

**Fermi National Accelerator Laboratory**

**FERMILAB-Conf-87/48**

2380.000

**Asilomar Conference on Managing Complexity  
in High Energy Physics:  
A Summary and Renaming of the Conference\***

T. Nash

Advanced Computer Program

Fermi National Accelerator Laboratory

P.O. Box 500, Batavia, Illinois 60510

February 1987

\*Invited Summary Talk at the Computing in High Energy Physics Conference, Asilomar State Beach, California, February 2-6, 1987.



Operated by Universities Research Association Inc. under contract with the United States Department of Energy

## Asilomar Conference on Managing Complexity in High Energy Physics:

### A Summary and Renaming of the Conference

Thomas Nash  
Advanced Computer Program  
Fermi National Accelerator Laboratory  
Batavia, IL 60510

#### Introduction

Computing, triggers and electronics in high energy physics are hardly new. In 1930, Rossi developed the first practical AND gate, not in order to start the business of digital electronics and computing, but to detect coincidences of Geiger counter signals (Fig 1). Shortly thereafter, this circuit was used to trigger cloud chambers so that they were expanded only when tracks were likely to have passed through. The struggle to remove technology constraints has always been an essential component of the research effort of experimentalists. If some technology constraint is not present, the experiment, if it is worth doing, has most likely already been done. In this environment, physicists seem to have evolved into personalities that enjoy having their abilities always near saturation, no matter what the sophistication of the prevailing technological environment. (This has the all too frequent side effect of encouraging novelty for novelty's sake, as in data acquisition systems which often differ in unnecessary ways.)

This predilection for saturating our collective ability to manage complex technological problems manifests itself today in the scale of complexity of the systems we are trying, or proposing, to manage. What is new is that this scale is now qualitatively different. UA1, CDF, HERA, D0, LEP, SSC all represent experiment detector systems of extraordinary complexity. This new scale of complexity is unquestionably today's key technological constraint. Managing it, which will require sophisticated state of the art (and beyond) computer techniques, is just as important to the progress of physics in the 1980s as Rossi's coincidence circuit was to the 1930s.

The organizers of this meeting recognized that many of the most serious and unsolved problems in computing for high energy physics are in software, and they focussed the attention of much of the conference on this. Software is always somewhere near the interface between physicists and complex detector systems. It is therefore not surprising that the theme that wound through just about all the discussions at Asilomar was *managing complexity* of one sort or another. This could easily have been named the *Asilomar Conference on Managing Complexity in High Energy Physics*.

Usually without anyone mentioning the word, we discussed managing the various *complex* aspects of our business:

1. **Detectors**  
Construction management reporting and fabrication plans
2. **Collaborations**  
Communication networks
3. **Events**  
Parallel processor farms, graphic workstations

4. **Software**  
Software development stations, software project management, new or refurbished languages, SASD
5. **Data**  
Data base management for constants and event data, physics analysis workstations
6. **Data Acquisition Systems**  
Expert system diagnostics
7. **Triggers**  
Real time program certification
8. **Theory**  
Lattice gauge theory, processors

It is striking to note in this list that of all these critical computing problem areas only theory is fundamentally a numerical calculation, the type of calculation for which computers were designed and for which they are really suited. Except for the magnitudes involved ( $10^8$  events, 500,000 detector channels, 100,000 lines of Fortran), these problems are better suited to processing by human brains than by computers. In fact, it is fair to say that the key goal of the active research field known as Artificial Intelligence (AI) is to figure out a sensible way to attack such non-numerical problems with machines. The extraordinary difficulties in AI are well known, and it is therefore not surprising that such problems are imposing in the physics environment as well.

### Managing Software and Data

It is now generally recognized that systems have gotten so complex in our business that there is no longer room for the old way of doing things. No more 20,000 line reconstruction programs, built of 10 unstructured subroutines, and containing no comment statements. No more genius graduate students single handedly responsible for all of an experiment's software. And no more foot thick paper outputs containing hundreds of obscurely labelled histograms and consuming the better part of a tree. Pleasantly such familiar horrors are fading into history, and the field is looking outside for professional approaches to managing software and displaying complex data.

A graph prepared by B.W.Boehm in 1976 shows how the cost of fixing a bug increases exponentially the later in a project it is identified. This plot was shown at least a half dozen times, replacing the Monterey Cypress as the conference logo. So why not reproduce it here again (Fig.2)? John Manzo, telling us how DEC manages large software projects, summed it up as "Pay now or pay lots more later." His important lecture showed how simple and obvious are the correct techniques for coherent management of complex projects. He emphasized functional modularity, frequent review, as well as organization structures and procedures that insure efficient communication between coworkers. We will have to learn how to apply all this in the fundamentally anarchistic physics environment.

At CERN Gottfried Kellner is pioneering use of structured analysis, structured design (SASD) software development techniques with his colleagues working on Aleph software. A full set of workstation tools for such structured software engineering is not yet available. Nonetheless, the power of this approach is so apparent that it has been readily accepted on Aleph and is also being used, at least in part, for D0 at Fermilab. Paolo Palazzi and colleagues have developed ADAMO, a data management system that couples entity relationships to the data flow diagrams of structured analysis. The data may be in a relational data base. ADAMO seems to tie together a wealth of sensible modern computer science software and data management techniques and applies them to high energy physics.

The immediate importance of this kind of software engineering to large experiment software efforts is apparent. It is also the first step toward solution of another issue which will become critical for SSC era experiments. SSC design reports anticipate  $10^8$  interactions per second with a 1-5 Hz event recording rate. Most of such a trigger reduction of  $10^8$  must involve complex high level language triggers. It is unlikely that physicists will soon become less inclined to adjust triggers in the middle of data taking. Reliable but changeable high level triggers require real time software verification and validation. Techniques for verifying that the software is correctly carrying out physicists' intentions will certainly depend on some kind of SASD-like foundation. Automatically validating that the intentions themselves, specified under time pressure, are correct given the parameters of the detector, the laws of physics, common sense, etc., is an unsolved fundamental AI issue. Solving this is critical to HEP (and other politically sensitive matters like SDI and reactor safety).

Good software development environments involving workstations have been a strong feature of languages used in the AI world such as LISP. One suspects that if physics ever migrates from its native Fortran to more "modern" languages it will be for better tools, not semantics. Affordable workstations are almost here. Valuable for software development, they are similarly important for data analysis workstations. The popularity of workstations is best explained by paraphrasing the remark, heard in America after World War I, "Once they have seen Paris, how are you going to keep them down on the farm?" Once they have seen the Macintosh, how are you going to keep them buried in HBOOK sub-routine calls and histogram output?

Toby Burnett's Interactive Data Analysis system (IDA) was originally written for his own use. It quickly became very popular among his SLAC colleagues because it allows quick and interactive physics analysis once a data base of potential variables has been prepared. The importance of such tools has been recognized at CERN where a significant effort led by René Brun is underway to develop a Physics Analysis Workstation system (PAW). Because of the enormous investment that has been made in such CERNLib programs as HBOOK, HPLOT, and ZEBRA, the PAW developers have been required, as a condition of their mandate, to base their work on the existing packages. PAW will support both a command mode "for experts" and a menu mode somewhat in Mac style, "for beginners".

The developers of PAW have spent a large amount of time considering the human interface, recognizing that physicists are not too good to have information communicated to them efficiently. One hopes they also recognize that context sensitive menus, search lists, dialogue boxes, etc., are for experts perhaps even more than for beginners. Obscure command languages are just the sort of thing many of us like to saturate our abilities with. (Is this, perhaps, as an escape from more difficult physics questions?) Silicon valley has learned the importance of good human interfaces and of what are now called "designer screens". It takes a special talent to design a display screen (like the Mac) or a new newspaper layout (USA Today, never mind the content) that communicates information efficiently. It is not inappropriate to hire out for such services. Work stations for physicists are so important that they require top down design, starting at the screen. Managers must accept that some major software may require rewriting.

Fermilab's ACP has identified some recent developments in video storage technology that will make it possible to process huge amounts of analysis data in a half hour, instead of a week or more. Two new and extremely cheap devices read and write at tape drive speeds on storage media that holds as much as ten 6250 BPI tapes. One of these, based on \$10, 8 mm video cassettes, already costs as little as \$2000 each in 100 quantities. Another uses CD size laser disks and is called a write once read many times (WORM) disk drive. It is possible to attach units of either type to a single (or small group of) ACP Multiprocessor CPU nodes through a standard SCSI interface. Each drive holds a portion of an experiment's reconstructed data in DST form. With a hundred or more being read and analyzed in parallel by the nodes, large experiment analysis data bases will be scanned in short order. This will put even more pressure on good human interfaces since it will no longer be acceptable to spend days setting up huge blocks of histogram calls.

## Build vs Buy

The concern about complexity in experimental physics has led to a new consciousness about the importance of some old issues. One is the question of build vs buy. The community no longer finds tolerable colleagues who insist on do-it-themselves designs, no matter how quick and clever the implementation, unless there is simply nothing available commercially at acceptable cost. When home built systems are justified, they must be designed to commercial standards, with marketing issues considered, so that it will be possible to arrange for ultimate commercialization.

Each time a decision to build is being considered, one must evaluate the field's competence compared to industry. For example, there has been some evidence that HEP can develop cost effective, commercializable processor systems. On the other hand, it is unlikely that there would be the same success if a competent HEP group attempted to design graphics work stations to compete with industry. With something like human interfaces the situation is less clear. One can hope that physicists can orchestrate expert help from industry to develop workstation software before it appears on the marketplace, driven by other motivations and thus only marginally acceptable to physics needs.

Just as those who we might call *build-bigots* need to be reminded of the problems with their philosophy if oversold, so are there serious problems with an excessive *buy-bigot* stance. The interests of industry marketeers just don't coincide with those of science. Marketing considerations all too often are an impediment to ready availability of technological advances. The client herd and the price-performance curve must be protected at all costs. Witness the resistance of the biggest computer company (and the lack of enthusiasm expressed by the second biggest) regarding establishment of the Fortran 8x standard. So long as private extensions to Fortran 77 are popular, why expose one's sheep-like herd to competitive wolves from small companies with cost effective products? Imagine also the reaction the marketing folks got when they asked the chip people to remove (!?) a feature from the CPU chip for a very popular super-micro so it would not (!) run a very popular operating system. Castrated in this way, the chip could be used for single board computers that would not violate the prevailing price-performance standards of the product harem. Sometimes, the commercial world's idea of innovative VLSI seems very foreign.

A very natural tendency of computer experts in HEP is to fall right into the marketing trap set around client herds. If one has become an expert in one operating system (or backplane standard or microprocessor family or etc.), the intellectual investment in learning all its complexities leads to a strong and defensive prejudice not to have to encounter another. This is really felt as an intellectual territorial imperative, and it must be avoided. *VM-bigots* and *VMS-bigots*, as they referred to themselves at Asilomar, must recognize that neither environment is satisfactory. Both operating systems contribute much too much irrelevant detail to the baggage of what must be remembered. They get in the way of more difficult and important science issues that really should be what is getting the attention of physics talent.

Vectorization of experiment software was discussed with a great deal of intensity. Clearly for extremely simple experiment configurations, enormous speedups can be obtained with reasonable effort. Clearly, also, vectorization will receive increasing attention with the arrival later this year of a Cray XMP at CERN. Large reconstruction codes and Monte Carlo with complicated geometries are notoriously difficult to vectorize. Given that parallel processor farms and emulators are far more cost effective for these programs than are super computers, there has not been a great rush to sweat out vectorizing them. Fermilab's ACP reported that Fortran reconstruction code benchmarks on a new generation of RISC chips are running 5 to 10 times faster than the existing "VAX-class" ACP multiprocessor CPU nodes. This reinforces the expectation that such multi microprocessor farms will continue to lead in cost effectiveness for some time to come. A likely direction that VLSI technology may take is to include vector instructions on a future generation of fully pipelined microprocessors. The cost effectiveness of parallel-vector farms built from these chips would certainly justify the effort at vectorization, and experience at this stage will prove useful in the long run.

Ken Miura taught the right way to go about vectorization when he presented his work with EGS, the shower program. Identical calculation sequences from different events must be gathered up and presented together to the processor when there are enough to fill the vector width. (A similarly clever approach to taking advantage of true parallel machines was presented by Martin Pohl for reconstruction programs. But here the motivation is not as clear because of the trivial event parallelism that makes looking beyond the simplicity of farms for more complicated solutions unnecessary.) The Miura approach will likely be used to bring up big canned programs like EGS and Geant in the new CERN vector environment. Heavy use of these programs can keep the big new Cray busy. The major remaining stumbling block is geometries which differ from user to user of these programs and so can't readily be vectorized by experts. Unless this is solved, the vectorization of the rest of the programs will be in vain because of severe scalar bottlenecks.

### Compatibility and Standardization

A long standing issue that has become recognized as acute in the modern high complexity HEP environment involves the difficulties surrounding system compatibility and standardization. These problems are among the pleasures at technology's frontier. Both at Asilomar and two years ago in Amsterdam, discussion of standards often seemed to dominate the conference. Standards have a habit of being introduced both too soon and too late. Too soon because all the considerations, wish lists, optimizations, etc., that will ultimately be recognized as appropriate for the standard will not be included. Too late because as soon as a technology appears that warrants a standard, the standard is needed to preempt conflicts with a second implementation of the technology. Inevitably some poor soul, usually a high energy physicist, must pioneer before the standard(s) appear. When a standard becomes associated with a particular large manufacturer, marketing considerations again come into play and often spawn support of competing standards.

There is a long list of compatibility and standardization issues that were presented and discussed:

1. **Networks (WANs)**  
In Europe the countries can't agree on standards for high level protocols (Coloured Books, DECNet, X400-EAN, etc.), or at the low level (Datapak, Finpak, Transpac, Datex-P, Helpac, usw.). In the United States it's the agencies (DOE, NSF, DOD). There appears to be a conservation of confusion.
2. **Networks (LANs)**  
Token rings vs. Ethernet: IBM vs. DEC and others.
3. **Networks (front end)**  
DECNet, Bitnet, each has something to offer. Will ISO satisfy all needs, and when?
4. **Closed vs Open Systems: Software Portability**  
Multiprocessors, experiments, front ends, minds.
5. **One Company's hardware/software**  
IBM VM variety requires HEPVM; Macintosh software; DEC's DMA interface timings: all are not as standard as one naively might have expected.
6. **Languages**

Extensions to Fortran 77; delays in getting Fortran 8x and its inadequacies for parallel processing; no plans for a standard for procedure (subroutine) calls between different languages.

7. **Software Development Methodologies**  
We know we should prepare software by a set of systematic rules and procedures, but whose?
8. **Graphics**  
GKS (linear, segmented, and here -- at CERN) vs PHIGS (hierarchical, desirable, and not quite here)
9. **Legal Documents**  
Just as every city has a standard apartment rental lease form that is generally accepted, we need standard software licenses, non disclosure agreements, etc., to make it unnecessary for lawyers to delay progress.

## Conclusions

As much as struggling with compatibility problems is mundane (though important), solving our critical non-numerical problems is very much within the mainstream goals of a glamorous research area. AI research problems have become a very acute part of our work. Human interfaces, expert systems, software/algorithm certification (off line and in real time), and symbolic manipulation all now have entered high energy physics' area of concern. We will have to buy, more likely order and specify, as much of this technology as we can. However, most of it will not be available as we need it without our research participation. Perhaps, the title of the next conference in this series should be *Artificial Intelligence in High Energy Physics*. And its location should meet the standards set by Asilomar (and Amsterdam, Guanajuato, and Padua).

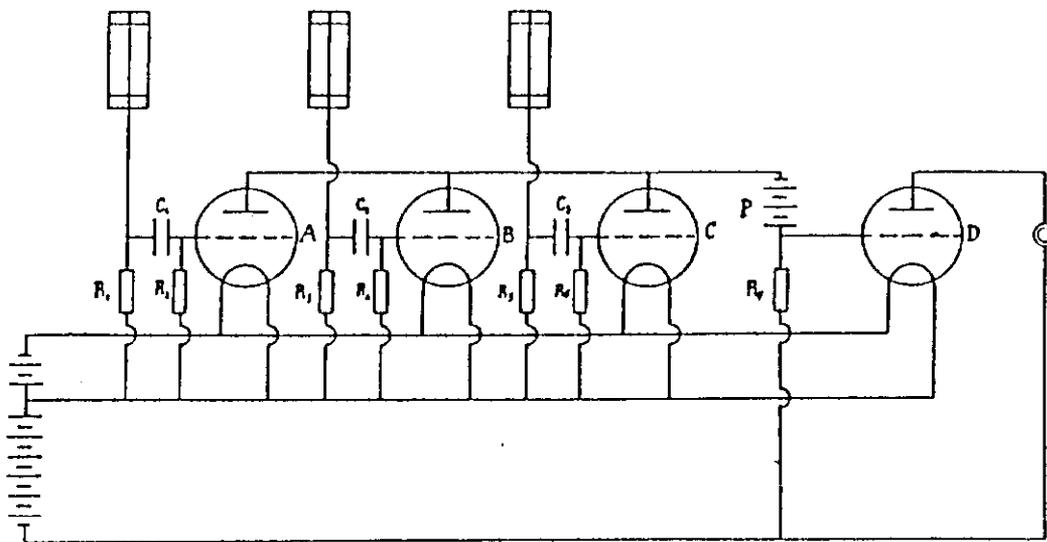


Figure 1. Particle physics computer technology in 1930: The first practical electronic AND gate, a Geiger Counter triple coincidence circuit. Pulses were “detected by a telephone” and scaled manually.

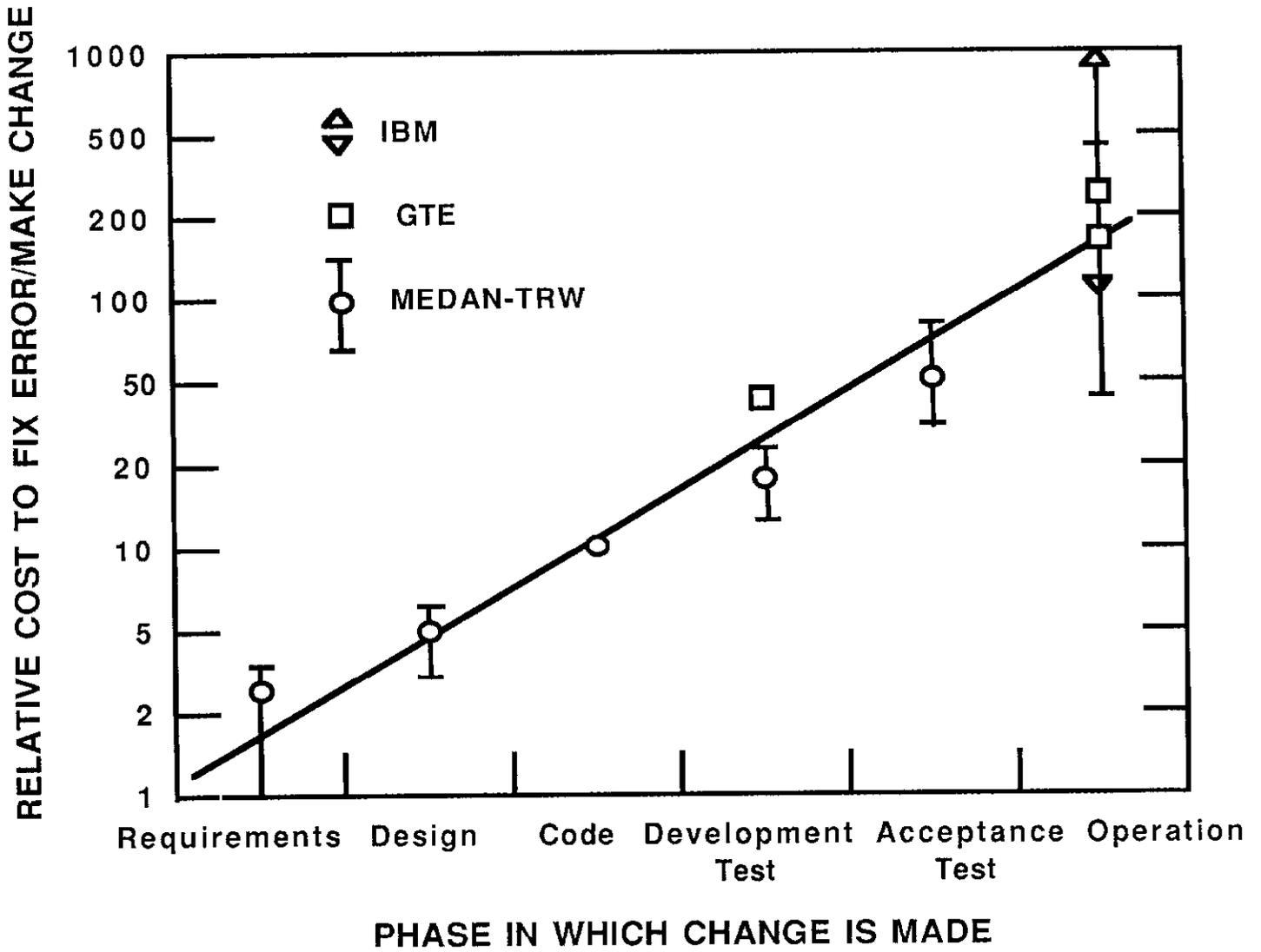


Figure 2. The need for software management.