THE PROCESSOR FARM FOR ONLINE TRIGGERING AND FULL EVENT RECONSTRUCTION OF THE HERA-B EXPERIMENT AT HERA

A. Gellrich, R. Dippel, U. Gensch, R. Kowallik, I.C. Legrand, H. Leich, F. Sun, and P. Wegner

DESY-Institut für Hochenergiephysik, Zeuthen, Germany

The main goal of the HERA-B experiment which will start taking data in 1998 is to study CP violation in B decays. This article describes the concept and the planned implementation of a multi-processor system, called processor farm, as the last part of the data acquisition and trigger system of the HERA-B experiment. The third level trigger task and a full online event reconstruction will be performed on this processor farm, consisting of more then 100 powerful RISC processors which are based on commercial hardware boards. The controlling will be done by a real-time operating system which provides a software development environment, including FORTRAN and C compilers.

1 Introduction

The HERA-B experiment at the 820 GeV proton ring of HERA at DESY in Hamburg is dedicated to B physics[1]. The B mesons are produced in the interaction of the proton beam with the nuclei of an internal target in the beam halo. The main goal of HERA-B is to study CP violation in the decay channel $B^0, \bar{B}^0 \to J/\Psi K_s^0$ with the subsequent decays of the J/Ψ into two leptons and of the K_s^0 into two charged pions. In the standard model CP violation originates from the interference of two complex amplitudes: The amplitude of the direct decay $B^0 \to J/\Psi K_s^0$ and the amplitude of the decay through mixing $B^0 \to \bar{B}^0 \to J/\Psi K_s^0$. Since $J/\Psi K_s^0$ is a CP eigenstate the original flavor of the b quark is unknown. It has to be determined by tagging the second b quark in the event through its semileptonic decay. The accuracy of measuring $\sin 2\beta$, a measure of the CP violation, depends very much on the $b\bar{b}$ cross section. Varying the total cross section and including trigger efficiencies, track finding efficiencies, the efficiencies of cuts to suppress background, and dilution due to mistagging, an accuracy for $\sin 2\beta$ down to 0.11 can be achieved in one year of running. The rate of $b\bar{b}$ and $c\bar{c}$ events at HERA-B is of the order of 1 Hz. The rate of unavoidable background is around 5 Hz. These events have to be extracted from an initial event rate of 10 MHz, corresponding to the HERA bunch crossing rate, with on average 4 interactions per event.

The HERA-B detector is a forward spectrometer which is designed to tag particles, flying in a narrow cone in the proton direction. The main part is a tracking system before, inside, and behind a magnet which allows to reconstruct secondary decay vertices and the momenta of charged particles even in areas of high particle densities ($50-100$ tracks/bunch crossing). The trigger is mainly based on the two leptons of the J/Ψ decay. The high transverse momenta p_T, the dilepton invariant mass $m_{l\bar{l}}$, and the dilepton vertex are used to reduce background. The mean decay length of the B^0 mesons is 9 mm. A typical event is shown in figure 1.

The HERA-B trigger system contains four levels (L1-L4). For time reasons the first and second level trigger systems can only access parts of the full event data. The reduction factors, necessary to store events for an offline analysis, can be achieved only by reconstructing the leptons of the J/Ψ decays using the entire tracking system of the detector with an optimal calibration. Therefore the third level trigger task will be performed on a multi-processor system which can make use of the full event data and of calibration constants. Due to the high event rates and the long running period of the experiment a complete event reconstruction and an almost final event selection is required to be performed online. The fourth level trigger task, running on the same multi-processor system, will perform the full reconstruction of the events online before storing them on a mass storage device. In table 1 the input rates, the average decision times, and the reduction factors for the four trigger levels are listed. It is planned to have the HERA-B detector fully operational in 1998. The aim of this paper is to describe the concept and the planned implementation of a multi-processor system called the processor farm.

Level	Input rate	Time	Reduction	Method of the trigger algorithm
L1	10 MHz	12 μs	200	lepton p_T, simple tracking, $m_{l\bar{l}}$
L2	50 kHz	20 μs	$25-50$	refitting, magnet tracking, vertexing
L3	$1-2$ kHz	30 ms	$20-40$	using calibration & full detector
L4	50 Hz	2 s	1	full event reconstruction

Table 1: Average input rates, decision times, and reduction factors of the four trigger levels.

2 The concept of the processor farm

In the current scheme which is based on a proper simulation of a real-time multi-processor system[2] the processor farm consists of more then 100 nodes of high performance processors. It is foreseen to distribute the data from the buffer of the second level trigger via a switching network to the nodes of the farm. Before the full reconstruction will be done for an event the third level trigger step will be performed. It will work with the full event data and will make use of the calibration constants from the data base. On average 30 ms processing time per event at an input rate of $1-2$ kHz is needed to achieve a reduction factor of $20-40$ to a rate of 50 Hz. Events which have not been rejected will then be fully reconstructed on the same node. The average time needed for the reconstruction will be 2 s. To handle events with event sizes, on average 150 kB, and processing times distributed widely, buffering of at least one more events on the nodes and a flexible high speed network for the system is necessary.

The tasks to be performed by the processor farm can be divided into controlling and processing tasks. These tasks were defined with the help of a simulation on an IBM SP2 machine, using the message passing system PVM. Processing tasks are the programs for the third level trigger and the event reconstruction and include in addition server tasks, such as delivering and decompressing event data, and fetching calibration constants from a data base. The whole system is controlled by a supervisor task that keeps track of the data requests of the processing tasks and a task to monitor the performance of the system.

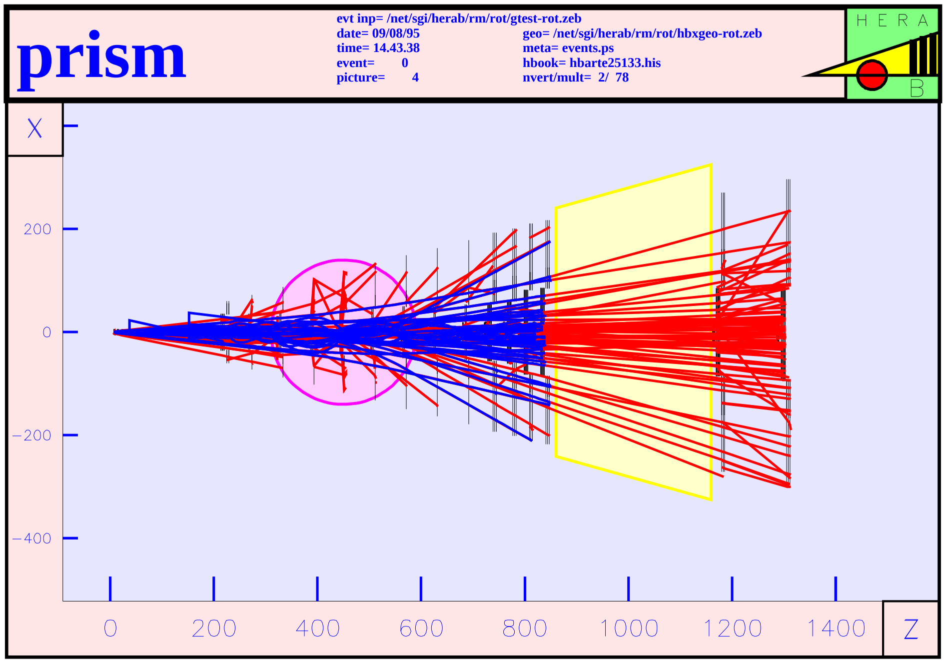

Figure 1: A typical HERA-B event with tracks reconstructed in the tracking system

All offline programs like Monte Carlo event generation, detector simulation, event reconstruction, and physics analysis are embedded in a frame program which takes care of the I/O and, in the offline case, allows interactive work. The frame program is written in FORTRAN and makes extensively use of CERN packages like ZEBRA and HBOOK. To keep the software used within the HERA-B experiment homogeneous it is foreseen to use main parts of the offline programs on the processor farm. An online version of the frame program, extracted automatically from the offline code with the help of a code manager, will be used on the processor farm. During shut down periods of the experiment without any data taking, the processor farm might serve as a production system for Monte Carlo events. For that purpose software packages like GEANT, PYTHIA, LUND, etc. are needed. All these requirements have to be met by the software development system running on the farm. It has to contain a FORTRAN compiler and must be able to handle the CERN libraries. A UNIX-like real-time operating system running on a host workstation and a simple kernel for embedded systems, controlling the running of the programs on the processor nodes, may be used for this purpose. A scheme of the data and control flow is sketched in figure 2.

3 The implementation of the processor farm

The consideration of different scalable and flexible real-time multi-processor architectures lead to a system of embedded powerful RISC processors based on com-

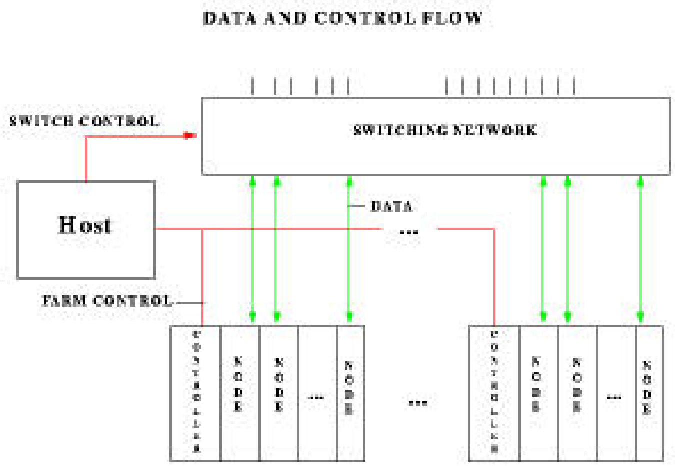

Figure 2: The event and data flow of the HERA-B processor farm

mercial hardware boards. Such a solution is favorable, because it meets the cost requirements and is, when housing the processor boards in a standard crate system, easy to maintain. It moreover allows to use high speed links for the data transfer. Although the considerations of the various hardware solutions have not been finished yet we favor processors of the PowerPC series. This series offers also in the future availability with increasing performance at good prices. Right now the 603/604 series is available from different vendors, in $1-2$ years the 604e or 620 series with considerably higher performance will very likely come into the market.
A big computer system with more then 100 nodes needs a real-time operating systems, allowing to develop the software for the controlling of the the farm and the running of the processing tasks. In High Energy Physics several such operating systems, containing a host system for the software development and a target system to be run on the embedded processors, are in use. Since none of these systems is tailored and fully tested for the PowerPC processors yet first tests have just recently taken place. Candidates for real-time operating systems are AIX compatible packages, LynxOS, Solaris, VMEexec, VxWorks, etc.. Hardware is offered by CES, Cetia, Motorola, Omnibyte, Parsytec, etc..
It is planned to build a small test farm until mid of 1996, consisting of only a few nodes, which contains already the main building blocks: Dedicated processing nodes, high speed data links, and a switching network. The test farm will be controlled by an IBM AIX host work station which runs the software development system. This system will be used to develop all the necessary software to distribute the data and to control the test farm, to evaluate different switching technologies,

and to simulate the architecture of a full scale processor farm. To perform benchmark tests, the offline programs for event reconstruction and third level triggering will be adapted to the test farm. In this context the implementation of CERN libraries, the usage of FORTRAN, and the handling of system calls on the embedded system is a main issue. The planned implementation is depicted in figure 3.

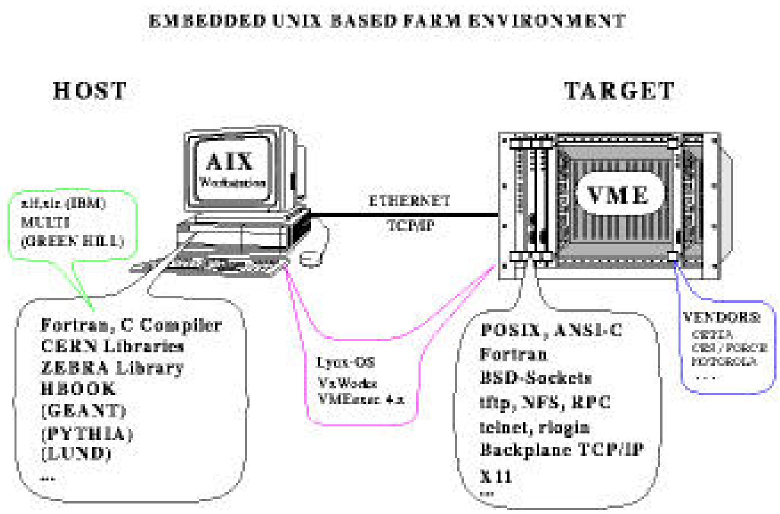

Figure 3: The planned implementation of the HERA-B processor farm

Acknowledgments

We would like to thank the electronics group and the computer center of DESY-Institut für Hochenergiephysik in Zeuthen and the whole HERA-B collaboration for their support.

References

1. E. Hartouni, et al., *HERA-B: An Experiment to Study CP Violation in the B System Using an Internal Target at the HERA Proton Ring*, Design Report, **DESY-PRC 95/01** (1995)
2. I.C. Legrand, A. Gellrich, U. Gensch, H. Leich, and P. Wegner, *Design and Simulation for Real-Time Distributed Processing Systems*, Contribution to the CHEP95, Rio de Janeiro, Brazil (1995)