The Architecture of the ZEUS Second Level Global Tracking Trigger

Mark Sutton
University College London, Gower Street, London WC1E 6BT, United Kingdom
On behalf of the ZEUS GTT Group

Abstract

The architecture and performance of the ZEUS Global Tracking Trigger (GTT) are described. Data from the ZEUS silicon Micro Vertex detector's HELIX readout chips, corresponding to 200k channels, are digitized by 3 crates of ADCs and PowerPC VME board computers push cluster data for second level trigger processing and strip data for event building via Fast and GigaEthernet network connections. Additional tracking information from the central tracking chamber and forward straw tube tracker are interfaced into the 12 dual CPU PC farm of the global tracking trigger where track and vertex finding is performed by separately threaded algorithms. The system is data driven at the ZEUS first level trigger rates $<$500Hz, generating trigger results with a mean latency of 10ms. The GTT integration into the ZEUS second level trigger and recent performance are reviewed.

INTRODUCTION

The ZEUS detector is a multi-purpose detector, for detecting the final state particles of collisions between 820GeV protons and 27.5GeV leptons at the HERA collider at DESY. It features an inner tracking system, consisting of forward, central and rear tracking detectors surrounded by a thin superconducting solenoid (1.43T), high resolution calorimeter, and muon detector [1]. The HERA collider operates with 220 bunches of electrons and protons with a bunch crossing every 96ns.

During the HERA luminosity upgrade [2] shutdown in 2001 several new components were added to the detector, including the forward Straw Tube Tracker (STT) and the silicon Micro Vertex Detector (MVD)[3, 4, 5] which consists of a cylindrical barrel section for central tracking and forward wheels for tracking in the forward direction.

Studies for a standalone MVD trigger at the second level found that the small number of MVD layers (3) meant this would not be feasable. Since the pre-upgrade trigger hardware is more than 10 years old, including data from the new tracking detectors in the existing (CTD) tracking trigger at the second level is not realistic so the GTT was envisaged to use the MVD data by using reconstruction of tracks in the CTD to help break the patterm matching ambiguity in the MVD. At the same time this would also allow more detailed reconstruction using the CTD data than was possible with the pre-upgrade CTD second level trigger algorithm.

A Forward algorithm is also being developed, which uses information from the STT to help resolve the pattern recognition ambiguities in the MVD forward wheels.

The primary aim of the GTT is to provide a significantly better primary z vertex reconstruction than the current CTD SLT trigger by using CTD axial and stereo hits and, since May 2004, the z-by-timing data, and including MVD hits, eventually providing secondary vertex reconstruction to identify heavy flavour decays. In addition, the inclusion of information from the STT and MVD forward wheels extends the acceptance of the trigger to the forward region.

Since the new GTT must function within the framework of the existing ZEUS DAQ and trigger systems it must provide a second level track trigger without compromising the experiment's DAQ performance.

ZEUS TRIGGER ENVIRONMENT

The maximum interaction rate of 10MHz together with detector data volumes up to 500kBytes put stringent requirements on the design of the trigger and DAQ systems [6]. The ZEUS trigger is a flexbible, three-level trigger system allowing progressively more time for event processing per level. It is essentially deadtime free ($\leq$1-2%), making use of trigger decision and data event pipelines, front end signal DSP processing and parallel processing using transputers (TP).

The First Level Trigger (FLT) system is a hardware, pipelined trigger designed to start processing of a new event at each 96ns bunch crossing. The Global First Level Trigger (GFLT) and the component FLT systems synchronise the DAQ to the cyclic HERA bunch crossing number. In the GFLT typical input rates of 10-200kHz, dominated mostly by background and noise events are reduced to $\leq$500Hz. The GFLT result is available after 44 bunch crossings (4.4μs). On GFLT accept components buffer their event data and send to the component based local second-level trigger (SLT) algorithms.

The Global Second Level Trigger (GSLT) [7] combines the results of the transputer based component local second level trigger algorithms to provide a rate reduction to $\leq$50Hz. Component SLT data is sent to the GSLT via serial links clocked at 1.8MHz. The GSLT operation requires that component SLT data be available with a mean latency of $\leq$10-15ms. A 15 event component side buffer reduces deadtime effects to $\leq$1%. On a GSLT accept component data is sent to the Event Builder (EVB) which sends completely built events to the Third Level Trigger (TLT) PC farm for reconstruction using a light weight version of the offline reconstruction. This reduces the rate of accepted events to $\sim$10Hz which are then archived.

Condition	Requirement
GFLT accept rate	$\leq$500 Hz
GSLT accept rate	$\leq$50 Hz
Mean latency of GTT result sent to GSLT	$\leq$15 ms
Mean latency of GTT data sent to EVB	$\leq$50 ms
CTD mean data size	$\sim$5kBytes
STT mean data size	$\sim$5kBytes
MVD mean cluster data size	$\sim$5kBytes
MVD mean strip data size	$\leq$15kBytes

Table 1: Trigger and DAQ requirements

The trigger and DAQ requirements placed on the GTT are summarized in Table 1. Note that the GTT mean latency at the GSLT is determined by the requirement that it must be within the existing CTD SLT latency envelope.

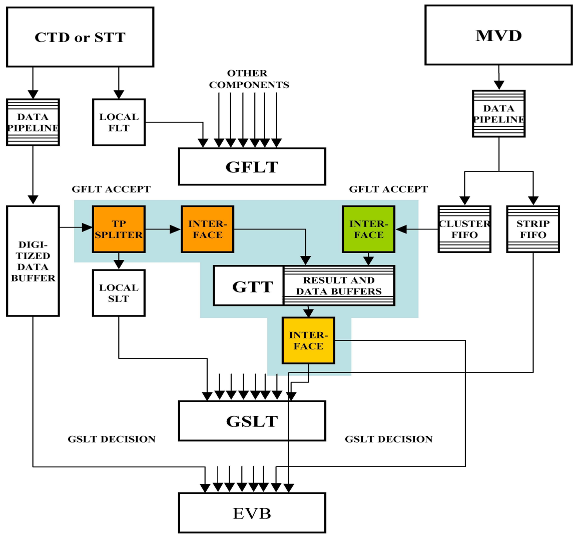

Figure 1: Schematic of ZEUS-GTT trigger architecture.

DATA AND TRIGGER INTERFACES

The GTT architecture is show schematically in Fig.1. On GFLT trigger accept, the MVD, CTD and STT component frontend systems send track hit data to their respective component interfaces which in turn send the data to a PC in the GTT processor farm running a reconstruction algorithm environment. The result of the reconstruction is then sent to the GSLT. The GSLT trigger result is sent back to the originating environment which, on accept, sends to the EVB the result bank and optionally, the MVD cluster data.

All sub-systems transfer data using TCP/IP via point-to-point Fast or GigaEthernet links.

MVD interface

The MVD readout system is described in detailed elsewhere [9] so is only briefly reviewed here.

On a GFLT accept data is transferred from the silicon detector HELIX readout chip pipelines to $\sim$30 ADC boards located in three 9U VME crates (upper barrel, lower barrel and rear wheel channels) where data is digitized and stored in strip (raw hits) and cluster (weighted hit position, bounds and pulse height) ring buffers. On completion of digitization an interrupt is generated in each crate and the cluster data is read out via VMEbus, by a Motorola MVME 2700 PowerPC/LynxOS readout cpu and forwarded to one of the GTT farm PC's running the GTT reconstruction environment. On GSLT accept the strip data is readout and transferred to the EVB interface.

CTD Interface

The existing CTD DAQ and SLT systems [8] are fully functioning components of the ZEUS readout and trigger systems. On a GFLT accept, digitized pulse height and drift time data from the 4608 sense wires are readout from custom FADC cards by 16 sector readout transputers which forward the data to the CTD SLT trigger and data merging TP networks. The SLT network performs segment finding and track reconstruction separately, in parallel, for each sector before combining the data to reconstruct the event as a whole. On a GSLT accept final data merging is completed and the data is sent to the EVB interface.

To enable the data from the CTD to be used at the GTT two sets of 16 additional data splitting TPs were added to the CTD DAQ system to allow data from the FADC and z-by-timing systems to be parasitically read out whilst causing minimal disruption to the data flow within the CTD network. Currently each set of 16 TP links are connected to two NIKHEF 2TP boards in two MVD VME crates – one for the FADC system, and one for the z-by-timing system – which write the incoming data into the 2TP's Triple Port Memory (TPM). PowerPC/LynxOS processors, one in each crate, are used to read out the TPM and forward the data to the GTT. As event data availability in TPM cannot be flagged at the PowerPC, the 2TP cannot generate a VME-bus interrupt, the PowerPC polls for data every 500μs.

The GTT receives FADC drift-time data from axial and stereo superlayers and the axial z-by-timing data. The pulse height (dE/dx) data are not sent. The pre-HERA luminosity shutdown FADC drift-time data volume generated per GFLT accept is $\sim$ 3-3.5kB with a readout latency, with respect to the GFLT accept, of $\sim$3.0ms at the PowerPC.

STT interface

The STT reuses the TP based readout system of the removed Transition Radiation Detector and is similar in design to that of the CTD so the STT interface is implemented using the system developed for the CTD, although only 8 input TP links to a single 2TP board are used.

The expected data volume per GFLT accept is $\leq$4kB with a readout latency similar to that of the CTD. At the time of writing the STT interface to the GTT is operational, and data is transferred, although the Forward Algorithm is not included during luminosity running.

GSLT interface

The GTT farm result is sent to a Motorola PowerPC VMEbus cpu which transfers the result, via NIKHEF-2TP and one of it's TP serial links, to the GSLT. As the order of the gflt number of trigger results arriving at the GSLT is strictly sequential the interface is required to buffer the results from the different GTTs before sending in order.

No special hardware is required to receive the GSLT trigger decision as this is sent via TCP and is received by an interface task before forwarding to the MVD data interfaces and the GTT algorithm enviroment which processed the event. On accept the MVD data interfaces send the MVD strip data. The algorithm environment sends the MVD cluster and algorithm result to the EVB interface.

EVB interface

The EVB interface waits for MVD and GTT data associated with GSLT accepted events which are formatted and sent to the TLT via TCP/IP.

GTT HARDWARE

Because the GTT runs in the ZEUS second level trigger layer the available combined event processing time and data transfer latency must be within about 15ms.

To enable sufficient processing for the GTT Algorithms, the GTT has been implemented as a PC-farm with network connections using TCP/IP protocol via point-to-point Ethernet links to network switches supporting one host per port. This decision has allowed newer, faster equipment (Gigabit Ethernet and PCs) to be integrated transparently into the GTT. The PC hardware and network switches used in the GTT were provided via a grant from Intel Corp.

The number of computing nodes of the GTT was estimated using a simple discrete event simulation with fixed GFLT accept rate, realistic random time intervals between triggers, a CTD processing latency envelope, and ignoring network transit times but with realistic data readout time. This provided measurements of the mean and maximum waiting times before data processing on a GTT node could start. Both Credit-Control and Round-Robin data distribution were studied with the former having the lower waiting times for small numbers of nodes and so was implemented. The estimated number of processing environments required was $\sim$10, and the actual number of nodes in the current system is 12.

GTT SOFTWARE

Each GTT node runs an algorithm processing task, a histogram pusher and a statistics pusher all controlled by an MVD daemon process. The daemon advertizes these processes to the MVD run control system [9] and, depending on run type, starts them during the SETUP transition of a run. After ACTIVATION the algorithm processing environment receives and processes hit data from all the data sources, and sends the result to the GSLT. The histogram and statistics pushers periodically forward monitoring information to monitoring tools.

The algorithm processing environment is a multi-threaded program with 1 thread per input data source (3$\times$MVD, 2$\times$CTD and STT), 1 thread per algorithm (CTD+MVD barrel, STT+MVD forward), and 1 time-limit thread. The trigger result bank is built and sent when all algorithm threads have finished the event. The time-limit thread terminates any excessively long processing sending a PASS trigger to the GSLT after, currently, 40ms. The free-for-next-event credit is sent only when all data and altorithm threads have finished the event.

A detailed description of the barrel algorithm is found elsewhere [10] and in these proceedings [11].

PERFORMANCE

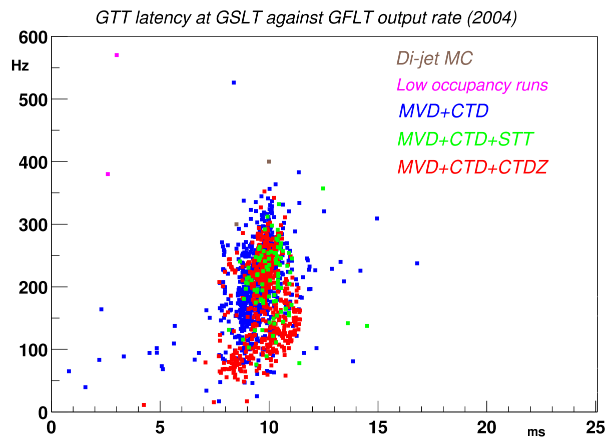

Figure 2: GTT run latency against rate

With the first HERA II test luminosity after the upgrade, the GTT was found to perform satisfactorily. Since then, the GTT has been running as standard in the ZEUS trigger chain under production luminosity conditions for the entire 2003-04 running period. Only the Barrel algorithm was in place during this period, for although the MVD hit matching is satisfactory for physics events, the large occupancy of beam gas events causes problems with the current implementation of the MVD hit matching which was therefore disabled during the 2003-04 run. The Barrel algorithm provided a reconstructed vertex position for the GSLT and is used to select events online in the physics filters. The GTT system and Algorithm have been found to be very stable with over 40pb^{-1} written to tape.

Figure 2 shows the mean GTT latency versus the GFLT output rate for the 2003-04 running period. The mean latency is around 9ms and is largely uncorrelated with the GFLT rate, showing that the system is coping well. Event rates of greater than 500Hz have been achieved with no appreciable increase in the observed deadtime.

In May and July 2004 the CTD transputer network was modified to allow the CTD z-by-timing data to be sent to the algorithm enviroment data. The required new component interface was also added, and the algorithm modified to use this data, resulting in a slight increase in the overall latency. Test runs reading out the data from the STT, but without running the Forward algorithm are also shown. Initially, the data size was larger than expected but after imposing a reasonable data cutoff it is seen that the increased data flow to the algorithm processor causes no degradation in the system performance.

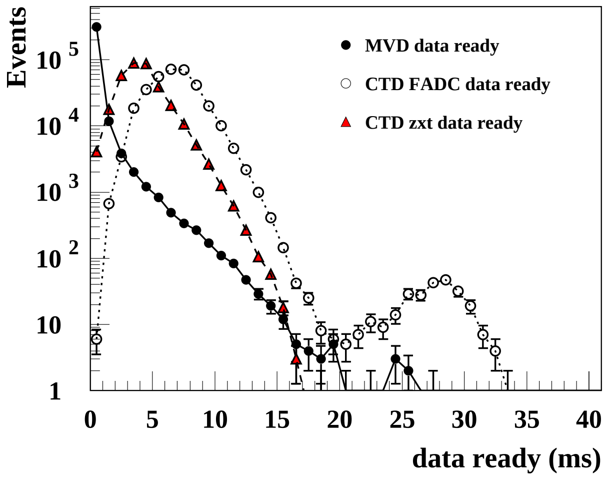

Figure 3: Distribution of data ready times at the algorithm process.

The distribution of arrival (data ready) times for the MVD, CTD and CTD z-by-timing data at the algorithm processor for a typical run is show in Fig 3. For this run the mean data ready times were 0.64ms for the MVD, 4.1ms for the CTD z-by-timing and 6.8ms for the CTD FADC data and the corresponding mean data sizes were 0.4kB, 0.8kB and 3.1kB respectively. The mean algorithm processing time for this run was 2.2 ms, with an overall system latency of 9.5ms, illustrating that the overall latency at the GSLT is dominated by the readout and transfer of the large CTD data volumes.

Detailed studies of the Algorithm performance are well underway and discussed briefly in these proceedings [11]. The vertex finding efficiency for events with a physics vertex wintin 60cm of the nominal interaction point is greater than 90% for vertices with more than 5 tracks.

For a small number of test runs near the end of data taking in 2004, the GTT was running online with a test version of the Forward algorithm enabled. The mean Forward algorithm latency for these runs, after the STT data cutoff, is seen to be within 1 ms with essentially no effect on the overall system latency. The online vertex distribution is seen in Fig. 4

Following on from the successful online tests, the forward algorithm will be included as standard for evaluation when data taking resumes in October 2004. Work is currently underway to include the hits from the MVD forward wheels into the forward algorithm.

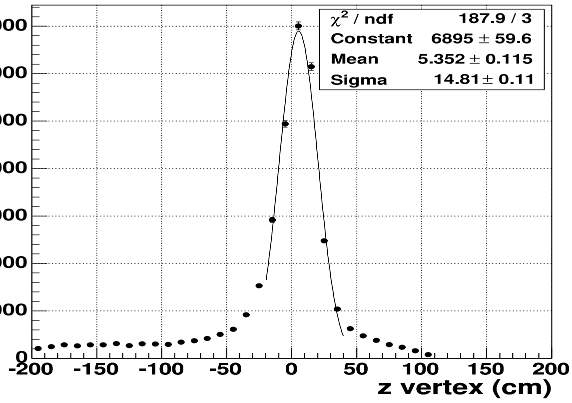

Figure 4: The online vertex distribution from the GTT Forward Algorithm

CONCLUSIONS

The GTT has operated successfully during production luminosity data taking at ZEUS since autumn 2002 without compromising the ZEUS trigger and DAQ performance. The latency is acceptable and design input rates of 500 Hz cause no additional deadtime.

The barrel algorithm is mature and it's vertex result is used to select events online in the ZEUS physics filters. A forward algorithm is ready and will be included for evaluation as standard in the next running period. Use of the MVD data is currently disabled in the online algorithms, although it is hoped that modifications to the algorithms to enable it's use will be complete before the next production luminosity run.

REFERENCES

[1] ZEUS Collaboration, Status Report, ZEUS 93.

[2] U. Schneekloth (Ed.). the HERA Luminosity Upgrade DESY HERA 98-05. 1998.

[3] A Microvertex Detector for ZEUS, ZEUS-Note 97-006, 1997.

[4] A. Garfagnini, Nucl. Instr. and Meth, A 435 (1999) 34.

[5] A. Coldewey, Nucl. Instr. and Meth, A 447 (1999) 44.

[6] C.Youngman, The ZEUS data acquisition system, in: C. Verkerk, W. Wojcik (Eds.), Proceedings of CHEP92 Annecy, CERN, 1992, p.222.

[7] H.A.J.R. Uijterwaal, The ZEUS global second level trigger, Ph.D. Thesis, University of Amsterdam, 1992.

[8] R.C.E. Devenish, Nucl. Instr. and Meth, A 438 (1999) 472-501.

[9] C.Youngman, ZEUS silicon vertex detector DAQ and SC, IEEE-RTC 2003 Montreal.

[10] M.Sutton, The GTT barrel algorithm, IEEE-RTC 2003 Montreal.

[11] M.Sutton, The GTT barrel algorithm, these proceedings.