CERN/DRDC 92-11
RD-11/Status Report
3 March 1992
(EAST note 92-05)

Status Report: Embedded Architectures for Second-level Triggering (EAST)

K.Bos, J.Vermeulen, L.Wiggers
NIKHEF-H Amsterdam

V.Buzuloiu, D.Coltuc, F.Folta, A.Gheorghe, I.Legrand, N.Mihai, A.Zamfir
Institute of Atomic Physics and Polytecnic Institute, Bucharest

E.Denes, J.Koch, T.Maroti, S.Molnar, G.Odor
Central Research Institute for Physics, Budapest

R.K.Bock (*), W.Krischer, R.McLaren, Z.Natkaniec, N.Tchamov
CERN, Geneva

Z.Hajduk, W.Iwanski, P.Kapusta, K.Korcyl, P.Malecki, A.Sobala
Institute of Nuclear Physics, Cracow

D.A.Belosludtsev, N.V.Gorbunov, V.Karjavin, S.V.Khabarov, V.S.Khabarov
LHSE, JINR, Dubna

B.J.Green, T.Medcalf, J.Strong, A.J.Wildish
Royal Holloway and Bedford New College, London

F. Klefenz, R. Männer, K.H. Noffz, R. Zoz
Institute of Computer Science V, University of Mannheim, Mannheim, and Interdisciplinary Center for Scientific Computing, Heidelberg

J.Badier, Ph.Busson, C.Charlot
Ecole Polytechnique, Palaiseau

P.Bitzan, M.Novak, E.Pelikan, V.Sedivy, I.Sklenar, J.Smejkalova
Inst.of Computer and Information Science, and Institute of Physics, Academy of Sciences, Prague

Y.Damatov, L.Levinson, M.Sidi
Weizmann Inst.of Science, Rehovot

W.Lourens, A.Taal
Utrecht University

H.Leich, U.Schwendicke
Institute of High Energy Physics, Zeuthen

(*) *Spokesperson*

1. Summary

The EAST Collaboration (RD-11) was formed to explore one of the aspects of the critical problem of event selectivity in LHC detectors, second-level 'intelligent' triggering. The assumed context, borne out by physics simulations, is this: after a reduction of the event rate by a first-level trigger, relying on calorimeter windows and muon identification, the event passing rate is of the order of ~100 KHz. At this rate, reduction 'algorithms' of some complexity will be required to reduce event rates further. They use physics features extracted from fine-grain local data in multiple detector windows, combining them into global decisions by correlating data from different subdetectors.

In particular, EAST explores the possibilities of introducing commercially available solutions at this stage. The collaboration uses as methods both studies with simulation, and prototype implementations of hardware.The members of EAST comprise physics institutes, university institutes in computer science and electrical engineering, and industrial partners. Close contacts are maintained with other R&D activities aimed at LHC, and with the activities in nascent LHC collaborations.

The key activities towards this goal were **defined at proposal time**:

- definition of characteristic second-level trigger tasks in terms of physics goals, detectors and triggering algorithms, and detector and first-level trigger electronics;
- 'benchmark implementation' of such algorithms on competing architectures, in simulation or prototype hardware;
- preparation of a test environment for architectures capable of a decision rate up to 100 KHz, by emulating local detector data output streams in special hardware.

Benchmarking demonstrates architectures' possibilities and limits, and permits to evaluate a first approximation of their cost. Practical implementations show the problems which will have to be solved in embedding architectures in the data flow of future detector electronics.

The EAST Collaboration has **achieved** all of **the goals** set out for 1991 in the Proposal and the Addendum (DRDC/90-56 and /91-13), except for external delays independent of our work. The main achievements are discussed in chapter 3 and 4, and in a detailed comparison with the proposal milestones in chapter 5; they are the following:

- Definition of trigger tasks, and the generation of corresponding simulated data sets. Together, they serve as a benchmarking exercise for different architectures, which is presently under way (detector models are the TRD tracker and a calorimeter);
- Design, construction and testing of a detector emulator capable of feeding target trigger architectures with program-controlled data patterns corresponding to detector readout, at the speed expected for second-level triggering at the LHC;
- Practical tests with a high-bandwidth point-to-point connection standard increasingly accepted by the computer industry (HIPPI);
- Simulation of the FERMI chip (RD-16), using as tool the hardware definition language VHDL.

In addition, much preparatory work in hardware and software has been done to achieve, in 1992 and 1993, hardware demonstrations of several architectures, using both the emulator hardware and detector prototypes.

The **goals for 1992** are discussed in chapter 6. In short, the collaboration expects in its second year to

- complete the running benchmark exercise, and draw conclusions on the suitability of the various architectural candidates;
- improve the benchmark data quality for a next round;
- focus on very few promising architectures and implement their essential parts in demonstration hardware, running with the emulator and/or detector prototypes in beams;
- include high-level decision processors in the retained architectures;
- initiate studies (using simulations) for moving trigger systems to realization in future LHC collaborations, where embedding in an overall data acquisition system will be necessary, and system aspects like error handling, maintainability, evolution, will be relevant factors.

2. Choices of principle in the second-level trigger

This chapter places the activities of our collaboration into a more general technical context. We believe that it will help in explaining some of the terminology used in the subsequent discussion.

To a large extent, work on second-level triggers depends on the progress of detector and frontend electronics developments, and hence ultimately on physics choices: the closer a detector moves towards (prototype) construction, the more meaningful become our trigger models. As requirements become more precise, pilot hardware implementations of trigger architectures gain in significance. We work in close collaboration with all R&D activities concerning electronics of R&D detectors (RD-12, RD-13, RD-16), and with electronics and simulation activities inside some of the detector development projects (RD-1/RD-3, RD-6). The existing close collaboration allows us to identify (and make) some basic choices concerning the embedding of second-level trigger architectures in the data flow of future detectors.

There are also important technological choices to be made in organizing the readout electronics, which are of capital influence on the second-level architecture and its embedding. We will briefly outline some general principles in the following.

Pipelines, decision frequency, and latency

One choice concerns the temporary storage in buffers or pipelines of data during the latency period of the second-level trigger. *Latency* is the time it takes for a given event to run through the execution of the trigger algorithm(s), as opposed to the *decision time*, measured between successive decisions. *Pipelines* are active or passive buffers of the FIFO (first-in-first-out) type, which events run through in monotonic sequence, possibly with a constant clock speed. General *buffers*, on the other hand, allow non-sequential access to events. Trigger processors most simply achieve speed by operating on multiple events in a pipelined sequence, decomposing the necessary operations (the 'algorithm') into multiple steps. The latency time may then be larger than the inverse decision frequency by a large factor, with the effect of imposing buffers on all collision-related data in the detector.

Pushing and pulling data

The buffering choice, then, is the following: in many data acquisition systems today, data is *pushed* by the frontend electronics into separate memories, typically remote from the detector. In LHC, data will mostly be kept on or close to the detector, in a pipeline, during first-level decision making. Only bunch crossings indicated as interesting by the first-level trigger are then passed through FIFO-s (to derandomize arrival time), and *pushed* typically through some multiplexing logic to a second-level buffer and/or trigger. This readout model lends itself to simplification by maintaining a (near-) synchronous

and monotonic data flow from the front-end into the buffer memories and requires a selective switch-and-copy unit (the 'Router'), which transmits and simultaneously copies selectively into the second-level architecture.

An alternative is being investigated by RD-16 (FERMI): buffering, in this architecture, is intended to remain on the detector until a second-level decision is available. The same physical buffer originally holds data from all bunch crossings, and later only the data belonging to first level-accepted crossings. A fast hardware memory manager, common to a group of FERMI chips, takes care of dynamically retaining only data still potentially useful. Rejection in level 1 results simply in reusing the corresponding buffer space. A first level trigger can identify the zone in which an interesting phenomenon has been observed, the 'regions of interest' (RoI). In our model, RoI-s are communicated directly to the second-level architecture, which has now to *pull* the relevant data by a 'read'-like operation. The intended gain is a vastly reduced bandwidth for all data except the fraction needed for forming the second-level decision, at the expense of software-like pointer management for the buffer, and an added task (selective reading) for the second-level device.

Occupancy, thresholding (zero-suppression), iconic processing

A classical way of reducing bandwidth and buffer requirements in data transfers consists of utilizing the fact that on any given event (bunch crossing), many channels have not been activated. One calls *occupancy* the probability of a channel being 'hit' by a physics-related phenomenon in an event. When the occupancy is low (i.e. a channel covers only a small physical region), *thresholding* of all signals allows to transmit only the significant information, which then has to be accompanied by an 'address', some additional information allowing the data to be associated to the active channel. Implied in such 'zero suppression' is, of course, that some downstream logic uses the address in processing the data. Easy neighborhood relations, as useful in feature extraction, are lost by the thresholding process. A general-purpose processor is obviously capable of dealing with lists of data items each of which carries an address, and it is then in a *symbolic processing* mode. Image processing equipment, on the other hand, much like fast data switches, are better at ease with streams of information in which the address is implicit in position and time of arrival; they typically perform *iconic processing*.

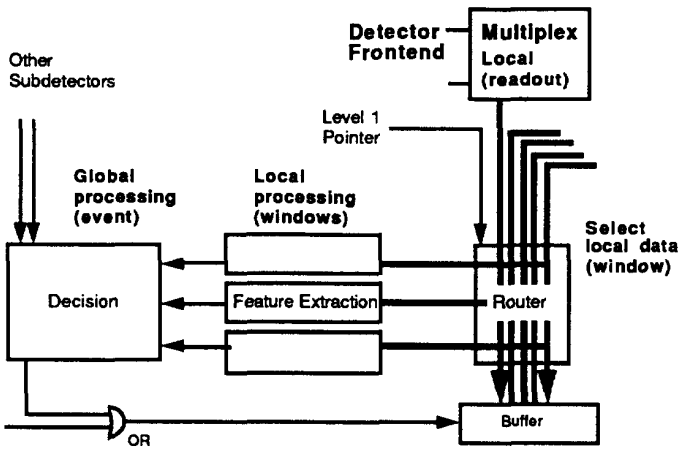
Communication

In most of our present experiments, digital data transmission uses the concept of a *data bus*: several data streams are possible over a single hardware line, whose capacity typically exceeds the required bandwidth of any of the connections it mediates. A bus standard is characteriized by a bus protocol including arbitration for conflicts, a clear master - slave relationship, and interfaces to the software running in processor components. Examples are CAMAC, VME, and Fastbus.

The generality of the bus approach sets limits to the achievable bandwidth, and makes them an uneconomic overkill for the simpler low-level communication tasks in future experiments. Most of these can be simply served by *one-directional point-to-point* links. In the computer industry, such connections are also required for high-bandwidth computer communication. In conjunction with switching equipment and higher-level protocols, the same links can also serve a more general high-bandwidth *networking* function, and again it will be mostly the computer industry that will provide us with the components for solving our data acquisition-related problems.

An example

Schematically, the data flow for a 'push' alternative without thresholding is represented in the following diagram, showing an implementation with data multiplexing as it is foreseen in the prototype transition radiation detector of RD-6.



The data flow in this diagram is assumed iconic, viz.constant event size, and could also be at constant clock rate. The diagram shows the important subdivision of the trigger task into the three phases data selection ('Router'), feature extraction, and global decision. It is the Router's task to extract the RoI(-s) from the total mass of data, whereas the feature extraction unit converts, locally, raw data into physics-related quantities like tracks or clusters. It should be understood that a second-level processor of the general-purpose type, fully programmed and with a global data memory, would not have to make this distinction: software can easily access data using selective addressing of a global memory, and feature extraction and decision making simply correspond to different phases (e.g. subroutines) in the program. Such a model, however, is largely unsuitable for the data rates and execution speeds required for our assumed environment.

Let us complete this diagram by some numbers. The envisaged full transition radiation detector of RD-6 could follow the above scheme, with its data flow determined by data volume and repetition rate: data come from some 500,000 individual drift chambers, each producing 2 bits of information, at an average rate of one first-level triggered event in 10 microseconds (calorimeter's first-level reduction). The total data flow into the decision system thus is 100 Gbit/sec (we ignore the potential necessity of additional information from successive time slices). The present working hypothesis is that this data flow can be maintained over a number of parallel HIPPI lines, each of which presently can run at up to 1 Gbit/sec (HIPPI is discussed in later chapters).

A fraction of this data volume, indicated by a pointer coming from the first level trigger, will be used for feature extraction and decision making in the second-level trigger. This fraction, the region(s) of interest, is a few %, which is a good guideline for most detector models. The sustained data flow into the feature extraction units thus is typically 1-5 Gbit/sec. Once converted to features, a short high-level description of the characteristics of every RoI will suffice: bandwidth for transmission thus is not an issue when communicating with the decision unit proper, although a very fast or regular interrupting possibility is important.

3. Activities of 1991/2

The EAST collaboration has initially concentrated on understanding the central issue of *feature extraction architectures* suitable for several pilot tasks. It is in this sector that our familiarity with commercial architectures and/or suitable components has to be built up largely from scratch. Experiments at colliders of the past have made a very clear breakpoint between time-critical (say below 100 microseconds) tasks, and tasks accessible to commercial systems or system parts. The faster tasks were implemented as custom-designed hardware. They include our definition of 'second-level trigger'. Commercial devices were confined to operations at slower speeds, sometimes piling up

many of them (processor farms) to obtain high average throughput, at the expense of long latency.

It is the intention of the EAST collaboration to push the limits of what can be done by commercial components towards the highest possible decision frequencies, hence far beyond what has been attempted in the past.

Understanding of architectures is achieved using benchmarks, viz. standard tasks which set a rigidly comparable environment. Architectural studies as a minimum use simulations; they may or not be backed up by measurements done on existing full or partial hardware. If sufficiently promising, some architectural models may even be built up to become significant hardware installations including data transmission functions with LHC characteristics. In order to demonstrate for some second-level trigger systems the detailed functioning of hardware, we have built a hardware test device, an *emulator*, called SLATE. This will permit implementations of feature extraction devices to be subjected to real-time tests. We have further invested in practical tests on the use of industrial *communication protocols* and associated hardware, for unidirectional point-to-point links of high bandwidth. The HIPPI protocol was selected to be at least a temporary standard, a definition necessary to stabilize the critical job of embedding.

We will expand on these points in the following, leaving several more technical details for separate discussion in chapter 4.

3.1 Benchmarking feature extraction algorithms

Benchmarking consists of the definition of a controlled task that can be solved using different hardware implementations, and of a systematic comparison of these implementations. In evaluating main frame computers, from where the term is borrowed, the task is usually given as a mix of characteristic (mostly large and long-lived) programs, which are then submitted to the system, evaluating simultaneously the computer hardware, the compilers, data I/O, the scheduler and other aspects of the operating system.

For the case of EAST, the task definition is that of the feature extraction part of second-level trigger algorithms, i.e. the conversion of raw data, over a given region of interest, into quantities meaningful for physics. The algorithms are simple enough to be reprogrammed or even hardware-implemented in dependence on the architecture to be studied. Although the algorithms are given, for a clear definition, in a high-level language, compiler performance is not one of the evaluation criteria. What is to be evaluated has been mostly defined during a workshop held at CERN in October 1991. The main criteria are

- algorithm execution time, for pipelined architectures separated into the two aspects *frequency* (inverse time interval between successive decisions) and *latency* (time interval for a given event between start of data input and output of results): this assesses the overall feasibility of a given architecture to contribute as a second-level trigger device (target numbers are 100 KHz for frequency, and of the order of 1 msec for latency);
- practical solutions to the *high bandwidth input*: this addresses a typical bottleneck for many commercial systems which are targeted at compute-intensive problems, and also challenges the flexibility of architectures or their manufacturers in interfacing to specific user constraints;
- possible constraints on the *order of input* data; this aspect is relevant as we deal with architectures that typically achieve performance by high parallelism with distributed memory, or by pipelining data in a certain sequence: we assess here how much of the data selection (the Router in the above figure) has to be loaded with tasks that would be, in a general-purpose device, part of the algorithm itself;
- *interfacing* to a high-level decision-making unit and to the (physicist) user: flexibility with respect to algorithm parameterization, and hardware possibilities of passing results (physics features) to a global device for overall decision making, are critical parameters in assessing the embedding difficulties of an architecture.

3.1.1 Target algorithms

We have defined two pilot tasks for our benchmarking of second-level trigger algorithms, both in close collaboration with R&D projects pursuing the detector developments. They are defined in internal EAST notes (see list at end of this report). The pilot tasks and algorithms reflect a certain state of detector development, frozen for our purpose. They refer to feature extraction in a single window, leaving the connection between possible multiple windows in the same detector, or the correlation between features in different detector subsystems, to the global second-level decision unit.

- a *calorimeter* algorithm based on an indication from a first-level trigger window of a 16x16 cell region. In our simulation this RoI corresponds to $\Delta\eta x \Delta\phi = .5x.5$. Each cell is assumed subdivided into several electromagnetic and hadronic volumes of cylindrical and wedge shapes, as explored in SPACAL (RD-1). The objective of the algorithm is to find features, i.e. decision variables. Second moments of clusters in two dimensions, or isolation in different ring-shaped zones could be such variables, likely to contain relevant information for electron/pion discrimination that can not be explored in a first-level algorithm, and to refine the positional resolution. Criteria must remain flexible to account for different physics goals (e.g. 'jets' from τ-leptons) or luminosity.

- a *tracking* algorithm based on a straw geometry as pursued in RD-6 (transition radiation detector), in a magnetic field. Again, the assumption is that a preceding step selects a RoI of TRD data for further treatment; the algorithm then recognizes patterns of digitizings that correspond to a high-momentum tracks, taking into account the pulse height distribution of digitizings for identification of electrons.

3.1.2 Target architectures

The above pilot algorithms are implemented, as systematically as possible, on different architectures. The architectures are discussed briefly below. A first comprehensive comparison of these architectures, using our benchmarks, will be made at a workshop foreseen for May 1992.

MaxVideo

MaxVideo is a fully commercial pipelined image processing device, which CERN has accumulated experience with since 1989: developed for quite different applications, like in television, aerial surveillance or medical scanning, image processing systems of this type are 'computers' made up from a variety of special-purpose architectural modules, that can be freely combined into one or several parallel pipelines using a proprietary data connection. All devices can be purchased off-the-shelf from industry; their performance and that of the connecting data lines is evolving over time with a rate similar to that of microprocessors, and their internal and external interfacing possibilities are better than on most commercial computing devices. Pipelined image processors typically execute two-dimensional neighborhood operations like convolutions or morphological operations, useful for feature extraction. They develop for these tasks an enormous equivalent of computing power.

DataWave (DAVIS)

This is again an image processing component, with an on-chip multi-processor mesh. Each processor is independently programmable, and their communication driven by the data. Data can be pipelined through these processors at very high speed, multi-chip blocks are possible, but at the price of reduced data speed. Presently, the DataWave architecture is under development in the context of HDTV (high-definition TV), first chips exist as prototypes, not yet available to us. We have installed a simulator under license, and hope to have access to first chips in 1992. The possible contribution of DataWave processors to the benchmarking of the calorimeter algorithm is under study, early results with a simplified algorithm are encouraging.

Commercial SIMD Machines

Single-instruction-multiple-data (SIMD) machines are considered by some to be the future architecture of choice for massively parallel problems. They are being marketed increasingly for some supercomputer applications (e.g. finite element calculations). Such systems could, in principle, be scaled with advantage to fit our real-time applications. SIMD devices are typically made of many, thousands or more, of mesh- or string-connected processors, each with its own memory, all working in lockstep under a single instruction issuing unit. Individual processors more often than not, are far from being general-purpose devices. 1-bit or 4-bit processors dominate, with the basic instructions implemented giving typically high weight to communication. Processors are fully programmable, however, and hence of considerable sophistication. For suitably parallelized applications, the equivalent of substantial computing power can be extracted. Connectivity is a key issue; typically, processor communication is the more efficient the closer two processors are, say, in a two-dimensional mesh. Such a SIMD system seems a priori to suit well the multi-channel parallelism naturally present in our detectors. It should be said, on the other hand, that commercially available devices (Connection Machine, MasPar, DAP) presently do not satisfactorily address the problem, critical for our application, of maintaining a permanent parallel data flow of high bandwidth into the machine, and of overlapping I/O with execution of the compute task (using a different window or event). A major embedding problem (custom-made I/O) thus has to be solved for our applications. Also, systems are packaged as stand-alone or attached systems for supercomputing, which results in low numbers of systems sold, and a pricing strategy that makes it difficult to imagine today that they could easily find their way into the challenging real-time environment of our future experiments. Nevertheless, three SIMD devices are under study: MasPar, one of the commercially available systems, (EAST has been granted access to an installed 4096-processor system), ASP, which is a system whose present prototype development proceeds with high-energy physics real-time criteria in mind, (but which is presently available only as a software simulator), and Blitzen, also a system existing as a prototype that could be adapted to our real-time constraints (and which again is being performance-evaluated using a simulator).

Neural networks

For some tasks with limited input data, neural networks (NNs) have been shown to present a possible solution path, even with decision frequencies fast enough for first-level triggering (if based on analog signals). NN implementations should, therefore, be systematically compared, using the EAST benchmark algorithms, to the performance of traditional electronics or processors with respect to the same tasks. As an early result, a comparison has been made and reported, using NN simulation, between feedforward neural networks with linear threshold units and normal logic gates, in their ability towards recognition of very local *binary* patterns, with the result that both are close competitors. This would be applicable to first-level triggers after thresholding, or for minimally guided second-level triggering. For input patterns with *grey values*, interesting features of NN-s have been demonstrated by the design of a neural trigger for a calorimeter, using a simple feedforward NN. It performs the identification of electromagnetic showers in the benchmark data for SPACAL as well as any other algorithm. A peculiarity of learning NNs is that the implemented 'algorithms' do not contain explicit physics cuts; the use of carefully tuned training samples is, therefore, of paramount importance. This aspect is under work for our benchmark data sets, and translates into massive Monte Carlo work. For the particular case of a calorimeter algorithm, it has been demonstrated that the neural network architecture and the weights can be determined without applying a learning algorithm. Another architecture, which needs preprocessing of the data, and for which the backpropagation learning algorithm was applied, also yielded good results.

Semi-custom designed processors:

Special-purpose processor architectures can be tailored to the needs of a given detector/algorithm, and it will then be difficult to outperform them by any device with more general applications in mind. This need not imply full-custom design of processors: field-programmable gate array devices can provide some flexibility in custom-designed solutions, and reduce substantially the development time. Implementations using this technology may take even less time than that needed to adapt a non-ideal commercial system to our problem. Using existing experience, there is now a development project in EAST to build a systolic array processor based on field-programmable gate arrays (Xilinx), for triggering the transition radiation detector of RD-6. This is called the Enable machine. It will be directly plugged into the TRD frontend readout electronics (prototype) via a HIPPI serial-link interface. The Enable array operates like a systolic processor array. The pixel image is shifted one column at a time into the rectangular mesh-connected array of processing elements, which execute on every operand the same function. According to the underlying deterministic space-time sorting and shuffling procedure, all data is routed to the corresponding histogram channels. The systolic data flow is maintained throughout, from data input up to the activation of the trigger lines. The system can be scaled by changing the number of standard boards used. One board handles images of 60 pixels height, and provides, according to the benchmark algorithm, m*k histogram channels, where m is the number of different bin offsets (=20) and k is the number of slope patterns (=15). A processor prototype board development is under way, the board will serve for testing and evaluating the TRD-emulation in SLATE (see 3.2 below). Simulation studies have shown, that the systolic machine will execute the TRD benchmarks with only minor deviations from the classification results shown off-line. The data flow has been tested on a clock-cycle basis with a hardware simulator. The prototype will be realized with modest cost and will be functionally testable in 1992.

General-purpose processor farms

Any solution introducing non-familiar types of processors and special-purpose architectures imported from other applications, possibly adaptable to high energy physics tasks only with difficulty, must be measured against the 'brute force' solution of using a massive number of general-purpose devices, and coping with the ensuing problem of long latency by event buffering. Such general-purpose devices, typically reduced-instruction-set-computers (RISC) or digital signal processors (DSP), are arguably on the fastest development edge of technology. They are fully programmable at high level and understood by the physics community, hence constitute a good platform for interfacing with physicist users. Their accumulation into large farms has already been solved in our laboratories in the past, albeit for quite different boundary constraints. Many recent general-purpose high-performance computing systems are based on the same idea.

RISC-s or DSP-s, particularly when offered with integrated communication controllers, are most likely the ideal processors for implementing third-level trigger tasks, i.e. at event input rates of 1000/sec. The same type of device also constitutes the best candidate for implementing the global decision part of second-level trigger systems based on special-purpose feature extraction devices. Some activity in EAST is, therefore, being devoted to understanding the limits of embedding RISC processors or modern DSP-s in the context of second-level triggering.

3.2 Data emulation and beam tests

Some benchmarking is being done by implementing the pilot algorithms in simulators, or on available hardware without investing the substantial added effort of connecting to detector-specific hardware. Missing parts of a full implementation will then have to be conjectured. The more promising solutions, though, should be demonstrated as more complete prototypes, connecting to hardware that corresponds to the (sub)detectors as expected to exist as parts of LHC experiments, and with rates corresponding to what is

predicted to be input to the second-level triggers. In the absence of suitable detectors (LHC-like frontend electronics must be part of trigger testing!), we have opted already at proposal time for building up a test environment for second-level trigger architectures that satisfies these conditions. A software-driven device called SLATE (for 'Second-Level Architecture Test Equipment') has thus been designed, built, and tested, which allows us to transmit a succession of bit patterns corresponding to the expected detector output, into the architectures that are meant to be subjected to high-speed hardware tests. Presently, this is divided into a mother board, containing an interface to an OS-9 system and the necessary event and trigger buffers with logic. A daughterboard translates the output into the HIPPI protocol (see below), using a commercial chip, and is, in principle, replaceable by other interfaces. Hardware testing of architectures will be done as soon as the interfaces exist (several are presently under construction).

While SLATE provides a platform for full laboratory testing of some of the target architectures, we believe there is no substitute for the more rugged requirement of testing in a particle beam. Conditions there may, of course, be less demanding than in emulation, due to bottlenecks in beam or transmission rates. Implementations around a detector prototype may also necessitate shortcuts because incommensurate developments become necessary in matching trigger and detector prototype. We foresee, nevertheless, to install and, where required, develop some demonstrations of second-level trigger architectures as part of the beam testing program planned for the detector prototypes of the transition radiation detector and of a suitable calorimeter. Originally, only non-interfering event tagging is foreseen, to avoid any risk for the other detector testing objectives. Ultimately, it will be from these installations that a final acquisition and trigger system will emerge.

3.3 HIPPI communication protocol

In connecting multiple components of the trigger and acquisition system, stability in and a clear definition of the conventions used for the various high-bandwidth connections are vital. We have opted for an ANSI standard for a unidirectional protocol, increasingly used in industry for connecting computers, the high-performance parallel interface (HIPPI). Traditional transmission protocols of high energy physics experiments (CAMAC, Fastbus, VME) do not provide the required bandwidth, or the simplicity on which high bandwidth can be built, and are not (apart from VME) supported by industry outside high energy physics - a prerequisite for economic mass implementations of the future. It is obviously not clear that HIPPI will remain the best choice for decisions to be taken around 1995, assuming that to be the time when LHC experiments get finalized; at present HIPPI seems a natural choice: many of the architectures under consideration (plus others that are not presently studied in EAST) have a commercial HIPPI interface available or under design.

General HIPPI work has started in 1989 by following the work in the ANSI working group for high performance interconnects. The first HIPPI chipset tests were done in 1990, commercially available components were purchased from 1990. In 1991/2 we have developed components not available from industry, and will in 1992/93 use them in test systems and in detector prototypes.

4. Technical details

In this chapter we give some details of the main work mentioned as achievements of 1991; it is to be considered an appendix to chapter 3 for the technically interested reader.

4.1 Benchmarks

4.1.1 Tracking (Transition radiation detector)

The purpose of the TRD benchmarks is to define a portable set of data for use in a comparison of different hardware implementations of a trigger based on this data. The specific task of the 'second-level' trigger is to recognize electron-like tracks in the TRD, based on hits aligned along a straight line in z-ϕ space, and distinguished by their pulse height distribution. The analysis is restricted to a RoI (region of interest) defined by a preceding 'first-level' trigger, which is assumed to use information from an electromagnetic calorimeter.

The agreed working hypothesis for the geometry of the TRD is descibed by the status report of RD-6 (CERN/DRDC/91-47, 22 Oct 1991). We have used it to simulate several hundred single electron and pion tracks, combining them with multiple minimum bias events to achieve realistic pileup situations (three scenarios: low, medium, high luminosity). Only the forward endcap part (pseudorapidity from 0.5 to 2.0) was considered.

Data sets were generated in track format, in hit list format and as images. Only the image information (pixels in RoI) is used for benchmarking. The RoI is chosen to represent the crude selection of data in a Router. This unit reduces the necessary bandwidth into the second-level architecture. Our RoI was chosen to be a rectangle of 80x240 pixels.

The principle of the algorithm is to find if at least one electron-like (by pulse height distribution) track can be identified, which ends in the calorimeter impact cell and is fully contained in the search window. We have expressed the algorithm in high-level language so that its intention can be expressed unambiguously, but this is not intended to be more than an example of a working implementation.

This algorithm has the following steps:

For suitable bins of impact point (end point, i.e. calorimeter), and charge/p_t
- establish a search road
- count separately all low and high pulse height hits in that road
- find the most electron-like track
- compare this track to minimum parameters for track quality (n_{low}) and for electron likelihood (n_{high}/n_{low})
- output the features:
  - yes/no above threshold
  - if yes, also the maximum bin's parameters (which indicate p_t, charge, impact point).

The criteria of judging a given implementation are these:

- statistical performance in discriminating the electron and pion sample
- statistical distribution of the number of hits (low/high threshold) on the track
- precision of the solution found for track parameters
- embedding problems: high-bandwidth input, feature output.

4.1.2 Calorimeter (a model based on SPACAL)

In our benchmark data we assume that the first level trigger indicates a window of 16x16 towers, spanning 0.5 in both η and ϕ. This 'image' is the input for the second-level trigger. Each of the 256 'pixels' contains several energy values: electromagnetic energy, hadronic energy, and (a peculiarity of our SPACAL model) a mixed e.m./hadronic energy deposition in a wedge-shaped element.

The particle discrimination algorithm makes use of the longitudinal segmentation of the tower (electromagnetic, hadronic and wedge volume) and of the lateral shower profile, to extract three features which provide a relatively good identification for e , π and jet, in different situations of pile-up (low, medium, high luminosity).

The main steps of the target algorithm are the following:

a) Find the center of gravity for the energy deposition in the window using all energies and the geometrical setup of the detector:

$$x_c = \frac{\sum_{i,j}^{16} x_{i,j} E_{i,j}^T}{\sum_{i,j}^{16} E_{i,j}^T} \qquad \text{and} \qquad y_c = \frac{\sum_{i,j}^{16} y_{i,j} E_{i,j}^T}{\sum_{i,j}^{16} E_{i,j}^T}$$

x_C and y_C are the coordinates of the center, $x_{i,j}$ and $y_{i,j}$ are the coordinates of each tower and

$$E_{i,j}^T = E_{i,j}^{em} + E_{i,j}^{ha} + E_{i,j}^{we}$$

is the total energy deposition in each tower (the sum of electromagnetic, hadronic and wedge parts).

b) Define a circular region around this center, which includes all neighbors of the central tower up to five towers away in radius (η,ϕ space). All towers outside this region will be ignored in the subsequent calculation. We have thus defined a moving window of diameter 11, fully contained in the original image. This local 'region of interest' (RoI) mimics the necessity to transmit for analysis a window larger than the interesting cluster.

c) In this RoI we calculate the following feature parameters:

- the total energy deposit:

$$E = \sum_{ROI} E_{i,j}^T$$

- the hadronic fraction:

$$f_{ha} = \frac{\sum_{ROI} E_{i,j}^{ha}}{\sum_{ROI} E_{i,j}^T}$$

- a tail shape discrimination function defined by:

$$f_{tail} = \frac{\sum_{ROI} \log(E_{i,j}^{ha} + E_{i,j}^{we} + 1)\, r_{i,j}}{\sum_{ROI} \log(E_{i,j}^{ha} + E_{i,j}^{we} + 1)}$$

where $r_{i,j}$ is the distance between each tower and the central one (tower centers are used). In this relation the energy is measured in MeV, hence is >>1. This function gives relatively good results for π/jet discrimination (EAST note 91-10, see list at end of this report). It corresponds to a 'first moment' of the radius, weighted by the logarithm of the energy. We also use a second moment of the radius defined by:

$$f'_{tail} = \frac{\sum_{ROI} E^T_{i,j}\ r^2_{i,j}}{\sum_{ROI} E^T_{i,j}}$$

Simple functions of these features have been shown to provide a good separation of e/π, e/jet, and π/jet, even for a high pile-up.

These algorithms are based on a preliminary analysis of our present benchmark data; they are open for future improvements. More realistic data for the second-level trigger task, filtered by the first level trigger, will have to be used, and probably newly optimized feature parameters will then have to be found for different tasks of particle or jet identification. Our hypothesis in discussing the feature extraction algorithms above, is that the characteristics of algorithms will not change dramatically, even with more realistic simulations.

4.2 The emulator (SLATE)

The task of the SLATE system is to provide data to second-level trigger architectures under test at a rate and of a complexity that allow a realistic appraisal of the performance of the architecture. With very few firm proposals for data moving systems the only reasonable assumption to be made is that data transfers will have to operate at the maximum rates possible for the buses and links being considered.

Where possible, for general purpose equipment, we have specified devices that are readily available and are current standards. If a choice exists, we have opted for flexibility and wide access. During the lifetime of the system, some current standards will be superseded and we expect to keep pace with these changes, but we believe that few modifications will be necessary to accommodate these changes.

To meet the data rates needed and the flexibility to cope with different detectors, different trigger architectures and different links between them, we have kept the design of the SLATE data module simple and performed the data generation and formatting in software. Basically, the SLATE data module consists of a memory into which pre-formatted data can be loaded, and a link from the memory to the architecture under test. Data can be loaded into the memory slowly and then read out at the rate needed for the test.

We have assumed that detectors will write information to the second-level on receipt of a first level trigger and as soon as they are free to do so i.e. the second-level system must be ready to accept data at all times and not expect to be able to request data as and when it wishes. This is typical for a 'push' architecture. At present, therefore, the SLATE data module outputs data to the architecture under test and does not receive flow control signals from the architecture. In order to increase the versatility of the system, we intend to incorporate some system of flow control into the next version of the data module.

This is how SLATE works:
Physics generators and detector simulations are used to produce standard data files. From these, further files are constructed containing data from regions of interest to the

second-level architecture being exercised and relevant data from the level one trigger. The data for a set of events are accessed by a workstation and formatted into a form suitable for the SLATE data module. Information from the user on data rates, link protocols and other parameters are added at this stage. A Unix based workstation running X-windows is used with a connection between the workstation and the SLATE data modules over Ethernet. Other systems could be used without significant software modifications.

The data for the test system are transferred over Ethernet to SLATE data modules housed in a VME crate (see Fig.1). VME operations are supervised by a resident CPU running OS9 with programs written in C. As it uses only the P1 backplane connector and functions only as a VME slave, a SLATE data module will work with any VME master .

Data and Trigger modules are constructed in two parts; the SLATE data module mother-board contains the memory section and a daughter-board contains the link electronics. The mother-boards for both modules are identical; the daughter-boards are those required for the links to the architecture under test and may be different for trigger and data.

The principal elements of the mother-board are the control and event memories. Fig.2 shows only the interconnections for data readout; the VME interface and the bus for loading and controlling the operation of the module are not shown. The event memory is 48 bits wide and contains event data and link protocol information; it is constructed from twelve 64k by 4-bit 25ns memories. The control memory contains a list with a start address and the time to the start of the next event for each event stored in the event memory. The control memory uses four 64k by 4-bit memories to store event start information and four similar memories to hold the delay times between events. The start address and delay memories use common address lines. For a data module the information stored is the simulated detector data; for the trigger module it is the information available to the second level from the first level trigger and includes event identification.

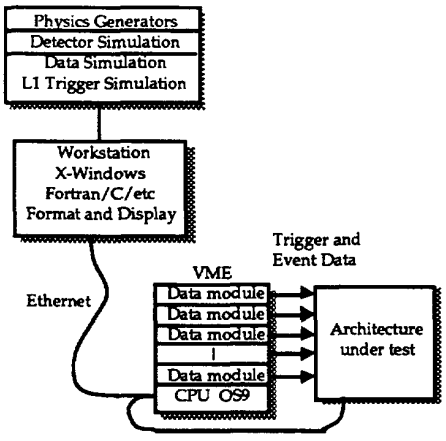


Fig. 1 TEST SYSTEM

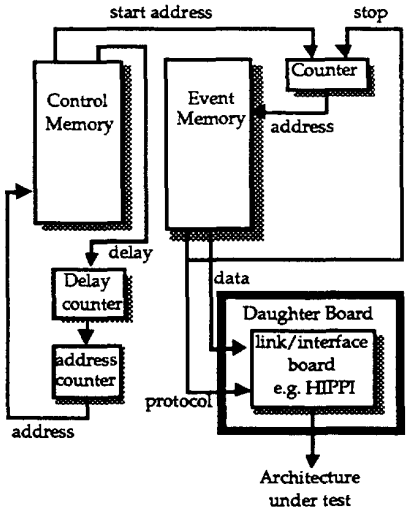


Fig. 2 TRIGGER / EVENT DATA MODULE

On receipt of a start signal, the first start address from the sequence memory is loaded into the event memory counter and data are read from the event memory until an end of block is found (bit 47 in the event memory) The event memory information for the event is passed to the daughter-board and translated to the format and timing sequences required by the link. The readout clock, distributed from the trigger SLATE data module, is used as the clock for both mother- and daughter-boards so that no change of

data format is required for a change of link operating frequency. The control memory is split into two parts, the sequence memory, which contains the start address of the events in the event memory, and the delay memory which contains the time between the start of consecutive events. At the end of the delay time the next address in the sequence memory is accessed and the process repeated. For simple tests, a partial or a single pass of the control memory can be made; for more complex or longer tests, the control memory can be programmed to cycle repeatedly through the event memory. The events may be called by the control memory any number of times and may also be cascaded so that there is no delay between consecutive events

Information from two detectors, SPACAL and TRD, indicates that regions of interest for second-level triggers would contain about 2 kbytes of data. Transfer of this data over 100 Mbytes/sec HIPPI links would need two links, and therefore two data modules, to achieve a transfer time of 10 μs. Each module would store about 250 words of 32 bits (1 kbyte) of data for each event, and 250 events would be stored in the data part of the event memory. A clock frequency of 25 MHz is required to achieve the 100 Mbytes/sec transfer rate with 32 bit words. With the addition of the 16 bits needed for the protocol information, the transfer rate between the mother- and daughter-boards reaches 150 Mbytes/sec. For data transfers from more than one region of interest, or for higher rate transfers, more data modules may be added and run in parallel.

Note that 250 events correspond to a time of only 2.5 ms at a 10 μs trigger interval. With the control memory of a similar size to the event memory, a single pass through the control memory expands this interval to 0.7 seconds and uses each event about 250 times. For longer test intervals, multiple passes through the control memory could be used. By setting the delay value to maximum, a single pass through the control memory can be expanded to take up to 86 seconds. To allow intervals longer than 2.5 ms between events, delays can be cascaded at the expense of the number of start address entries.

Status and Control registers on the mother-board with connections to the daughter-board are used to set the mode of operation of the module (e.g. Load, Stop, Go, cycle etc.) and monitor its status (e.g. Reset, Stopped, Waiting for start etc.).

4.3 HIPPI

We complement here what was said in chapter 3 on the HIPPI standard.

The 'High-Performance Parallel Interface' is a standard for the physical layer of an efficient simplex (one-directional) point-to-point interface at a speed of 100 Mbytes/sec. It is defined for twisted-pair cables of 32 bits width, operating at 25 MHz, over a maximum distance of 25m. Data transfers are mediated by bursts of 256 words (of 32 bits), multiple bursts can be transmitted without handshake delays, once connection has been established. A simple look-ahead flow control allows sustained data transfer rates for large files not much below peak rates.

Presently, HIPPI is under development, or on market, as interface to the following systems: multiple workstations (DEC, IBM, Sun, Silicon Graphics), mainframes (IBM 3090, Cray), MasPar, iWarp, VME, Ultranet, and various peripherals, e.g. high-performance disk arrays. For MaxVideo, we are developing an interface as part of the EAST program, to allow demonstrations in real time. We also have under design, in EAST, a HIPPI interface at full bandwidth to the Sparc chip's S-bus, preparing tests for a general-purpose farm concept.

5. EAST status with respect to proposal milestones

In the following, the milestones are literally quoted from the addendum to our proposal (DRDC 91-13), and are compared to the present status of the individual activities.

5.1 *Trigger test system*

We intend to build a hardware box as test system for generating LHC-like data input for second-level trigger studies. The system hardware consists of a memory based on VME, and is connected to a workstation running programs for simulating partial detector data. Multiple events are entered into the memory. A trigger pulse generator running at typical 2nd-level speeds initiates data transfers into the triggering architecture under study. The memory is not primarily detector-dependent, different front-end or transmission protocols are implemented as individual daughter boards.

Milestones for the first year:
Design by middle of first year, prototype by end of first year.

Status end-February 1992:
Design completed, prototype successfully tested. This includes a daughter board implementing the HIPPI protocol .

5.2 *SPACAL test system based on MaxVideo*

We intend to build a full model implementation of a SPACAL second-level trigger based on the emulator (see above) and the MaxVideo image processing architecture. MaxVideo will run algorithms resulting in detailed cluster descriptions suitable for optimal electron/hadron discrimination or for jet identification. SPACAL test event samples in form of simulated calorimeter data for a fractional detector, are fed into the hardware emulator, the interfacing to MaxVideo is via several parallel MaxBus links. The algorithms can then execute under conditions approaching those of LHC both for data content and time sequence.

Milestones for the first year:
- provide representative data samples on a SPACAL subdetector by middle of first year;
- design the interface from the emulator to the MaxVideo20 system, and implement as an emulator daughter board, by end of first year;
- prepare the triggering algorithms and compare their performance with what the SPACAL collaboration achieves in off-line analysis by middle to end of first year ;
- study the implications of a distributed fully digital front-end approach which includes a 2nd-level pipeline, as proposed by the FERMI collaboration (DRDC/P19), in close collaboration with FERMI.

Status end-February 1992:
Data samples have been generated, and circulated for general benchmarking on different architectures. They have also been transcribed for use with the emulator. The interface from the HIPPI daughter board (part of the emulator) to MaxVideo is under test, an interface test system has been built up. Optimized algorithms are under study, and simulated data samples for second-level triggering are being prepared (using physics and detector assumptions as used in the proto-collaborations). Interaction with FERMI has resulted in joint simulation studies of both the FERMI chip and its embedding in a trigger/data acquisition system, and in the definition of a feature extraction and trigger structure to be implemented as prototype in 1992/93 (see section 6.5 below).

5.3 *Transition radiation detector trigger test using the ASP*

We intend to implement a model of a second-level trigger for the TRD, using ASP cards as provided by the MPP collaboration as part of the trigger. The trigger architecture, besides the ASP-s, will comprise a fast data selection phase (pre-ASP) and a final high-level decision device (post-ASP). Again, the emulator hardware will serve to inject data samples under conditions close to LHC into the architecture under test.

Milestones for the first year:
- provide representative data samples for TRD output, by middle of first year ;
- interface to the front-end electronics (emulator) via the HIPPI protocol, as model for high-bandwidth point-to-point connection, by end of first year ;
- study the architecture for implementing a flexible selection mechanism to reduce substantially the data rate transmitted to the ASP boards; detailed study completed by end of first year;
- high-level decision mechanism (post-ASP) using standard processors:
this amounts to an optimization between SIMD and MIMD architectures;
study based on simulated data by end of first year.

Status end-February 1992:
Data samples have been generated using the latest design of RD-6. The interface HIPPI/ASP has been delayed, because of lack of definition of the ASP side. As ASP hardware will not be available at the original time scale, alternatives to that SIMD system have been studied; parallelized algorithms have been developed using the MasPar machine (EAST has access to an installed machine with 4096 processors). A hardware test of the TRD trigger system, as part of the RD-6 prototype, is foreseen using the MaxVideo image processor, and the Enable machine (see section 6.4 below), as soon as available.

5.4 *Further studies*

- build-up of experience with potential muon second-level trigger architectures. Paper study finished end of first year
- preparation of muon data for the emulator - comparative study of neural network possibilities for the SPACAL case (using the same data as MaxVideo for benchmarking). Advanced paper study by end of first year.
- a similar study on SPACAL data, using the DataWave processor architecture (simulator). Study completed by end of first year.

Status end-February 1992:
Muon work on second-level trigger has been delayed, because more fundamental detector and trigger questions dominate the muon discussion. Neural net possibilities for simple local calorimeter triggers have been explored, first status reports on simplified problems have been given (e.g. at the 'Workshop on Neural Networks in HEP' at CERN on 11 Dec 1991). Algorithms of the full complexity required in second-level triggering still have to be demonstrated. The DataWave (DAVIS) simulator has been installed, and first results on algorithms of reduced complexity have been obtained.

6. Program of work and milestones in 1992/3

In the following, we outline activities to which we are driven naturally in pursuit of our original goals, and which either have already started or are foreseen to start in the course of 1992. Beyond completion of already initiated studies, and continued work on demonstrations as outlined above, we have set ourselves several natural new goals. We will explain these in the following paragraphs.

6.1 Completion of current benchmark exercise

The benchmark algorithms have been described above. The work to map them onto various architectures, and to study the possibilities and limitations of possible implementations, including cost, are presently under way. A workshop is planned (11/12 May 1992, at CERN) at which reasonably advanced results are expected. It is our intention to reduce the number of candidate architectures for subsequent further studying and for pilot implementations substantially. Lack of suitable interfacing possibilities, high cost, unsuitability for (some of) our algorithms, or delayed availability, all are possible causes for deciding to put some candidate architectures on hold for the future.

The present benchmark exercise is expected to terminate with at least one comparative publication. Some architectural studies will, however, continue, such as adaptation of algorithms to newly emerging candidates on the commercial market. Also, the neural network-inspired algorithms for recognizing energy depositions in calorimeters, will be pushed further, based on the present and improved benchmark data to be generated (see next section), and on Monte-Carlo and real data from the ZEUS calorimeter. We further plan system design studies of a second-level trigger implementation using these NN algorithms, with an array of DSPs (the Texas Instruments TMS320C40).

A different implementation of the neural network approach for electron/jet discrimination is already under test, using the L_Neuro chip, designed specially to be a co-processor of a T800 transputer. Also this activity will continue beyond the present round of benchmarking. This chip embeds 16 elementary processors performing in parallel 16 multiplications of two 8-bit integer numbers, and a final summation, in a time less than 0.6 μs. L_Neuro 1.0 also contains a RAM (random-access memory) for storing up to 1024 signed synaptic weights of 8 or 16 bits. A future generation of L_Neuro in 0.8 μm CMOS technology is in preparation (present feature size is 1.6 μm). For our type of NN application, LEP (Laboratoire Electronique Philips, Paris) built special boards using four L_Neuro chips connected to a transputer T800. We have assembled from four of these boards (plugged into a PC), a parallel machine with 16 L_Neuro chips. Our aim is to evaluate and validate the concept of coprocessing tasks executed by neural chips coupled with a local DSP-like processor, an architecture much like those used in other feature extraction tasks described above.

Further studies of neural network-inspired algorithms are foreseen for reconstructing tracks and for determining the bunch crossing times associated with these tracks, for data from a muon detector constructed from honeycomb strip chambers.

Milestones 1992/93: *Current benchmark exercise concluded 3Q92 by publication(s). Subsequent focus on a substantially reduced number of architectural candidates. Continuation of neural network studies in architectures and implementation possibilities, throughout 1992, with the goal of defining a demonstration with SLATE and/or detector prototypes.*

6.2 Improved benchmark data

The ongoing benchmarks were initiated at a time when Monte Carlo activities in LHC proto-collaborations had not started, and detector parameters had to be 'educated guesses'. EAST had made no attempt to bridge this gap, and our benchmark data, while requiring relevant feature extraction algorithms, do not reflect correctly the situation of a

second-level trigger, which is dominated by fluctuations in physics background. This situation is constantly converging towards a more coherent approach, and we have started recently to participate in the Monte Carlo calculations of physics background events (QCD jets) that pass a 'reasonable' first-level trigger criterion in the calorimeter. This will give us both a chance to repeat our benchmarking exercise with more realistic data sets, and to verify that the algorithms do indeed perform a valuable function in future experiments. We consider this a vital activity for the overall credibility of second-level triggering; it is not likely nor intended, however, that this new benchmarks should change our conclusions about architectures.

Absolutely essential for neural network learning algorithms, realistic calorimeter data will also be used for optimizing cluster definition algorithms (as in 4.1 above) in 'off-line' studies. On the TRD, we hope to be able to show, with improved input data, that algorithms as implemented on the remaining candidate architectures can provide reasonably precise track parameters, beyond establishing the existence of electrons.

Milestones 1992/93: *New benchmark data, most importantly QCD jets filtered by a first-level electron trigger, available for calorimeter and TRD algorithm studies by mid-1992.*

6.3 Hardware implementations and demonstrations

A small number of promising architectures will be implemented in hardware, and demonstrations of the trigger algorithms will become possible either with SLATE (the hardware emulator) or with detectors in beam test setups.

SLATE has been discussed in sections 3.2 and 4.2 above. Its HIPPI format output stream is presently being interfaced to the data memories contained in the MaxVideo image processing system. More precisely, a HIPPI destination module and supporting software are under test, that interface SLATE (i.e. HIPPI) to the MaxScan digital pixel port of the MaxVideo system. The VME module (called HiMax) provides 20Mbytes/sec input rate, which could be doubled by adding another MaxScan unit. Using HiMax, the algorithms already running in MaxVideo can be made to run at LHC speed.

As for tests with detector modules, we are severely limited in our choice by the availability of sizeable systems with LHC-like readout electronics. We are participating actively in defining a trigger role in the TRD prototype (RD-6), foreseen for 1993. For this purpose, we now have a Router (data selection and reordering unit) under design, which ensures that full data on HIPPI lines will pass to the data acquisition system, whereas selected data, in the arrangement expected by MaxVideo/Enable, will be branched off to these systems. The remainder of the test, feature extraction device and decision unit, will be identical to those used in the SLATE test.

For both SPACAL and the accordion calorimeter (RD-1/RD-3), the planned testing make second-level trigger additions not a meaningful exercise: the speed limitations in the ADC-s would teach us little more than an off-line test using recorded data (such tests have been done already several times). A second-level trigger device integrated into a calorimeter thus has to wait for the availability of a suitable detector.

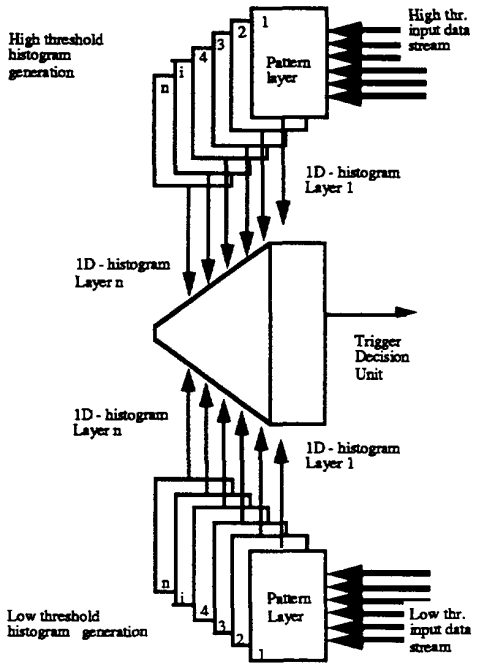
Milestones 1992/93: *A demonstration system integrating SLATE and MaxVideo via the 'HIMAX' interface board, will be under test before the middle of 1992. Once operating, both benchmark algorithms can be demonstrated on MaxVideo. Tests with the TRD prototype will be possible on the timescale of RD-6 (1993). The existence of a Router unit, i.e. a hardware unit arranging raw TRD data in suitable iconic form, is probable for that test, but can not yet be guaranteed (as fallback can be mimicked by software).*

6.4 The Enable Machine

The Enable machine is a flexible special-purpose processor implementation tailored especially to the needs of the given detector/algorithm, in our case the TRD feature extraction algorithm. An implementation in field-programmable gate array (FPGA)

devices has been chosen because it provides flexibility inside a custom-designed solution, and reduces substantially development time. These properties have been borne out in our past experience, in other high energy physics experiments. FPGA devices are becoming increasingly popular, and are available at modest cost.

The Enable machine is a SIMD machine consisting of an array of identical processing elements realized with field programmable gate arrays. In order to identify electron and/or pion tracks in binary images of limited size (regions of interest), the Enable machine is composed of two main functional building blocks: a histogram generation unit and a trigger decision unit (see figure).



The high and low pulse height images are sent through two separate but identical histogram generation units. The subsequent trigger decision unit analyzes the content of the histogram channels and, by comparing the bin content from the two images, achieves 'electron'/'pion'/'no high p_t track' classification in hardware.

The Enable machine will be equipped with a HIPPI serial-link interface to maintain the dataflow bandwidth from the TRD frontend and readout electronics, arranged into an image representing the natural straw coordinates z and ϕ (azimuth) by the Router unit. Note that curved tracks appear as straight lines in this projection. The pixel images are shifted one column at a time into the mesh-connected array of processing elements of two separate histogramming units. All predefined search patterns are scanned for in parallel, and the matching hits of each pattern are histogrammed. Each histogram channel corresponds to a search road (or pattern) of a given slope and ϕ value (note that the slope $d\phi/dz$ defines the curvature of the track, or p_t. For every ϕ value there is a fixed number of search patterns, with maximal positive and negative slopes. The total number of search roads amounts to n*k (n is the number of ϕ values, and k the number of different slope patterns. For our benchmark data, n=20 and k=17).

After histogramming the corresponding histogram channels from high/low threshold images are compared, in a trigger decision unit, to derive the e/π classification result. The trigger decision unit has a maximum-finder subunit, which searches the global low+high threshold hit count maximum. The ratio of (number hits low thr.)/(number of hits high thr.) for this maximum computed by table-look-up serves as a probability measure for electron identification/pion rejection. The table look-up value is compared to the threshold, and a trigger signal enabled, if the threshold is exceeded.

Milestones 1992/93: *A prototype of the Enable machine will be available before the end of 1992, for demonstration with SLATE. An installation in the 1993 test runs of the TRD is foreseen, with Enable being a feature extraction alternative to the MaxVideo system.*

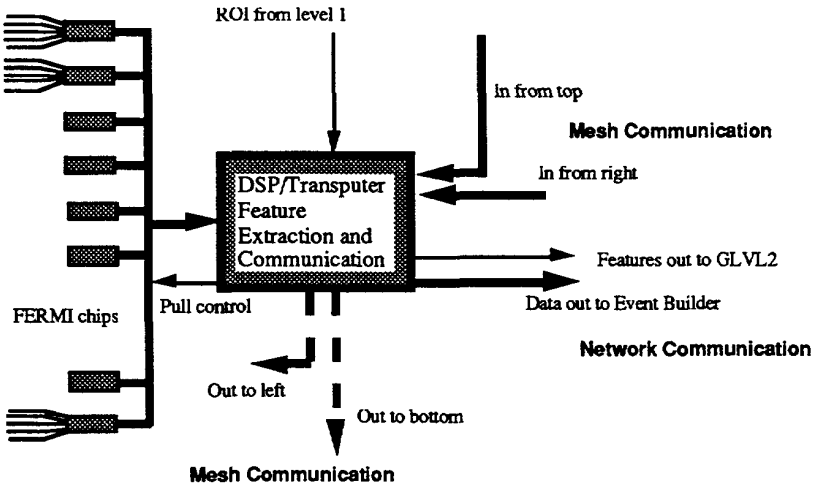
6.5 FEAST: Feature extraction in the FERMI calorimeter readout hypothesis

The FERMI project (RD-16) aims to build fully digital front end electronics for calorimetric detectors. The intended chip will include high speed analog-to-digital converters, a fully programmable pipeline/digital filter chain, and local buffering over the period of both first and second-level triggering, using an intelligent buffer manager.

In the choice of the second-level trigger processor, the EAST project has so far concentrated on processor systems which work serially on events, under the hypothesis that the frontend electronics 'pushes' the data into the acquisition system after first level selection. The possibility of using a farm of microcomputers to process multiple events in parallel is now being studied.

The aims of this FERMI/EAST (FEAST) sub-project is to link the FERMI chips to the farm of processors being evaluated within EAST. Appelquist and Iwanski (see list of EAST notes at the end of this report) have simulated a board containing the FERMI chips, and Kapusta has simulated methods of routing the data from the FERMI chips to the farm of second-level trigger processors. Hansen (in an informal EAGLE note) has also proposed a data acquisition system based entirely on transputers.

We are now proceeding to design a more complete trigger system around the FERMI chip. The present, preliminary ideas are the following: A FERMI board contains 36 FERMI chips each of which is connected to 9 cells. Optical links provide high-bandwidth communication from the FERMI board to a FEAST board which contains the Local Level 2 Trigger Processor (LLVL2) and has local links to LLVL2 processors on adjacent FEAST boards. The boards also communicate via a switching network with the Global Level 2 Trigger Processor (GLVL2) and, for second-level retained events, with the event builder of the data acquisition system.



The operation of the system is as follows. The region of interest (RoI) is identified by the Global Level 1 Trigger Processor (GLVL1) which broadcasts it to the FEAST boards via the Local Level 2 links. The cells identified by the RoI is assumed distributed over four neighboring boards; one FEAST board (lower left) is defined by the RoI broadcast as being the "master", the three neighboring boards (right and top in the mesh) as "slaves". The slaves collect data from their FERMI chips and transmit it to the master.

When the master has received all the data in the RoI, it performs the local functions (e.g. feature extraction), and transmits results to the GLVL2. The output from the Feast board must then be routed through a network of switches to arrive at one of the processors in the GLVL2 farm.

To implement the FEAST part of the system using available technology, it is proposed to use a combination of Digital Signal Processors and transputers. The DSP-s are seen as the Local Level 2 Trigger Processors and provide capacity for local feature extraction and LLVL2 links. The transputers serve as communication processors for the connection to the routers, which forward the RoI data to one of the processors in the farm.

Milestones 1992/93: *The following workplan is a joint plan for the EAST and FERMI collaborations, for 1992: The FERMI Board will be further simulated. Tests for the high-bandwidth links from FERMI to FEAST must be done, defining first the protocol which would run over the links, and designing and building prototype optical links running at 1 GHz. The ideas how DSP communication and processing possibilities will be optimally used, must be clarified by hardware evaluation and simulation. VMEbus-based DSP cards and a development system for the DSP-s must be purchased, the communication link speed and the latency must be understood. A study of the DSP/ transputer interface must be done, the already initiated simulations of the switching network have to be finalized.*

In 1993, if available, an early sample of transputers and switches will be purchased, to gain hands-on experience. Detailed system requirements including monitoring and controls will be defined.

Longer-range planning: *If after first tests of all components the basic ideas have not been basically modified, several prototype FEAST boards will be designed and built, each board containing an optical link interface and a DSP/transputer combination. Software will have to be written to allow communication between FEAST boards, data transfer from/control dialogue with the FERMI board, and the output of data to the transputer network. A test environment with multiple FEAST boards, a transputer network, and some hardware emulation of FERMI chips, will have to be put in place.*

6.6 High-level decision units behind feature extraction devices.

The MaxVideo system is a typical low-level feature extraction device. It is capable, for instance, of convolving an image with a 16*16 convolution mask in 100 ns per output pixel (corresponding to a processing speed of 5 Gops) and (in parallel) with a 2*256*64 mask (corresponding to 640 Gops). The achievable speed per pixel increases continuously; there exist today designs for image processing chips at 250 MHz (DAVIS, in 0.4 μm technology).

SIMD arrays like MasPar or ASP, or special-purpose processors as the Enable machine, have similar characteristics: They work in fixed point arithmetic and with limited precision, and their programmability is restricted by their architecture, at least if high throughput is required.

For one- or two-dimensional feature extraction, such modules may turn out to be just what we need. For higher level decisions, based on the output of several of the low-level systems, more flexibility is required. Three-dimensional *geometrical correlation* of physics features, from different subdetectors, or for kinematical test quantities like effective masses, *random access* to feature data and higher precision, possibly floating point, are needed. Processors with full high-level programmability and floating point hardware become necessary. In the following, we describe some realizations that will allow us to implement a full second-level trigger system, including multiple windows in different subdetectors.

a) DECperle

The Enable machine is described above as an implementation in field-programmable gate arrays (FPGA-s). As such, it is on our schedule for prototype delivery later this year, so that a demonstration with the RD-6 prototype Transition Radiation Detector becomes possible in 1993. Digital has announced, in 1991, a Research Initiative around a generalized concept based on FPGA-s (Xilinx), here called Programmable Active Memories (PAM-s). In the 'DECperle' system, multiple FPGA-s are surrounded by additional memories, switching elements, and high-bandwidth I/O ports that can be interfaced to HIPPI. Also, significant software tools are foreseen to speed up development at the gate level, and to make system changes more transparent. Beyond that, DECperle boards are interfaced to a standard DECstation 5000. Anticipating an implementation of the Enable machine using DECperle, we can envisage the PAM-s of DECperle to run some of the Router and all feature extraction functions on the PAM-s, and the DECstation to take the role of the high-level decision unit. The fact that the PAM-s possess ample and fast memory coupled to active switching elements, qualifies them for the Router task; the fact that they are closely linked, in hardware and software, to a high-level processing unit, makes the solution particularly attractive as a rather homogeneous trigger system.

Milestones 1992/93: *A DECperle system will be installed and evaluated as soon as available (present statement for availability is 2Q92). Only then can a decision be made which functions and algorithms should be implemented first on this system. Candidates are the routing and feature extraction functions of the TRD trigger, but also cluster definition algorithms, e.g of the neural network type, for calorimeter data.*

b)Max860 board

Among the suitable processors for collecting features from lower-level devices, in particular from MaxVideo, we have chosen the RISC processor i860 (Intel). Packed into what is called a SuperCard (by CSPI), and in the future interfaced to multiple MaxBus links (by the MaxVideo manufacturer, Datacube), this 'Max860' device allows ideally to act as 'physics concentrator' for multiple feature extraction devices, giving physicists control in a high-level program over the decision process. This new system combines now the advantages of a pipelined image processing system with the capabilities of modern RISC processors. An i860 is presently under evaluation as part of the MaxVideo activity at CERN. We should note that the i860 is also a target processor in high-performance heterogeneous systems as designed by GP-MIMD and other manufacturers of high-performance computing systems.

Milestones 1992/93: *Installation and first demonstration of the Max860 board as soon as available (target date April 1992). If available, demonstration as high-level decision unit with the MaxVideo system fed by SLATE in 3Q92, and use in the TRD prototype test (1993).*

c) transputers

The capabilities of transputers as communication controllers and, possibly, as general processors for feature extraction and/or global decision making, are under study as part of the FEAST project. We participate in simulation and studies, and will learn about the next generation of transputer hardware through this participation and from the GP-MIMD project. We do not consider a separate transputer activity as part of EAST.

d) general model

We believe it necessary to understand fully the practical limitations of a feature extraction/global decision system based on general-purpose processors alone. It can be argued that despite long latency, such a system is ideally scalable and flexible, probably

also economically attractive. Manufacturers have on market or in preparation a multitude of processor components that point in this direction. Note that inter-processor communication is not an expected problem in such a system, as each processor is fed with data for one region of interest, and subsequently needs no high bandwidth communication to the global decision unit.

In preparation of a possible model implementation, we have under development a full-bandwidth HIPPI interface to the internal bus of the Sparc chip. This will allow us later to demonstrate the limits in a system based on this processor. Definite plans to design a multi-processor system do not exist at present, although preliminary studies of the RISC-farm solution are under way.

Milestones 1992/93: *Working interface HIPPI/S-bus in 3Q92. A definite plan for a small multi-Sparc system before the end of 1992.*

6.7 Fibre optics high-bandwidth connections

The HIPPI standard is a useful interim standard for low-level connectivity, and seems well accepted by industry. The standard has, however, its limitations. For small test setups, the use of 10 HIPPI cables is reasonable. For the final application in LHC, hundreds of HIPPI cables would be required. This number of cables would take up a very large volume. This and the present maximum length of 25m could be a problem in a future experiment.

The general developments around HIPPI, currently in progress, will be terminated by the end of 1992. We plan to study two new standards which may be the future accepted way for HIPPI to run over optical fibres. The ANSI committee which specified HIPPI has defined two possibilities: Serial HIPPI and FIBERchannel.

Serial HIPPI will not be a formal standard but is an "implementers agreement" which defines the components used at the physical layer. Serial HIPPI runs over two optical fibres and "extends" HIPPI links from a maximum of 25m to 2000m. FIBERchannel, which will be an ANSI standard, would also allow HIPPI to be extended. In addition, FIBERchannel defines a switching fabric which could be very useful in data acquisition systems. First commercial products are on the market for both standards.

Milestones 1992/93: *Test of Serial HIPPI and FIBERchannel chipsets in 1992. Purchase of commercially available components end 1992/1993.*

Only later: Development of components not available from industry and use in test systems. The development work in 1993 will include the building of a FIBERchannel or Serial HIPPI daughter board for the SLATE emulator.

6.8 Formal system modelling using VDM++

Together with other partners from outside the High-energy Physics community, the Utrecht and CERN partners of the EAST collaboration have been accepted as 'application partners' in an Esprit III project AFRODITE. The objective of this project is the practical application of tools for formal specification of mixed hardware/software systems. The tool to be used is an object-oriented specification language called VDM++, based on the standard VDM. The objective in EAST is the modelling of a partial acquisition system, with at least the second-level trigger steps included, under the constraints of fault-tolerance, maintainability, and reusability that characterize such large systems. It is in this activity of modelling that the specification possibilities offered by VDM++ are expected to be of major impact, allowing easy adaptation over the evolution of both technologies on the market and particle detector design. We expect that an acquisition system specification shown to work in an abstract model, subsequently

successfully implemented in software and hardware prototypes, would be naturally accepted in the eventual large-scale implementations of the experiments. We expect thus to effect a transfer of a new technology from professionals to the high-energy physics community.

Milestones 1992/93: *This activity is in a very early state of definition, and no definite milestones can be set out before a detailed discussion with the Esprit authorities and the other project partners.*

7. Composition of the collaboration, institute responsibilities, and resources

Since proposal time, the collaboration has had several changes in its **composition**. An active collaboration with CGS Amsterdam and with the University of Rome, in the direction of our activities, has not materialized, and these institutes are no longer members of EAST. Newly joined have Dubna, Mannheim, Prague, and Rehovot (see cover page).

The individual **responsibilities** of institutes are as follows:

NIKHEF Amsterdam: Neural Network algorithms and their implementation (with Utrecht University);
IFA and Polytechnic Bucharest: Participation in MaxVideo and SLATE, benchmark data for calorimetry and MasPar implementation of algorithm;
CRIP Budapest: SLATE software, benchmarking on the ASP simulator;
CERN: MaxVideo, DAVIS simulations, hardware demonstrations, high-bandwidth communications, SLATE daughter board;
INP Cracow: Benchmark data for TRD and implementation on MasPar, FERMI and FEAST simulation (VHDL) and participation in FEAST board design;
JINR Dubna: implementation of the Router unit for the TRD beam tests (with Universität Jena);
RHBC London: SLATE design and motherboard, FEAST board design;
Universität Mannheim: Design and prototype construction of the Enable machine;
Ecole Polytechnique Palaiseau: Monte Carlo for calorimetry benchmark data, neural network implementations;
Academy of Sciences, Prague: High-level decision processors;
Weizmann Institute, Rehovot: Interface hardware/software connecting HIPPI and MaxVideo;
Utrecht University: Neural network algorithms (with NIKHEF), formal system specification with VDM++;
Institute of High Energy Physics, Zeuthen: Interface board connecting HIPPI and the Sparc chip (S-bus), for a possible trigger farm based on Sparc processors.

Resources: The 1991 estimate for material resources had been 600KSF total, of which 200 KSF from CERN. That sum has been allocated and spent. For 1992, with a much increased activity leaning more strongly towards hardware demonstrations, we estimate increased spending, at the level of 850 KSF total, of which 350 KSF from CERN. Manpower involvement has increased substantially over the 30 full-time equivalent (FTE) of the first year, and we now estimate over 40 FTE involved in EAST.

We will continue to perform tests with beam only in conjunction with other R&D activities (e.g. RD-6), and hence do not require any beam time allocation.

We have followed, in the past, the same idea of latching onto other ongoing activities for computing purposes, but would find it practical and a speedup to have our own computing budget, for Monte Carlo calculations, at the level of 100 hours CERN.

8. Acknowledgements

We gratefully acknowledge the participation of individuals and institutes that are not formally part of the EAST collaboration. V.Dörsing, A.Reinsch, H.U.Zühlke from the Institut für Angewandte und Technische Informatik, Universität Jena, Artur Thielmann, Jacek Kosiec, from GEVA Ltd., Warszawa, and G.Klyuchnikov from IHEP Protvino, are participating in different parts of our work. We also gratefully mention the excellent collaboration from other R&D projects and from the activities in the LHC (proto-) collaborations. We further want to express our thanks to the following sponsors, without whom we could not have reached the present results, nor reached the breadth of our ongoing and future activities: Digital, Eloisatron project/LAA, Esprit III, GSI Darmstadt, MasPar Distributor Switzerland, Laboratoire d'Electronique Philips.

9. List of internal notes generated in connection with EAST activities

Benchmarking Architectures with Spacal Data,
J.Badier, R.K.Bock, C.Charlot, I.C.Legrand, 25 November 91, EAST note 91-10

Benchmarking with Data from the Transition Radiation Detector,
P.Bialas, J.Chwastowski, P.Malecki, A.Sobala, 2 December 91, EAST note 91-11

The HIPPI Source Daughter Board of the Detector Emulator, T.Anguelov,
30 Sept 91, EAST note 91-12

Second-level Test System Status, B.Green, N.Mihai, J.Strong, February 1992,
EAST note 91-13

Status of the SLATE Software, Ervin Denes, 15 September 91, EAST note 91-14

Architecture and Algorithm Workshop (Copies of transparencies)
14/15 Oct 1991, EAST note 91-16

Level_2 Trigger System in LHC Calorimeter, W.Iwanski, P.Kapusta,
4 Oct 1991, EAST note 91-17

Simulations of a FERMI Board using VHDL, G.Appelquist, W.Iwanski,
November 1991, EAST note 91-18 (also FERMI note #6)

The Enable Machine, F.Klefenz, R.Männer, K.-H.Noffz, R.Zoz,
1 February 1992, EAST note 92-01