

SEP  
CERN-LHCC  
95-45

CERN LIBRARIES, GENEVA



B00005935

CERN/LHCC 95-45  
LCRB Status Report/RD11  
15 August 1995  
(EAST note 95-09)

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

J.Vermeulen (*NIKHEF-H Amsterdam*)  
F.Constantin, A.Gheorghe  
(*Institute of Atomic Physics and Polytecnic Institute, Bucharest*)  
E.Denes (*Central Research Institute for Physics, Budapest*)  
R.K.Bock (\*), J.Carter, F.Chantemargue, R.Hauser, W.Krischer, L.Lundheim,  
R.McLaren, S.Muñoz-Frutos (*CERN, Geneva*)  
J.Renner-Hansen (*NBI Copenhagen*)  
D.Botterill, R.Hatley, J.Leake, R.Middleton, I.Newman, F.J.Wickens  
(*Rutherford Appleton Laboratory, Didcot*)  
D.Belosludtsev, S.V.Khabarov (*LHSE, JINR, Dubna*)  
V.Dörsing, W.Erhard, P.Kammel, A.Reinsch  
(*Institut für Informatik, Universität Jena*)  
Z.Hajduk, W.Iwanski, K.Korcyl, P.Malecki, Z.Natkaniec  
(*Institute of Nuclear Physics, Krakow*)  
B.J.Green, J.Strong (*Royal Holloway and Bedford New College, London*)  
P.Clarke, R.Cranfield, G.Crone (*University College, London*)  
R.Hughes-Jones, D.Mercer (*Manchester University*)  
H.Högl, A.Kugel, R. Männer, K.H. Noffz, S.Rühl, R. Zoz  
(*Department of Computer Science V, University of Mannheim*)  
J.Jirina, M.Kucera (*Inst. of Computer and Information Science, Prague*)  
L.Levinson (*Weizmann Inst. of Science, Rehovot*)  
L.Caloba, M.Seixas (*Federal University, Rio de Janeiro*)  
C.Balke, J.Haveman, W.Lourens (*Utrecht University*)  
U.Gensch, H.Leich, U.Schwendicke, P.Wegner  
(*Institute of High Energy Physics, Zeuthen*)

(\* ) Spokesperson

## *Summary*

Over the period of reporting, EAST has focussed all activities towards the proposed proton-proton experiments at LHC. Collaboration with ATLAS is close in all areas; a majority of members of EAST participates in ATLAS working groups; they take responsibilities in ATLAS, and have contributed substantially in writing the relevant parts of the ATLAS Technical Proposal. EAST also contributes to some aspects of trigger implementations of CMS. Contact (and personnel overlap) exist also with the Hera-B project, which has similar timing constraints as LHC experiments, and where some of EAST's ideas may be implemented on a shorter time scale, albeit with quite different detectors and hence algorithms.

The milestones set for EAST at the time of the last Status Report (CERN/DRDC 94-20) have been met:

- the data-driven elements of a trigger (Router, Enable and DecPerle) have successfully been demonstrated in beam tests (June 1994) jointly with RD6/RD13; proposals for a generalisation of these components have been made to ATLAS, and implementations have started;
- an improved version of the SLATE emulator has been designed and three identical units have been built and tested; these and previously built modules will be heavily used for laboratory testing at high frequency, and are already booked out for preparing beam tests;
- prototype hardware for a DSP-controlled L2 buffer (based on TMS320C40 digital signal processors from Texas Instruments) and a small switching network have been tested in the H8 beam in September 1994; algorithms have been benchmarked; new DSP choices are under study and may be implemented in the coming year;
- DecPerle has now been reprogrammed for all algorithms of feature extraction and global decision making, running faster than 100 kHz and thus demonstrating its suitability of principle for an implementation that is data-driven throughout (ATLAS fallback solution);
- the problem definition for all data manipulations from front-end electronics to feature extraction (based on ATLAS detectors) has made substantial progress, but is incomplete due to the convergence process in the experiment;
- an overall model for a farm-based solution of level-2 triggering has been derived based on the Modsim II language, a more detailed model targeting the C40/C104 solution has been written;
- SCI interfaces from the C40-s (L2 buffer, feature extractors) and to Alpha processors have been built, and are now being demonstrated executing the global part of the trigger;
- a benchmark suite has been used for the mapping of level-2 triggering onto an MPP machine (the CS-2); this work has been completed, successes and limitations are documented; communication parameters using various commercial processors, interfaces, and an FCS (Fibre Channel) switch were measured and documented;
- the AFRODITE project has followed its time scale, a level-2 system has been expressed in the VDM++ language, for which an evaluation paper is available; a part of the data-driven trigger implementation (data converter) has been translated into VHDL and compared to an engineer's version of the same unit.

## 1. *Second-level triggering, short overview*

The main issues dealt with in EAST have been described in detail under different viewing angles in earlier status reports to the DRDC (DRDC/94-20, DRDC/93-12, 4-year Status Report DRDC/94-48). Summarily, the work of the collaboration has resulted in a widely accepted problem decomposition and a set of remaining choices of implementation. The view on the issues is shared by the groups in both ATLAS and CMS, although the confidence in technological developments over the coming years do differ; the decisions will ultimately have to be made in function of the offerings of the industry of computers and communications industry.

A key assumption is that second-level algorithms have to be implemented for a level-1 output frequency of around 100 kHz, using *full-granularity, full-precision* data for *some detectors*. Detailed physics and detector simulation, as yet incomplete in both proton-proton experiments, will have to determine which data are effectively needed in level 2, they may well be dependent on the target physics. Obviously, reducing the requirements on the amount of data and their precision in the trigger will result in architectural simplification and thus cost savings, possibly substantial. Risk factors in marginally adequate implementations will have to be evaluated by the collaborations.

Another common assumption concerns the availability of all detector data, in digitized form, after a level-1 'accept' in buffer memories ('level-2 buffer'). Each buffer will contain data for a fraction of one subdetector, in a format largely dictated by the bandwidth constraints between the frontend electronics and the buffers. The granularity of these buffers, the detailed grouping of data, and their possible preprocessing upstream will be of paramount importance for the implementation of a second-level system; again, the working hypotheses in both experiments have not yet converged.

Two main 'pure' alternatives can be seen in implementing level-2 triggers:

- If the level-2 trigger is considered a general real-time *computing* problem, one can schematically connect all buffer memories via a single switching network to a farm of general-purpose processors, each of which thus has access to any part of the event information. This provides perfect freedom in implementing trigger algorithms of any desired complexity, but puts stringent constraints on the network, on the supervisor and scheduler, and on the available computing capacity. A distinction between level-2 and final computer filters ('level-3') is not necessary in this alternative. Diagram 1a below shows this system schematically, exemplified by the CMS Technical Proposal. Present commercial components offered in the communication and computing market make it an optimistic view that an idealised computer network of the simple form above, can be built and operated economically in a few years' time.

- If, on the other hand, the level-2 trigger is considered a *trigger* problem to be implemented via parallel components with specialised tasks, a natural decomposition is offered by the likely structure of algorithms and the substantial simplification of using, wherever possible, the information provided by the first-level trigger. Algorithms can be assumed to operate locally on subdetectors only, and the 'Region-of-Interest' (RoI) concept can be applied in order to restrict detailed investigations to areas indicated by the previous trigger stage. Schematically, this

option results in the functional diagram 1b below, which is at the base of the 'data-driven' implementation of the ATLAS Technical Proposal:

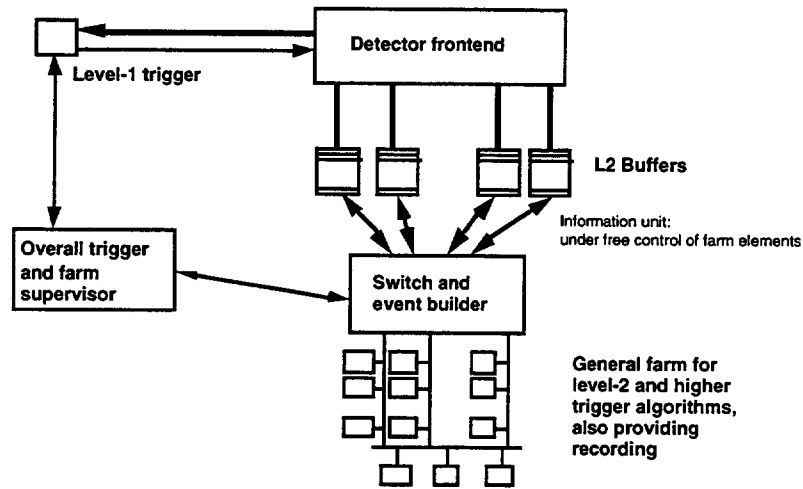


Figure 1a: Pure farm-based solution, with a level-2/level-3 farm

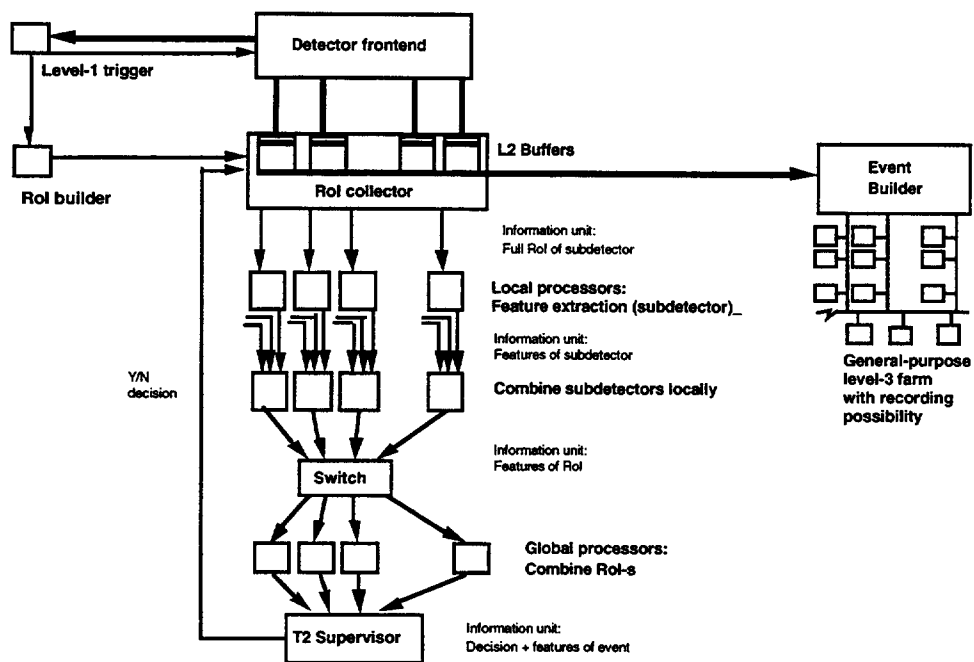


Figure 1b: Data-driven solution for level-2 feature extraction, with a small farm for global decision

The EAST collaboration working closely with ATLAS has defined this detailed decomposition of second-level triggering algorithms, in order to take advantage of the natural parallelisms in the problem, and to avoid implementation bottlenecks. Fig. 1b shows, in fact, a hybrid system maintaining a much reduced idea of 'farm' for the implementation of global decision making, where the number of input ports

and the transmission parameters for a switch are much less demanding than in the generalised full farm solution, where all buffers are connected to a (large) farm of processors. The technologies for transmission and processing still leave a wide variety of implementation choices; for some of these, EAST has been able to demonstrate pilot implementations.

## *2. Work in EAST 1994/95*

In the following we discuss in more detail how EAST activities have explored critical parts of the above architectural choices, based on technological components available today or in the foreseeable future. The unchanged intended end result of the EAST collaboration is a short catalogue of options, sufficiently understood to be assigned cost and risk factors. At the time when LHC experiments will make definitive choices, around 1997/98, this catalogue, the familiarity with products, and the contacts with industry, will be of paramount importance for making informed decisions. In parallel, the close involvement with experiments will ensure (and has already done so) that relevant early choices in detector and readout design are made with the possible constraints of level-2 triggering in mind, e.g. the grouping of data and their formats.

Many conclusions of the past work in EAST have already been documented in the ATLAS Technical Proposal; a trigger frequency of 100 kHz can certainly be maintained by data-driven pipelined systems, without introducing event parallelism in farms. The alternative of a possible implementation based on switches and computer farms has made progress, and first demonstrations based on digital signal processors (DSP-s) have been made. Farm-based systems are sufficiently attractive to explore their components as they become available, jointly with industry and other collaborations, and we believe this work has to continue, with increased urgency.

### *Beam test demonstrations of Enable and DecPerle*

Together with project RD6, EAST has shown in beam tests (in H8) a demonstration of triggering at LHC speed using the previously developed TRT-specific RoI collector (Router) and two alternative data-driven implementations for feature extraction, Enable and DecPerle. Both systems are Xilinx-based. DecPerle is a general-purpose board, marketed during 1992/93 in a small series for computer architectural prototyping and data pre-processing. A hardware interface from the detector (Router) to DecPerle was provided by a commercial HIPPI/Turbochannel board. Enable is a TRT-specific parallel architecture for fast histogramming (trackfinding). In contrast to Decperle, Enable is not only a small-scale prototype, but meets the entire known requirements of the TRT in I/O and processing.

Both devices were shown in the laboratory to run faster than 100 kHz for the TRT's trackfinding task, and were limited by transmission rather than processing. Both devices were also shown to cope with the conditions at the TRT prototype in the H8 beam (see fig.2 below). Their results were recorded and shown to be bit-by-bit compatible with off-line programs.

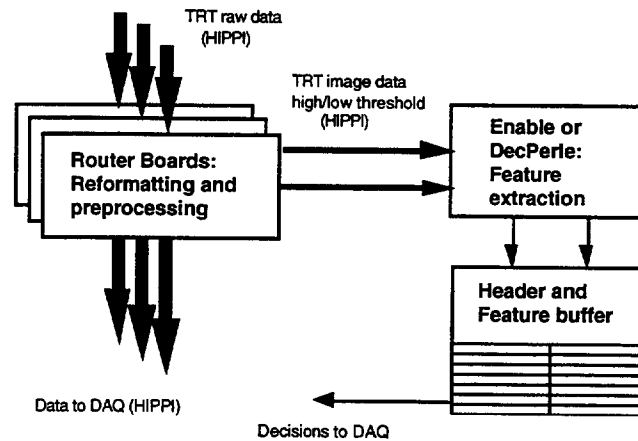


Fig.2: Schematic setup for data-driven beam tests in 1994

### *Field-programmable gate arrays as a general device*

Beyond the TRT application, DecPerle was also demonstrated in the laboratory to provide data-driven execution of algorithms for quite different detectors, by reprogramming only. An implementation of a calorimeter trigger algorithm was already shown in 1993 to run at frequencies faster than 100 kHz on this device. A trigger algorithm for the thresholded (zero-suppressed) data of the Silicon tracker of ATLAS has also been designed and tested, and again processing speed is determined by the speed of transmission; for thresholded data the event frequency depends on the data volume, and is typically much faster than 100 kHz. Successful implementations were also made for algorithms representing the combination of features for different subdetectors, and for the final decision-finding. These global algorithms are of limited complexity (although they do include effective mass calculations), and their input consists of small data sets; they use only a small fraction of DecPerle's potential and execute in the microsecond range.

It is an important result that significantly different algorithms, in fact all presently conceivable algorithm parts needed for a decomposed second-level trigger, can be implemented on a single programmable hardware platform by only modifying 'software', and that such a device performs at the required maximum trigger speed. This underlines earlier claims that a data-driven solution can be regarded as an established fallback for implementations of triggers at 100 kHz.

### *Level-2 buffers*

A pilot implementation of a level-2 buffer with associated DSP (digital signal processor TMS320C40 from Texas Instruments) has been tested in the H8 beam together with the TRT prototype in September 1994. It used boards from LSI (Loughborough, UK), containing multiple TIM modules (replaceable daughter boards), of which some had been individually modified for our purpose. The C40 showed to be an acceptable buffer manager at the expected rates, up to a probability of 50% for a single buffer memory to be part of an 'active' region of interest.

The buffer memory has now been modified to use embedded C40 processors, providing a solution of the same performance at lower cost. A small series of these buffers is planned to be tested in the joint ATLAS tests at the H8 beam in September 1995, using ad hoc interfaces to the various detector prototypes that will be running during this period.

A first design for a level-2 buffer based on the ideas of the original Router unit for the TRT is now also available. In this design, a 'buffer' is assumed to contain a large fraction of or an entire subdetector, and substantial intelligence is added in order to extract a full region of interest and preprocess the data. A high-bandwidth local bus spans multiple individual (partial) buffer modules with many input fibres, and several parallel units based on FPGA-s allow the local transformations that are necessary between the data in the buffer (that go to level 3 and to data acquisition) and the data that are needed for a streamlining of the level-2 algorithm. Typical operations of preprocessing are the reformatting of thresholded data, or the signal extraction from multiple bunch crossings which are deemed necessary in recorded data, but a harmful complication for level-2 algorithms. It is intended to demonstrate a prototype buffer memory of this type in 1996.

### *Alpha/SCI*

The implementation of a partial global level-2 system based on DEC Alpha processors connected via SCI has now been finished. Beyond the demonstration of principle of September 1994, more work can now be foreseen in beam and laboratory tests, in conjunction with the C40 DSP tests.

The existing SCI interface for the Alpha workstation is based on a design by RD24, with the original GaAs chip replaced by the CMOS version, and the host DECstation replaced by an Alpha. The SCI connection to the Alpha is through the PCI bus; the interface has been implemented in a modular way. OSF/1 is used as operating system on the Alpha workstation and the VxWorks real-time kernel on the Alpha SBC-s.

### *Commercial switches and processors*

In the desire to stay close to manufacturers' development in the area of 'off-the-shelf' components for a data collection system into a farm, performance parameters for commercial fibre channel equipment (switch, interfaces, related software) and connected processors from Hewlett-Packard and IBM were measured. The results are not encouraging for level-2 trigger implementations: any single communication requires an overhead delay of 300 to 500  $\mu$ sec, determined mostly by low-level (commercial) software. This makes them acceptable candidates for event building, but seems to imply an unacceptable base for the multiple short packets transmitted at high frequency for level-2 trigger tasks. We did not think that an investment in the main limiting factor, software which is highly dependent on the characteristics of hardware and operating system, would be a defensible investment; also, our industry partners were not forthcoming in this area.

## *Mapping trigger structures onto HPCN systems*

The standard 'benchmark suite' for the L2 trigger has been ported in parts onto high-performance (parallel) computers and network (HPCN) systems, starting with the CS-2 at CERN, IBM's SP-2 system, and a Convex Exemplar SPP1000. Presently, the high cost of HPCN systems may make this activity look somewhat academic; however, the market situation is predicted to evolve, and these systems may end up as the only truly commercial solution for level-2 triggers: with an HPCN system, a manufacturer could indeed supply large system parts that leave only an interfacing task for the experiments. They are furthermore fully scalable, will follow (the vendor's) technology evolution, and can be monitored with ease, this being part of their normal application.

HPCN systems have versatile high-bandwidth switches and general-purpose processors much like any farm-based solution requires. In fact, HIPPI, ATM, Fibre Channel and SCI switches all are among candidates to be used in such systems, albeit not switches in the open form which might allow vendor interchanges. Just as important, HPCN systems may be the only ones to come with the software playing a critical role in achieving high communication speed, in particular low latencies.

The benchmarking results so far are encouraging in the sense that no other commercially available switch with processor interfaces has been demonstrated to achieve similar performance (bandwidth and communication latency). On the CS-2, processors can perform the global tasks at a 25 kHz rhythm, if these are executed in a fixed processor configuration. Dynamic processor assignment implies, on the CS-2, the use of different software, and results in substantial slowdown, like by a factor of three. Preliminary results on the SP-2 (as yet less fully understood in its software details) are also in this category of 10 kHz for the global task.

Feature extraction, instead, has been shown to be at the limit of parallelisation with already very few processors: single feature extraction tasks typically execute (fastest single processor) in 300 - 1000  $\mu$ sec (i.e. at 1.0 - 3.3 kHz). For TRT feature extraction, it has been shown that this time can be brought down by a factor of no better than three by using groups of several (up to eight) processors; the addition of more processors results in no further gains (even in losses), due to overheads.

## *Continued work on algorithms*

Algorithms for benchmarking have to be continually adapted to the evolving definitions of the detectors, to maintain their credibility with the physics community. EAST has defined feature extraction and global decision making algorithms, which are being used as 'benchmarking suite'. They are reasonably up to date for feature extraction in three characteristic detectors (calorimeter, Si tracker, barrel and endcap TRT) of ATLAS in the 'Cosener's House' definition of the detector, and for a simple model of global decision. Over the period of reporting, however, their value has decreased. They do not include pre-processing tasks, which are now recognised to be of comparable if not longer duration, if left to high-level processors.

In order to update the benchmark 'suite', EAST also participates actively in ATLAS groups, hoping to provide large samples of L1-filtered signal and background data, using the detector design of the Technical Proposal or later; based on such data of training samples, algorithm studies can proceed. In parallel, members of EAST have taken responsibility (in ATLAS) for defining relevant parameters beyond purely algorithmic definitions: raw data formats, transmission



characteristics and other detector-dependent parameters needed for modelling; triggering is not primarily a question of general computing, but one of data transport and communication. For meaningful comparisons, such a complete model has to be accepted as significant by the collaboration, and kept stable over a lengthy period of time - criteria not easy to satisfy, and not under our control, but of utmost importance for future work in the collaboration.

### *Formal system modelling using VDM++*

The Esprit project AFRODITE has provided a translation of a high-level specification of a feature extractor and a full L2 system in the formal specification languages VDM and VDM++. A part of the data-driven RoI collector (Router) was also described in VDM++, and submitted to a semi-automatic translation into VHDL. A specific implementation of VDM++ will soon be marketed by the partners of the Esprit project, embedded in familiar OMT tools.

Evaluation reports on the experience acquired in this work have been written. The short conclusion is that the use of formal methods in software design is a likely variant for future work, whereas hardware design may not be able to take much advantage of the formal abstraction. The specifics of the VDM++ language or of the implementation resulting from the AFRODITE collaboration may or may not be successful on the market for such tools.

This conclusion has been brought up for discussion in the MOOSE collaboration (RD41). The AFRODITE project has come to its scheduled termination in August 1995.

### *3. ATLAS/EAST activities in 1995/96*

As R&D projects move closer to the LHC collaborations, to the point of becoming indistinguishable from the collaborations' developments, it is relevant to separate collaboration-internal work from possible remaining projects with a strong generic component. We describe here summarily the activities which are being absorbed by ATLAS. Their future resources will have to be decided in the collaboration, and no milestones are presented here. The only remaining generic goal for RD-11/EAST consists of investigations on HPCN systems; this subject will be dealt with in more detail in paragraph 4 below, and is at the base of EAST's request for continued funding in 1996.

#### *General*

The two major architectural options still open, data-driven and farm-based, clearly need further work. Pipelined data-driven systems of the first generation have been demonstrated; successor solutions are under development and will be demonstrated in 1995/1996 (Enable++, intelligent data collector with pre-processing).

Farm-based solutions, on the other hand, are of equal interest to both proposed collaborations for proton-proton physics, and concern questions like commercial availability, standardized components, high-level programmability, scalability,

structural flexibility, homogeneity including level 3, reconfiguration in case of malfunction of components, etc. In ATLAS, it has been proposed to continue the investigations based on DSP-s and C104 switches, and of SCI switches and rings coupled to Alpha processors. The characteristics of farm-type solutions and their foreseeable technological evolution, are far from being understood in detail. Full-scale pure farm-based systems seem not implementable in practice with truly commercial components, even a few years from now. It seems relevant, therefore, to continue and even expand the work that has been started in EAST towards this end, in close collaboration with other R&D projects (RD31, RD24 in particular) and the two collaborations. Substantially more exploratory work in the relevant R&D projects and in the collaborations (and a better understanding of the general market evolution) will be required before defining farm-based architectures and possibly discarding non-farming options.

The joint tests with ATLAS detector prototypes in the H8 beam will continue, to the extent that common runs of multiple detectors are scheduled. A run in September 95 will comprise most elements of two different farms: a supervisor, switches, and processors. There will be a set of DSP-controlled L2 buffers, connected via DS links and a C104 switch to a feature extraction farm (also implemented as DSP-s). Data will be transmitted from the feature extractors into a global decision farm of Alpha processors connected by an SCI ring, which is interfaced via C40/SCI interface units. An input signal from a calorimeter 'level-1' trigger will be available from simultaneous tests of RD27. The system will be run as push architecture, i.e. the supervisor, realised as another DSP, connects to all buffers. It will use, for simplicity, a round-robin scheduling algorithm.

ATLAS testing of the level-2 trigger in the H8 beam will provide important hands-on experience, but triggers will not be given a high priority in these tests, designed for high-priority detector issues to be settled. Detector interfaces will be ad hoc, and event speeds will be dictated by the available front-end modules, data acquisition limits, and the SPS cycle. For complementing the beam tests, there will therefore be a testing period in the laboratory towards the end of 1995 or in early 1996, in which a more complete test setup using emulator units as input will perform the full level-2 trigger task, based on the same units (DSP-s, C104, Alphas, SCI). In the lab setup, the speed limitations of such a system can be readily studied by varying the input frequency.

In parallel, the modelling activity around an ATLAS farm structured level-2 system will continue, and will be fed with the results established during the test sessions and benchmarking. Modelling has resulted in a software layer (SIMDAQ) above the basic commercial Modsim language, in which the various implementation possibilities have already been expressed. Particular effort went into models of the C104 switch and the C40 implementations (buffer and feature extractors).

### *RoI collection and pre-processing*

*a) Data-driven solution.* The generalization of the existing and demonstrated 'Router' for the TRT is under design, a first proposal has been made to ATLAS. The target is a detector-invariant buffer, covering a wide geometrical area or entire subdetector. This is achieved by high-bandwidth busses. Programmable units cater for the necessary preprocessing of raw detector data and can adapt to different transmission protocols or data formats used between detector fronted and L2 buffer. A scaled-down prototype implementation will be suggested for 1996.

*b) DSP-controlled buffer.* The successful test of a prototype L2 buffer controlled by a C40 DSP has shown that simple RoI collection algorithms and buffer management can be handled in the DSP, at the 100 kHz rate. The original idea of connecting DSP-s via its links to a network of DSP-s which implement RoI collection and feature extraction, has now been replaced by a routing switch, the C104 (supplied by the Macrame project). This switch uses the DS (datastrobe) line protocol, the necessary interfaces have been implemented. Data travel through this switch to a farm of feature extractors, equally realised as C40 DSP-s. An extrapolation from the measured transmission numbers of this small-scale network to a full-scale system will be made by a model in Modsim.

Several institutes in ATLAS and EAST have created a working group for common development of the C40/C104 network and farm. The group provides common software development tools, and coordinates laboratory and beam tests.

### *Processing and switching*

*a) Data-driven.* Demonstrations with boards based on field-programmable gate arrays (FPGA-s) as feature extractors have successfully been concluded in the past. DecPerle has meanwhile also been shown to be programmable for the global tasks, viz. combining all features in a RoI and final decision making. The benchmark algorithms typically only use a fraction of the board, and execute at cycle times below 1  $\mu$ sec, so that a fully pipelined implementation of the trigger via DecPerle can be considered demonstrated. To overcome some limitations in memory size and parallelism (encountered for the TRT only), the University of Mannheim has defined a successor board (Enable++) and associated software, that can cater for *all* presently defined feature extraction algorithms without restrictions, viz. cluster definition in calorimetry, pattern finding (Hough transform, histogramming) for the TRT and for the sparse trackers (e.g. Silicon strips or muon chambers).

There are two main improvements in Enable++, compared to DecPerle and Enable. A flexible I/O concept allows the use of multiple I/O subsystems that meet the requirements of different subdetectors of ATLAS. And spC, an extension of C, provides parallel and pipeline constructs, and allows to write Enable++ programs in a high level language. A source code debugger is also under development. Enable++ developments are supported through the INTAS program of the European communities.

An Enable++ prototype will be ready for demonstration in fall 1995. For the Enable++ I/O system a new interface to the HIPPI spy (for TRT data) will be tested for the test beam in September 95. When Enable++ is available, the TRT and the calorimeter algorithm will be implemented in a high level language (spC). Functional tests in the lab are planned in December 95/January 96. Before the next common run of ATLAS detectors in April 96, high-speed measurements as well as tests for the embedding of the Enable++ subsystem into the calorimeter environment are planned. In April, Enable++ will take part in the common ATLAS run to get first hands-on experience in working with more than one detector.

*b) Farm solution.* Benchmarking RISC and other processors with the algorithms selected for the feature extraction, had already been reported in the past, and no major new results can be added. Beyond pure feature extraction simulation, we intend to address also the impact of other factors on processing time: the farm-based solution must include, in addition to the execution of pure processing algorithms,

the interference of processing with the interfaces to data switching, kernel delays, and processor scheduling. If L2 buffers, due to local bottlenecks, send RoI-s in a format unsuitable for processing, data formatting aspects and address manipulation in the feature extractor will also have to be considered. We believe that simple economics will dictate the implementations of such parallelisable tasks in the L2 buffers, most likely in low-level software (DSP-s, FPGA-s).

A small-scale farm demonstrator based on C40, SCI and Alpha processors will start to be evaluated at the ATLAS test beam later this year; due to ATLAS constraints, the test will comprise only data from at most two detectors (TRT and L1 data from the LAr calorimeter). Early in 1996 the same setup will be speed-evaluated in the laboratory, and subsequently run with multiple detectors in the H8 beam. Eventually, this demonstrator will provide a full pilot implementation of a farm-based architecture including all stages from RoI collection to global decision. It will be interesting to use this small system and its measured speed for tuning the Modsim description, and for checking the reliability of architecture modelling. Several critical components running at 100 kHz and necessary for any farm-based solution (RoI builder, processor scheduler, RoI distributor) will be tested along with this setup, for the first time.

### *Global decision algorithms*

Up to now, studies of global decision making in L2 are still based on an intuitive data model after feature extraction, that has been based on statistical data, but ignores correlations in data between subdetectors. EAST participates in the forthcoming detector definition and trigger simulation in ATLAS; there is now a compelling deadline for ATLAS to have a coherent training set for the User Requirement documents, so that the definition of algorithms and benchmarking of all phases of the level-2 trigger should be possible over the coming year.

## **4. HPCN investigations**

All areas of EAST activity over the last years have been driven increasingly with the specific constraints and desires of the experimental proposals in mind, to the extent that the collaborations have built up their internal structures to cope with problems of defining and implementing triggers, in particular level-2 triggers. To the extent at which ATLAS has adopted a more detailed decomposition of the triggering problem, EAST has contributed more to the discussion in that proposal than to CMS, and most surviving activities of EAST are now firmly anchored in ATLAS.

There is one area in which EAST has made some initial steps and which concerns equally the two proton-proton collaborations, and possibly not only those: the possibility of using fully integrated parallel communication and processing systems for our purposes of high- to intermediate-frequency triggering. We propose here to continue an RD11 activity in this area for at least one more year, with CERN as leading laboratory, and both ATLAS and CMS defining the boundary conditions and control. The systems we refer to are called HPCN (for High Performance Computing and Networking), of which different implementations are offered on the commercial market by all major computer manufacturers.

## *Motivation*

The coming two, at most three years will require both collaborations to define in sufficient detail their respective trigger problems at what is called level 2, so as to make their final architectural and technological choices to begin implementation. The level 2 trigger must be able to deal with the frequencies left over by the fast level-1 40 MHz trigger, and must be sufficiently selective to allow in the next level the collection of full event information and high-level processing; generally the level-1 output rate is assumed to be 100 kHz or less, and the gathering of full event information is expected to cope with a rate of at least 1 kHz.

In the implementations of level-2 triggers (as shown in paragraph 1 above) a recurring idea is to use 'farms' of high-level programmable devices, with a flexible high-bandwidth and low-latency interconnection scheme. Multiple activities in the collaborations and in separate R&D projects (RD24, RD31), concern specific possible processing or switching implementations; the eventual hope is that the components under study will end up as commercial and reasonably priced off-the-shelf items, easily interface among each other and maintained by industry over many years.

While processor boards for small-scale applications are in rapid evolution, and switching components for Telecom applications are becoming available, there is presently no possibility for such equipment to be combined into systems with limited effort, and to be demonstrated at the performance numbers required for our purposes of 100 kHz triggering. Nor is there a user community emerging with performance requirements similar to those of LHC triggers which could push industry to develop open systems in the desired direction.

There is then a fair chance, if we exclude the option of major in-house developments of hardware, that the only supplier of high-rate parallel farms will be the computer industry, with integrated systems of the HPCN type. Despite the fact that present pricing policies make such systems look like a financial stumbling block, we believe that the rapid evolution of the customer base could well make them a much more realistic choice during the next few years.

## *Proposal*

We propose to assess systematically the capabilities and characteristics of existing HPCN systems, starting where work in EAST over the period 1993 - 95 (see paragraph 2 above) has left us. Presently, some knowhow has been obtained on Meiko's CS-2 (Sparc-based), on IBM's SP-2 (Power-2-based), and to some extent (ongoing work) on Convex SPP1000 (HP-based). The communication characteristics of these machines are quite different, and they are likely to show deficiencies and virtues in function of the problem definition. Past work has been done in close collaboration with the respective manufacturers, and this is expected to continue. All machines have an upgrade track which must be followed closely, and other vendors' candidates may be added (Digital, Intel).

Our intention is to define with the collaborations a systematic list of desired characteristic performance measurements, possibly including benchmark algorithms running on multiple nodes, and to obtain these numbers. To the extent that CERN's CN division installs and upgrades such systems on site, we would want to access those; more frequently, we intend to use systems installed at manufacturer or application sites other than CERN. In some cases, small configuration changes

and/or hardware interfaces could become necessary; most frequently, simple access to a multi-node partition of a system without interference from other users is expected to be sufficient.

CERN with its variety of installed systems is the present and proposed future group to pursue these investigations. The substantial knowhow existing in CN division on parallel systems of this general type is essential in achieving fast progress.

### *Milestones*

- defining characteristic problems, performance parameters and ways to measure them, jointly with ATLAS and CMS, between now and the end of 1995;
- constituting a small group (2-3) at CERN to execute the measurement program, from now until March 1996;
- discussion of access with manufacturers, measurements on several systems, iterations on problem definition with collaborations, iterations on measurements with manufacturers, ending with a report (or a series of reports): between early 1996 and mid-1997.

## **5. Resources and Collaboration Composition**

As a collaborative effort, RD11 proposes to cease its activities, with the exception of the HPCN investigations described in the preceding paragraph. For these, CERN (CN/ECP) will be responsible.

For 1995/6 we estimate next year's total spending at CERN to require 100 Ksf, for travel, interface equipment, possibly boosting the SP-2, and for visitors. There will be no beam tests except possibly in conjunction with ATLAS or CMS. For some flexibility and for general access (visitors!), we require our own computing budget at the modest level of 300 hours CERN.

## **6. Acknowledgements**

We want to acknowledge the efficient and pleasant collaboration with members of ATLAS, RD6, RD24, GPMIMD-2, MACRAME and AFRODITE.

## **5. Publications**

### **Publications and conference contributions:**

*Noffz K.-H., Zoz R., Kugel A., Klefenz F., Maenner R.: Results of On-Line Tests of the ENABLE Prototype, a 2nd Level Trigger Processor for the TRD of ATLAS/LHC; Proceedings Comp. in High Energy Phys., San Francisco, CA (1994) 53-56*

*A.J.Borgers et al.:Execution-time Reduction for Triggering Algorithms by accelerating TI's TMS320C40 DSP, RTD 94.*

*Noffz K.-H., Zoz R., Kugel A., Klefenz F., Maenner R.: Die Enable Machine -Ein Echtzeitmustererkennungssystem auf FPGA-Basis; Abstr. GI/ITG-Workshop Architekturen fuer hochintegrierte Schaltungen, Schloss Dagstuhl (1994) 67-70*

*R.K.Bock et al.:A commercial image processing system considered for triggering in future LHC experiments ; Nucl.Inst.Meth.A 356 (1995) 304-308*

*D.Belosloudtsev et al.: Programmable Active Memories in real-time tasks: implementing data driven triggers for LHC experiments , Nucl.Inst.Meth.A 356 (1995) 457-467*

*L.Lundheim, I.legrand, L.Moll: A programmable active memory implementation of a neural network for second-level triggering in ATLAS, Fourth International Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics, Pisa, 1995.*

*J.M.Seixas, L.P.Caloba, B.Kastrop A neural network second-level trigger based on calorimetry and principal component analysis, Fourth International Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics, Pisa, 1995.*

*A.C.Balke, J.Carter, J.Haveman: Experience using formal methods in High-energy Physics, Fourth International Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics, Pisa, 1995.*

*Högl H., Kugel A., Ludvig J., Männer R., Noffz K.-H., Zoz R.: Enable++: A Second Generation FPGA Processor; Proc. 3rd IEEE Symp. on FPGAs for Custom Computing Machines, Napa, CA (1995)*

*B.J. Green, J.A. Strong, R. Cranfield and G. Crone: A second level data buffer with LHC performance, Nucl.Inst.Meth. A 360 (1995) 359-362*

*H. Högl, A. Kugel, J. Ludvig, R. Maenner, K.H. Noffz, R. Zoz: Enable++: Ein Hochleistungs-Prozessorsystem auf FPGA-Basis mit komfortabler Entwicklungsumgebung; Proc. GI/ITG-Workshop Anwenderprogrammierbare Schaltungen, Karlsruhe, Germany (1995).*

*R.Hauser and I.Legrand:A Real-Time Application for the CS-2, Proceedings of HPCN Europe 1995 (May 1995, Milano) Lecture Notes in Computer Science 919, Springer, p.684.*

*H. Högl, A. Kugel, J. Ludvig, R. Männer, K.H. Noffz, S. Rühl, R.Zoz: ENABLE++: A General-Purpose L2 Trigger Processor; accepted for publication in Proc. First Workshop on Electronics for the LHC Experiments; Lisbon, 11-15 September 1995*

*H. Högl, A. Kugel, J. Ludvig, R. Männer, K.H. Noffz, S. Rühl, R.Zoz: ENABLE++: A General-Purpose L2 Trigger Processor; accepted for publication in Proc. IEEE Nuclear Science Symposium; San Francisco; October 95*

## **EAST Notes since the last Status Report:**

EAST note 94-15 *Using a field-programmable gate array as a dedicated coprocessor for the Texas Instruments' TMS320C40 DSP* (A.J.Borgers et al.) June 1994

EAST note 94-16 *Status Report EAST (DRDC 94-20)* (R.K.Bock) 9 May 1994

EAST note 94-17 *Modelling of local/global architectures for second level trigger at the LHC experiment* (Z.Hajduk et al.) 20 February 1994 (also ATLAS DAQ Note #13, and RD-13' note TN-116)

EAST note 94-18 *Understanding derandomization* (Jos Vermeulen) 19 May 1994

EAST note 94-19 *EAST collaboration meeting #12, Minutes* (R.K.Bock) 20 June 1994

EAST note 94-20 *TRT on DecPeRLe-1* (Lars Lundheim, Laurent Moll, Philippe Boucard, Jean Vuillemin) revision 2, 16 August 1994

EAST note 94-21 *Implementation of a pattern recognition algorithm for the Si tracker on DecPeRLe-1* (W.Krischer and L.Moll) 11 October 1994

EAST note 94-22 *Communication benchmarks with the Ancor Fibre Channel Fabric*, (Fabrice Chantemargue) 28 July 1994

EAST note 94-23 *4-year Status Report to the DRDC* (R.K.Bock and W.Krischer) 20 August 1994

EAST note 94-24 *Communication benchmarks with HPs and IBMs connected around an ANCOR Fibre Channel Fabris* (F.Chantemargue), draft, 29 August 1994

EAST note 94-25 *Meeting on C40 digital signal processor activities, Minutes* (R.K.Bock and H.Leich) 8 September 1994

EAST note 94-26 *A commercial image processing system considered for triggering in future LHC experiments* (R.K.Bock et al.) 20 September 1994 (published in Nucl.Inst.Meth.A 356 (1995) 304-308)

EAST note 94-27 *Programmable Active Memories in real-time tasks: implementing data driven triggers for LHC experiments* (D.Belosludtsev et al.) 7 October 1994 (published in Nucl.Inst.Meth.A 356 (1995) 457-467)

EAST note 94-28 *Global Decision on the CS-2* (R.Hauser, I.C.Legrand) 5 Oct 94

EAST note 94-29 *Router Input Board* (D.Belosludtsev and S.Khabarov) 23 Sept 94

EAST note 94-30 *Data collection and preprocessing for the ATLAS second-level trigger* (R.K.Bock, J.Carter, R.hauser, I.Legrand) Draft 0.2, November 1994

EAST note 94-31 *3rd Meeting on DSP developments, Minutes* (R.K.Bock and H.Leich) 22 October 94

EAST note 94-32 *EAST Collaboration meeting #13, Minutes* (R.K.Bock) 22 Oct 94

EAST note 94-33 *ATLAS, data for trigger studies in a portable text format* (J.Carter et al.) revision 1, 21 Oct 94

EAST note 94-34 *SCI with DSPs and RISC processors for LHC 2nd level triggering* (P.Clarke et al.) November 1994

EAST note 94-35 *A second-generation FPGA processor for second-level triggering* (H.Högl et al.) 05 December 94 (also ATLAS DAQ Note 26)

EAST note 94-37 *Algorithms in second-level triggers for ATLAS and benchmark results* (R.Hauser and I.Legrand) 14 Dec 94 (also ATLAS DAQ Note 27)

EAST note 94-38 *Local processing for a farm-based second-level trigger at LHC* (J.Strong) December 94 (also ATLAS DAQ Note 21)

EAST note 94-39 *A second level buffer with LHC performance* (B.J.Green et al.) Dec 94 (also ATLAS DAQ Note 20)

EAST note 95-01 *A neural network for global second level trigger - a real-time implementation on DecPeRLe-1* (L.Lundheim, I.C.Legrand, L.Moll) 31 March 95



EAST note 95-02 *A data collection and preprocessing unit for the LVL2 trigger of ATLAS* (P.Kammel, A.Reinsch, V.Dörsing) draft, 7 April 1995

EAST note 95-03 *Simulation and design of the on-line trigger and reconstruction farm for the Hera-B experiment* (U. Gensch, I.C. Legrand, H. Leich, P. Wegner) April 95

EAST note 95-04 *EAST Collaboration Meeting #14, of 8 May 95, Minutes* (R.K.Bock) 18 May 1995

EAST note 95-05 *Fibre Channel performances with IBM equipment* (C.Miron, M.Bernaschi, F.Chantemargue, S.Muñoz, G.Maron) 22 June 95

EAST note 95-06 *Fibre Channel performances measured with a Fibre Channel tester* (F.Chantemargue, S.Muñoz, E.Denes, G.Rubin) (draft)

EAST note 95-07 *MA-16-Based Global L2 Trigger Sensitive to Invariant Mass* (R.Odorico) 7 June 1995

EAST note 95-08 *A Data Driven Implementation for Invariant Mass Computation in Level 2 Global Triggering* (L.Lundheim, R.K.Bock) 29 June 95

