, and widespread use (from less than 200 systems on networks fifteen years ago, there are now millions). Current trends are that:

Table 1: Estimated rates and event sizes for various experiments

| Experiment | L1 Rate | Event size | Data rate (peak) | Rate to storage (mean) |
|---|---|---|---|---|
| CDF | | | | 1 TB/yr |
| KTeV | 500 kHz | 4 kB | 2 GB/s | 3.3 MB/s |
| KEK B | 200 Hz | 30 kB | 6 MB/s | |
| OPAL | 3 Hz | 100 kB | 0.3 MB/s | 0.3 MB/s |
| DØ | 150 Hz | 400 kB | 60 MB/s | 1 MB/s |
| GEM | 100 kHz | 300 kB | 30 GB/s | 1 MB/s |
| LSND | 100 Hz | 2 kB | 0.2 MB/s | 0.2 MB/s |
| CMS | 50 kHz | 500 kB | 25 GB/s | 20 MB/s |
| FNAL Run 3 EOI | 5 kHz | 300 kB | 2 GB/s | 60 MB/s |

- The power of a single CPU increases by a factor of 1.4–1.6 per year.
- CPU power drops in cost by a factor of 1.6–1.8 each year.
- Fast memory drops a factor of 1.3 per year.
- Slow memory drops by a factor of 1.4 per year.
- Magnetic disks drop a factor of 1.2–1.7 per year.
- The cost of bandwidth into mass storage drops by a factor of 1.1 per year.
- The speed of communication links is expected to change in steps: perhaps a factor of three every four years.

These trends mean that a constant expenditure of funds on equipment will buy double the power every two years. Measuring computing power in VAX-11/780 equivalents (VUPs), a million VUPs in workstations cost about $130M in 1993, and will cost $10M in 1999. The cost of bandwidth to mass storage is declining more slowly; nevertheless, from 1993 to 1999 a factor of three decrease can be expected. I/O bandwidth will likely evolve in a more step-wise fashion, *e.g.*, by ATM or Fibre Channel replacing Ethernet and SCSI.

These trends appear to put the CPU power required, and storage for pre-LHC era experiments, within an acceptable cost range. Even for LHC, computing hardware costs may be a small fraction of the total detector cost, provided that volume-produced systems continue to use chip area for CPU's and caches, and not for signal processors for special tasks such as speech recognition or image processing. Network technology appears to be advancing at an appropriate rate. The area of data base technology is also advancing, but LHC experiments may have difficulty in meeting their cost and performance requirements in this area.

### *4.3 Information Systems and Communications*

With the large collaborations and international character of HEP, providing consistent data and rapid communication will be of prime importance for our field's productivity in future years. During the SSC era, physicists were added to the U.S. HEP community, and there were new university groups, resulting in more groups, smaller groups, and scattered groups, increasing the need for effective communication. Several collaborations have found that publication of results is delayed by the pace of communication within a widely dispersed group.

Communications within collaborations will need to be supported in the following areas:

- videoconferencing at every location;
- systems to support "virtual meetings" across widely separated time-zones'
- document standards, including graphics, to permit electronic sharing, and group editing, of documents; compatible document processing software licensed at all HEP sites; and
- inter-regional and international networks to pull together world-wide collaborations.

Bringing communications to the desktop of the active participant is critical to get the most productivity out of a widely-dispersed community. Currently, the Internet is a bottleneck in fully exploiting computing capabilities. Trends in information systems (World-Wide Web, satellite systems, desktop videoconferencing, BISDN and SONET) are favorable, as the commercial world develops similar needs. There are still national and regional barriers to fully open inter-regional and international communications, since communications are now so linked to national and regional competitiveness that progress may be slower than purely technical considerations would require. As recently as last year, the primary link for HEP between the U.S. and Europe was connected with a specific experiment, and links to Russia could not be direct, and went via a 9600 baud connection. These arrangements are being changed but are indicative of the problems that must be addressed.

There are also very significant network needs for data serving and data analysis, and for management of distributed computing facilities. With petabytes of data being contemplated, the issues of centralizing data and production computing, versus having regional centers are inextricably tied up with those of network capacity, reach, and cost. Within the U.S, the HEP community has benefited from having connectivity provided without direct cost on ESNET and the Internet; it is very likely that increasing commercial exploitation will change this, especially as the bandwidth needs increase.

Much of the international connectivity needed for participation in the CERN and DESY physics programs is directly funded by U.S. HEP. Extrapolation of the historical downward evolution of international tariffs (10–15% per year) yields a highly constrained model for the collaborative analysis of petabyte data samples. Deregulation and a massive growth of commercial traffic could result in one or two orders of magnitude more bandwidth per unit cost.

## 5 Recommendations

The computing environment has changed greatly over the past decade, in ways that dictate new approaches for our discipline: HEP is not a driver of general computing technology, we must position ourselves to move with industry and to take advantage of industry developments for other customers. Overall, the field must find ways to do more with less; this to us implies greater coordination of computing efforts, but also an investment in new software and hardware to maximize physicist productivity. Particularly in the U.S., computing efforts have been fragmented and have not had the usefulness or scope of the coordinated effort in Europe. In the future, it is clear that the development of computing systems for HEP will be driven by the needs of LHC, and hence will be focused at CERN. Nevertheless, the U.S. HEP community must participate in these developments, to provide training for students and support for U.S. efforts in HEP.

There are several paths that could be followed. One would be to make the bulk of computing support a professional activity, by employing software and hardware professionals to do most of the work (this has budget implications, but may increase physicist productivity). Another approach, in which physicists would continue to develop computing systems, would be to raise the status of contributions to HEP computing, and define satisfying career paths for those who choose this area, partly by tying the work to a coherent body of HEP developments (thus in effect professionalizing the work inside HEP). That is, either regard it as engineering to be paid for and managed, or as part of the discipline to be valued as it builds a productive technology within a coordinated framework.

- Our main recommendation is that the U.S. HEP community should formulate a coherent approach to the use of computing technology. A mechanism should be created to foster these discussions, involving the HEP and Nuclear Physics laboratories and user groups.

  Specific recommendations that we make here, and that we expect would be part of a coordinated approach, are:

- Make policy choices lying between the extremes of
  - contracting out all computing and software responsibilities and
  - assigning these responsibilities to university and laboratory personnel.

  The policy must be formulated with the agreement of funding agencies. Where responsibilities are assigned to universities and laboratories, appropriate career structures may have to be created, particularly if the universities are to retain a major rôle in this key area of HEP.
- Re-capitalize U.S. HEP institutions (universities and laboratories) in a consistent way, compatible with current and anticipated industry development toward open systems. This means consistent hardware, consistent operating systems, consistent application software, and good networks. If necessary, provide training in workshops and summer schools to speed transition. The goal is to provide access to adequate CPU power, good networks, and videoconferencing.
- Provide centers of CPU power and storage for simulation, reconstruction, and analysis, accessible to any HEP institution.
- Fund R&D in computing:
  - software systems and tools: data bases and data access methods, software methodologies;
  - computing hardware: data handling and analysis systems (parallel I/O, disk farms, robots), trigger and DAQ systems; and
  - information systems and communications: desktop video conference, document preparation systems.

  In view of the rapid changes in computing and software technologies, R&D projects should be regularly reviewed by experts from within and from outside HEP.

There are questions which must be debated, areas where there is advantage to be gained in taking a specific path. As one example, one operating system could be chosen and full use made of the services provided, or we could continue the CERN approach of providing applications that run on top of any OS.

We conclude by emphasizing increasingly important to the HEP community, and a coherent and well-designed approach is needed to maximize the productivity of a distributed community with high demands.