

# DATA ACQUISITION AND ANALYSIS IN EXTREMELY HIGH DATA RATE EXPERIMENTS

The goal of elementary particle physics is to unravel the properties of matter at the deepest level—that is, to answer questions such as, What are the basic constituents of matter and how do they interact with each other? The search for the answers to such questions has led us to probe the structure of matter at ever-smaller length scales—from atomic to nuclear and now subnuclear scales. Investigating very small distance scales (below  $10^{-13}$  cm) has required very high-energy particles and, consequently, the development of gigantic particle accelerators for producing such particles. (See the box on page 55.) Although these high energy accelerators are marvels of technical and scientific ingenuity, the massive amount of data they produce could not be collected, analyzed and reduced to physics results without another marvel of modern technology: the digital computer!

Computing technology has always been essential to realizing the potential of the giant accelerators, and the amount of computing, data storage and input/output (I/O) bandwidth available strongly influences the way high-energy physics (HEP) experiments collect and analyze data. (See the article by Robert Seidel on page 33 of this issue.) In the next 10 to 15 years, experiments will generate several petabytes ( $10^{15}$  bytes) of data per year, all of which will have to be analyzed and made available to physicists who can extract physics results from the data.

## Data acquisition and triggering

When an interaction occurs in a detector, the data acquisition (DA) system records the electronic signals from the detector elements in a buffer, or temporary storage medium, reduces the data rate to a manageable level, records the remaining interactions on a permanent storage medium and controls and monitors detector performance. Computers span the whole DA chain.

Conceptually, the DA system for a typical HEP experiment can be viewed as a multistage pipeline, like that illustrated in figure 1. Within the detector, interactions, or events, occur at a rate far too rapid to be accommodated by the data analysis and data storage components of the

Although computers will be essential in coping with the petabytes of data generated each year by next-generation particle physics experiments, perhaps their greatest role will be coordinating the efforts of truly global collaborations of over a thousand researchers.

Joel N. Butler and David R. Quarrie

experiment. Most of these events, however, are physically uninteresting or contain data that may not be usable. While the information for the event is stored in the temporary storage, or latency, buffers, trigger processors perform rapid, crude pattern recognition algorithms to eliminate the uninteresting events. The extensive use of buffers throughout the DA pipeline limits the effect of statistical

fluctuations in the interaction rate.

Because the level-1 trigger system is confronted with the highest data rates, it typically uses information from only a fraction of the experiment's detector channels and is composed of high-speed, special-purpose electronics modules; as such, it requires only a few microseconds to eliminate uninteresting events. Higher-level triggers face much lower rates, and so can use commercial high-speed processors to perform more sophisticated event-rejection algorithms.

The event builder, which assembles all the detector information about the event for the first time, has benefited as much from advances in commercial network-switch technology as from those in processor power. This DA pipeline element used to be a dedicated piece of hardware and thereby caused a single-point bottleneck. Switching technology now makes it possible to do relatively sophisticated event-rejection analyses by routing fragments of each event to one of several computers.

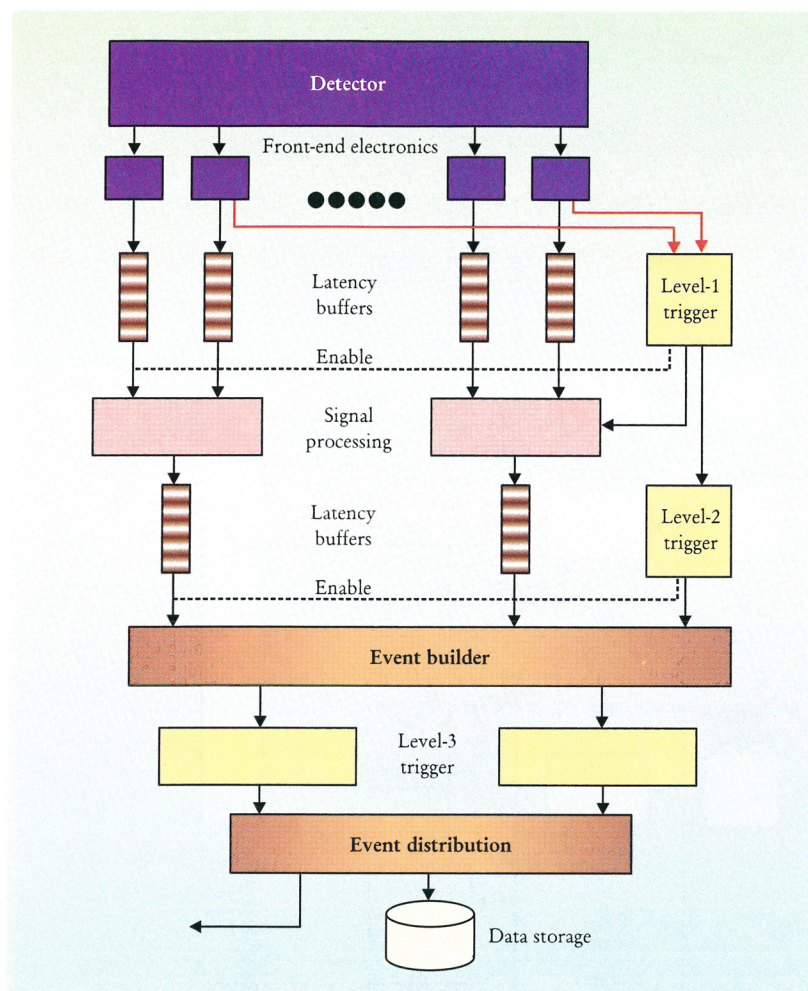
## An example: BABAR's DA system

The BABAR detector<sup>1</sup> at the Stanford Linear Accelerator Center is a typical HEP detector and demonstrates how the conceptual system outlined above may be realized. The whimsical name of the experiment is derived from the fact that the physics of interest—an asymmetry in the decays of the B and  $\bar{B}$  (pronounced "bee-bar") mesons—resembles the name of the 'elephant king' in a famous series of children's books. The detector, which is currently under construction and is scheduled for completion at the end of 1998, consists of five major subsystems surrounding the interaction region. These subsystems have a total of over 200 000 channels of electronics, which attempt to provide complete information about the event. Because the accelerator bunches the electron and positron beam intensities with a frequency of about 250 MHz, this is the basic interaction rate in the BABAR detector. For comparison, the rate for physically interesting events is about 30 Hz.

The BABAR DA system (shown in figure 2) should be

JOEL BUTLER is a physicist on the research staff, and currently the head of the computing division, at Fermilab in Batavia, Illinois. DAVID QUARRIE is a research scientist at Lawrence Berkeley National Laboratory in Berkeley, California.





**HEP DATA ACQUISITION SCHEMATIC.** A detector in a HEP experiment sends information about the particles passing through it by means of front-end electronics to latency buffers, which hold the information for a few microseconds. During this time, high-speed data lines (red) transport information from a portion of the detector's electronics channels to the level-1 trigger, which performs rapid and simple pattern recognition and decides to accept or reject the event based on this information. An accept signal from the trigger sends the detector signals through signal processors and then to another set of buffers, where the event awaits the decision of the more sophisticated level-2 trigger processor. An accept at level 2 sends the detector signals to the event builder, which assembles the complete event information. The event builder then sends complete events to a farm of processors, which perform a final, relatively sophisticated level-3 event rejection based on this complete information; events that pass are stored permanently for later analysis. **FIGURE 1**

compared to the conceptual system in figure 1. Information from the charged particle tracking chamber and the electromagnetic calorimeter (about 10% of the total electronics channels) is directed over dedicated links to the level-1 trigger, while the remainder of the data is stored in a digital latency buffer. The latency buffers store the information for about  $12 \mu\text{s}$  while the level-1 trigger uses a set of highly specialized processors, including programmable logic arrays and programmable microprocessors, to make a rapid decision about whether the event fulfills such simple criteria as a minimum number of charged particle tracks or a minimum energy deposited in the calorimeter. This reduces the event rate from 250 MHz to about 2 kHz; information is transferred at that rate from the latency buffer to about 250 readout modules. These modules compress the event by suppressing uninteresting channels and by performing feature extraction (such as waveform analysis to determine the peak and width of a pulse shape). The resulting fragments are then partially combined into 20 readout controllers.

Both the readout modules and readout controllers in this experiment are single-board computers based on commercial microprocessors. The compression reduces the event size to about 25 kbytes, which, given the 2 kHz triggering rate, results in an aggregate bandwidth of 50 Mbytes/s. This bandwidth is small enough that no level-2 trigger is needed to further reduce the data rate. Instead, a commercial switch network fabric routes the event fragments to a set of conventional workstations, such that

each workstation receives complete events. These workstations carry out the level-3 trigger analysis, performing more complex event filtering based on complete event information and reducing the data rate to 100 Hz. These events are sent to permanent storage at this rate for later off-line analysis. Although this experiment does not currently utilize a level-2 trigger, the design allows one to be implemented if accelerator performance upgrades make it necessary.

A set of control and monitoring computers provides the user interface for the physicists and ensures that the detector is performing correctly. Performance monitoring involves checking not only the quality of the current event, but also the values of many detector-environment electronics channels. These channels measure quantities ranging from power supply voltages for the electronics racks to the current in the superconducting magnet that forms much of the detector structure and allows charged particle momenta to be determined from the curvature of their trajectories in the magnetic field.

This typical experiment therefore depends on the correct functioning of several hundred computers, ranging from conventional workstations to highly specialized units designed and built for a particular role within the DA system.

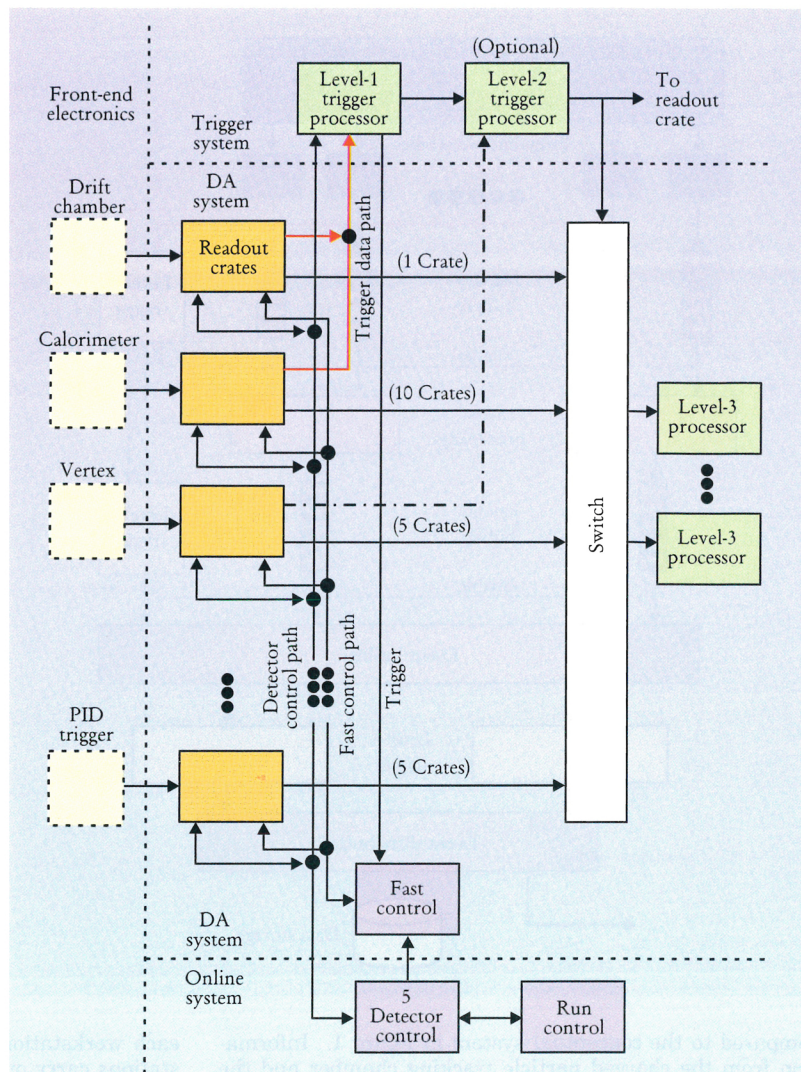
### Data acquisition: the next generation

Although the BABAR experiment described above is typical of current HEP experiments, it is dwarfed by those already



IN BABAR'S DATA ACQUISITION SYSTEM, a dedicated line (red) transfers information from the tracking system (the drift chamber) and the calorimeter to the level-1 trigger, which decides whether the event is interesting based on a series of simple criteria, such as the number of tracks it contains or the energy deposited in the calorimeter.

Within 12  $\mu$ s, the trigger relays an accept or reject signal to the fast control unit, which, in turn, notifies the readout crates (buffers) whether to flush (reject) their contents or forward them, by means of a commercial switch network fabric, to a level-3 trigger processor farm. The level-3 trigger processors perform relatively complex filtering based on complete event information. Events satisfying the level-3 trigger are stored for later off-line analysis. The computer system also provides a control interface for the five detector subsystems, the trigger and the overall running conditions of the experiment. Along the control pathways, information flows both ways, as indicated by the directions of the arrow heads. **FIGURE 2**



being designed for future accelerators, especially those at CERN's Large Hadron Collider (LHC). The number of electronics channels, the bandwidth requirements and the complexity of the interactions at the LHC will exceed those in current experiments by orders of magnitude. Instead of the 50 Mbytes/s required by BABAR, LHC detectors will require aggregate bandwidths of well over 1 Gbyte/s and will involve the coordinated operation of thousands, rather than hundreds, of computers.

To meet these challenges, the ATLAS collaboration<sup>2</sup> at the LHC is developing a multilevel trigger similar to the idealized system described above. Beginning with an interaction rate on the order of  $10^9$  Hz, the level-1 trigger would consist of hardwired processors and would require about 2  $\mu$ s to perform event rejection based on information from calorimeters and from the muon tracking system. The level-2 trigger would consist of one or more farms of several hundred fully programmable processors, and, although it would have access to information for the entire event, it would perform event rejection (in about 10 ms) based on more detailed analysis of detectors in the regions where the level-1 trigger had been satisfied. The level-3 trigger would perform rejection based on completely reconstructed event information; with a 1000-processor farm, the trigger could take up to 1 s per event. The

level-1, 2 and 3 triggers would supply rejection factors of about 10 000, 100 and 10, respectively, resulting in an event rate of about 10–100 Hz, or a permanent storage rate of about 10–100 Mbytes/s.

The Compact Muon Solenoid (CMS) detector group<sup>3</sup> at LHC opts for a slightly different strategy. The hardwired level-1 trigger would be similar to that of ATLAS and would reduce the event rate to less than  $10^5$  Hz. A farm of 4000 1000-MIPS (million instructions per second) processors would then perform a "virtual" level-2 rejection based on partial event information—"virtual" in the sense that if the event passes the trigger, full event information would be sent to the same processor, which would then perform the level-3 rejection to reduce the event rate below 100 Hz. In this scheme the event builder would have to handle a data rate of about 200 Gbytes/s, although the permanent storage rate would be less than 1 Gbyte/s.

Researchers are confident that the pace of advances in commercial processors and high-bandwidth networking technology will continue to be sufficiently rapid that only hardware specific to demands of the LHC experiments will have to be specially built—specifically the front-end electronics that read out the detectors and the fast, hardwired level-1 trigger processors capable of rendering accept or reject decisions within a few  $\mu$ s.



## Wallet card requirements for some future experiments.

| Experiment  | level-1 input/s | level-3 output/s | Event size, kbytes | Mbytes/s | MIPS           |
|-------------|-----------------|------------------|--------------------|----------|----------------|
| BABAR       | 250 000 000     | 100              | 25                 | 2.5      | 20 000         |
| D0/CDF      | 10 000 000      | 50               | 250                | 12.5     | 50 000–100 000 |
| PHENIX/STAR |                 | 1–44             | 20 000!!           | 20       |                |
| ATLAS/CMS   | $10^9$          | 100              | 1 000              | 100      | ~ 1 000 000    |

However, the enormous increase in the complexity of experiments is causing a corresponding increase in software complexity. To avoid swamping the DA system, more and more processing needs to be performed closer to the electronics channels, and the activity of the increasing number of processing elements has to be orchestrated.

Another area of concern is providing efficient access to the enormous quantities of recorded data for the large and geographically dispersed user community. BABAR is expected to generate approximately  $10^{14}$  bytes of information per year; each of the major LHC experiments will generate about  $10^{15}$  bytes per year. The BABAR user community includes over 500 physicists representing over 80 different institutions in ten countries. The LHC user communities are significantly larger. For example, the CMS collaboration currently has 1789 members from 138 institutions in 38 nations spanning the alphabet from Armenia to Uzbekistan. Table 1 briefly summarizes the approximate characteristics for some experiments scheduled to take data during the next decade.

### The reconstruction problem

The output of the data acquisition and triggering system is a list of “raw data”—quantities such as the times and positions at which the particles cross the various detectors and the pulse heights of signals the particles produce in the detectors. These data must be processed so that physicists can analyze them and try to learn something new about nature. This enterprise requires an enormous amount of processing, and the resulting amount of data is quite daunting.

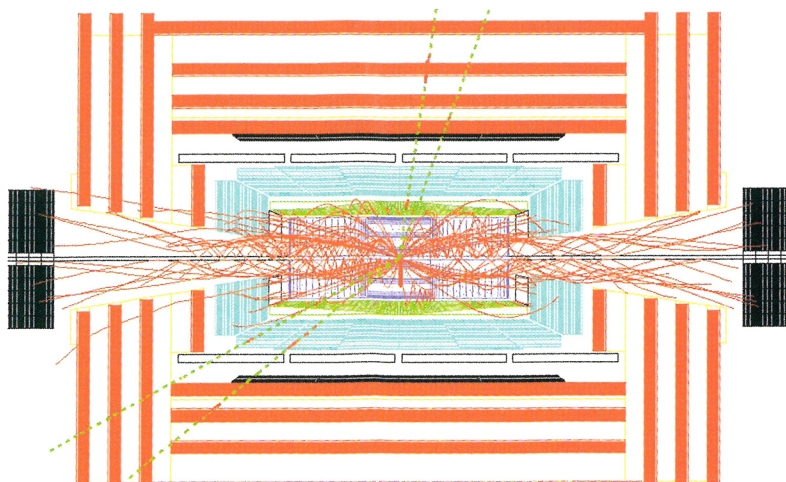
The data analysis programs reconstruct the apparently uncorrelated collections of signals in the raw data into objects of interest to the physicist, such as

▷ tracks of charged particles, along with the particles’ vector momenta and type (for example, pion, kaon, proton, electron or muon), and

▷ electrically neutral objects such as photons or neutral hadrons (neutrons or neutral K mesons).

These objects may be further aggregated into other objects of interests, such as “jets” and “vertices.” Jets are collimated bundles of particles that form as quarks or gluons become mesons or baryons. Vertices are groups of tracks emanating from a single point in space. Most tracks in an event originate from the primary interaction vertex. However, some tracks are associated with vertices removed from the primary interaction. Such secondary vertices result when a parent particle travels some distance from the primary vertex and decays, perhaps by the electroweak interaction, an area of intense interest at present.

To identify these objects, a variety of special pattern-recognition programs must analyze the raw data. For example, charged track reconstruction programs use information from a small number of the available detectors—for example, “hits” in a few planes of the tracking system—to form a hypothesis about the possible existence of a track and then check to see whether the remaining (redundant) detectors confirm the hypothesis. This process is time consuming because there are many more track candidates than true tracks, because the detectors are imperfect and because hits unrelated to the event (caused, perhaps, by particles from the accelerator) can confuse the analysis program. Care must be taken—using goodness-of-fit criteria, for example—to assign only the proper hits to each track candidate so the track’s parameters can be determined without degrading the resolution. Equally complex programs are needed to reconstruct data from



**A SIMULATED HIGGS-PARTICLE EVENT** in LHC’s Compact Muon Solenoid detector shows how complicated will be the task of developing reconstruction algorithms in the next generation of HEP experiments. The Higgs particle is a key prediction of the Weinberg–Salam–Glashow electroweak theory. The Higgs decay produces very energetic muons (green) that penetrate through the detector at high angles to the main axis of the event (indicated by the majority of particle tracks shown in red). (Courtesy of Fermilab.) **FIGURE 3**



RECONSTRUCTED DATA provides several tools to help researchers understand the physics of this event from the CDF experiment at Fermilab's Tevatron.

**Bottom left:** The tracking view shows a very energetic lepton (red) and the loose clusters of tracks that identify jets. These characteristics identify the event as a top candidate. **Upper right:** The lego plot shows the combined results of calorimetry and momentum determination analyses, with angular information in the plane of the plot and the transverse momentum (or energy) component of each jet and lepton shown on a perpendicular scale. The large energies of the jets and lepton identified in the tracking view and the large "missing transverse energy" (green), derived from balancing total momentum for the event, strengthen the top quark hypothesis.

**Bottom right:** Using detailed tracking information from vertex views, one sees evidence for the decay of heavy quarks, again strengthening the top quark hypothesis. (Courtesy of CDF.) **FIGURE 4**

## e + 4 jet event

40758\_44414

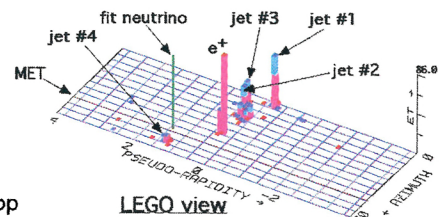
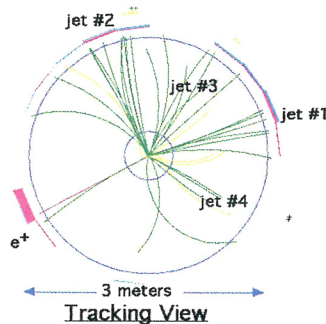
24-September, 1992

### TWO jets tagged by SVX

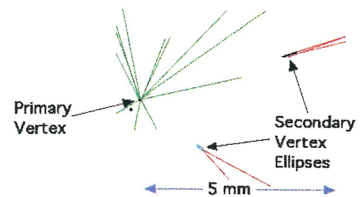
fit top mass is  $170 \pm 10$  GeV

$e^+$ , Missing  $E_T$ , jet #4 from top

jets 1,2,3 from top (2&3 from W)



### Two Vertex Views (note scales)



other systems in the detector. Figure 3, which shows a simulated LHC interaction in the proposed CMS detector, gives an idea of the complexity of event reconstruction.

This event reconstruction phase of the data analysis can take a long time. For example, an event in E687, a typical fixed-target experiment at Fermilab, may take 0.25 s on a modern workstation with a processing power of 100 MIPS, and the new version of the experiment expects to write about 500 events per second. An event from the Collider Detector at Fermilab (CDF) requires about 5 s on such a processor, and the experiment expects to write more than 100 events per second in the future. (See figure 4.) An event recorded by the STAR experiment at the Relativistic Heavy Ion Collider (RHIC) will be 10–20 Mbytes long and will take more than 10 s to analyze on a very-high-end RISC processor.

## The reconstruction solution

As recently as 1988, obtaining adequate computing power to cope with this onslaught of data was not easy. The reconstruction phase, which is by no means the end of the data analysis, often took months or even years. It was often difficult to tell whether the detector was even taking data properly. The need to monitor detector performance provided strong motivation to speed up the reconstruction process. Competitive pressures also pushed investigators toward analyzing data almost as soon as they were recorded. Competition between the two Fermilab experiments searching for the top quark was particularly intense. Similarly, several teams are pursuing the observation of the asymmetries that are a main objective of the BABAR detector. Thus, it has become a goal to analyze the data in quasi-real time—that is, at the same rate the data are taken, with at most a slight time lag for the extraction of various calibration and alignment constants for the detector.

The embarrassingly parallel nature of high-energy physics analysis has been crucial in dealing with the growing amount of data. In principle, each event is a separate computational problem and may be computed on

a separate processor. When the main processor options were expensive mainframe computers, this parallelism was not very helpful. However, the invention and widespread adoption of microprocessors, especially in personal computers and workstations, has made it possible to exploit the intrinsic parallelism of the HEP reconstruction problem by employing many cheap commercial processors. Although this approach to the problem seems obvious today, it faced significant obstacles in the mid-1980s. One major obstacle was the rather limited program development environment of the early microprocessors, particularly in FORTRAN, the language in which most HEP code was written. Linkers and debuggers were also barely adequate for the large, complex HEP analysis programs. Many programs did not easily fit into the memory available on those machines, and support for peripheral devices was limited.

One early project that successfully overcame some of these problems was the Advanced Computer Program (ACP) at Fermilab,<sup>4</sup> which developed single-board computers based on Motorola 68020 processors to do HEP reconstruction. (See figure 5.) The system worked well for smaller codes, but physical memory limitations and inadequacies in the development environment limited its usefulness for the most complex codes.

Many of the problems with the ACP system were solved with the advent of the much more powerful Unix-based RISC microprocessors, which are aimed at the scientific, visualization and server market, and so come supplied with excellent code development environments, capable compilers, debuggers and linkers, good peripheral support and large memories. With the commercial success of these systems, buying a workstation became more economical than building customized single-board computers. The current Fermilab system,<sup>5</sup> which was acquired between 1991 and 1993, consists of 180 Silicon Graphics Inc and 140 IBM workstations (without the keyboards, mice or monitors). Larger SGI and IBM I/O-server computers read data from tapes and transport events to the worker nodes over several Ethernet networks. When a



node finishes the computation of the event, it sends the results back to the I/O server so it can be recorded on the output tape. The system can do about  $10^{10}$  computations per second.

The ultimate test of such a system is how well it carries out its intended task—the analysis of HEP data. For the recent Tevatron run, during which the top quark was discovered by the CDF and D0 collaborations, the data could be reconstructed as fast as they were taken.

Variants of the approach described above use symmetric multiprocessors (now available from several vendors), rather than the loosely coupled networks of workstations described above, and processors running the Windows NT or Linux operating systems. We believe that HEP can continue to take advantage of the falling cost of commercial processors to meet the challenge of future experiments far into the LHC era.

## Data mining

Although the general consensus is that the event reconstruction problem has been solved, HEP data analysis does not end there. Many hundreds of terabytes or even petabytes of data may be produced during the reconstruction phase of the analysis. Once reconstruction is completed, a researcher must follow a long, multistep path before performing the kind of statistical studies required to extract from the data an important result such as evidence for the top quark or CP violation.

The current method for this analysis phase is to classify the reconstructed events according to their potential physics interest and then split them into separate data sets, or data streams, by physics topic. Further selections are performed to eliminate background events and get purer samples of potentially interesting events. Unnecessary data are eliminated at each stage, and the sample is compressed whenever possible to speed the analysis and ultimately allow a physicist to perform in a timely manner the kinds of explorations and iterative studies that are required to extract signals and make quantitative statements about them.

This procedure has five main problems. First, it is very I/O intensive, consuming a huge amount of bandwidth between disk and memory and on the network, while using only a small fraction of available power of the central processing unit. Advances in I/O and network bandwidth, while significant, have been nowhere near as spectacular as those in processing power. Second, it is quite labor intensive, generating at each data reduction stage a very large number of data sets, all of which must be verified and tracked. Whether one can easily scale these techniques to deal with the challenges of the next decade is uncertain. Third, the procedure is error prone. A bad choice of an analysis cut can eliminate important classes of events. A decision to throw away some information may mean that some problem cannot be thoroughly investigated at a later stage of the analysis. Fourth, the procedure can be biased. One builds some physics prejudices into the sorting and selecting that is required to achieve the data reduction. By using features of a model—such as the Standard Model of particle physics—to decide how to reduce the data, it is certainly possible to inadvertently eliminate events that are evidence of phenomena that lie outside the model or even contradict it. Although physicists are well aware of this problem and are constantly working to avoid it, there is always a certain discomfort. Fifth and finally, the procedure does not scale well as the amount of data, the number of physicists and their geographic dispersion increases. Access patterns to the central data repositories are difficult to control and may lead to inefficiencies, and remote resources at inves-

## Tools of Particle Physics: Accelerators and Detectors

Several different kinds of accelerators are needed to study the important questions in particle physics. Head-on collisions between beams of electrons and positrons are made to occur at the Large Electron-Positron (LEP) Collider at CERN (near Geneva, Switzerland), at the Stanford Linear Accelerator Center (SLAC) in Stanford, California, at the National Laboratory for High-Energy Physics (KEK) in Tsukuba, Japan, and at the INFN facility in Frascati, Italy. At Fermilab (Fermi National Accelerator Laboratory) in Batavia, Illinois, beams of 1 TeV antiprotons collide with beams of 1 TeV protons. The Relativistic Heavy Ion Collider (RHIC) being built at Brookhaven National Laboratory in Upton, New York, will collide two beams of heavy nuclei, such as gold. The German Electron Synchrotron (DESY) in Hamburg, Germany, can collide a beam of protons with a beam of electrons (or positrons). The Thomas Jefferson National Accelerator Facility (formerly known as CEBAF) in Newport News, Virginia, Fermilab, BNL, SLAC, DESY and CERN can also create collisions between extracted beams and stationary targets (in so-called fixed-target experiments). The Large Hadron Collider (LHC) at CERN, scheduled to come on-line in 2004 or 2005, will produce head-on collisions of two proton beams, each with an energy of 7 TeV; it will also produce collisions between beams of nuclei.

To study the collisions produced by these accelerators, physicists must construct massive detectors having hundreds of thousands or even millions of channels of electronics and costing many tens of millions of dollars. The detectors at the LHC will cost more than \$500 million each.

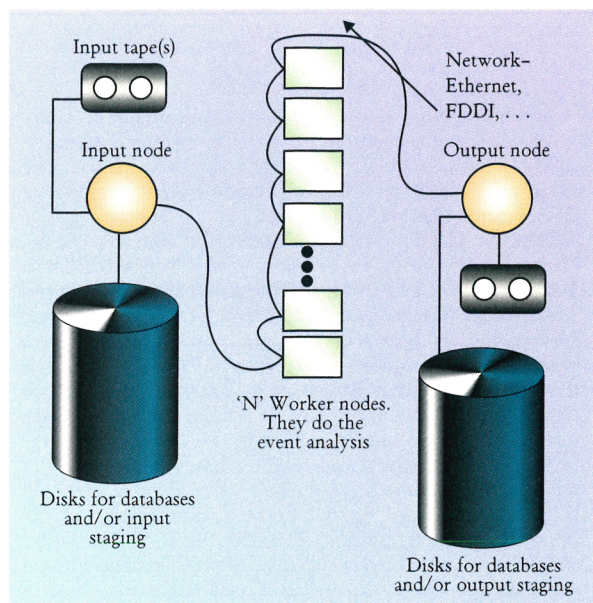
Within such detectors, when a particle in a primary beam from the accelerator collides with either a stationary target or, in a colliding beam accelerator, with a particle from the other particle beam, many thousands of the electronic elements record information about the “secondary elementary particles” produced in the collision. The tracking system records the time and position at which a charged particle passes near one of the detector elements—information that is used to reconstruct the particle’s trajectory, or track. Calorimeters measure particle energies. Vertex detectors provide very precise tracking near the collision region to allow determination of the points from which various groups of tracks originated. Still other elements can be used to identify what kind of particle (electron, pion, kaon and so on) made the track. Together, information from these elements allows particle physicists to reconstruct what occurred in a particular event.

tigator institutions may not be employed to full effect.

Future experiments will produce massive amounts of data. During the next collider run at Fermilab, each experiment will produce on the order of a terabyte of data per day. Around the year 2005, each LHC experiment will begin to churn out data at a rate of several petabytes per year! Some investigators feel that taking advantage of every advance in commercial data storage and data access technology will still not solve all the problems inherent in dealing with such large data sets.

Considerable effort has been directed toward understanding how best to organize HEP data to facilitate rapid and efficient retrieval. One approach would be to identify and store on quickly accessible media the parts of events that are most likely to be needed by the analysts and to store less used, but still possibly interesting parts of the events on slower, less rapidly accessed media. In this way, all the data would remain available, although the





**STRUCTURE OF FERMILAB'S ACP SYSTEM.** The ACP project exploited the embarrassingly parallel nature of HEP event reconstruction by using a farm of single-board computers to do event reconstruction. Raw, unreconstructed data from input tapes and data from disks were fed into a MicroVAX host, which sent data over the Fermilab-developed branch bus to nodes in the farm. These processors reconstructed the events and then sent them back over the branch bus and an Ethernet network to a second MicroVAX host, which wrote output tapes of reconstructed events. At its height, ACP used over 500 worker nodes, organized into a half dozen systems, each with its own input/output host, to do the event reconstruction for Fermilab data runs from 1985 to 1989. The system was retired in 1993. **FIGURE 5**

time to access the information would vary depending on how "hot" it is. Several approaches based on these principles appear to have some promise, but a proof of principle for any of them will require much more work.

The term "data mining," which is often applied to this process, is particularly apt in HEP, in which only one interaction out of every 100 billion has an observable top quark! However, particle physicists are fortunate that their data mining problems are not unique, but rather are shared with many other scientific and commercial activities. If solutions arise in other fields, HEP may be able to borrow them.

## Networking: tying it together

Computers have also proved invaluable in coordinating the efforts of the huge collaborations that are now the norm in HEP experiments. Although effort tends to increase linearly with the number of collaborators, communications overheads increase as the square of that number (or worse!). There is great difficulty attached to maintaining the coherence of a large collaboration and ensuring efficient use of all its resources, including those remote from the site of the experiment. Computers and networks such as the World Wide Web play an increasingly important role in all aspects of an experiment—from planning the detector to gathering and analyzing data to the ultimate publication of experimental results.

In HEP experiments, extensive networking can help to alleviate the burden of coordinating the efforts of the central laboratory and remote labs at universities and other institutions, all of which may be developing and building substantial parts of the detector. Videoconferencing improves communication and helps reduce the likelihood of major misunderstandings between the central and remote-site personnel. Drawings, budgets, schedules and status reports can be transmitted over the network to make sure that detector components are built to the correct dimensions, within budget and on schedule.

Networking has also played an important role in the success that HEP groups have had in maintaining central code bases while allowing distributed software development and distributing the software and associated databases to remote sites. Computer-aided software engineering (CASE) tools are now widely used, allowing software

designs to be distributed throughout the developer community and discussed in greater detail than was previously possible.

Because it is impractical to have all the experts permanently located at the experiment site, efforts are now in progress to use modern networking to decentralize the traditional central control room. Even within the control room, most critical displays are transmitted over a network, making it relatively easy to provide the displays to remote sites. Control of most detectors is also now done by networked computers, and permitting control over the World Wide Web from remote sites is a logical and easy extension of this trend.

Finally, modern networking can help ensure the effective use of remote data analysis resources by facilitating the successful distribution of data sets. Although it is probably impractical to distribute full data sets to remote institutions, it is very feasible to distribute more selected data sets from the central sites to remote sites. This procedure allows the remote sites to use their own resources to do the final stages of data analysis and diminishes some of the very user-intensive activity at the central site. Although data may be, and are now, distributed to remote sites by copying the data onto tape at the central site and shipping the tapes to the remote site, remote data analysis will work best if the data can be easily transported over the network. Acquiring sufficient network bandwidth to permit rapid transport of data to remote sites is a major issue for the future.

*This article was written at, and the work was supported by the Fermi National Accelerator Laboratory (FNAL) and the Lawrence Berkeley National Laboratory (LBNL). FNAL is operated by Universities Research Association, Inc., under contract DE(-)AC02-76CH03000 with the US Department of Energy. LBNL is managed by the University of California, Berkeley, under contract DE-AC03-76SF00098 with the US Department of Energy.*

## References

1. A. Eisner *et al.*, SLAC BABAR Collaboration: Technical Design Report, SLAC, Stanford, Calif. (1995). G. Wormser, Nucl. Instr. Meth. A **351**, 54 (1994). P. Harrison, Nucl. Instr. Meth. A **368**, 81 (1995).
2. The ATLAS collaboration, *ATLAS Technical Proposal*, CERN Publication CERN/LHCC/94-43 LHCC/P2, CERN, Geneva, Switzerland (1994).
3. The CMS collaboration, *Compact Muon Solenoid Technical Proposal*, CERN Publication CERN/LHCC 94-38 LHCC/P1, CERN, Geneva, Switzerland (1994).
4. I. Gaines, T. Nash, Annu. Rev. Nucl. Part. Phys. **37**, 177 (1987).
5. F. Rinaldo, S. Wolbers, Computers in Physics, **7** (2), 184 (1993). ■