The Computing Model for HERA-B

Hartwig Albrecht

DESY Hamburg

The computing model of the experiment HERA-B is presented. Based on the physics goals of the experiment, it will be shown that demands on computing and software are higher than for the present generation of experiments. Most challenging: all reconstruction including calibration and alignment procedures must run online.

1 The Experiment HERA-B

The experiment HERA-B is currently being built at the *ep* storage ring HERA at DESY / Hamburg. It is a fixed-target experiment which uses a system of eight target wires positioned in the beam halo of the HERA 820 GeV proton ring.

The goals if HERA-B are concentrated on B physics:

- Measurement of CP violation of the $B^0\overline{B}^0$ system in the decay channels $B^0 \to J/\psi K_S^0$ and $B^0 \to \pi^+\pi^-$
- Measurement of $B_s\overline{B}_s$ mixing
- Spectroscopy of B hadrons
- Investigation of rare B decays

The physically most interesting decay channel is $B^0 \to J/\psi K_S^0$, $J/\psi \to \ell^+\ell^-$, $\ell = e$ or μ, $K_S^0 \to \pi^+\pi^-$. This is expected to allow for the first observation of CP violation outside the Kaon system.

The challenge of the experiment comes from the requested event rates and the necessary background suppression. Both requirements exceed by far corresponding requirements of present and past detectors and impose strong impacts on the demands on computing and the quality and reliability of the software.

The available CMS energy, about 40 GeV, is not too far above threshold for $B\overline{B}$ production ($\mathcal{O}$ (10GeV)). This leads to the fact that the signal-to-

Table 1
Rates

HERA bunch frequency	8 MHz
Interaction rate	40 MHz
Interactions / bunch	5
Physics output rate:	
$B \to J/\psi X$	0.05 Hz
Semileptonic B decays	0.5 Hz
Signal / background	$< 10^{-7}$
compare to other experiments:	
B-factory	0.2
LHC	10^{-4}
output on mass storage	< 20 Hz
background suppression	
in the trigger	$> 4 \cdot 10^5$

background ratio, *i.e.* number of events containing B-hadrons / number of events with light quarks only, is worse by a factor 10^3 compared to hadronic experiments at sufficiently high energies. The comparison with B-factories at e^+e^- storage rings running on the $\Upsilon(4S)$ gives even a factor 10^6. The trigger system for HERA-B which provides the required selectivity is described in section 2.

Moreover, in order to collect high statistics, the target is operated in a way that, in average, 5 interactions take place per bunch crossing. This leads to an extremely high track density with 100 charged tracks at the target and 200 tracks in front of the calorimeter. As a consequence, the reconstruction program must be able to deal with detector occupancies of up to 25%. The rates are summarized in Table 1 and Figure 1; some numbers on the detector are given in Table 1.

The Detector HERA-B has two independent pattern recognition areas: the Silicon Vertex Detector between target and magnet and the Main Tracker behind the magnet. In both areas tracks form straight lines due to the absence of a magnitic field. This is a precondition for an efficient and fast pattern recognition under the conditions given at HERA-B (see section 7). In addition, particle identification is provided by an Electron calorimeter, a Transition Radiation Detector, Muon Chambers, and a RICH.

Table 2
The detector HERA-B

electronic channels	700,000
mean charged multiplicity:	
at target (5 interactions)	100
in front of calorimeter	200
maximum occupancy	25%
mean occupancy main tracker	12%
mean occupancy vertex det.	$< 3\%$
event size	
on mass storage:	
raw data and reconstruction	100 kb
on disk:	
selected reco. data	5 kb

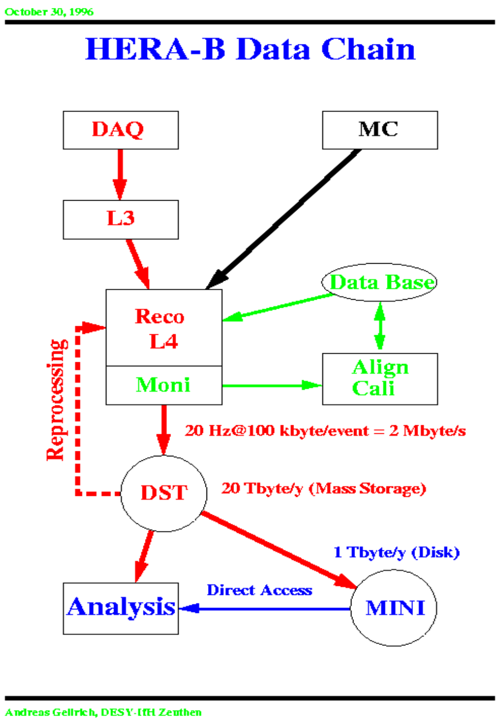

Fig. 1. Data Chain

2 Triggers and Farms

The required background rejection capability is achieved with an elaborate trigger system consisting of four levels.

Level 1: analyze lepton candidates. It consists of a pretrigger which searches electron and muon candidates in the calorimeter or muon chambers, respectively, followed by a tracking through selected layers in the Main Tracker. With the origin of the leptons approximately known by the position of the target, momenta and finally invariant masses of the leptons are calculated.

Level 2/3: analyze lepton candidates; trigger level 2 is restricted in time consumption to 7 ms due to the level 2 event buffer. Now, tracking is performed in all layers of the Main Tracker. The lepton tracks are swum through the magnet, thereby improving the momentum determination. Finally, the the Silicon Vertex Detector is used to find a common vertex of the leptons which indicates the decay $B^0 \to J/\psi K_S^0$, $J/\psi \to \ell^+\ell^-$. Note that the vertex determination at this early stage is indispensible for a sufficient background rejection.

Level 4: analyze full event. It starts with the full event reconstruction: track reconstruction, vertex finding for weakly decaying particles, and particle identification. Using the full reconstruction information, events are physically classified according to specific decay channels. A final event selection is performed according to this classification.

Trigger level 2, 3, and 4, the latter one including full event reconstruction, will run on three distinct processor farms with several 100 nodes. The choice is to use PC's with LINUX for level 2 and 3 and most likely also for level 4. The data flow above level 3 and the correponding rates are shown in Figure 1.

3 Specific Requirements for HERA-B

For the experiment HERA-B, the classical distinction between online and offline does not exist any more. Trigger level 4, and consequently the final event selection, depends on the full event reconstruction. Later event reprocessing is useful in exceptional cases only, and will not recover events rejected by the trigger. All triggers depend on good calibration and alignment. In particular, the vertex determination in level 2 and above needs a perfect Silicon Detector alignment. All calibration and alignment constants have to be monitored permanently. In case these constants must be modified, reliable tests of the modifications have to be performed and then the new constants immediately have to be made available for all farms and for offline analysis.

4 Tools

The foollowing tools are used within the software for HERA-B:
Program frame: ARTE (see next section)

Source code management: CVS, C-preprocessor (also for Fortran code)
User interface: KUIP; Tcl/Tk
Memory management: "minimal" Zebra (the original idea was to abandon Zebra at a later stage)
Data model: ideas borroughed from ADAMO (tables, entities, relationships)
Event generators: Jetset, Pythia, Fritiof; modified for $pN \to$ *heavy quarks* reactions (non-trivial)
Detector simulation: Geant 3.21
Graphics: HIGZ / HPLOT
Statistics: HBOOK / PAW

The HERA-B "offline" software runs on the following platforms: SGI with IRIX_mips2, HP with HP-UX, IBM RS6000 with AIX, ALPHA with OSM1, and PC with LINUX. This large variety of platforms is necessary because at least Monte-Carlo jobs have to run on as many machines in different labs as possible. Testing programs under different operating systems also helps efficiently to find program bugs.

5 ARTE: The Program Frame for HERA-B

ARTE is the common frame for all program modules above trigger level 3: Monte-Carlo event generation, Monte Carlo detector simulation, reconstruction (tracking, vertexing, particle identification), event display, physics analysis. It provides the general program steering by using a user interface or command files. It controls input / output of events as well as of calibration / alignment constants.

The most important rôle of ARTE, however, is that it is the place where all data structures of more than local importance are centrally declared. This has the following advantages: (i) all program developpers are forced to restrict themselves to a well-defined set of data structures and (ii) the structure declarations can be provided in a language-independent way as described in the next section.

6 Language Independency

All global data structures are declared by the "DD", a language-independent Data Declaration. A specific DD program derives language-specific include (header) files from these DD files:
Fortran: Type declarations, COMMON specifications
Fortran: Statement functions which allow to use intellegible names in Fortran

C: Structure declarations
C++: Class declarations including special treatment of relationships
Zebra: Automatic initialization of all necessary banks
Manual: A readable description of the data structures.

By means of the DD both, Fortran as well as C++ programmers have fast and efficient access to all data structures. The underlying Zebra memory management is complitely hidden from the user. The preferred language is clearly C++. Given the tight time scale for completing HERA-B, however, it was decided that there is no possibility to abandon Fortran completely. ARTE constitutes a compromize: old Fortran programs can still be used; new program developments are written in C or C++. In particular, it is foreseen that tools for physics analysis which really demand an objective language will be provided in C++ only.

7 Track Reconstruction

A precondition for the approval of the experiment HERA-B (1995) was the proof that tracks can be reconstructed at all under the given hostile environment. The pattern recognition and track fit program "Ranger" for the main tracker uses the Kalman-Filter technique which takes into account effects from multiple scatttering etc. The existing program reconstructs events with typical occupancies of up to 25 %, charged multiplicities around 200 with a track reconstruction efficiency of 97 %. The time consumption for those events is 4 s on a HP 9000 which meets the demands on time consumption on a PC farm. A typical event is shown in Figure 2 and 3.

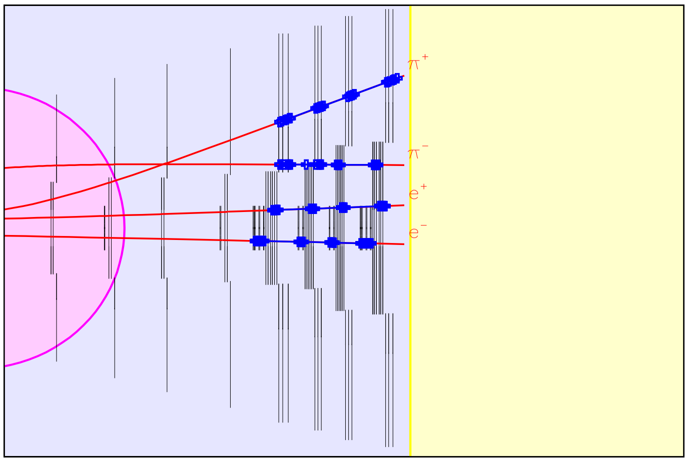

Fig. 2. HERA-B event display. Shown are only the tracks originating from the decay $B^0 \to J/\psi K^0_S$.

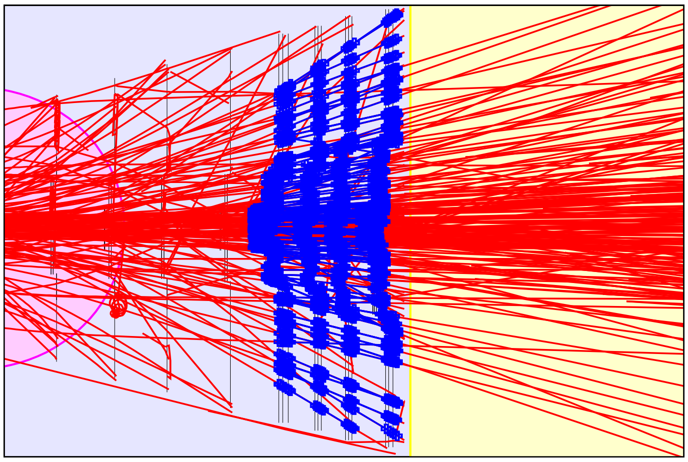

Fig. 3. HERA-B event display. The same event as in Figure 2 with all tracks.

8 Summary

The basic aspects of the computing model for HERA-B are:

- Farms: Trigger level 2, 3, and 4. Level 4 also contains the full event reconstruction.
- Arte: One common frame for all software above trigger level 3
- Language independence: Smooth transition from Fortran to object-oriented programming
- Track / vertex reconstruction well advanced.

The unprecedented challenge for the software of HERA-B is the fact that event reconstruction and calibration / alignment procedures must run online.

Test data with parts of the detector have already been taken in 1996. The detector is expected to be almost completed in May 1998, and the software must be ready at this time. Physics running will start in 1999.