CERN/DRDC 94-20
RD11 Status Report
9 May 1994
(EAST note 94-16)

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, G.Odor *(Central Research Institute for Physics, Budapest)*
R.K.Bock (*), J.Carter, F.Chantemargue, R.Hauser, W.Krischer, I.Legrand, L.Lundheim, R.McLaren *(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, N.V.Gorbunov, V.Karjavin, S.V.Khabarov *(LHSE, JINR, Dubna)*
V.Dörsing, P.Kammel, A.Reinsch, H.U.Zühlke
(Institut für Angewandte und Technische Informatik, Universität Jena)
Z.Hajduk, W.Iwanski, K.Korcyl, P.Malecki, Z.Natkaniec
(Institute of Nuclear Physics, Krakow)
R.Nobrega, J.Varela *(LIP, Lisbon)*
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)*
F. Klefenz, A.Kugel. R. Männer, K.H. Noffz, R. Zoz
(Institute of Computer Science V, University of Mannheim)
P.Bitzan, M.Kucera, J.Smejkalova
(Inst.of Computer and Information Science, Prague)
L.Levinson, M.Sidi *(Weizmann Inst.of Science, Rehovot)*
L.Caloba, M.Seixas *(Federal University, Rio de Janeiro)*
C.Balke, A.Borgers, J.Haveman, W.Lourens, A.Taal *(Utrecht University)*
U.Gensch, H.Leich, U.Schwendicke, P.Wegner
(Institute of High Energy Physics, Zeuthen)

(*) Spokesperson

Summary

EAST has continued, over the last year, to focus most activities towards questions that still need better understanding for the experiments now in preparation. Collaboration with ATLAS is close in all areas; EAST participates in ATLAS working groups and depends on the contributions from trigger-oriented physicists in ATLAS. Indeed, there is substantial overlap in personnel that take responsibilities in ATLAS and EAST. Contacts in some areas exist also with CMS, where the detailed study of second-level triggering has only recently been given serious attention. Some members of EAST also participate in setting up triggers for 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 93-12) have largely been met:

- the Router and two pipelined feature extraction devices (Enable and MaxVideo) have run in October/November 1993 jointly with the RD6/RD13 beam tests;
- FEAST prototype hardware (now called the L2 buffer) has been prepared with C40 digital signal processors, and will be tested in the H8 beam later this year, including an interface from the C40 to SCI;
- DecPerle has demonstrated a realistic calorimeter algorithm running faster than 100 kHz, and will run in June with a tracking algorithm in beam tests;
- with some modifications to the original plan, an SCI interface from the L2 buffer to an Alpha processor is being built, and will be demonstrated at least partly later this year in conjunction with the L2 buffer;
- overall simulations for the global decision part have been done, and have led to the understanding of the relevance of pilot implementations of critical components; this has induced joint and larger-scale modelling activities directed by ATLAS;
- the possible mapping of (part of) level-2 triggering onto an MPP machine has been pursued, although the delivery of the CS-2 to CERN has been delayed;
- the work on optical multi-fibre links has not progressed further, as no manpower was available; due to the slow development in industry, we may also consider this as a lower-priority investigation for the moment;
- the AFRODITE project follows its time scale, a level-2 system has been expressed in the VDM++ language and an evaluation paper is available.

1. Second-level triggering, overview

The past activity in EAST has resulted in a rather detailed decomposition of second-level triggering. This decomposition has become widely accepted and a working hypothesis in ATLAS; part of it is an option in CMS. The achieved problem structure leaves a wide variety of implementation choices, i.e. of expressing existing possibilities of parallelizing the architecture in different ways. Even more choices exist in the technologies for transmission and processing, and these are premature to discuss at this stage. Based on these choices and extrapolations, a final system will eventually have to be costed.

Increasingly, in a joint effort mostly with ATLAS, EAST has been working on reducing the architectural options to as few as possible. We expand the architectural properties in Appendix A, but give here a few introductory remarks, mostly for defining the terminology of the presentation.

EAST has developed the quite fundamental *Region-of-Interest (RoI) concept* for level-2 triggering, at least at high luminosity. RoI-s are spatially limited areas ('roads') in the detector in which the level-1 trigger has identified candidates for phenomena to be triggered. Using data only in RoI-s alleviates the bandwidth requirements in transmission from data buffers to processors executing the algorithms (data outside RoI-s are not transmitted to the trigger structure), decomposes the problem (RoI-s can initially be analysed independently), and simplifies algorithms (only a specific phenomenon must be looked for). The RoI concept relies on the level-1 trigger to identify those parts of the detector containing candidate features (electrons, photons, muons, jets).

The goal of EAST is to have solutions for implementations of level-2 algorithms on all characteristic detector components, *using full-granularity, full-precision* data. The extent to which data with full precision are effectively needed in level 2 will continue to be the subject of detailed simulation studies, and may eventually be dependent on the target physics. In general, reducing the requirements on precision in the trigger will result in architectural simplification and thus cost savings.

The overall functional decomposition of the level-2 problem as used in EAST identifies three phases:

Phase 1: Front-end buffering and collection of regions of interest

The full raw detector data for level-1 (L1) triggered events are collected in local non-overlapping memory modules (structured into chips, boards, crates), in a detector-dependent modularity. The raw data pertaining to regions of interest (RoI-s) have to be selected by some mechanism, which we term **'RoI collection'**. This operation is guided by a device realized outside the L2 data stream, which indicates the whereabouts of RoI-s. This L1-guided unit is called a **'RoI-builder'**. We assume RoI collection to proceed independently and in parallel for different subdetectors and for different RoI-s.

Phase 2: Feature extraction: Local processing of data in a RoI of a subdetector

Algorithms in the concept of RoI have the initial task to convert a limited amount of local raw data from a single subdetector into 'features', variables containing the relevant physics information, like cluster or track parameters that can be used to corroborate or disprove the different physics hypotheses. This phase is called **'feature extraction'**. Feature extraction algorithms have a locality that permit to exploit the natural double parallelism of RoI-s and subdetectors. Simulation will have to show if and to what extent this simple concept has to be diluted in order to avoid physics losses e.g. in regions of overlap, where each individual detector party only has a weak signal. Feature extraction algorithms are expected to be closely coupled with the respective subdetectors. While they will have to be suitably parameterized, it is not clear to what extent conceptual flexibility (i.e. profound changes in algorithms) will have to be foreseen.

Phase 3: Global decision: RoI and event processing

Physics features have to be collected from all subdetectors and from all RoI-s, for forming an overall decision on the entire event. If one looks for further decomposition, the natural and efficient order of processing is to combine first all subdetectors that relate to the same physics 'object', by an algorithm which is limited to an RoI. This is then followed by an algorithm combining information from all RoI-s into an event decision. Both steps together are termed '**global decision**'. It is in the global decision that algorithms predictably will have to be adapted to evolving physics understanding, and where detector changes will invariably result also in algorithm changes.

Hardware options

Basic choices exist in expressing the implementation of the three phases as architectures. For simplification, only two overall options are discussed in this note, although hybrid architectures are a very possible result of the studies yet to be completed. We call the two options the farm-based and the data-driven architecture respectively (see the appendices for more details and for a discussion).

The "**farm-based**" approach uses standard commercial devices, general-purpose processors and network components. In the simplest scheme, there are two layers of processors performing local feature extraction and global decisions respectively. Local processors receive their data from intelligent devices in the L2 buffer, through a switching network. The global processors are organized as a general farm; local processors may be run as one farm per subdetector, but groups of processors may also be permanently assigned to regions of the detector.

The "**data-driven**" approach uses low-level devices for ROI collection (sometimes called 'routers'), directly coupled to (or inserted upstream of) the L2 buffers. Typically, these would be implemented as local busses with control through field-programmable gate arrays. Devices based on the same principle of low-level programmability, are also used as feature-extraction processors. Solutions based on field-programmable gate arrays have been shown to operate in a pipelined mode and cope with the level-1 rate of 100 kHz. Although data-driven pipelines of digital signal processors (DSP-s) have also been successfully explored for the global decision, the network/farm approach is kept for this low-bandwidth problem, for reasons of flexibility.

2. *Work in EAST*

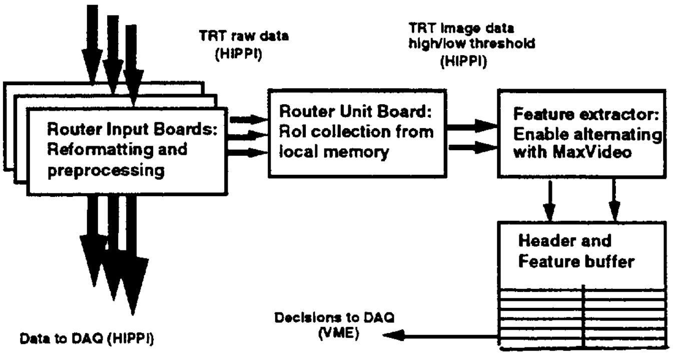
In the following we will discuss in some detail how EAST activities have explored critical parts of the above architectural choices, using the technological possibilities available today. The collaboration's intended end result is a catalogue of remaining options, that can be costed in detail at the time when LHC experiments will have to make definitive choices. We should note that some originally attractive options have already been discarded by the work in EAST over the last two years, and our goal is to reduce the discussion to only few viable choices.

The conclusion so far can be summarized as follows: work in EAST has shown that by proper choice of hardware, for RoI collection and feature extraction a frequency of 100 kHz can be maintained by data-driven pipelined systems, without introducing event parallelism in farms. This result should be interpreted as no more than a demonstration of one definitely possible implementation. Although presently not demonstrated, and presently somewhat at the leading edge of technology, farm-based systems have solid arguments going for them on some accounts, and we want to explore their components as they become available, jointly with the collaborations and the two projects specifically interested in fast switching devices, RD24 and RD31. We also believe that digital signal processors (DSP-s) with their combined communication and processing capability, have potentially a major rôle in building up trigger systems, and want to continue exploring them.

2.1 Activities and milestones of 1993/4

Beam test demonstrations of Router, Enable, MaxVideo

Together with project RD6, EAST has shown in beam tests (in H8) a demonstration of triggering at LHC speed using a TRT-specific RoI collector and two alternatives (Enable, MaxVideo) for feature extraction.



Shown in the laboratory to run faster than 100 kHz, and limited by transmission rather than processing, both devices have followed the much lower speed at the experiment (determined by the data acquisition). Results were recorded and shown to be bit-by-bit compatible with off-line programs. The more general pattern recognition capabilities of Enable will lead to further development of this device (see below), whereas the fully commercial and somewhat limited MaxVideo implementation will not be pursued further.

DecPerle

A DecPeRLe1 system produced by the Paris Research Laboratory (PRL) of Digital, has been at CERN since January 1993. As a general Xilinx-based board, marketed in a small series for computer architectural prototyping and data pre-processing, DecPerle in our application is intended to demonstrate the data-driven execution of algorithms for quite different detectors, on the same hardware. An implementation of a calorimeter trigger algorithm has been shown to run faster than at 100 kHz on this device. A TRT trigger algorithm has been designed and tested by PRL, and again keeps processing at the speed of transmission. A hardware interface from the detector passes through a commercial HIPPI/Turbochannel board. All testing and integration is imminent at the time of writing, DecPerle is part of the planning of the June 1994 beam tests with RD6.

Successful implementations of significantly different algorithms on DecPerle (and on Enable, which has a similar scope of development) have a major significance: On a single programmable hardware platform, totally different feature extraction algorithms can be implemented by 'software'. Further, a small number of boards can be locally interfaced to a single standard high-level processor (for DecPerle presently the Decstation, via Turbochannel). General-purpose FPGA-based feature extraction boards connected in this way to the same general-purpose node, result in a natural multi-detector feature extraction system for a single region of interest, obviating the need for a switching network at this level.

L2 buffer

A pilot implementation of a level-2 buffer with associated DSP (digital signal processor TMS320C40 from Texas Instruments) does now exist as prototype. It uses boards from LSI (Loughborough, UK), containing multiple TIM modules (replaceable daughter boards), of which some have been individually modified for our purpose. The buffer memory is planned to be tested in the joint RD6/RD11 beam tests, using the same interface (Router Input Board) as used by the pipelined processors (Enable, DecPerle).

Alpha/SCI

Design and implementation of a prototype global level-2 system based on DEC Alpha processors and SCI is now well under way. The first system is expected to comprise 3 or 4 Alpha-based single-board computers (SBC-s) and a standard DEC Alpha workstation, all connected by SCI. This will form part of a prototype farm-based level-2 system also incorporating TMS320C40 DSP-s from Texas Instruments (a figure in 6.2 below shows the full demonstrator system now being designed and assembled, including the RoI collection and feature extraction parts).

First commercial implementations of the SCI protocols in CMOS have recently become available. Although the original bandwidth target for the first CMOS chips was 125 Mbyte/s, a rate as high as 180 Mbyte/s has been demonstrated by the manufacturer. We have one such chip under test and are expecting delivery of some more very soon.

The SCI interface for the Alpha workstation is based on a design by RD24, but with the original GaAs SCI interface replaced by the CMOS version, and the host DECstation replaced by an Alpha. Fabrication of the interface is nearly complete and commissioning should begin in a few weeks.

The SCI connection to the Alpha SBC-s will be through the PCI bus. The interface has been completely specified and will be implemented in a modular way. The SCI part has been designed and is currently in manufacture, whilst the design of the PCI part is now well under way. OSF/1 will be used as operating system on the Alpha workstation and it is planned to use the VxWorks real-time kernel on the Alpha SBC-s.

The originally proposed Alpha, Futurebus+ based SBC solution has been replaced, in collaboration with the Digital Joint Project, by an evaluation board targeted at OEMs in the PC market. This board is more suited to our needs, is much more attractively priced and has a higher level of on-board integration brought about by the direct incorporation of PCI bus onto the Alpha processor and other support chips.

Continued algorithm work

Algorithms have to be continually adapted to the evolving definitions of the detectors. 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 inner detector, and for a simple model of global decision.

In order to update this 'suite', EAST also participates actively in ATLAS groups providing large samples of L1-filtered signal and background data, for the same detector design, based on which 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.

Global decision, comparative evaluation by small-scale implementation and modelling

The existing decision algorithms, and simple data sets generated for six different physics channels, plus QCD background, have been used for extensive benchmarking for potential transmission and processing components. As indicated earlier, for this implementation only farm-based solutions are seriously considered for this phase: due to a much reduced number of connections and small packet sizes after feature extraction, switching seems quite realistic with the technology available soon, and algorithms in this phase need the flexibility and scalability only offered by general-purpose processors. Pipelined implementations of global decisons based on DSP-s and transputers were also investigated, though, and found to be a realistic alternative (e.g.EAST note 94-12).

Modelling requires well-understood input parameters: first results exist on algorithm execution times and switching latencies (first numbers are given in Appendix B, but for the moment are preliminary and have to be taken 'with a grain of salt'). The compute nodes under consideration are Texas Instrument TMS320C40 DSP-s, Digital Alpha chips, Power PC (IBM RISC 601), Sparc 10, HP 750. They connect in various ways among themselves and to feature extraction processors, using their own fast links (C40) or networks or switches expected to become available commercially (ATM, SCI, fibre channel, C104 transputer switch). All of these components compete potentially for their place in 'farm-based' solutions, but more than isolated algorithm benchmarking and transmission measurements have not been done yet. Algorithm execution and basic communication times are but a fraction of the problem understanding; their interplay and additional network/switch latencies and interrupt delays in the real-time kernel (minimal operating system) play a substantial role.

Other architectural possibilities

EAST has also used its 'benchmark suite' for the L2 trigger in 1993 for an investigation to port the trigger in parts or as a whole onto high-performance (parallel) computers and network (HPCN) systems, starting with the CS-2. This exercise is somewhat academic, because of the high cost of HPCN systems, at least for the task of RoI collection with its hundreds or thousands of parallel inputs. The implementation should nevertheless give us better understanding of the interplay between communication and processing. HPCN systems have versatile high-bandwidth switches and general-purpose processors much like the farm-based solution requires to build. In fact, HIPPI, FCS and SCI switches are among candidates to be used in such systems. It is intended to include other HPCN systems in the benchmarking, with the help of their manufacturers: contacts exist with Thinking Machines (CM-5) and Intel (Paragon).

Possibilities to implement the trigger on other devices than pertaining to the models outlined above, are not being pursued further. The low-level MIMD chip from ITT, originally found very attractive, has not reached the vicinity of market. Neural-net-inspired solutions have been found helpful in some areas of algorithm optimization, but not as a possibility of implementation on hardware: Neural network-oriented special hardware, which is given much publicity in some areas, is not considered a viable option at the present time. In earlier work in EAST, architectural solutions based on single-instruction-multiple-data (SIMD) hardware had already been discarded.

Optical links

The work on optical multi-fibre links has not progressed further, as the foreseen manpower at CERN was not available.

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++. An evaluation report on experience acquired in this work has been written.

2.2 Activities in 1994/5

This section starts by an overview: as R&D projects move closer to the LHC collaborations, and their results are used in writing the Technical Proposals, it appears increasingly relevant to discuss the planning in terms of 'collaboration options'. The milestones of individual projects are, therefore, given under those headings.

Work in EAST in 1994/95 must concentrate on providing a maximum of information towards deciding about the two major architectural options still open , data-driven and farm-based (see chapter 1 and appendix A). We already noted that pipelined data-driven systems have been or are being demonstrated; they may arguably provide a simple and cost-effective solution, because using best the existing parallelisms and minimizing the number of processors. Farm-based solutions are usually defended with the arguments of commercial availability, high-level programmability, flexibility etc. Certainly, arguments like standardizing components, homogeneity, scalability, technological evolution, ease of programming, reconfiguration in case of malfunction of components, all are cost factors in their own right, and will have to be considered. The characteristics of farm-type solutions are, however, not yet understood in detail, and full-scale pure farm-based systems may not be implementable in practice with truly commercial components.

We list some of the results obtained in the course of the past farm-based work in Appendix B, and expand the argument there. It is a conclusion also agreed in ATLAS and CMS that 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 for good. EAST proposes below to

- work on farm-based pilot implementations in detail,
- continue and possibly terminate the pure FPGA-based data-driven studies,
- invest work in DSP-based systems; DSP-s can be seen as powerful hybrids which can be used both in a pipelined (data-driven) and farm-based architecture.

The following plan of work for the coming year is submitted suggesting that this is likely to be more or less last year of RD11 operating outside the collaborations.

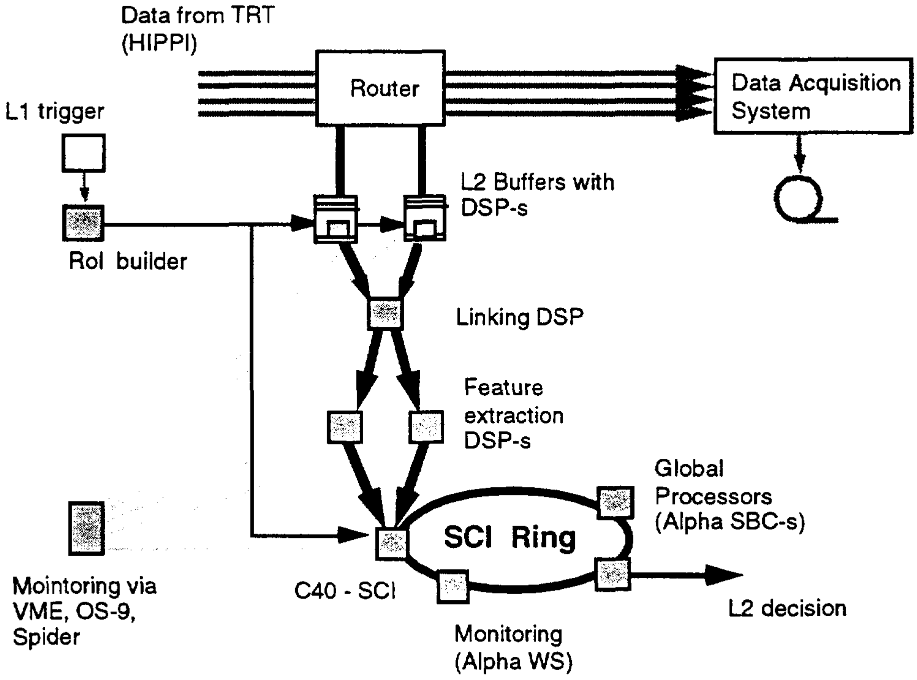
Beam tests

The joint tests with RD6 in the H8 beam will continue, comprising in the period June-September 94 a more complete test of data-driven (and TRT-specific) elements: Router, Enable, and (new) DecPerle. In September, first tests with the L2 buffer including the C40 DSP will be added, and possibly data will also be routed from there into the Alpha processor via SCI. Note that due to an entirely new front-end electronics chain on the TRT (drift time will be read for the first time), a redesign of the Router Input Board became necessary. It is now executed in programmable gate arrays, and hence more likely to be adaptable to future changes.

The ATLAS testing in the H8 beam is foreseen to develop into a multi-detector activity. Presently, event building from different detectors is done in software and is painfully slow. Members of EAST are eager to use the collaboration's various architectural components to execute a multi-detector trigger (single RoI) in real-life conditions; they believe this to be the best way of providing meaningful parameters for later inclusion in designs of final data acquisition systems. In this spirit, a longer-range test of serious trigger options for multiple detectors is being proposed as part of the ATLAS Barrel Sector Prototype (not operative before 1996/97).

Milestones: mentioned below

demonstrator will provide a full pilot implementation of a farm-based architecture including all stages from RoI collection to global decision.



This demonstrator will be an important existence proof for a farm-based architecture. It will also be interesting to simulate this small system as a means of checking the reliability of architecture modelling. There is further an independent activity to boost C40-s by an FPGA-based coprocessor, so that these DSP-s can themselves compete as feature extractors. In a similar spirit, the implementation of algorithms based on artificial neural networks is pursued, on DSP-s.

Note that several critical components running at 100 kHz and necessary for farm-based solution (RoI builder, processor scheduler, RoI distributor, see Appendix A) have not presently been specified; EAST, however, considers these to be beyond the scope of its activities, and does not foresee any resources for completing the model.

Milestones: problem definition with all algorithm parts and transmission (4Q 94), modelling (EAST participates in an ATLAS-driven activity), and pilot implementations: commissioning of the SCI interface to the Alpha workstation (3Q 94), integration of the SCI-PCI interface with the Alpha SBC-s (4Q 94), integrated testing and demonstration SCI/Alpha (4Q 94), implementing and testing a C40 with coprocessor (4Q 94), evaluation of a C40-based switching and feature extraction network (3Q 95).

Global decision farm

Present studies of global decision making in L2 are 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 trigger simulation under way in ATLAS; as there is now a compelling deadline for ATLAS the foreseen level-2 buffer, and a continuation of the DSP-oriented work. Several institutes in EAST have already created a working group for common development of C40 networks, which provides common software development tools and has built a C40 link implementation, pipelined over optical links and allowing a substantially longer link than foreseen with the original C40 link (limited to less than 50cm). Several EAST notes describe the work of these groups.

For serious testing of multi-buffer situations, a new version of the existing emulator has also been defined, which allows flow control and thus correctly tests synchronized or asynchronous equipment, without access to detectors.

Milestones: Provide an updated version of the SLATE emulator (4Q 94), benchmark RoI collection by the C40 (1Q 95), operate a pilot network of C40-s and evaluate (4Q 94), study new DSP choices like the AD 21060 (1Q 95).

Feature extraction: data-driven solution

Tests with boards based on field-programmable gate arrays (FPGA-s) as feature extractors have already been done in the past, and will continue in 1994 (Enable, DecPerle). While this establishes the possibility to run some algorithms on these boards, in the case of DecPerle even to run several on the same hardware, the question of defining the bottlenecks and limits in generality of these solutions, programmable at the gate level, must be addressed. EAST has, therefore, started to define a board that can cater for *all* presently defined feature extraction algorithms. This will comprise few classes of algorithm types, cluster-oriented (calorimetry) and pattern-oriented (Hough transform, TRT). Work on algorithms for the sparse trackers (Silicon or GaAs strips, microstrip gas chambers, muon chambers) is still proceeding to define if these can fall into the same class as the TRT, with thresholded ('zero-suppressed') data as input; possibly, classical algorithms with permutations will also have to be considered. The new board will be based on principles of Enable and Decperle. It should be noted that some support has been obtained for a generalized Hough transform design through the INTAS program of the European communities.

Milestones: define the generalized problem rigorously (3Q 94), project and simulate solutions using FPGA-s (1Q 95), prototype parts of it on DecPerle (1Q 95), implement and test a prototype board in FPGA-s (not before end 1995).

Feature extraction: general-purpose processors

Benchmarking various RISC and other processors with the algorithms selected for the feature extraction, has already been done in part, as noted above. Beyond pure feature extraction simulation, we want to address also the impact of other factors on processing time: the farm-based solution must include, beyond 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.

A small-scale demonstrator based on C40, SCI and Alpha processors will start to be evaluated at the ATLAS test beam later this year. Eventually, this

RoI collection: data-driven solution

The generalization of the existing and demonstrated 'Router' for the TRT is presently under study, and first proposals are under discussion. The target is a hardware solution, programmable to cater for all detectors and adaptable to different models for data transmission between detector fronted and L2 buffer. It appears that a largely detector-independent solution can be implemented in principle. The constraints this solution imposes on transmission and on the functioning of the RoI builder can be specified, e.g. how many crates will have to house the Router, how many RoI-s can be extracted, which latency and asynchronicity is allowed in frontend-to-buffer transmission. These constraints will then have to be compared with realistic estimates of how ATLAS detectors plan to link to the L2 buffers, information which is presently not available.

These studies will continue, in close collaboration with ATLAS. A scaled-down implementation may be suggested in the course of the year, but being strictly detector-oriented, this should be funded out of an ATLAS budget.

Milestones: test data-driven Router for TRT in beam (June/September 94); submit a proposal for a generalized data-driven Router to ATLAS working groups (3Q 94).

RoI collection: intelligent buffer

The test of a prototype L2 buffer later this year will allow to measure execution times of RoI collection algorithms in the DSP, for different detectors. The DSP can, in principle, be connected via its links to a network of DSP-s which mediates RoI collection, possibly even feature extraction. The DSP will also be connected to an SCI ringlet, from where data can be sent to Alpha processors (see below). Having an intelligent device attached to a memory allows a serious evaluation of the limits of data selection for RoI collection and of memory management in general.

Measurements will also be made for various switching networks: C40-based (where possibly RoI collection and feature extraction can be mixed), ATM (done jointly with RD31), fibre channel (done with CERN's CN division), and SCI (done jointly with RD24). If available, it is also intended to include C104 switches based on transputer links. An extrapolation from such measured numbers to a full-scale system will then be made by a model (modelling work was initiated jointly with ATLAS). From this work it should become apparent, to which extent geographical subdivision of this task is necessary, and what technology extrapolations are necessary (if any) for the DSP and the associated switching network, to satisfy the overall performance criteria in ATLAS.

Also RD13 is considering extending their DAQ studies with commercial readout cards (CES RIO-s) and HIPPI interconnects, to include trigger studies in the RISC processor of their system. The CMS collaboration is following similar ideas, based on fibre channel. It would seem desirable to establish better communication between the efforts of different groups currently working independently in this area.

EAST works on the assumption that the present joint approach between different R&D projects, outside institutes and CERN-based groups, and the LHC collaborations continues in its present form. For EAST, this means an evolution of to have a coherent data set for the Technical Proposal, we can hope to also derive a set of data useful for first defining and then benchmarking the global decision phase of the level-2 trigger.

The pilot implementations based on ATM, fibre channel (FCS), or SCI on the network side, Power PC, Alpha and others for processors will serve not only for an evaluation of RoI collection and feature extraction, but allow also to judge their suitability for the lower requirements of global decision making in L2, as defined by the benchmark algorithms and data sets. The pilot implementations have to be completed and fully understood, so that extrapolation to a full-scale system becomes meaningful. Potential demonstrators are based on the Alpha/SCI interface described above, an FCS set-up with IBM Power PC-s, and ATM when it becomes available with processor interfaces.

Milestones: Define a benchmark suite and test data derived from more realistic Monte Carlo studies in ATLAS (3Q 94), provide pilot implementations on commercial FCS switches and interfaces with Power processors, possibly switching to board products with integrated interfaces (3Q 94). See also the previous paragraph for work on SCI/Alpha, which may serve both in a RoI collection and a global decision farm.

Overall simulation and modelling

EAST participates actively in ATLAS-driven activities around physics simulation to study optimal algorithms and provide data sets, and in modelling the farm solutions for feature extraction and for global decision. This work is presently focused on providing meaningful results on the time scale of the Technical Proposal. The EAST contribution to this activity will be kept at the highest manpower level we can afford.

Milestones: No EAST-specific milestones defined.

VDM++

The continuation of EAST involvement in the AFRODITE project foresees to provide an architectural simulator based on a formal specification written in VDM++. The translation into a simulator language will be non-automatic, but the objects used in VDM++ will be reused in simulation. We also intend, with other AFORODITE partners, to establish a definition of the data-driven Router in VDM++. This will then be translated by an existing code translator into a hardware description expressed in VHDL. In order to disseminate formal specification experience gained in AFRODITE, and to confront it with the realities in the design of large physics programs, members of the team have also joined in the proposal P55.

Milestones: L2 architecture simulation based on VDM++ objects (3Q 94)

3. Collaboration administration

3. 1 Composition

In the past year, University College London and the Federal University of Rio de Janeiro have joined the collaboration. Ecole Polytechnique have reoriented their activities and are no longer members of RD11.

3.2. Work attribution over the next year

The subdivision of work inside EAST is as follows:

Router beam tests, design of generalized Router:	Jena, Dubna
L2 buffer with intelligence (DSP):	RHBNC, UCL, Krakow
Enable beam tests:	Mannheim, Dubna
Decperle design and beam tests:	PRL, CERN
Definition of trigger algorithms:	CERN (with much help)
Generalized Hough transform board:	Mannheim, CERN, Dubna, Copenhagen (+ non-EAST)
SLATE emulator, Mark-2:	RHBNC, CERN, Budapest
Pilot for intelligent L2 buffer:	RHBNC, UCL
DSP networks and feature extraction:	Utrecht, Zeuthen, NIKHEF, RHBNC, CERN, Rio
Pilots for commercial global decision farm:	CERN
Pilots for Alpha/SCI-based farms:	RAL, Manchester
Modelling of farms:	CERN, RAL, NIKHEF (in ATLAS working group)
VDM++:	Utrecht, CERN

3.3. Resources

For 1994/5 an unchanged level of activity is planned, and further pilot implementations and hardware demonstrations are foreseen. The possibility of spending some ATLAS money may alleviate budget restrictions, and forecasts will be somewhat lower than in the last year. We estimate next year's total spending at 600 KSf, of which 200 KSf for CERN. The overall EAST manpower involvement is around 30 full-time equivalent.

We will continue to perform beam tests only in conjunction with ATLAS or other R&D activities, and hence do not require a beam time allocation of our own.

The policy for CERN computing is to participate, for Monte Carlo work and modelling, in the budget of the large users (mostly ATLAS). This situation will not change. For some flexibility and for general access (visitors!), we again require our own budget at the level of 300 hours CERN.

4. *Acknowledgements*

We owe many of our results to the effective collaboration with partners in industry and in other projects. We want to express our thanks to Jean Vuillemin, Mark Shand, Philippe Boucard and Laurent Moll from the Digital Paris Research Laboratory, without whose help we would not have a DecPerle demonstrator. The contribution of Gena Klyuchnikov (Protvino) to the DecPerle developments is also acknowledged. We are further grateful for the efficient help provided by Tim Bell (IBM) and Ben Segal (CN) in starting up the measurements on the fibre channel switch. We appreciate the good collaboration with Alberto Guglielmi (DEC Joint Project) in building up demonstrators and interfaces for our pilot implementations of Alpha processors and SCI. We also want to acknowledge the efficient and pleasant collaboration with members of ATLAS, RD6, RD13, RD24, RD31, GPMIMD-2 and AFRODITE.

We acknowledge further gratefully the support given to projects in EAST by the following funding agencies and institutes:

Bundesministerium für Forschung und Technologie (BMFT), Germany
Particle Physics and Astronomy Research Council (PPARC,
formerly part of SERC), United Kingdom
Fondamenteel Onderzoek der Materie (FOM), The Netherlands
The Polish State Committee for Scientific Research
The Grant Agency of the Czech Republic
Commission of the European Communities DG XIII (Esprit)
CERN
GSI Darmstadt

5. Publications, conference contributions, EAST Notes 1993/4

Publications and conference contributions:

H.M.A.Andree, G.T.Barkema, W.Lourens, A.Taal, J.C.Vermeulen: *A Comparison Study of Binary Feedforward Neural Networks and Digital Circuits,* Neural Networks 6 (1993) 785 - 790.

Klefenz F., Noffz K.-H., Zoz R., Männer R.: *ENABLE - A Systolic 2nd Level Trigger Processor for Track Finding and e/p Discrimination for ATLAS/LHC;* acc. for publ. in Proc. IEEE Nucl. Sci. Symp., San Francisco, CA (1993)

Männer R., Gläß J., Klefenz F., Kugel A., Noffz K.-H., Zoz R., Baur. R.: *Real-Time Pattern Recognition by Massively Parallel Systolic Processors;* in: Adams M., Eck C. Hilf W. et al. (Eds.): Proc. Open Bus Systems '93, München, Germany (1993) 298-296

J.M.Seixas, L.P.Caloba, M.N.Souza, A.L.Braga, A.P.Rodriguez, H.Gottschalk: *A Second-level Trigger System Based on Calorimeters and Using Neural Networks for Feature Extractionmand Electron/Jet Discrimination ;* presented at 'Calorimetry in High-energy Physics', Elba, September 93

J.Renner-Hansen: *FEAST: A possible second-level trigger for ATLAS;* presented at 'Calorimetry in High-energy Physics', Elba, September 93

Contributions to the Third International Workshop on **Software Engineering, Artificial Intelligence and Expert Systems** for High Energy and Nuclear Physics, Oberammergau, October 1993, "New Computing Techniques in Physics Research III", ed. K.-H Becks, D. Perret-Gallix, World Scientific, 1994:

H.M.A.Andree, W.Lourens, A.Taal and J.C.Vermeulen, *Feedforward Neural Networks as an On-Line Pattern Recognition Tool*

R.K.Bock, J.Carter, I.Legrand: *Real Time data-driven architectures simulated in concurrent C++ for the LHC second-level trigger*

J.M.Seixas, L.P.Caloba, M.N.Souza, A.L.Braga, A.P.Rodriguez, H.Gottschalk: *Neural Networks applied to a second-level trigger based on calorimeters*

J. Haveman, A.C. Balke, W.Lourens: *Formal Methods in the design of a second-level trigger*

J.C.Vermeulen: *Simulation of Data-Acquisition and Trigger Systems in C++*

Contributions to the 1994 Conference **Computing in High-Energy Physics,** San Francisco, April 1994 (CHEP94):

Noffz K.-H., Zoz R., Kugel A., Klefenz F., Männer R.: *Results of On-Line Tests of the ENABLE Prototype, a 2nd Level Trigger Processor for the TRD of ATLAS/LHC*

P.Wegner, U.Gensch, H.Leich: *A second-level trigger concept based on communicating digital signal processors*

J.Strong: *Local processing for a farm-based second-level trigger at LHC*

A.J.Borgers, F.B.T.Golbach, A.W.Lodder, W.Lourens, A.H.D.Ockeloen, A.J.M.de Vries: *Using a Ti Tms320c40 DSP-FPGA-combination based system for the implementation of a real-time feed-forward neural network*

W.Lourens, A.W.Lodder, A.Taal: *On the mapping strategy of a feed-forward neural network and a Bayesian classifier onto mesh-connected machines for a calorimeter pattern recognition task*

Z.Hajduk, W.Iwanski, K.Korcyl, J.Strong: *Modelling of local/global architectures for second-level triggers at LHC experiments*

J.M.Seixas, L.P.Caloba, A.L.Brage, E.Duarte: *Global decisions with a neural second-level trigger system*

W. Lourens, A.C. Balke, J. Haveman:*The formal developement method VDM++*

EAST Notes since the last Status Report:

EAST note 93-01 *Test data for the global second-level trigger* (R.K.Bock, J.Carter, I.C.Legrand, J.Varela) 28 January 1993

EAST note 93-02 *EAST collaboration meeting #8, Minutes* (R.K.Bock) 9 February 1993

EAST note 93-03 *Modelling of L2 global decision structures, Revision 1* (R.K.Bock, J.Carter, I.C.Legrand, M.Novak) 21 April 1993

EAST note 93-05 *The specification of a pipelined feature extractor in VDM* (J.Haveman) 4 March 1993. Also AFRODITE/UU/JH/DOC/V1

EAST note 93-06 *Results of Second Level Trigger Algorithm on Spacal Calorimeter Data using the Blitzen Parallel Machine* (S.Centro, E.W.Davis, P.Ni, D.Pascoli, E.Siliotto) March 1993 (also DFPD 93/EI/114)

EAST note 93-07 *HIPPI to TURBOchannel Interface* (T.Anguelov) 22 March 1993

EAST note 93-08 *Status Report EAST 1 May 1993* (also CERN/DRDC 93-12)

EAST note 93-09 *EAST collaboration meeting #9, Minutes* (R.K.Bock) 5 July 1993

EAST note 93-10 *Input data specification and other requirements for building regions of interest* (R.K.Bock and J.C.Vermeulen) 28 November 1993

EAST note 93-11 *The FEAST Project* (P.Clark et al) 1 February 1994

EAST note 93-12 *Implementation of the RoI task on the Intel i860* (Mats Jirstrand) 27 August 1993

EAST note 93-13 *The Router Problem* (P.Kammel, A.Reinsch, V.Dörsing) Rev.0, 10 October 1993

EAST note 93-14 *Data transfer rates using the communication ports of the Texas Instruments TMS320C40 digital signal processor* (A.J.Borgers, F.T.Golbach, and W.Lourens) Version 1.0, September 1993

EAST note 93-18 *EAST Collaboration meeting #10, Minutes* (R.K.Bock) 15 October 93

EAST note 93-19 *Notes on the FEAST meeting: Cracow, 1 Oct 93* (J.Strong) 18 October 1993

EAST note 93-21 *Data transport with the Texas Instruments TMS320C40 DSP* (J.C.Vermeulen) 18 October 1993

EAST note 93-22 *Simulation of Data-Acquisition and Trigger Systems in C++* (J.C.Vermeulen) 18 October 1993

EAST note 93-24 *A readout model for some ATLAS detectors* (R.K.Bock and J.C.Vermeulen) revision 1, 29 Nov 1993 (also ATLAS-DAQ #10)

EAST note 94-01 *Interactive communication and debugging tools for a multiple TMS320C40 system architecture* (A.Borgers et al.) January 1994

EAST note 94-02 *A TRT Algorithm Implementation suggested for the Beam Test in June 1994* (L.Lundheim) 4 February 1994 (rev.1)

EAST note 94-03 *Results of On-line tests of the Enable Prototype* (K.H.Noffz et al.) 3 February 1994

EAST note 94-04 *Experience with HIPPI from Beam tests* (S.Khabarov et al.) February 1994

EAST note 94-05 *Theory and Technical Implementation of a Neural Network with Switching Units* (P.Bitzan et al.) 5 February 1994

EAST note 94-06 *East Collaboration Meeting #11, Minutes* (R.K.Bock) 18 February 1994

EAST note 94-07 *Specification of second-level trigger systems for LHC experiments* (Ch.Balke, R.K.Bock, J.Carter, J.Haveman, I.Legrand) (Also AFRO/CERN/JC/DOC/V3, April 1994)

EAST note 94-08 *What can artificial neural networks do for the global second-level trigger* (R.K.Bock, J.Carter, I.C.Legrand) 22 March 1994 (also ATLAS DAQ #11)

EAST note 94-09 *Data Format for the TRT Prototype in 1994* (S.Khabarov and P.Lichard) 20 February 1994

EAST note 94-10 *A calorimeter feature extraction algorithm adapted for a DSP network running in a data-driven model* (I.Legrand) 20 March 1994

EAST note 94-11 *Algorithm benchmarks with Power PC processors* (F.Chantemargue) 6 April 1994 (draft)

EAST note 94-12 *Specification for the Transition Radiation Tracker on DECPeRLe1* (Jean E.Vuillemin) 1 April 1994

EAST note 94-13 *Second-level trigger feature extraction algorithms* (R.Hauser, I.C.Legrand) April 1994 (draft)

EAST note 94-14 *Discussion of a generalised programmable Router* (P.Kammel et al.) 28 April 1994 (draft)

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.) March 1994 (draft)

APPENDIX A: Basic concepts of level-2 triggers

EAST has developed the fundamental *Region-of-Interest (RoI) concept* for level-2 triggering, at least at high luminosity. This alleviates the very stringent bandwidth requirement on providing data for algorithms from distributed buffers, at high frequency. The RoI concept relies on the level-1 trigger to identify those parts of the detector containing candidate features (electrons, photons, muons, jets). Only a small fraction (of order a few %) of the data has then to be moved to the level-2 processors.

Modifications to the principle of RoI may become necessary for B-physics, where the example of low-p_t electrons may well result in a need for an unguided scan of a full detector. Presently, however, the B's triggering problem and the expected rates are not yet sufficiently understood for undertaking a serious study.

EAST studies solutions for implementations of algorithms *using full-granularity, full-precision* data on all characteristic detectors. The extent to which data with full precision are effectively needed in level 2, will have to be the subject of detailed simulation studies, and may even be dependent on the target physics. In general, reducing the requirements on precision in the trigger will result in architectural simplification and thus cost savings.

In all implementations, it must be a guiding principle that any proposed or demonstrated *partial solution* can be envisaged to *hold for the entire detector,* or at least for an entire detector component, including the usual constraints of flexibility, robustness, ease of control and maintenance, and readily embedded in the overall data acquisition system. As we are dealing with a market that evolves fast and is not driven by our application, we will have to ensure that technology improvements can be absorbed as easily as possible in the system, the limit being that some new technologies will, obviously, require architectural adaptation. These criteria translate into a maximum use of commercial components, standard interfaces, and in particular the introduction of as few different components as possible (viz. detector-independent solutions).

The following paragraphs outline the overall functional decomposition of the level-2 problem, into three phases.

Phase 1: Front-end buffering and collection of regions of interest

Raw detector data, after a level-1 (L1) trigger has occurred, are naturally transmitted via cables or fibres and collected in local non-overlapping memory modules (chips, boards, crates). The collection of these memories needs to hold full information over the duration of level-2 (L2) operation, the so-called latency. Data are accumulated according to a detector-specific granularity in the **L2 buffer.** It is a serious goal to commence the 'standardization' (viz. detector-independent hardware) for the entire data acquisition system with the L2 buffer.

The concept of RoI results in algorithms that convert local data from a subdetector into variables containing the relevant physics message ('features'). This conversion ('feature extraction') is achieved by an algorithm working on local subsets of data only as indicated by the results of L1, the RoI-s. For this data set to be available, two functions must be implemented:

- The data pertaining to regions of interest (RoI-s) have to be selected by some mechanism, which we term **'RoI collection'**. Detector-dependent differences will exist in the implementation, as substantial variety exists in the modularity of collecting data in L2 buffers. All intended implementations could make use of the RoI parallelism, i.e. deliver data for different RoI-s simultaneously. This parallelism does not, however, match directly the parallelism of readout or L2 buffers: RoI-s do, in general, extend across the boundaries of buffers.

- A device realized outside the L2 data stream has to indicate the whereabouts of regions of interest (RoI-s). This L1-guided unit is called a **'RoI-builder'** and 'drives' the RoI collection. Note that this is true even for RoI-s that do not themselves actively participate in the L1 decision, like lower-threshold calorimeter clusters.

We assume that RoI-s are collected independently for different subdetectors.

Phase 2: Feature extraction: Local processing of data in a RoI of a subdetector

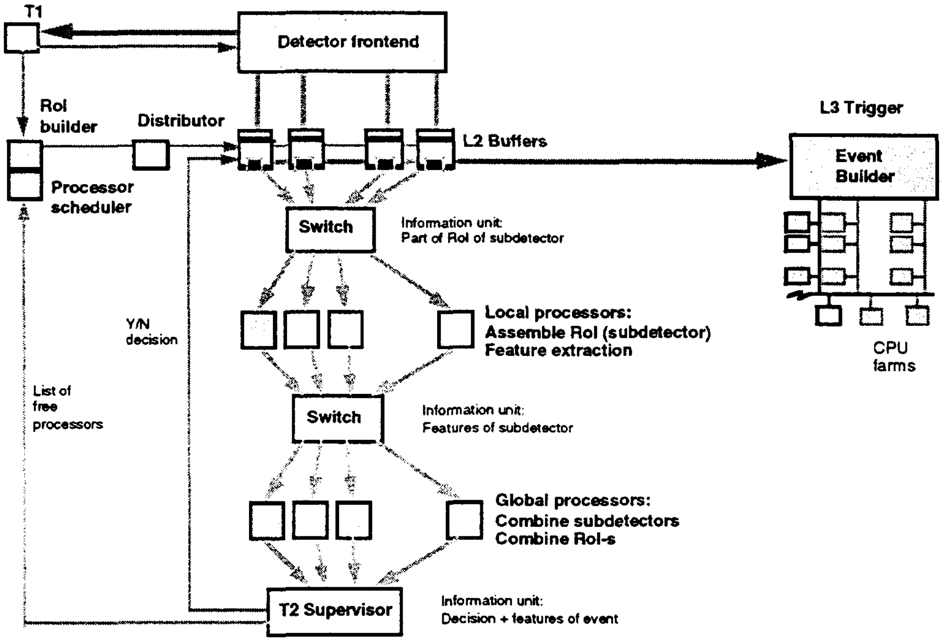
Algorithms in the concept of RoI have a clear and simple task: to convert a limited amount of local raw data from a single subdetector into 'features'. Features are variables containing the relevant physics message, like cluster or track parameters that can be used to corroborate or disprove the different physics hypotheses. This phase is called **'feature extraction'**, a term taken from image processing. Feature extraction algorithms have a locality that permit to exploit the natural double parallelism of RoI-s and subdetectors. Simulation will have to show if and to which extent this simple concept has to be diluted in order to avoid physics losses e.g. in regions of overlap of detector parts (e.g. barrel/end cap), where each subdetector only has a weak signal.

Phase 3: Global decision: RoI and event processing

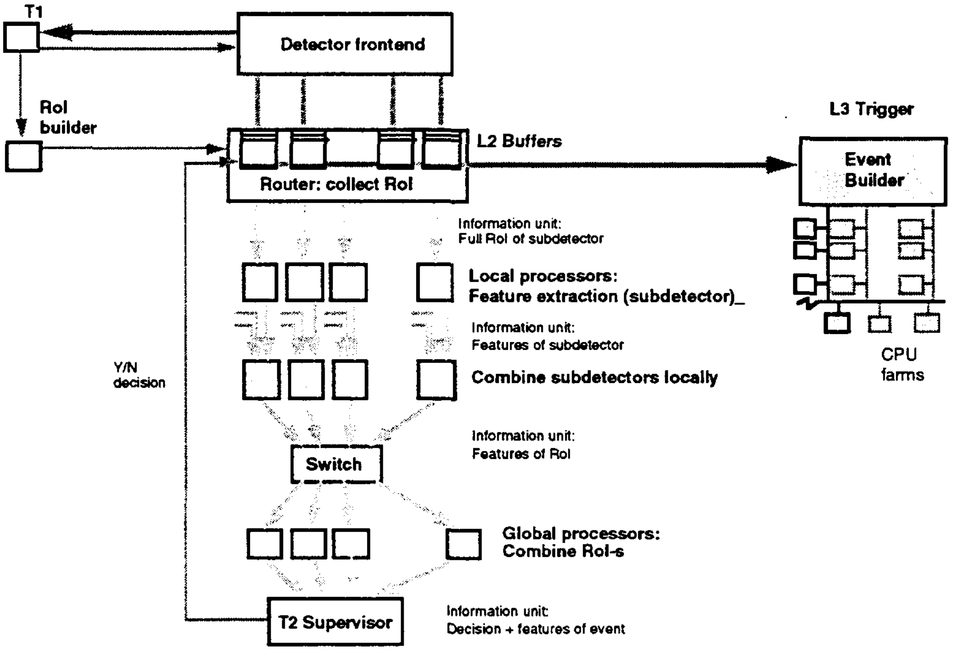
Once raw data have locally been converted to physics features, these have to be collected from all subdetectors and from all RoI-s, for forming an overall decision on the entire event. The bandwidth requirements are substantially reduced by feature extraction, so that implementations of this phase have less of a problem with data transmission, although the trigger frequency of 100 kHz is, of course, unchanged and constitutes a stumbling block for many high-level devices. It is quite apparent, if one decomposes further the part of the algorithm dealing now with multiple sets of features, that the natural and efficient order of processing is to combine first all subdetectors that have 'seen' the same physics 'object', into decision variables which are again local (same RoI), and follow this by combining all RoI-s into an event decision. While this seems the most inviting algorithm strategy, it is by no means necessary (although possible and possibly cost-effective) to express the strategy also by a corresponding implementation in separate and parallel processors. If this is the adopted architecture, one may want to take advantage of another natural parallelism, that of testing multiple physics hypotheses against the same set of data. That data set, i.e. all features from all RoI-s, amended possibly by quantities derived from them, must then be shared by all physics processes. This data-driven decision architecture has been modelled in EAST, but is not presently proposed for implementation, because of manpower limitations.

Implementation options

We have, so far, avoided discussing the hardware on which these algorithms or algorithm parts are implemented. In fact, basic choices exist that lead to quite different architectures. We will discuss here two overall options showing the basic choices in a pure way. Both of these are still under study and have specific advantages that seem to make it worthwhile studying them in more detail, before potentially discarding one for good reasons. We call them the farm-based and the data-driven architecture. A more detailed discussion of these options follows next. Note that practical implementations are unlikely to follow such a pure solution for all detectors; hybrid implementations are quite possible and potentially the most cost-effective overall.



The "**farm-based**" approach uses standard general-purpose processors and commercial network components. In the simplest scheme, there are two layers of processors performing local feature extraction and global feature combination respectively. Local processors receive their data from intelligent devices in the L2 buffer, through a switching network. The global processors are organized as a single general farm; local processors may be run as one farm per subdetector, but groups of processors may also be permanently assigned to regions of the detector.



The "**data-driven**" approach uses low-level devices for ROI collection (sometimes called 'routers'), directly coupled to (or inserted upstream of) the L2 buffers. Typically, these would be implemented as field-programmable gate arrays. Devices based on the same principle of low-level programmability, are also used as feature-extraction processors. Solutions based on field-programmable gate arrays seem to be adequate to satisfy the characteristic constraint applied to this solution, which is that processors operate in a pipelined mode capable of coping with the level-1 rate of 100 kHz, and hence obviate the necessity of event parallelism in a farm. The network/farm approach is, however, kept for the low-bandwidth problem of global decision making, as above.

We will now discuss the characteristics of these models in more detail.

RoI collection

In the farm-based approach, data arrive in the multiple L2 buffers and remain stored locally. They can be 'written' by an intelligent programmable device attached to the buffer, to the relevant feature extraction processor, provided the device knows which data and to which processor to send. Data can also be 'read' from the L2 buffer by the device implementing the feature extraction algorithm, if that has knowledge about the buffer addressing scheme. In both cases, the foreseeably large number of local buffers requires many potential connections to the feature extraction devices. A switching device of some generality, custom-made or commercial, becomes then a necessary and critical element. Bandwidth requirements for this device will be an important constraint.

In the data-driven hypothesis of implementation, good nearest-neighbor connectivity among buffers is assumed, so that specialized devices can extract and format, possibly preprocess data before sending them to feature extractors. To achieve this broad access to buffers may be difficult and lead to complications. In fact, access to the (deep and well-managed) L2 buffer is not mandatory, if data come sufficiently synchronized to pass through a shallow buffer with the connectivity needed for RoI collection.

Feature extraction

In the data-driven concept, feature extraction algorithms are executed on suitable hardware units in a pipelined way (viz. in functional decomposition), maintaining the imposed overall decision frequency of 100 kHz. This approach is particularly indicated if algorithm execution times depend little or not at all on the data content. If data determine to an important extent the algorithm execution time, then processors that keep up with overall frequencies of decision have to take 'worst-case' scenarios into account and/or smooth out the variations by buffering. Re-synchronizing will then become necessary.

In the farm-based approach, the execution of feature extraction algorithms is spread onto a farm of processors, which can be scheduled according to availability. A general switching network is then required which connects all possible sources (L2 buffers) to all possible processors (the number may be large). If numbers get too large, and the resulting communication bandwidth too high, geographical segmentation (inside a subdetector) can be envisaged.

Global decision

In both the data-driven and the farm-based concept, the processing after feature extraction is based on the hypothesis of farming. The argument has been made earlier: due to much reduced number of connections and lower bandwidth requirements after feature extraction, switching based on commercially available equipment seems definitely realistic. An added argument is that algorithms in this phase can be less rigorously defined before implementation, and are in need of the flexibility and scalability only offered by general-purpose processors.

APPENDIX B: Data-driven versus farm-based architecture

EAST has invested substantial efforts in demonstrating the feasibility of data-driven solutions for RoI collection and feature extraction. However, these solutions might not be optimal with respect to

flexibility
scalability
upgradability
commercial availability and maintenance
reconfigurability for new concepts, like RoI-free B-physics
constraints on synchronicity in data transmission

(the list is from an ATLAS discussion, but CMS bases its resolutely farm-oriented approach based on very much the same arguments). The wish list is, of course, the same for any implementation; the proposed solutions are different in interpreting the list. Questions that need much more studying are :what flexibility or scalability or upgradability are needed? at which price do we stop needing them? are components commercial when produced by a small company according to an experiment's specifications? Which performance degradation (e.g.rate) is acceptable if a trigger designed for high luminosity is also used for B-physics?

We will in the following present what knowledge is available, today, about the elements of farms in the various parts of the L2 trigger, based on available components, and will attempt to discuss possible extrapolations. We believe that we show that the commercial offerings today do not permit to differentiate solutions in a clear way, and extrapolations into the future are necessary. We do not believe, therefore, that the time has come to confront these quite different architectures, let alone discard one or the other. Instead, much more detailed understanding has to be acquired over the next two or three years, so that these economically relevant choices will be taken with all information available in the collaborations.

What information is available today?

EAST has taken its **benchmark** suite including both feature extraction and global decision algorithms, onto most of the market-leading **processors**, and measured execution times. These algorithms are conceptually optimized (e.g. making extensive use of look-up tables), and contain no system kernel overheads or interference with communication, nor the necessary 'service functions' like unpacking information or address manipulation (from 'local in readout module' to 'local in RoI' or 'global'). The results are preliminary and documented in recent EAST notes. Feature extraction algorithms execute typically in around a millisecond, first estimates for the neglected overheads are in the hundreds of microseconds. Global decisions take of the order of 200 microseconds on average; all times are given for the best available processor. No optimization to specific processor architectures was attempted. Note that for feature extraction, results are for a single RoI and subdetector combination; an added 100 microseconds in execution time translate into an additional 200 processors, if we assume four subdetectors and five RoI-s on average, at a 100 kHz event rate.

On **communication and switching** devices, theoretical and measured information on commercial equipment is available for HIPPI, used for transmission and switching between a few ports. Of the theoretically possible 100 Mbytes/s of HIPPI, 70 to 80% have been realized for transmission and over the switch (reported by NA48). In a different real-life implementation (RD13's data acquisition system in the H8 beam), 30 Mbytes/s have been quoted as limit for large blocks and sustained connection. This degradation is largely due to the use of a VME-programmed processor which takes data from the HIPPI destination to a local buffer. HIPPI does not look like a possible candidate due to the bulky 32-bit-parallel cables limited to 25m length.

As mentioned above, a fibre channel (FCS) switch with interfaces has also been made available recently (but no results meaningful for our context were extracted yet). From its specifications, there is likely to be a switching latency problem, at 100 kHz (switching latency limits frequency, as opposed to transmission latency, which is of little concern). Early transmission results exist also for the prototype chips of SCI (from RD24), and for one ATM realization (NetComm DV2). The NetComm switch runs at a theoretical maximum of 12.5 Mbytes/s, of which a (strongly packet-size dependent) fraction only is achieved. Interpreting the measured NetComm rates for small packet size, a switching latency in excess of 100 microseconds and substantial processor overhead may be inferred, but better understanding and more measurements are required. Hardware for ATM switches up to about 20 Mbytes/s have been realized and will be available soon. In principle, the standard also defines higher bandwidth options, but manufacturers have not realized those. Both SCI and ATM for faster theoretical speed have been simulated in quite some detail.

Needed for RoI collection into feature extractors are typically multiple blocks that together constitute up to 1 kbyte of data, to be switched into each processor serving a RoI/subdetector combination. Their spread into geographical areas and in time has not been defined, and will need careful simulation. This modelling activity is about to start in ATLAS. It is probably not unfair to state that none of the devices available commercially today can with confidence be called a candidate for a viable switch mediating RoI collection. Even for the global decision task a clear statement can not be made on the suitability of existing commercial devices. After feature extraction, we deal with fewer ports, and substantially smaller data packets, but the required frequency remains at 100 kHz and poses potential problems to switching components.

What extrapolation is possible to the future?

On the processor side, there are decades of development both of architecture and semiconductors to extrapolate from. Assuming we are allowed to continue the exponential curve (a factor two every two to three years, for no increase in Si size or price), and that algorithm optimization will contribute, then we can be confident that the thousands of processors that seem necessary today, will reduce to a number below one thousand, and thus become quite manageable. This evolution of capacity/price has been driven by the explosive use of PC-s and workstations in private homes and in engineering applications. Note that the argument refers to bare chips, whereas board products have not evolved in price the same way, and are not readily available commercially. Given the demands on the high-volume markets, that situation may remain unchanged in the future.

Our problem, however, is not primarily one of compute power, but one of communication (as an aside: statements of the same content are being made, perhaps surprisingly, in the high-performance (parallel) computing and networking community today). An evolution nearing the performance curve of processors can not be predicted easily for communication and processor interfaces: there is not much past to extrapolate from. Although it is not excluded that the introduction of data superhighways and high-definition television will leave us with Telecom-oriented products that satisfy our requirements, such extrapolation can not be made with confidence today. It is quite likely that the switching requirement for RoI collection at 100 kHz is beyond the requirements in the future Telecom market, primarily due to different requirements in latency, flow control, and error correction. Our specifications could be much closer to the switching capabilities that are built into modern parallel supercomputers (HPCN systems), which are neither open (accessible to any vendor) nor easily available commercially. Interfacing them to a large number of processors may remain prohibitive in cost. It should also be noted that for the final LHC systems, the numbers of boards (processors, interfaces, switches) are such that they will have to be mass-produced by a company, regardless if designed by a collaboration or taken from commercial designs. This 'commerciality' will, of course, be accessible to components of any architecture, and is not restricted to farm-based solutions.

And digital signal processors?

The above discussion was kept very polarised to general-purpose processors and switches on the one, field-programmable gate arrays on the other side. We have not done enough justice to the combined power inherent in digital signal processors (DSP-s), which have typically compute performances approaching that of a RISC processor, with added and properly integrated communication capability. DSP-s seem to develop their characteristic performances at least as fast as RISC-s; they can be arranged into farms, but also into pipelined architectures, and their links allow constructing intelligent networks from them. They can constitute an ideal intermediate link between the 32-bit general-purpose processor and FPGA-s.

In EAST, several groups cooperate to explore, how the switching mechanism could be mediated by a network of DSP-s. Other activities concern their linking to FPGA-s as co-processors, and an interface to a (long-distance) optical link. The groups involved mostly have past experience with transputers, an early device that combined computing and communication. If today's favorite model is the TMS320C40, tomorrow interest may focus on the AD21060, announced for July 94, or on even newer products.

Presently, no simple and cheap general solution has been proposed based on DSP-s, but the EAST activities and milestones including DSP developments intend to make a major and generic contribution to understanding the eventual role of these devices in the L2 trigger. In particular, they may constitute the necessary bridge between 'farm-based' and 'data-driven', allowing to incorporate the best features of both. Efficient interplay between communication and processing in DSP-s needs thorough understanding of the device's architecture and internal pipelines (see several EAST notes on the C40). Superficially, DSP-s can be seen as devices that present difficulties of implementation somewhere in between the 'easy' general-purpose processors (except that in practice also these will be optimized 'to the bone', in all likelihood), and the 'difficult' lower-level devices like programmable gate arrays, from which data-driven devices are made.

What can we conclude?

If we accept that the extrapolation for 'affordable, robust and commercial' communication and switching equipment will leave us with a possibility that the Telecom market offers only low-bandwidth and modest-reliability (because devoid of flow control) devices, and HPCN proprietary switches remain confined to efficient general-purpose message passing in a specific machine, then it is important to continue exploring alternative avenues and to build up further the potentially relevant know-how. The time for confronting the possible architectural choices has simply not come.

EAST has explored and proposes to continue exploring and generalizing data-driven processing; this work must be completed, documented, and the know-how must remain available to the high-energy physics community. EAST further proposes to continue spending effort for good understanding and implementing DSP-based switching and processing. EAST finally proposes to step up its activities on farm-based solutions, jointly with other interested groups and with the more specialized projects RD24 and RD31, staying as close as possible to market developments.