The charged trigger system of NA48 at CERN

S. Anvar, F. Bugeon, P. Debu, J.L. Fallou, H. Le Provost, F. Louis, M. Mur,
S. Schanne, G. Tarte, B. Vallage*

DAPNIA, CEA Saclay, F-91191 Gif-sur-Yvette Cedex, France

Abstract

The NA48 charged trigger is a mixed hardware and software real-time processing system intended to detect the interesting configurations of K^0 charged decays. It achieves real-time event building, track reconstruction and kinematics computation on drift chamber data at an event rate of 100 kHz and within a maximum decision latency of 100 μ s. The system uses data driven, FPGA-based coordinate builders, a hardware event builder based on a crossbar switch, and a farm of up to 16 event processors for its software part. It has been installed and operated at CERN since 1995. After a description of the constraints and architecture of the various subsystems, the paper gives an account of the results and performance of the system based on the 1996/1997 runs. More specifically, the replacement of the present DSP-based implementation of the processing farm by RISC processors will be discussed. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

NA48 is a particle physics experiment aiming at the precise measurement of direct CP violation in K^0 decays into 2 pions [1]. The experiment comprises two main detectors: a calorimeter for neutral decays ($K \rightarrow \pi^0 \pi^0$) and a spectrometer for charged ones ($K \rightarrow \pi^+ \pi^-$). The spectrometer consists of four drift chambers and a magnet. The hits left by the charged particles in the drift chambers allow the system to reconstruct their trajectories (tracks), and the track deflections caused by the magnet lead to the kinetic parameters of the decays.

The primary rate of charged decays is about 200 kHz. However, among all possible charged decays, only $K \rightarrow \pi^+ \pi^-$ is relevant for the proposed measurement and is considered as the signal (good events) the others being therefore considered as background. The raw signal/background ratio produced by the front end data acquisition system is of the order of 10^{-4} . The function of the so-called charged trigger is to achieve, in real time, a substantial increase of the signal/background ratio by eliminating as much background data as possible, in order to improve the statistical significance of the recorded data and so reduce the bandwidth and storage needs of the experiment.

As shown in Fig. 1, the charged trigger system is made of two subsystems: the level 1 trigger (L1C), which reduces the primary event rate down to

*Corresponding author.

Fig. 1. The charged trigger inside the experiment.

100 kHz, and the level 2 trigger (L2C) which reduces that rate down to about 2 kHz. The L1C is a fast logic trigger, based on several simple criteria, which achieves a first selection of the charged events data and injects them into the L2C. The L2C is a parallel processing system mixing hardware and software elements; for each event, it computes the coordinates of the particle in the drift chambers, reconstructs tracks, calculates the kinetics and flags the event as signal or background. This paper describes the salient features of the L2C.

2. Basic concepts

2.1. The charged detection principles

The axis of the Kaon beam defines the z -axis of the coordinate system of the experiment. Each drift chamber (DCH) has 8 parallel sense wire planes, all perpendicular to the z -axis. They are grouped by staggered pairs to form 4 coordinate views – X, Y, U, V VIEW's – (Fig. 2). Each VIEW is essentially made of 512 parallel wires perpendicular to

Fig. 2. $K \rightarrow \pi\pi$ decay in the spectrometer elements used by the L2C (not to scale).

the axis of the coordinate it measures. Whenever a charged particle crosses a VIEW, it necessarily goes through two neighboring wires, leaving an electric pulse on each. The coordinate of the crossing point (also called space-point) is computed by combining the coordinates of the wire pair with an analysis of the timing difference between the two pulses.

In principle, two VIEWs (x and y) would be enough to determine the space-point of a particle inside a chamber. But since we are interested in *pairs* of particles ($K \rightarrow \pi^+\pi^-$), a typical event comprises a pair of x coordinates and a pair of y coordinates; a *third* coordinate ($u = (x + y)/2$) corresponding to a fixed linear combination of x and y is then needed to determine which x is associated to which y . Finally, since each VIEW has an inefficiency close to 1%, a fourth plane ($v = (y - x)/2$) is added to each DCH to improve the overall trigger efficiency.

The L2C uses all the hits produced in DCH1, DCH2 and DCH4 to compute the coordinates, tracks and kinetics of each event it receives.

2.2. The trigger principles

The whole NA48 experiment is synchronized by a 40 MHz clock, that is, time is defined for all

subsystems by a reference clock in units of 25 ns. All data are associated with a Time Stamp (TS) value, representing the time at which the corresponding event (decay) was detected in the charged hodoscope at a given z . For the system, an event is a set of data detected around a TS. When the L1C spots a candidate event around a particular TS, it prompts the DCH readout system to send the corresponding data to the L2C. The L2C processes that event and produces a single Trigger Word (TW) containing the TS and several flags which summarize the main characteristics of the event. The TW is then transmitted to the trigger supervisor of the experiment which, depending on the trigger patterns coming from all trigger systems, decides to readout the corresponding data or not.

3. The L2C requirements

NA48 being a precise measurement experiment, all biases introduced in the acquired data by the instrumentation (especially triggers) must be known to a high degree of precision ($\sim 10^{-4}$).

The L2C initial requirement was to meet an input event rate of 100 kHz. It must also process events within a latency of about 100 μ s: if the processing of an event takes more than 100 μ s, that event is flagged as Not Computed and the system discards it. Therefore, the rate of Not Computed events must be sufficiently low so that not more than a few percent of good events are lost. The relative inefficiency of the system – that is the loss of good events – for the two types of Kaon decaying (K_S and K_L) must be known with a precision of below 10^{-3} .

The L2C is an asynchronous queued system. Since the statistical distribution of events in time is Poissonian, the queues will fluctuate around an average value, provided that the average processing rate is faster than the event rate [2]. The actual processing latency of an event depends on its intrinsic complexity and on the time lost waiting in queues. Therefore, reducing queues increases the system efficiency. In order to control the queuing levels, the L2C includes an X_{on}/X_{off} mechanism: when queues reach a critical level, the L2C asserts an L1OFF signal, turning off the L1C; once the

queues are absorbed by the system, the L2C releases the L1OFF, allowing the L1C to resume its activity. For all practical purposes, this mechanism reduces inefficiency by increasing dead time, which is preferable since statistical biases due to dead time are better controlled. However, the dead time generated by this mechanism must not exceed 1%.

The L2C must be sufficiently flexible to allow for small algorithmic changes. It must be scalable in order to adapt to rate modifications. It also has to provide all tools and probes needed for the optimization and fine-tuning of the overall system.

4. The L2C core structure

As shown in Fig. 3, the L2C system is made of four subsystems: the Coordinate Builders (CB), the Event Builder and Dispatcher (EBD), the Event Workers (EW) and the Event Worker Farm Manager.

4.1. The coordinate builders (CB)

Each VIEW has its own CB board, in charge of producing coordinates by analyzing wire hits

Fig. 3. The level 2 charged trigger system.

timings. The raw data are injected by the DCH readout system into the input FIFOs of the CB. These data are then processed through a 40 MHz pipelined algorithm implemented on firmware (Xilinx): all neighboring wire hits are matched and, for each pair of hits, the event time is subtracted from the hit times yielding two drift times which are then used as indexes in a 2-D lookup table to retrieve the corresponding coordinate value. The implemented version of this algorithm exhibits $O(3n)$ dependency (where n is data multiplicity) whereas the original nested loops algorithm is $O(n(n+2))$. For a typical event, the processing latency in the CB is 1.2 μ s, which allows for an event rate of about 800 kHz. All computed coordinates are then sent through an optical link (Fiber Channel 266 Mbits/s) to the Event Builder and Dispatcher – see next section – in variable-length packets.

4.2. The event builder and dispatcher (EBD)

Like the CBs, the EBD's design is based on firmware (Altera). It is a programmable switch with 16 input ports, 16 output ports and 16 internal data channels. A specific additional input FIFO contains the current pool of available output port id's. The event structure can be programmed as a fixed subset of the Input Ports and a number of packets per port. For each event, the EBD reads the id of the next available Output Port, then reads and serializes the data packets from the specified Input Ports in order and writes them in the Output Ports. Up to 16 concurrent transfers can be active in the switch, which brings the EBD's theoretical internal capability to a throughput of ~ 400 Mbytes/s.

Since the L2C uses the data coming from DCH 1, 2 and 4, a complete event is made of 12 planes. We therefore use 12 of the 16 Input Ports and consider 1 packet of coordinates per port, each produced by a particular CB. The EBD builds the whole event by putting together the 12 coordinate packets and sends them to an Event Worker (EW) for further processing. When the EW associated to a EBD output port finishes the processing of an event, it sends back the port's id into the available Output Port FIFO to inform the EBD that the port is available. This setup would allow for a sustained

event rate of ~ 700 kHz, if the EWs were able to absorb it.

4.3. The event worker (EW) farm

The function of an EW is to receive the 12 coordinate packets of an event and use them to compute the particle space-points, tracks and magnetic deflections in order to determine whether the event is compatible with a $K \rightarrow \pi^+ \pi^-$ decay. The principle of the algorithm is to assume that the event is a $K \rightarrow \pi^+ \pi^-$ decay and compute the corresponding invariant mass; if the calculated mass is not compatible with the Kaon mass, the event is flagged as background.

A major complication, mainly due to the high luminosity of the Kaon beam, is that many events are not clean 2-particle events:¹ depending on various conditions, up to 20% of events are liable to have a coordinate count greater than 2, which introduces a combinatorial problem both in the space-points computation (XYUV) and in the mass computation (MASS) algorithms. For instance, if we have n (resp. p) space-points in DCH1 (resp. DCH2), the number of possible pairs of tracks is $n(n-1)p(p-1)/2$, which means that the number of combinations in the MASS algorithm increases like n^4 , where n is the number of particles per event. This is the main reason for introducing MIMD (Multiple Instructions Multiple Data) parallelism inside the present implementation of EWs.

Presently, an EW is a cluster of 4 totally connected TMS320C40 (C40) DSPs [3] by Texas Instruments [4]. The C40 DSPs are implemented on industrial octal boards by MIZAR [5].

As represented in Fig. 4, DSP1 receives the data and dispatches the packets of DCH2 (resp. DCH4) to DSP2 (resp. DSP4) keeping the packets of DCH1. Then, the three begin XYUV computation in order to associate x and y coordinates and produce the space-points of each DCH. DSP1 and DSP2 share their results between themselves and DSP3. DSP3 and, a little later, DSP1 and DSP2

¹ e.g. overlay of stray particles accidentally crossing the detector at nearly the same time.

Fig. 4. Data flow between the DSPs of one EW.

analyze together all possible 2-tracks combinations, eliminating those which – either physically or geometrically – are not compatible with an interesting $K \rightarrow \pi^+ \pi^-$ decay. All three send compatible track pairs to DSP4 which, upon reception, extrapolates the tracks to DCH4, spots possible deflections by comparing the extrapolated space-points to DCH4 data and computes the associated invariant mass. If it finds *one* mass compatible with the Kaon mass, it flags the event as signal. The Trigger Word bears this flag as well as other bits summarizing the characteristics of the event. When DSP4 finishes all its computations, it sends the Trigger Word to the Farm Manager and then proceeds with various table purges and housekeeping tasks. When that is done, the EW signals that it is available for another event by sending (through DSP3) its associated Output Port id to the Farm Manager which, in turn, writes it to the available Output Port FIFO of the Event Builder and Dispatcher.

Fig. 5 shows the chronological processing distribution between the 4 DSPs, in a case where there are 3 (resp. 3, 5) particles crossing DCH1 (resp. DCH2, DCH4).

4.4. The event worker farm manager

The main functions of the Farm Manager are:

1. to receive the Trigger Words from the EW's Trigger Word daisy chain and send them to the experiment's Trigger Supervisor,

Fig. 5. Chronology of the various processes in the four DSPs of one EW.

2. to receive the Output Port ids of available EWs from the corresponding daisy chain and write them back into the available Output Port FIFO,
3. to manage the L1C X_{on}/X_{off} mechanism which regulates queue levels.

It also implements specific signaling functions for system debugging and emulation.² The Farm Manager is firmware-based and controlled by a Quad MIZAR board [5] through C40 communication links. The firmware has recently been reprogrammed in order to implement a hardware control of the X_{on}/X_{off} mechanism, allowing the system to stand event rates well over 100 kHz.

5. The L2C software and control

The L2C is installed on 4 VME crates and one workstation over distances of up to 100m: 3 crates for the CBs associated to DCH 1, 2 and 4, and one

² For debugging and fine-tuning purposes, the L2C system is equipped with a full emulation mode which allows it to run in standalone using either real or Monte-Carlo data at a maximum rate of 200 kHz.

crate containing the EBD, the EWs and the Farm Manager. Each crate is controlled by a SPARC VME SBC computer running UNIX. The whole system is monitored by a Sun workstation through a private Ethernet network. The core software of the L2C consists in the DSP programs and the libraries developed for the control of specific VME boards. In order to maximize performance, the DSPs do not run any multitasking system. A higher-level distributed program achieves the synchronization and monitoring of the whole setup.

5.1. The core software

The Coordinate Builder's and the Event Builder and Dispatcher have both a slave VME interface which allows for their full configuration and setup. Complete libraries have been developed in C. A complete *IO* and communication software has also been developed for the DSPs so that engineers and physicists may easily program and fine-tune the MIMD code running on the EW and Farm Manager boards without suffering performance losses. The EW code, especially, has been designed through an in-depth algorithmic analysis making a wide use of the C40 communications and computing features in order to optimize the parallel structure of the EW cluster (Figs. 4 and 5).

5.2. The control software

The control software implements various functions, namely configuration, process synchronization, online monitoring and communication with the NA48 experiment's run control program. It uses ISIS, an off-the-shelf communication software [6] which allows for an easy development of distributed software. The visualization modules are based on PAW [7], a powerful data analysis toolkit developed by CERN.

The running conditions of the NA48 experiment depend on the CERN-SPS proton accelerator. Two phases alternate during the run: burst and interburst. The burst phase corresponds to the actual release of the Kaon beam and lasts about 2 s. The interburst phase corresponds to the filling and acceleration phase inside the circular accelerator and lasts about 12 s. Therefore, the actual acquisi-

tion and triggering takes place during the burst while the interburst is used to reinitialize the L2C components and retrieve monitoring data from the system. The Event Workers, the Farm Manager and Coordinate Builders accumulate the monitoring data in on-board static RAM or FIFOs during the burst and transmit them to the control software during the interburst.

6. Some results

The first run of the L2C took place in 1995 in a reduced configuration, allowing a thorough debugging and setting up. In 1997, the full system ran, complying with all the requirements, the results being even better than predicted. The online K^0 mass resolution is 5 MeV/c² (Fig. 6) and the rejection power close to 60.

The main constraint in the L2C system is the strict latency limit of 100 μ s beyond which an event is lost (cf. Section 3). The overall average latency is presently of 80 μ s, which is in agreement with design estimations. The latency fluctuations, due to complex events and queuing, account for a loss of 1% of candidate events (Fig. 7).

There are currently eight EWs in the L2C, of which four are based on 40 MHz and four on 50 MHz C40s. This setup stands event rates of 70 kHz with less than 1% dead time.

7. Projected upgrade of the EWs

7.1. New requirements for the L2C

The NA48 collaboration has requested that the charged trigger be able to stand an increased event rate of about 150–200 kHz. At such a rate, the present implementation would see its dead time and event loss increase to prohibitive values. Meanwhile, we have studied the possibility of replacing the Event Workers based on DSP clusters by mono-processor EWs based on state of the art general purpose RISCs: the straightforward, sequential software of a mono-processor EW would indeed be much more easy to maintain than the

Fig. 6. Online $\pi\pi$ mass resolution seen in a K_S beam.

present MIMD code as it would get rid of much communication and synchronization code.

7.2. Benchmarks

We have run the L2C core software on various processors and obtained very encouraging results. The processors we tested were a 147 MHz Ultra Sparc, a 133 MHz PowerPC 604, a 200 MHz Alpha and a 200 MHz PowerPC604. The latter holds the best results since it displays a mean computing time of $\sim 12\mu\text{s}$. From our present measurements, we estimate that the time required to transfer the event data to the processor memory is $\sim 20\mu\text{s}$ at worst. This brings the mean processing time per event to $\sim 32\mu\text{s}$ for the 200 MHz PowerPC, to be com-

pared to the $\sim 80\mu\text{s}$ mean processing time achieved by the 4-DSP cluster.

7.3. RISC versus DSP

This tremendous leap in performance can be explained by the fact that the architecture of present-day general purpose processors is particularly adapted to the L2C algorithm. The amount of data per event is very small (typically ~ 160 bytes) while the track-reconstruction algorithm is relatively complex (though not bulky). The first access to the data by the processor practically loads the whole event into the cache of the processor and then the algorithm proceeds at full speed. This also entails that any increase of clock frequency

Fig. 7. Processing time distribution for $K \rightarrow \pi\pi$ decays.

on a processor means a gain of the same proportion in computing time.

7.4. Implementation inside the existing system

Presently, the communication between the EWs and the rest of the system goes either through C40 links for fast data transfers or VME for monitoring and control. Any upgrade of the EWs requires a hardware development, which is to interface the new RISC boards with the EBD and the Farm Manager through the C40 link protocol, and a software change, which involves the porting of the DSP

Fig. 8. Data transfers through the PMC.

software on a RISC platform and the development of the drivers of the hardware interface.

Since most of today's RISC VME boards are made to receive at least one PCI mezzanine card (PMC), the most straightforward way of interfacing the mono-processor EWs with the rest of the system is to develop a specific PMC that will handle the different data transfers on firmware, using the C40 link protocol (Fig. 8).

It will also have to take on some low-level functions (e.g. timers) which are present on the DSP chips but will not be available on the RISC processor.

8. Conclusion

The L2C is a successful implementation of a high rate, software, scalable trigger system. Its strong features include a small scale, efficient and compact switched-based event builder, as well as a scalable processing farm of commercial CPUs which allows for easy upgrades.

It is to be noted that capabilities for standalone emulation and testing are crucial for the development of such a complex system: without these features the integration of the system would have proven much more difficult.

The importance of control software should also be underlined: monitoring a 100 kHz or more event processing system requires a strong software backend which has to be robust.

The NA48 charged trigger may be considered as a small example of what will have to be done for future LHC experiments: software trigger systems, statistical performance, massive parallelism, commercial electronics. The experience acquired through the design of this particular system is precious for the design of such future systems.

References

- [1] Proposal for a Precision Measurement of ε in CP Violating K^0 to 2π Decays, CERN/SPS/90-22 SPSC/P253, 1990.
- [2] James Martin, System Analysis for Data Transmission, Prentice-Hall Inc., Englewood Cliffs, NJ, 1972.
- [3] TMS320C40 User's Guide by Texas Instruments.
- [4] Texas Instruments Inc., 8505 Forest Lane, M/S 8671, Dallas, Texas 75243, USA.
- [5] MIZAR, 2410 Luna Road Carrollton, TX 75006, USA.
- [6] Isis Distributed Systems, Inc. 111 South Cayuga Street, Ithaca, NY 14850, USA.
- [7] CN/ASD Group, PAW Users Guide, Program Library Q121, CERN, October 1993.