# PHENIX On-Line Systems


S.S. Adler[a] M. Allen[m] G. Alley[m] R. Amirikas[d] Y. Arai[h]
T.C. Awes[m] K.N. Barish[b] F. Barta[a] S. Batsouli[c] S. Belikov[g,m]
M.J. Bennett[i] M. Bobrek[m] J.G. Boissevain[i] S. Boose[a]
C. Britton[m] L. Britton[m] W.L. Bryan[m] M.M. Cafferty[i]
T.A. Carey[i] W.C. Chang[q] C,Y. Chi[c] M. Chiu[c] V. Cianciolo[m]
B.A. Cole[c] P. Constantin[g] K.C. Cook[g] H. Cunitz[c]
E.J. Desmond[a] K. Ebisu[k] Y.V. Efremenko[m] K. El Chenawi[j]
M.S. Emery[m] D. Engo[a] N. Ericson[m] D.E. Fields[l] S. Frank[m]
J.E. Frantz[c] A. Franz[a] A.D. Frawley[d] J. Fried[a] J. Gannon[a]
T.F. Gee[m] R. Gentry[m] P. Giannotti[a] H.-A. Gustafsson[j]
J.S. Haggerty[a] S. Hahn[i] J. Halliwell[m] H. Hamagaki[r]
A.G. Hansen[i] H. Hara[k] J. Harder[a] X. He[e] F. Heistermann[a]
T.K. Hemmick[o] M. Hibino[s] J.C. Hill[g] K. Homma[f] B.V. Jacak[i]
U. Jagadish[m] J. Jia[o] F. Kajihara[r] S. Kametani[r]
Y. Kamyshkov[m] A. Kandasamy[a] J.H. Kang[t] J. Kapustinsky[i]
K. Katou[s] M.A. Kelley[a] S. Kelly[c] J. Kikuchi[s] S.Y. Kim[t]
Y.G. Kim[t] E. Kistenev[a] D. Kotchetkov[b] K. Kurita[f]
J.G. Lajoie[g] M. Lenz[a] W. Lenz[a] X.H. Li[b] S. Lin[a] M.X. Liu[i]
S. Markacs[c] F. Matathias[o] T. Matsumoto[r] J. Mead[a]
R.E. Mischke[i] G.C. Mishra[e] A. Moore[m] M. Muniruzzamann[b]
M. Musrock[m] J.L. Nagle[c] B.K. Nandi[b] J. Newby[p]
J. Nystrand[j] E. O'Brien[a] P. O'Connor[a] H. Ohnishi[f]
A. Oskarsson[j] L. Osterman[j] K. Oyama[r] L. Paffrath[a]
C.E. Pancake[o] V.S. Pantuev[o] A.N. Petridis[g] R.P. Pisani[a]
T. Plagge[g] F. Plasil[m] M.L. Purschke[a] S. Rankowitz[a] R. Rao[m]
M. Rau[a] K.F. Read[m] S.S. Ryu[t] T. Sakaguchi[r] H.D. Sato[n]
R. Seto[b] T. Shiina[i] D. Silvermyr[j] J. Simon-Gillo[i]
M. Simpson[m] W. Sippach[c] H.D. Skank[g] S. Skutnik[g]
G.A. Sleege[g] G.D. Smith[i] M. Smith[m] P.W. Stankus[m]
P. Steinberg[c] T. Sugitate[f] J.P. Sullivan[i] A. Taketani[n]
M. Tamai[s] Y. Tanaka[k] W.D. Thomas[g] R. Todd[m] F. Toldo[a]
G. Turner[m] T. Ushiroda[k] J. Velkovska[o] H.W. van Hecke[i]

M. Van Lith [a] L. Villatte [p] W. Von Achen [a] J.W. Walker [m] H.Q. Wang [b] S.N. White [a] A.L. Wintenberg [m] C. Witzig [a] L. Wood [g] W. Xie [b] G.R. Young [m] W.A. Zajc [c] C. Zhang [c] L. Zhang [c]

(The PHENIX Collaboration)

[a] *Brookhaven National Laboratory, Upton, NY 11973-5000, USA*

[b] *University of California Riverside, Riverside, California, 92521, USA*

[c] *Columbia University, New York, NY 10027 and Nevis Laboratories, Irvington, NY 10533, USA*

[d] *Florida State University, Tallahassee, FL 32306, USA*

[e] *Georgia State University, Atlanta, GA 30303, USA*

[f] *Hiroshima University, Kagamiyama, Higashi-Hiroshima 739-8526, Japan*

[g] *Iowa State University, Ames, IA 50011, USA*

[h] *KEK, High Energy Research Organization, Tsukuba-shi, Ibaraki-ken 305-0801, Japan*

[i] *Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

[j] *Lund University, Box 118, SE-221 00 Lund, Sweden*

[k] *Nagasaki Institute of Applied Science, Nagasaki-shi, Nagasaki 851-0193, Japan*

[l] *University of New Mexico, Albuquerque, NM, USA*

[m] *Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*

[n] *RIKEN(The Institute of Physical and Chemical Research), Wako, Saitama 351-0198, Japan*

[o] *State University of New York, Stony Brook, NY 11794, USA*

[p] *University of Tennessee, Knoxville, TN 37996, USA*

[q] *Institute of Physics, Academia Sinica, Taipei 11529, Taiwan*

[r] *Center for Nuclear Study, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, 113-0033, Japan*

[s] *Waseda University, Advanced Research Institute for Science and Engineering, 17 Kikui-cho, Shinjuku-ku, Tokyo 162-0044, Japan.*

[t] *Yonsei University, IPAP, Seoul 120-749, Korea*

The PHENIX On-Line system takes signals from the Front End Modules (FEM) on each detector subsystem for the purpose of generating events for physics analysis. Processing of event data begins when the Data Collection Modules (DCM) receive data via fiber-optic links from the FEM's. The DCM's format and zero suppress the data and generate data packets. These packets go to the Event Builders (EvB) that assemble the events in final form. The Level-1 trigger (LVL1) generates a decision for each beam crossing and eliminates uninteresting events. The FEM's carry out all detector processing of the data so that it is delivered to the DCM's using a standard format. The FEM's also provide buffering for LVL1 trigger processing and DCM data collection. This is carried out using an architecture that is pipelined and deadtimeless. All of this is controlled by the Master Timing System (MTS) that distributes the RHIC clocks. A Level-2 trigger (LVL2) gives additional discrimination. A description of the components and operation of the PHENIX On-Line system is given and the solution to a number of electronic infrastructure problems are discussed.

## 1 Introduction

The PHENIX detector at the Relativistic Heavy Ion Collider (RHIC) is designed to search for the phase transition from ordinary nuclear matter to a deconfined state of quarks and gluons called the Quark-Gluon Plasma (QGP). In order to achieve this goal the detector has been designed to efficiently detect leptons and photons in order to search for the relatively rare interactions that provide direct probes of the QGP that are not degraded by interactions with the surrounding nuclear matter. PHENIX is designed to make measurements on a variety of colliding systems from p-p to Au-Au. The occupancy in the detector varies from a few tracks in p-p interactions to approximately 10% of all detector channels in central Au-Au interactions. The interaction rate at design luminosity varies from a few kHz for Au-Au central collisions to approximately 500 kHz for minimum bias p-p collisions. The PHENIX DAQ system was designed to seamlessly accomodate improvements in the design luminosity. This was accomplished through the pipelined and deadtimeless features of the detector front ends and the ability to accomodate higher-level triggers. The calculated relations between colliding system, data rate and data size are shown in Fig.1. An overview of the PHENIX detector program is given elsewhere in this volume[1].

The wide range of event sizes shown in Fig.1 and luminosities present special challenges for triggering and data acquisition. In PHENIX it is necessary to measure low-mass lepton pairs and low $p_T$ particles in a high-background environment. In order to preserve the high interaction-rate capability of PHENIX

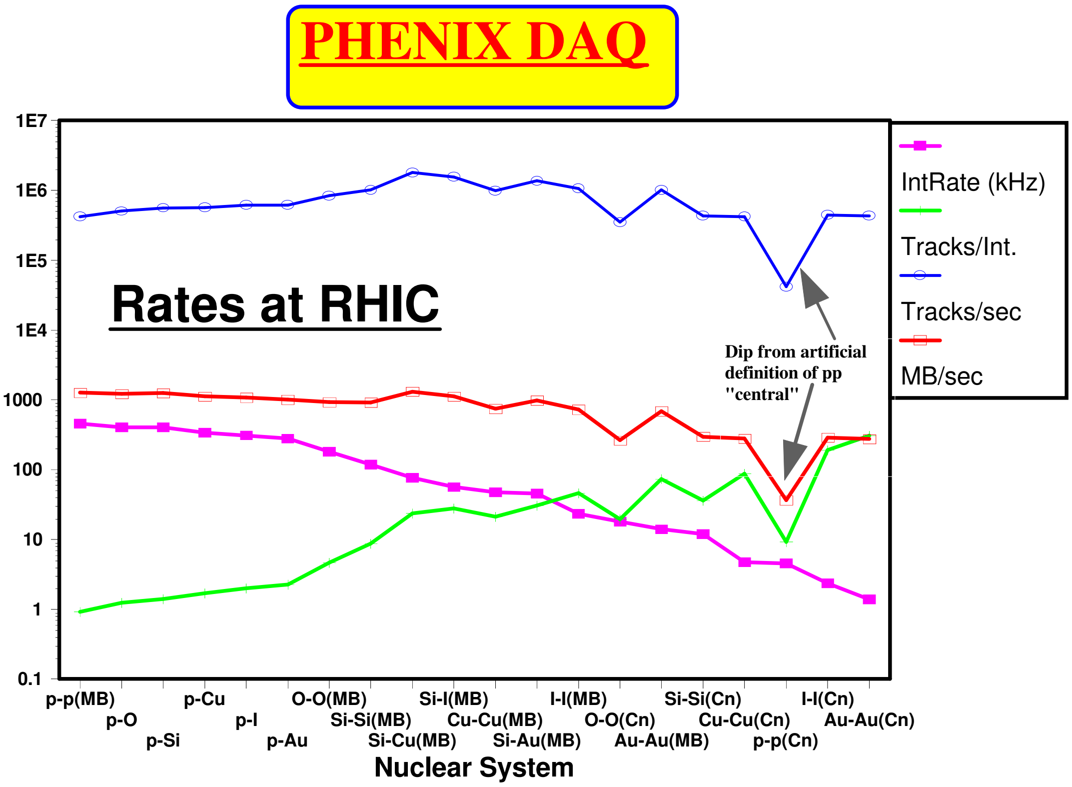



Fig. 1. The relations between colliding system, data rate and data sizes. The data is given for PHENIX design luminosities.

a flexible triggering system that permits tagging of events was constructed. The On-Line system has two levels of triggering denoted as LVL1 and LVL2. The LVL1 trigger is fully pipelined, therefore the On-Line system is free of deadtime through LVL1. Buffering is provided that is sufficient to handle fluctuations in the event rate so that deadtime is reduced to less than 5% for full RHIC luminosity. The LVL1 trigger and lower levels of the readout are clock-driven by bunch-crossing signals from the 9.4 MHz RHIC clock. The higher levels of readout and the LVL2 trigger are data-driven where the results of triggering and data processing propagate to the next higher level only after processing of a given event is completed.

The general schematic for the PHENIX On-Line system is shown in Fig.2.

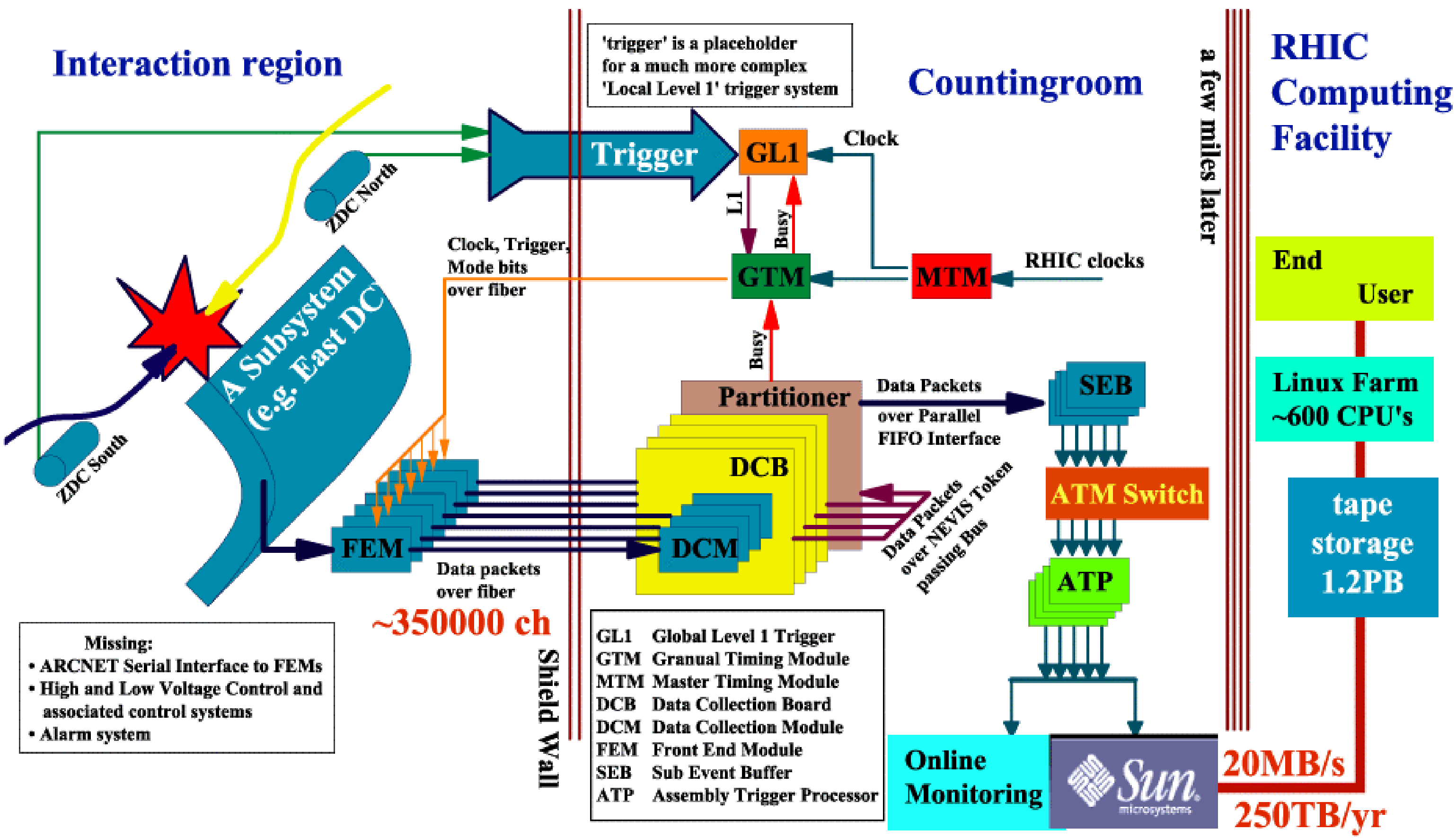



Fig. 2. Schematic diagram of the PHENIX On-Line system

Signals from the various PHENIX subsystems are processed by Front End Electronics (FEE) that convert detector signals into digital event fragments. This involves analog signal processing with amplification and shaping to extract the optimum time and/or amplitude information, development of trigger input data and buffering to allow time for data processing by the LVL1 trigger and digitization. This is carried out for all detector elements at every beam crossing synchronously with the RHIC beam clock. The timing signal is a harmonic of the RHIC beam clock and is distributed to the FEM's by the PHENIX Master Timing System (MTS). The LVL1 trigger provides a fast filter for discarding empty beam crossings and uninteresting events before the data is fully digitized. It operates in a synchronous pipelined mode, generates a decision every 106 ns and has an adjustable latency of some 40 beam crossings.

Once an event is accepted the data fragments from the FEM's and primitives from the LVL1 trigger move in parallel to the Data Collection Modules (DCM). The PHENIX architecture was designed so that all detector-specific electronics end with the FEM's, so that there is a single set of DCM's that communicate with the rest of the DAQ system. The only connection between the Interaction Region (IR) where the FEM's are located and the Counting House (CH) where the DCM's are located is by fiber-optic cable. The DCM's perform zero suppression, error checking and data reformating. Many parallel data streams from the DCM's are sent to the Event Builder (EvB). The EvB performs the final stage of event assembly and provides an environment for the LVL2 trigger to operate. In order to study the rare events for which PHENIX was designed, it is necessary to further reduce the number of accepted events by at least a factor of six. This selection is carried out by the LVL2 triggers while the events are being assembled in the Assembly and Trigger Processors (ATP) in the EvB. The EvB then sends the accepted events to the PHENIX On-line Control System (ONCS) for logging and monitoring. The technology used to control the many components that must work together to successfully accumulate the data is the Common Object Request Broker Architecture (CORBA) system. CORBA makes it possible to transparently access objects on remote computers of various types throughout the network. The main control process called Run Control (RC) accesses and communicates with remote objects which in turn control a given piece of hardware. The RC process determines the configuration of the whole DAQ front-end. The CORBA and RC systems are described elsewhere in this volume[2].

Detailed descriptions of the subsystems that together form the PHENIX On-Line system are given below. In addition, there are descriptions of a number of PHENIX electrical monitor and control systems needed to distribute power, protect the environment and insure safe operation of the PHENIX detector.

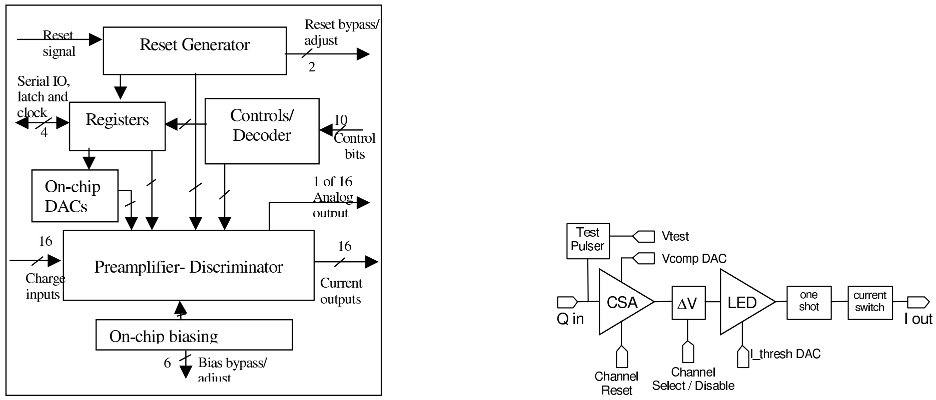
## 2 Front End Electronics and Digitization

The purpose of the Front End Electronics (FEE) is to digitize analog signals from detector elements and buffer the data to allow for LVL1 trigger decisions and data readout latencies. This involves analog signal processing with amplification and shaping to extract the optimum time and/or amplitude information from the signal and development of trigger input data. The data is buffered for up to 40 beam crossings to allow for the time needed to make the LVL1 trigger decision. If the LVL1 trigger accepts an event, a signal is transmitted to the Granule Timing Module (GTM) which generates an ACCEPT signal that is transmitted to the detector FEM's in the IR. The FEM's process the data from the individual subdetectors and send it to the DCM's for assembly. Zero suppression occurs outside the IR in the DCM's. In this section the process of digitization of the data is discussed in more detail and special features of particular subsystems are presented. In the next section the function of the FEM's and their role in preparing the data for communication to the DCM's is outlined.

The PHENIX detector subsystems use two basic approaches to collect and digitize data for transmission to the DCM's. In the first approach the data is digitized in real time during the latency period of 40 beam crossings. The Beam Beam Counters (BBC)[3], Zero Degree Calorimeters (ZDC)[4], Drift Chambers (DC)[5], Pad Chambers (PC)[5], Time Expansion Chamber (TEC)[5] and Muon Identifiers ($\mu$ID)[6] use this approach. In the second approach the data is sampled and stored in analog form in Analog Memory Units (AMU) and is only digitized after receipt of an accept from the LVL1 trigger. The Time of Flight (ToF) detectors[7], Ring Imaging Cherenkov Detectors (RICH)[7], Electromagnetic Calorimeter (EMCal)[8], Muon Trackers ($\mu$Trk)[6] and the Multiplicity Vertex Detector (MVD)[3] use this approach. All electronics include embedded firmware contained in field-programmable gate arrays (FPGA's). Two common FPGA functions are 1) timing, trigger, and general control, usually designated as the "heap manager" function, and 2) formation and transmittal of raw data packets, usually designated as the "data formatter" function.

### *2.1 Systems Employing Digitization in Real Time*

The PC's, TEC, DC's, BBC's, ZDC's and $\mu$ID's digitize their data in real time prior to receiving an accept from the LVL1 trigger. The PC and TEC systems are similar in that the data is buffered in a Digital Memory Unit (DMU). The PC's use the TGLD chip[9] for the initial charge amplification and discrimination. A block diagram for the TGLD chip and a block diagram for a single TGLD channel is shown in Fig.3.




TGLD chip Block diagram

Single Channel Block Diagram of PC TGLD

Fig. 3. A block scheme for the TGLD chip is shown to the left and a single TGLD channel is shown to the right

The DMU's are located next to the TGLD's on the readout cards which in turn are on the chamber and link in parallel with the FEM[10] that also includes the heap manager and data formater. The TEC uses a TEC-PS Application Specific Integrated Circuit (ASIC) that amplifies and shapes the pulses. They are then sent over about 8 m of twisted-pair cable to the TEC-FEM. The DC's use ASD8 chips[11] for the initial amplification, shaping and discrimination of the signals. The resulting signals then go to Time Memory Cells (TMC)[12] that hold the results of beam crossings spanning the previous 6 $\mu$s. A more detailed description of the PC, TEC and DC electronics with block diagrams illustrating the circuit logic is found elsewhere in this volume[5].

Both the BBC's and the $\mu$ID's electronic readout chains were constructed from commercial parts. The BBC uses shaping amplifiers and leading edge discrimination. The timing and pulse height signals are converted to voltages with Time-to-Voltage (TVC) and Charge-to-Voltage (QVC) Converters, respectively. The data is digitized during each beam crossing, stored in buffer memory and sent as input to the LVL1 trigger. A similar arrangement is used for the ZDC. The BBC electronics are described briefly elsewhere in this volume[3]. The signals from the $\mu$ID are amplified on the panel and drive differential signals over 30 m of twisted-pair cables to a crate based processing system consisting of a number of Readout Cards (ROC) that synchronize the many signals from the $\mu$ID's Iarocci tubes and contain constant-fraction discriminators to obtain digital data without contributing added timing slew due to risetime. The FEM's interface with the On-Line system and manage the transfer of the data from the ROC's after a LVL1 trigger accept signal has been received. The $\mu$ID's electronic readout chain is described in detail elsewhere in this volume[6].

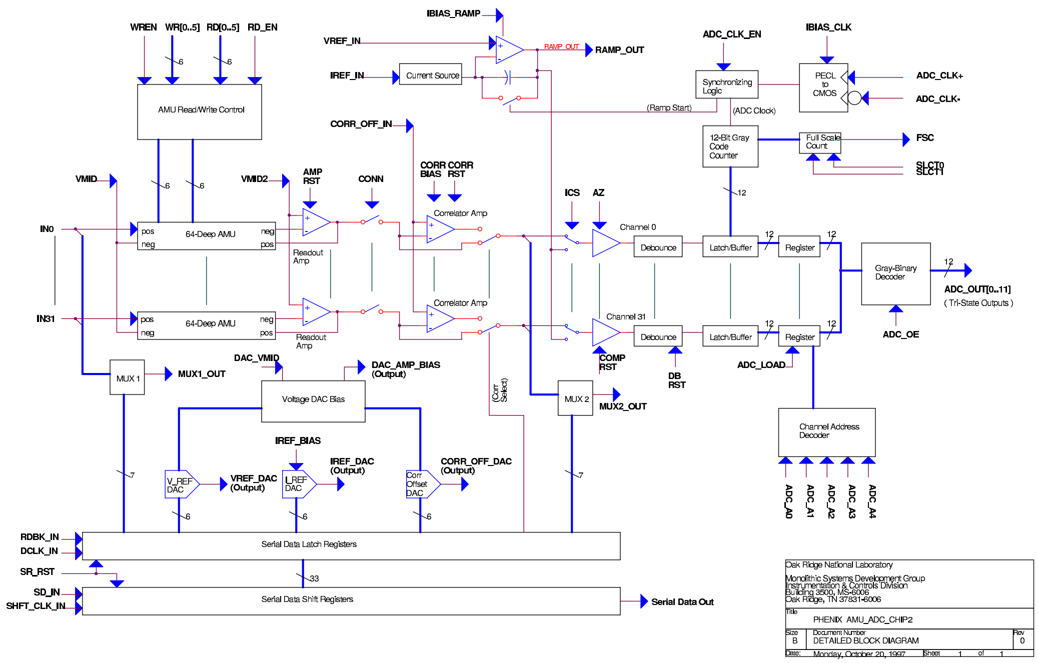



Fig. 4. Detailed block diagram for the AMU/ADC32 ASIC.

### *2.2 Systems Digitizing After Receipt of LVL1 Trigger*

The $\mu$Trk's, EMCal, RICH, MVD and ToF do not digitize their data in real time but store the data in a switched capacitor array called an Analog Memory Unit (AMU) and only digitize the data after a LVL1 trigger accept signal has been received. All of the above five units except the ToF use the AMU/ADC32 ASIC developed at Oak Ridge National Laboratory (ORNL) which is a 32 channel system that stores the data in the AMU[13] and digitizes the data with an 12-bit ADC[14] after receipt of the LVL1 accept. A detailed block diagram for the AMU/ADC32 ASIC is shown in Fig.4 and a photo of the same ASIC with the various components labeled is shown in Fig.5.

The above four systems process the analog data by different means before storage in the AMU's. The $\mu$Trk's use a Cathode Readout Card (CROC) to process the tracker data. The CROC's contain two ASIC's, namely the Cathode Preamps (CPA) and the AMU/ADC32's. Because of space constraints the FEM's are located 60 cm (45 cm for station 1) from the detectors. The $\mu$Trk's

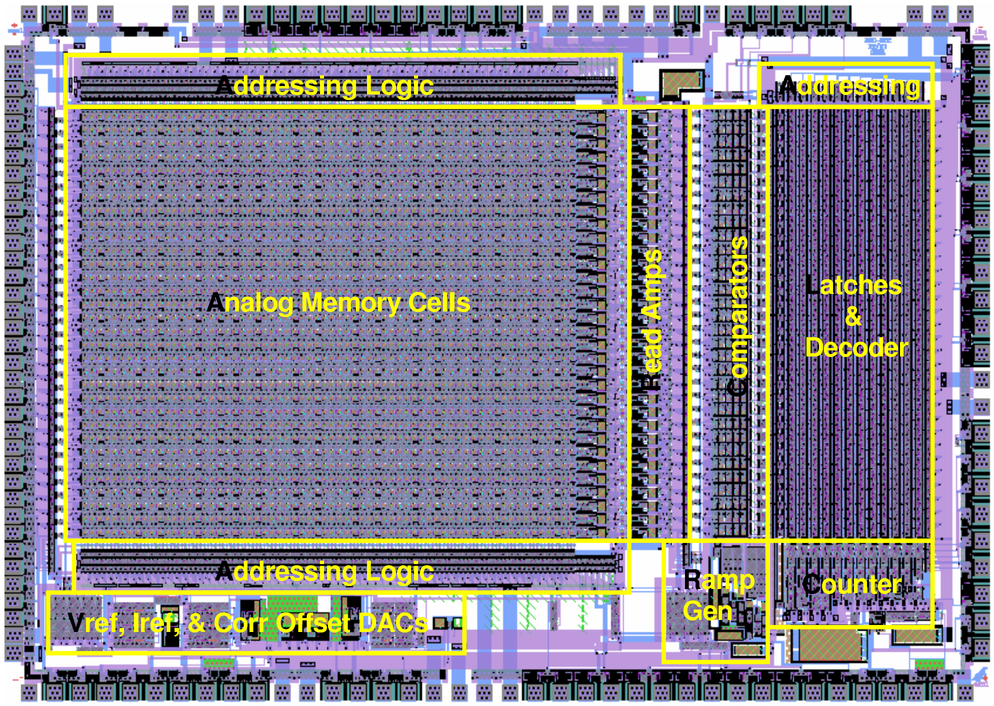



Fig. 5. Photo of AMU/ADC32 ASIC with functional blocks labeled.

electronics are described in detail elsewhere in this volume[6]. The electronics for the EMCal involves processing of both timing and pulse height signals from photomultiplier tubes in both the Pb scintillator and Pb glass detectors. The electronics chain contains amplifiers, constant-fraction discriminators, Time to Amplitude Converters (TAC) and the formation of trigger sums. A specialized 4-channel ASIC was developed to do all of this.. A special constant-fraction discriminator was developed at ORNL for the EMCal[15,16]. The EMCal electronics employ many specialized features and are described in great detail elsewhere in this volume[8].

The RICH detector employs photomultiplier tubes that produce only small amounts of charge so it is necessary to mount the preamps[17] in the RICH vessel. They are relatively inaccessible so the outputs of the preamps are routed to the FEM's which are composed of several boards located in a 9U VME crate outside the RICH vessel. This consists of Transition, AMU/ADC, Controller, Trigger and Readout boards. The FEM's provide further amplification, process timing and pulse height information, and form trigger sums, again using a specialized ASIC[17]. Details of the RICH electronics are given elsewhere in this volume[7]. The MVD detector uses an electronics chain similar to that used by other AMU/ADC type systems but it is highly specialized due to the fact that it is very important to have a minimum of mass near the beam interaction region. A specialized ASIC was developed to process pulse heights and provide a discriminator output. The pulse height is buffered and digitized by the AMUADC32. The power requirement for each of these ASICs was 1 mW per channel. The MVD electronics are described elsewhere in this volume[3].

The ToF system uses an AMU to store both timing and pulse height information but does not use the AMU/ADC32 chip like the other AMU based systems. The readout chain consists of discriminators, ASIC's, Time to Voltage Converters (TVC), Flash Analog to Digital Converters (FADC) and AMU's with the standard PHENIX output functions. The ToF electronics are briefly described elsewhere in this volume[7].

### *2.3 Organization of Data by Beam Crossing*

All triggered data is "time-stamped" so that all the data for a particular beam crossing from different subdetectors can be subsequently collated. The "time stamp" is in units of beam crossings. For analog systems the Address List Manager (ALM)[13] supplies the proper addresses. It operates at the beam-crossing rate of 106 ns and provides read/write control and data overwrite protection for delays associated with both the LVL1 trigger and the digitization latency. It also reorders the AMU addresses following conversion. The ALM is a recirculating loop composed of LVL1 Delay, ADC Conversion and Available Address FIFO's. The LVL1 Delay FIFO delays a given address for up to 48 beam crossings, the number being programmable. At the end of the delay the address is placed in the ADC Conversion FIFO to await digitization of the data from that beam crossing for those detectors digitizing after the receipt of a LVL1 trigger.. If that crossing is not qualified by the LVL1 trigger, the address is passed to the Available Address FIFO to be reused.

## 3 Front End Modules

FEE modules in a physics experiment require a high level of control. In physically large detectors such as PHENIX with multiple subsystems and a very large number of readout channels, it is necessary to exert localized control. This is accomplished for PHENIX using a large number of microprocessor-based Front End Modules (FEM) located on the detector. Even though the FEM's differ in detail for the various subsystems, it was possible to develop a modular topology so that many of the functional blocks are identical for a num-

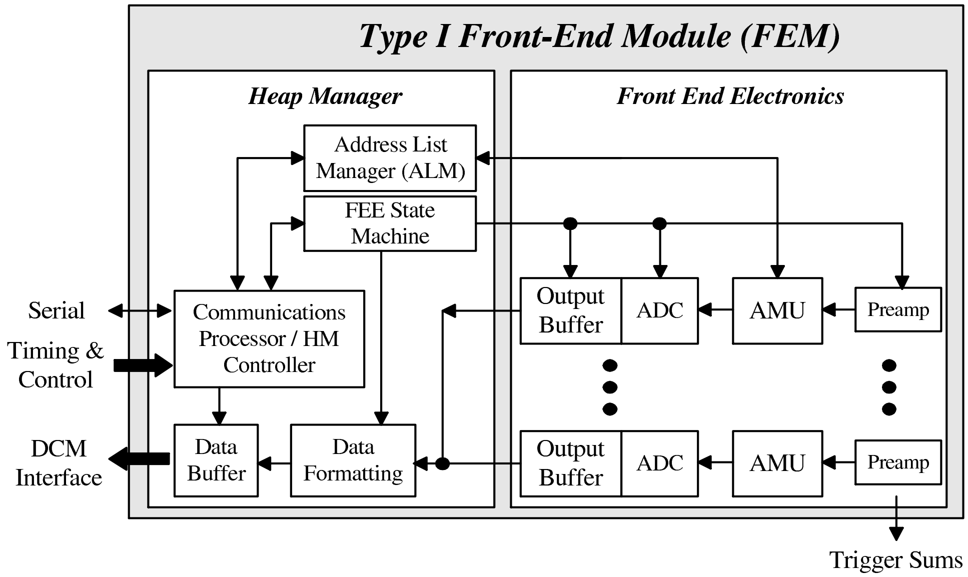



Fig. 6. Block diagram for the type of FEM used with systems where digitization occurs after receipt of a LVL1 trigger.

ber of subsystem types. Flexibility in programming is maximized by the use of flexible modular functions and implemented by the use of Field Programmable Gate Arrays (FPGA). The functions of the FEM's are controlled by a Heap Manager (HM)[18] that performs all the functions associated with FEE control such as mode interpretation and execution, FEE timing and control, receipt of LVL1 trigger and generation of test pulses, collection and formatting of the data, packet forming and communication of LVL1 accepted data to the DCM's. It also receives readout requests and provides a means of slow-controls input. The FEM's function synchronously with the beam crossings for preamp, DMU and AMU control while digitization, packet forming and transmission occur at a rate consistent with that required for data throughput. Thus each FEM has a five-deep FIFO for events accepted by the LVL1 trigger. A block diagram for a typical FEM used with AMU's is shown in Fig.6.

Each data channel has an integrated, programmable calibration element. It is required for FEE maintenance, diagnostics and calibration and can be selectively enabled or disabled. It also has the capability of injecting both constant levels and detector-like pulses of variable amplitude and delay into the signal processing input so that calibrations accounting for channel-to-channel response variations can be measured. This element also provides important survey and fault-tolerance capabilities.

### 3.1 FEM Interfaces

Serial control of the FEM's is carried out via ARCNet[19] which communicates with the FEM's by twisted-pair cable. The timing signals for the FEM's are transmitted from the Granule Timing Modules (GTM) over GLink based fiber optic cable. Data needed to produce a LVL1 trigger decision is communicated to the Local Level-1 (LL1) systems by fiber optic cable. After the receipt of a LVL1 accept via the GTM's the subsystem data is communicated to the DCM's via fiber optic cable. The extensive use of fiber optic cable for signal transmission between the IR and CH eliminates a large number of noise, cross-talk and grounding problems.

### 3.2 Mode Bit Subsystem Control

In PHENIX it is desirable to control and/or change the operation of each subsystem for each separate beam crossing. This control is carried out using mode control bits that are supplied to each FEM. The mode bit is supplied by the scheduler and sampled by the FEM on the rising edge of a beam clock tick and implemented on the rising edge of the next clock tick. Eleven bits for mode control are available and are discussed in detail elsewhere[18]. Three are used for timing signals, namely a signal at the beam crossing frequency, a signal at four times the beam crossing frequency and the LVL1 accept. One bit (run/halt bit) defines the run status of the FEM, two bits are used to control resets and another two are user defined. The remaining three bits are reserved for each subsystem and can be used or ignored.

### 3.3 System Timing

After every RHIC beam crossing, the FEM's corresponding to the BBC's, ZDC's, $\mu$ID's, EMCal and RICH send data to the LVL1 trigger. The trigger is pipelined and after a latency of 40 beam crossings issues an accept if the event is to be saved. This accept is transmitted to the GTM's that generate an accept signal that is transmitted to the FEM's in the IR. At this point the HM initiates readout and digitization of data from the AMU's. The digitized data from both the DMU's and AMU's is prepared for forwarding to the DCM's as discussed in the next subsection.

### *3.4 Data Collection and Formatting*

A primary function of the HM is to collect digitized data from the FEE's and prepare it for transmission to the DCM's. In the first step digitized data is retrieved from the DMU's or the AMU/ADC's. The HM then generates a data packet consisting of header and trailer words, event number, ADC data, module address and for analog type data the AMU cell number. The data packet is then transmitted to the DCM's. The time required to process a LVL1 trigger and completely transfer the event data packet to the DCM's is about 36 $\mu$s. The output of four FEM's are multiplexed into an individual DCM. The LVL1 accept rate is thus limited to 12.5 kHz to avoid pileup.

### *3.5 Serial Control via ARCNet*

Slow control and monitoring of the FEM's is carried out by ARCNet. The commands are sent from a workstation via optical fiber to hubs that communicate via twisted-pair cable with a total of 240 FEM's per hub acting as ARCNet nodes. The FEM's use a serial interface that provides for programming the FPGA's during FEM initialization and reading back the DONE bit indicating successful programming. Also the interface is used to set Data Acquisition and Control (DAC) thresholds, masks and gains, and run FEM diagnostic routines as well as access HM and FEE parameters. Values of various parameters can be downloaded and confirmed by serial readback. The serial interface also is used to determine if the GLinks are in lock during data collection.

## 4 Master Timing System

The main purpose of the PHENIX timing system is to distribute a harmonic of the RHIC accelerator beam clock to all the FEM's which participate in the data collection process in the PHENIX detector. The technical challenge was to perform this clock distribution over a 900 MHz optical serial link using HP optical transceivers, the optical component common to gigabit Ethernet technology. Other tasks which the timing system hardware performs are the synchronization of data taking with the beam crossing number, a task which is most important for the polarization program, the distribution of the LVL1 trigger accept signals and the management of the FEM's through the transmission of the mode bits to these devices.

Fig.7 is a block diagram illustrating the design of the PHENIX timing system. The RHIC clock is provided to PHENIX by the Accelerator Control group via

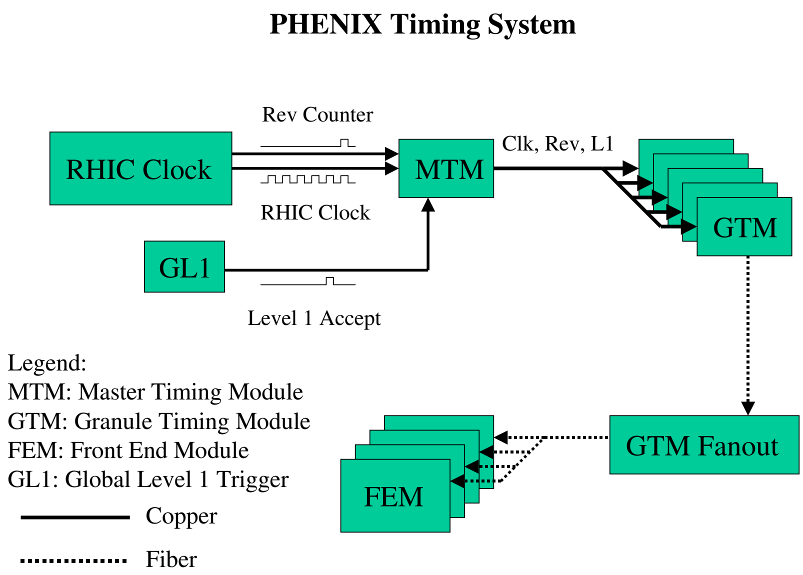



Fig. 7. Block diagram for PHENIX Master Timing System

optical serial links through a VME module called the v124 board. This board provides two key signals. The first one is the harmonic of the base accelerator clock and the second is a first bunch crossing fiducial marker which fires every time the first bunch buckets of each ring cross at the 4 o'clock intersection region. There is a copy of these two signals for each accelerator ring which are referred to as the Blue and Yellow rings. These four signals are fed into the Master Timing Modules (MTM).

The MTM's perform the first timing distribution by fanning out a copy of the RHIC clock (either the blue or yellow ring clock), selectable through a register setting on the GTM's. The fan out is performed using standard differential ECL as the signal transmission mode. The signals are received by the GTM's through a transition card in the back of the 9U VME crate where the GTM resides. The bunch 1 fiducial signal is fanned out as well to all the GTM's by the MTM but LVDS is used as the transmission media through CAT-5 twisted pair cabling. The MTM has an internal phase lock loop designed by the Collider-Accelerator department which minimizes the clock jitter. It provides 100 ps FWHM clock jitter filtering. This same phase lock loop component is replicated on the GTM's to provide for further jitter filtering introduced through the MTM to GTM fanout process. Finally, the MTM provides a "Global Start" function to the GTM's. This global start function enables sequencers on the GTM's to start executing on the next rising edge of the bunch 1 fiducial signal. This is the method by which all the schedulers on each

GTM are synchronized resulting in the synchronization of the initialization and data collection of all the FEM's distributed throughout PHENIX.

The GTM is the second stage in the clock distribution system. The GTM's provide the clock, mode bits and LVL1 trigger accepts to the FEM's. The GTM's also provide the path by which the granule busy signals are returned to the GL1 from either its internal count N busy state or from the DCM busy signal which is fed into the GTM through a transition card which sits in the rear slot of the VME crate. The clock and mode bits are converted from a 20 bit parallel stream being clocked at 9.4 MHz (the beam clock rate) up to 4 times the beam clock rate and then serialized into a 900 MHz stream.

The final stage of clock distribution is the 1 to 8 fiber fanout on a 6U board. A Semtech SK1904AFT fanout component is designed to operate at a rate of up to 2 GHz. The optical signal is received through the HP transceivers. The serial analogue signal from the transceiver is then fed into the clock fanout chip. The eight outputs of this chip are fed into the sending end of the transceiver where the analog clock is then converted back into photons. This was the preferred method of clock fanout since it was possible to set up the GLink fanout crates in the collision hall. Therefore, only 1 fiber per granule (or GTM) was needed to transport the signal from the IR to the CH. This shortened the total length of the fiber cables needed to run the DAQ system by a substantial amount. The timing system for the PHENIX detector was fully tested during the the first year of RHIC running. During that data taking run the performance of the MTS exceeded design specifications.

## 5 Level-1 Trigger

The responsibility of the Level-1 trigger (LVL1) is to select potentially interesting events for all colliding species and provide event rejection sufficient to reduce the data rate, for example by as much as 50 for p-p collisions, to a level that the PHENIX DAQ can handle and generate seed data for higher level triggers. Also the LVL1 trigger supports the partitioned running of the PHENIX detector to facilitate commissioning of new detector subsystems.

### *5.1 Level-1 Architecture*

The PHENIX LVL1 trigger is a parallel, pipelined system designed to provide a highly configurable trigger capable of meeting the demands of the PHENIX physics program. It consists of two separate subsystems. The Local Level-1 (LL1) system communicates directly with participating detector systems such as BBC[3], MuID[6], ZDC[4], EMCal[8] and RICH[7]. The input data from these detector systems are processed by the LL1 algorithms to produce a set of reduced-bit input data for each RHIC beam crossing. The Global Level-1 (GL1) system receives and combines this data to provide a trigger decision. In addition, busy signals (both global and trigger specific) are managed by GL1.

The PHENIX detector system is divided into two sets of elements: granules and partitions. A granule is the smallest detector element that communicates with the PHENIX timing and control system via a GTM. Thus a granule may be a subsystem or a portion of a subsystem. The GTM's distribute the local 9.4 MHz RHIC beam clock as well as control bits and event accepts to the granule. A partition is a programmable collection of granules that share both busy signals and LVL1 triggers. This makes it possible to run individual or small collections of subsystems in parallel for detector commissioning.

Both the LL1 and GL1 systems are implemented in nine layer 9U VME-P format boards using Actel[20] and Xilinx[21] FPGA logic. In order to minimize design time and effort, all LVL1 boards share a great deal of common infrastructure logic (for example, VME communication). All boards incorporate the ability to inject test patterns at their inputs under both VME and GTM control. In addition, the data flow on the board can be monitored at the beam clock level via a set of synchronous FIFO's that store 1024 beam crossings worth of data from all stages in the board algorithm. The combination of test patterns and monitor FIFO's allows real-time monitoring of the LVL1 trigger and permits users to quickly pinpoint hardware failures or misconfigurations. A separate set of synchronous FIFO's store additional information for inclusion in the data stream for accepted events.

### *5.2 Global Level-1*

The GL1 system is shown in Fig.8. The input data from the LL1 systems are placed on a custom P3 backplane in the 9U VIPA GL1 crate by a set of transition cards called 6Rx. The 6Rx transition card can be configured to direct data from its six input connectors to different locations on the P3 backplane. The Standard transition card receives the beam clock and control data from the GTM for GL1 and distributes this data to all cards in the GL1 crate, as well as providing a readout path for accepted event data to the DAQ. The 4x1 transition card provides a readout path for LL1 data and is discussed in detail below. The LL1 electronics often produce bit sums or energy sums. The number of resulting bits is larger than is useful for GL1 decisions and is thus judiciously reduced in number to a more manageable address size. The resulting smaller bit space is called "reduced bits" in PHENIX.

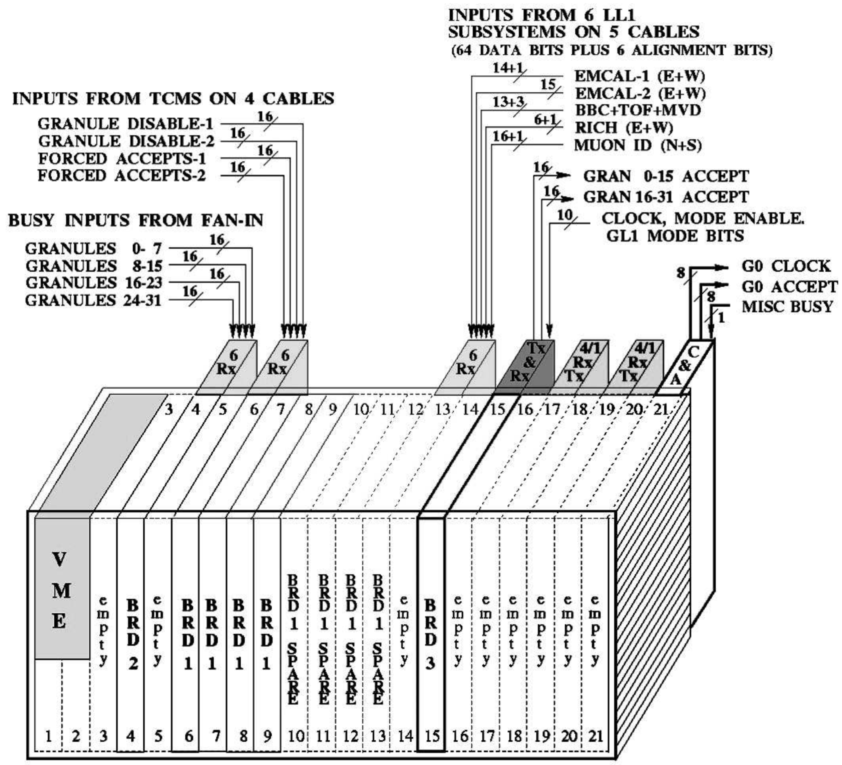



Fig. 8. The GL1 trigger system, showing the GL1-1, GL1-2 and GL1-3 boards, as well as the accompanying transition cards.

The heart of GL1 consists of three 9U VME-P boards. GL1 Board 1 (GL1-1) receives the reduced-bit input from the P3 backplane and combines the 130 bits of trigger input into eight groups of twenty bits using two programmable ICube IQX-240B crossbars[22]. Each of the eight groups of twenty bits is used as the input address to a 4-bit static RAM. Each output bit of the static RAM is called a raw GL1 trigger. Thus, by programming the crossbar and lookup table (LUT) static RAM, 32 separate triggers can be configured on a GL1-1 board.

The GL1 system is designed for four GL1-1 boards, for a total of 128 triggers, however, the GL1 crate can accommodate up to eight GL1-1 boards for a total of 256 triggers. The generation of a raw trigger from the input trigger vector takes one beam clock. After a raw trigger is generated, it is compared to a trigger mask and the partition busy vector from the P3 backplane (as generated by the GL1-2 board, described below). The partition busy vector is mapped to raw triggers using a second programmable crossbar. If a raw trigger passes this test, it is called a live trigger and is compared to a programmable scaledown counter. If a trigger passes the scaledown test it is called a scaled trigger, and is sent through a third crossbar to drive the partition accept line on the GL1 backplane that owns the trigger. The second crossbar is contained in an IQX-

128B crosssbar chip, while the third is implemented in custom Actel logic due to the required "many-to-one" mapping that was not possible in our configuration with the ICube devices. The time required from raw trigger to a partition accept is an additional beam clock. The combination of programmable lookup tables and crossbar mappings within GL1-1 provides a highly configurable system that can support the full range of possible PHENIX configurations, from one trigger and one granule per partition (as is likely during detector commissioning) to all granules in one large physics-data partition.

The GL1 Board 2 (GL1-2) monitors the partition accept bus and holds the corresponding partition busy line on the P3 backplane high for four clock ticks. This "dead-for-4" restriction is required by the readout and conversion electronics for the detector systems and is determined largely by signal collection times in the tracking detectors. In addition, granule busy inputs from the P3 backplane are mapped to the partition busy vector on GL1-2 using the same Actel logic crossbar used in GL1-1. It is important to note that the busy bus is used as an input to GL1-1, so that the feedback loop between GL1-1 and GL1-2 must take less than a single beam crossing. GL1-2 also distributes two global busies that can hold off all GL1 triggers.

The GL1 Board 3 (GL1-3) monitors the partition accept section of the P3 bus and converts the partition accept vector into a granule accept vector using a crossbar mapping. The granule accept vector is sent to the PHENIX MTM for distribution to the GTM's controlling each granule to initiate readout of the detector. In addition, GL1-3 initiates readout of the LL1 event data from the LL1 subsystems via the 4x1 transition cards which multiplex multiple readout streams of data onto a single cable. A complete readout of the PHENIX detector systems will take a minimum of 40$\mu$s. This time may be longer depending on the number of DCM's in the PHENIX DAQ. In order to keep the event rate to within the limitation of the PHENIX DAQ, GL1-3 keeps count of the accepted events waiting for readout and imposes a programmable minimum time for each readout cycle (40$\mu$s in the full PHENIX) to limit the LVL1 accept event rate. If the number of events reaches a programmed maximum (up to five), a global busy is asserted, holding off further triggers from GL1-1 until the accepted event count reaches a programmable lower limit. This governor prevents the system from oscillating between busy and non-busy states when the DAQ is run close to its maximum data rate. In addition GL1-3 contains a real-time clock for generating a time stamp, a 64-bit beam crossing counter, a bunch crossing counter and 32 16-bit granule accept counters, all of which are sent to the accepted event data stream.

An additional board (GL1-1P) has been designed for luminosty monitoring on a bunch-crossing basis. This board is similar to GL1-1 and contains the same crossbar and LUT logic to define trigger conditions from reduced-bit inputs, but does not generate scaled triggers. Instead, the trigger results are used to increment luminosity scalers that are read out and cleared for each accepted GL1 physics trigger. This board includes 120 sets of such scalers for recording trigger counts separately for each of the up to 120 beam bunches in RHIC. The polarization direction of each individual bunch can be specified at injection during polarized p-p running.

### *5.3 Local Level-1*

The majority of LL1 systems communicate with the detector systems over fiber data links to the detector FEM's. An example is the BBC system[3]. For the BB LL1 this input TDC hit data is received on an 8Rx transition card, which can accept data from eight fiber connections for a combined throughput of 9.6 Gbit/s. The BB LL1 system consists of two types of 9U VME-P boards. A BB Board 1 (BB-1) receives data from the 8Rx transition card and performs some preliminary trigger processing by counting the number of hit tubes in one arm of the BBC and calculating the mean time for all hit tubes above threshold. A BB Board 2 (BB-2) receives the BB mean times from each arm, and calculates the mean time and time difference (vertex location) for the interaction. These data are passed through a LUT and reduced to two bits that are sent to GL1 for use in making a trigger decision.

In the case of the $\mu$ID[6] LL1 a much larger amount of data must be collected and processed in parallel to facilitate the use of an adaptive road finding algorithm to identify muons that penetrate the $\mu$ID steel to different layers. Each projection of the $\mu$ID (horizontal and vertical) is processed on a single board, which receives 20 fibers for an aggregate data rate of 19.2 Gbit/s. The input data for a beam crossing is demultiplexed from the fibers and processed through the tracking algorithm entirely within Xilinx FPGA logic. The high logic density and reprogrammability of the Xilinx logic allows the implementation of tracking algorithms that are specific to the physics of interest. One of the two boards in the $\mu$ID LL1 is selected through software configuration to send the resulting data to GL1.

For p-p running additional LVL1 triggers are needed due to the high interaction rates. A EMCal/RICHtrigger (ERT) will use the same LL1 hardware as the $\mu$ID LL1 system with a different set of algorithms programmed into the Xilinx logic. These algorithms will be used to combine EMCal and RICH information to provide a number of different triggers for photons, electrons and pions.

There are also LL1 systems which communicate with GL1 without a dedicated LL1 system. An example of this is the ZDC[4], signals from which are used to form a coincidence between the two arms of the ZDC in NIM electronics. This coincidence is suitably delayed and placed on the GL1 backplane by a special transition card that receives a TTL signal, resynchronizes the signal with the local GL1 clock and places the signal on the GL1 backplane. Once properly timed these input bits can be used in the trigger decision in the same manner as all other LL1 information.

The LVL1 primitives are included in the data stream for accepted events and can be used as seeds for LVL2 trigger processing. For example, a LVL2 muon trigger can use the LVL1 $\mu$ID primitives as part of a combined muon trigger and avoid the need to recalculate these in LVL2.

The PHENIX LVL1 trigger provides a highly configurable, pipelined, parallel trigger system. The system consists of LL1's which processes detector information into reduced bit data and GL1 which makes trigger decisions, coordinates busies and issues a granule accept vector to initiate detector readout. Together GL1 and LL1 provide the PHENIX experiment with triggers that select potentially interesting events, reduce the trigger rate to conform with the rate limitations of the PHENIX DAQ and provide seed information for higher level triggers.

## 6 Data Collection System

The DAQ system for the PHENIX experiment is designed as a pipelined system with simultaneous triggering and readout. The maximum average LVL1 trigger rate is 25 KHz and the RHIC beam crossing clock runs at 9.4 MHz. At the maximum LVL1 trigger rate, the FEM's send over 100 Gbytes of data per second. The PHENIX Data Collection Modules (DCM) are designed to receive this large uncompressed event fragment data volume, provide data buffering, perform zero suppression for all detector system data, perform error checking, perform data formatting and output the compressed data to the PHENIX Event Builder (EvB). They provide buffering for up to five LVL1 events. The DCM's have been described in an earlier paper[23]

One of the main design constraints for the DCM's is that they must be able to handle the full LVL1 trigger accept rate for very different detector occupancy conditions. The calculated relations between colliding system, data rate and data size are shown in Fig.1 in the Introduction section of this article. The FEM's design was heavily influenced by considerations of p-p running where the interaction rate is expected to be approximately 10 MHz (at 10 times design luminosity). In these interactions the detector is relatively quiet with a maximum of tens of tracks. In Au-Au collisions the maximum interaction rate is 13 kHz with a 2.5% estimated detector occupancy (with hundreds of tracks). The DCM data receiving part of the system must be able to process this large data volume. The heavy ion colliding systems present conditions that large scale high energy experiments do not encounter. These considerations influenced our design by requiring the DCM's to have a large amount of processing power and a buffered pipeline system.

### *6.1 Data Collection Modules*

The experiment has roughly 375,000 channels of electronics and the FEM's continuously sample the detector signals and store them in either digital or analog memory. After they receive a LVL1 trigger accept signal, and after a 4 $\mu$s latency, the corresponding event is transferred to the DCM system. In order to minimize any deadtime, the system is designed to continuously record signals while digitizing and transferring the previously accepted events. The FEM's act as slaves and do not perform zero suppression on the accepted LVL1 event. The data passing out of the FEM's is transmitted at rates up to one Gigabit/s per fiber connection. Each DCM is connected via fiber-optic cable to four specific FEM units. The DCM system resides in the PHENIX CH. It consists of the order of 200 DCM boards with over 16 VME crates in multiple electronics racks. The DCM boards reside in custom crates and communicate over custom VME64X backplanes. An additional partition module (PAR) is used to extract data from the custom dataway to send event fragments to the EvB for further assembly.

The DCM data flow logical block diagram is shown in Fig.9. There are five major logical units: the compressor, the DSP block, the VME interface, the LVL1 interface and the data output port. Each DCM board has four parallel data input streams with fiber receivers, data compressors and digital signal processors. A fifth DSP is used to merge data from the four compressors and do alignment checking.

The DSP's are used to handle data formatting, alignment and error detection after zero-suppression. The DSP's can also calculate trigger primitives, for example by applying gain corrections to convert ADC's into energy values. However, it was felt that DSP-base zero suppression algorithms would not be able to handle the maximum anticipated rates. Therefore, we chose an FPGA as the computer engine for that step.

#### *6.1.1 Compressor Ports*

The DCM's perform zero suppression by comparing data with preset threshold values. During normal running conditions, only data above the threshold is kept. The data transfer rate between FEM and DCM is two or four times the beam crossing clock rate, i.e. 18.8 or 39.6 MHz. The data words are either 16

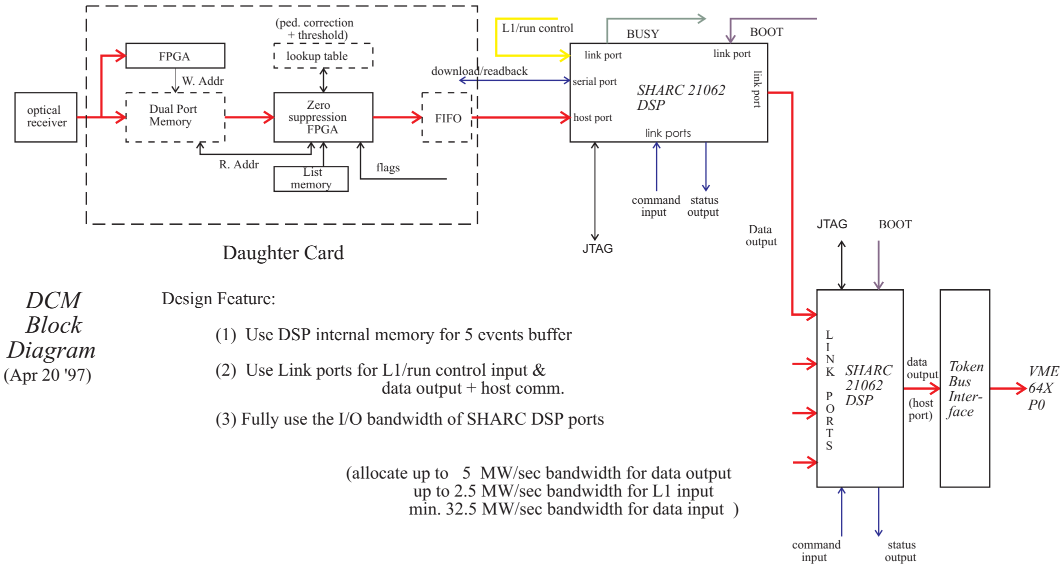



Fig. 9. Logical block diagram of the DCM data flow.

bits or 20 bits wide. This data transfer rate is faster than most commercial DSP's can handle, however, it is still within the ability of FPGA technology. Note that technology choices were made around 1996-7.

The compressors are built as daughter cards to achieve uniformity of the remainder of the DCM board and easy upgradeability. AMP 0.8 mm free height 100 pin connectors are used to connect the daughter cards to the main board.

The DCM's use four different algorithms to zero suppress data from over ten different detector types in the PHENIX experiment. One category of the detectors are those using AMU's to record input signals in an array of capacitors at the FEM's. These detectors include the EMCal[8], ToF[7], RICH[7], $\mu$Trk[6] and MVD[3]. The array serves both as a LVL1 trigger delay and as a buffer for accepted LVL1 data while waiting for digitization. After receiving a LVL1 trigger, the charge on the corresponding capacitor is digitized. Different capacitors in the AMU array will have slightly different pedestals. The PHENIX AMU system has 64 capacitor cells in each array. In the DCM individual AMU cell corrections need to be applied before threshold values can be compared. The capacitor cell numbers are part of the FEM data packet and can be used for a cell dependent pedestal lookup table.

The compressor port in the DCM consists of an optical receiver, HP GLINK receiver, dual-port memory and compressor FPGA. One list memory and one pedestal table are also built into the compressor daughter card. Data is stored in the dual-port memory after being received from the optical link. The dual-port memory has 4k memory and is divided into 2 pages, one for the incoming data while the previous event, that is stored in the other page, is being processed. An 8k by 16 bit list memory is used by the compressor FPGA to regroup data in the dual-port memory. The lowest 12 bits of the 16 bit data field serve as the dual-port memory address. The upper 4 bits are used as a label to classify the data. The reordering feature is important in order to provide decoupling of the FEM's and DCM's in the exact data order. The label in the list memory merged with the data from the dual-port memory classifies the data into the compressor FPGA as well as the DSP.

The compressor FPGA uses channel number, through an incremented counter, and the AMU cell number, contained in the data stream, to address the pedestal memory table. The pedestal value is then removed from the data. The result is compared to the threshold value as a function of channel number. An ALTERA FLEX 10k20 208 pin package is used as the zero suppression engine [24]. The internal memory of the FPGA is used for the threshold memory table with a maximum size of 512k by 12 bits. The pedestal memory carries 32k by 8 bits. The system runs fully pipelined at 20 MHz.

The second category of detectors output patterns of ones and zeros indicating hits on a detector element or signal in a given time bucket. For these detectors, including the DC[5], TEC[5], PC[5], BBC[3], $\mu$ID[6] and ZDC[4], the compressor board is simpler and does a pattern comparison in order to decide whether to pass through a given data word.

### *6.1.2 Digital Signal Processors*

We use the Analog Devices SHARC Digital Signal Processor (DSP) that has the advantage of high speed host ports, six 4 bit wide link ports and two serial ports [25]. The total I/O bandwidth is 40 Megawords per second through all I/O ports. We also take advantage of the large internal memory, both for program and data buffers. The 21062 SHARC DSP was chosen which has 2 Megabits of dual-port memory. The internal DMA controller and dual-port memory allow the computing unit to operate independently of the I/O. The computing unit runs at 40 MHz. In order to run the compressor port at speed, there is no backflow control. Careful calculations and the development of board testing stations were important to check this feature.

The SHARC DSP's also provide buffering for up to five completely un-zero suppressed FEM data events. In normal (zero suppression) mode, this amounts to a buffer for more than 50 events. This enables us to avoid providing high speed memory on the compressor board. A ring buffer and an automatic DMA chain are used inside the DSP. Busies are raised once the DSP has less than six uncompressed event buffers available as up to five more events may be pending digitization and readout in the FEM at any given moment. The DSP's have four user defined flags. One of these flags is used as the "busy".

The link ports are 4 bits wide and run at a maximum speed of 40 MHz. This translates to 5 Megawords/s for a 32 bit word transfer. The input link port speed is controlled by the sender and the output speed is controlled by the DSP clock. Link port 2 is used for receiving LVL1 information. The VME interface uses ports 3 and 4 to read/write data into the DSP. Link port 4 is used through an FPGA interface to boot the DSP. Both the LVL1 data link port and VME input link port are slowed down to reduce ringing. Link port 5 is used to transfer data between the compressor port DSP and the extra DSP which merges data from the four compressor port DSP's. This limits the average compressor output port to 5 Megawords/s.

### *6.1.3 LVL1 Port*

The LVL1 trigger system described earlier in this paper generates a summary data packet for every accepted event. This data packet contains information on the LVL1 triggers that fired, LL1 trigger primitives, clock counters and scalers. This information is transmitted to the DCM's via LVDS cable on every LVL1 accepted event. We have designed a cable fan-out system so that the LVL1 trigger packet information is sent to all DCM boards and thus is available to all DCM DSP's. This feature allows for future trigger primitive calculations in the DSP's to be selectively performed on specific events of interest.

The LVL1 data is received via LVDS cable by a partition module. This module then sends the data to the DCM's through the VME64X backplane by a token passing method. The maximum backplane transfer rate is 20 MHz. The DCM's use an ALTERA 7k 128 cell FPGA + 64Kx1 SRAM to receive the data with a short FIFO to back it up.

The LVL1 data is 16 bits wide with one address tag bit. This data packet is divided into many sub-packets separated by an address word. The address word is recognized by the address tag bit. The upper 8 bits of the address word contain sub-packet identifier information. The lower 8 bits contain the lower 8 bits of the global event counter. The DCM FPGA is able to selectively pass the LVL1 data to the DSP's on a packet-by-packet basis. The filter algorithm uses the address word of the packet plus the word counter in the sub-packet as the address word of the 64Kx1 SRAM table. The output of the table, one bit YES/NO, will determine whether the data is passed out or not.

The filtered data is then passed to a FIFO and broadcast to the four DSP's on the DCM board using a 4 bit wide DSP link port. One FPGA is used as an interface between the FIFO and the four DSP link ports. Fig.10 shows a block diagram of the DCM-VME and LVL1 data interface. One of the LVL1 data sub-packets contains the event counter. By comparing this event counter with the FEM event counter contained in the data packet, the DSP's are able to check data alignment within the system. The DSP's use compressed FEM data

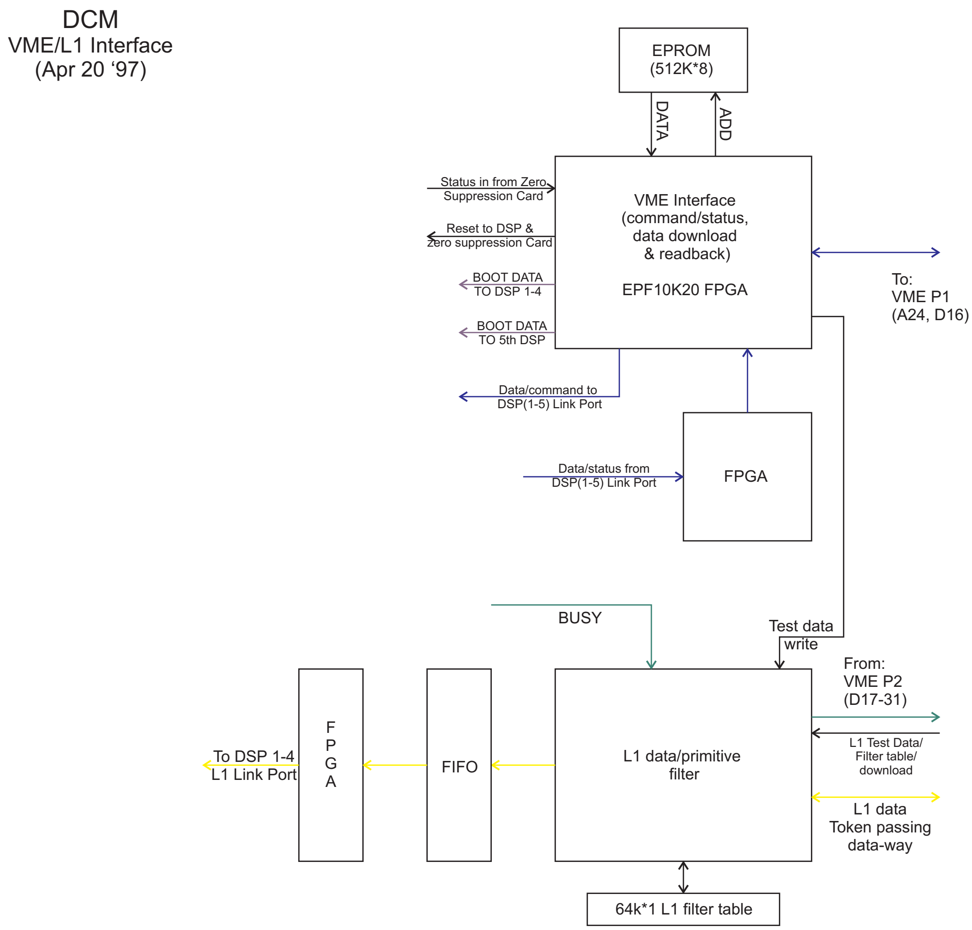



Fig. 10. Block diagram of the DCM-VME and LVL1 interfaces.

and LVL1 data in order to renormalize the FEM data and calculate trigger primitives for the higher level triggers.

### *6.1.4 VME Interface*

The DCM's use VME as the path for downloading and slow control/read back. The host interface on the DSP's is used for the event data path so the communication from the VME controller to the DSP's has to use the link ports. On each DSP, one port is set up to be the read port and one port to be the write port. One 240 pin ALTERA FLEX 10k20 chip is used to interface between the VMEbus and the DSP link ports. The DCM interface uses VME A24 D16 transfers with a user defined address modifier code. The 24 bits on the address field are decoded into three portions: bit 23 is the broadcast bit, bits 22-18 are the geometric address word, and bits 17-14 are used as a command code. Commands fall into three classes: read the board status, define to which DSP the interface is connected and read/write data to a DSP. The SHARC DSP's use 32 bit data and 48 bit program instruction words. It takes two or three 16 bit operations to communicate one DSP word from the controller.

Unlike most commercial VME boards, the crate controller and the DSP exchange information through an interface FPGA. The real communication between the DSP's and the VME controller is through the FPGA program. This feature provides flexibility but it lacks the ability to directly probe the DSP memory. During power-up the VME interface FPGA also connects to a 64kx8 EPROM which contains programs for all DSP's on the board.

#### *6.1.5 Output Port*

The fifth DSP collects data from the compressor port DSP's (1-4) through link ports. It reformats the data before sending it out. An ALTERA 7K 192 cell FPGA is used to interface the DCM board with the VME64X backplane P0 dataway. The token passing method is used to pass data in the dataway. This allows output data to get built in the right order with the minimum overhead. The DSP host bus is used to connect to the FPGA. Because of the bandwidth sharing between link ports and host bus, the maximum bandwidth is around 20 Megawords/s, if the input link port is running at full speed, with 40 Megawords/s burst speed.

#### *6.1.6 DCM Crate*

The DCM boards have a 6U x 280 mm VME form factor. The P0 connector on the VME64X backplane is used as a token passing 32 bit wide user pin defined dataway. Eight signals including Strobe (clock), Valid, Token and Hold are used in controlling data from one DCM to the next DCM. The dataway is designed to run at 40 MHz. The token passing scheme enables the events to be built in the right sequence and at high bandwidth. A similar structure is also constructed in the P2 user defined space for passing LVL1 trigger primitives. Due to insufficient power provided to the VME backplane, 7 additional power pins are placed in the P2 user defined space. Busy signals are accomodated in a similar fashion. A partition board is used to interface the crate to the EvB and LVL1 trigger system. This board determines how the DCM's are partitioned in terms of data flow. Fig.11 shows a block diagram of the DCM crate.

### *6.2 Recent Progress*

In the first round of fabrication, we built 110 DCM boards in 8 VME crates with custom backplanes. The system was used to readout the PHENIX detector in the RHIC year-1 run in the summer of 2000. Since that time we have carried out two more fabrication runs, thus more than doubling the number of DCM boards. The zero suppression algorithms for all detector subsystems

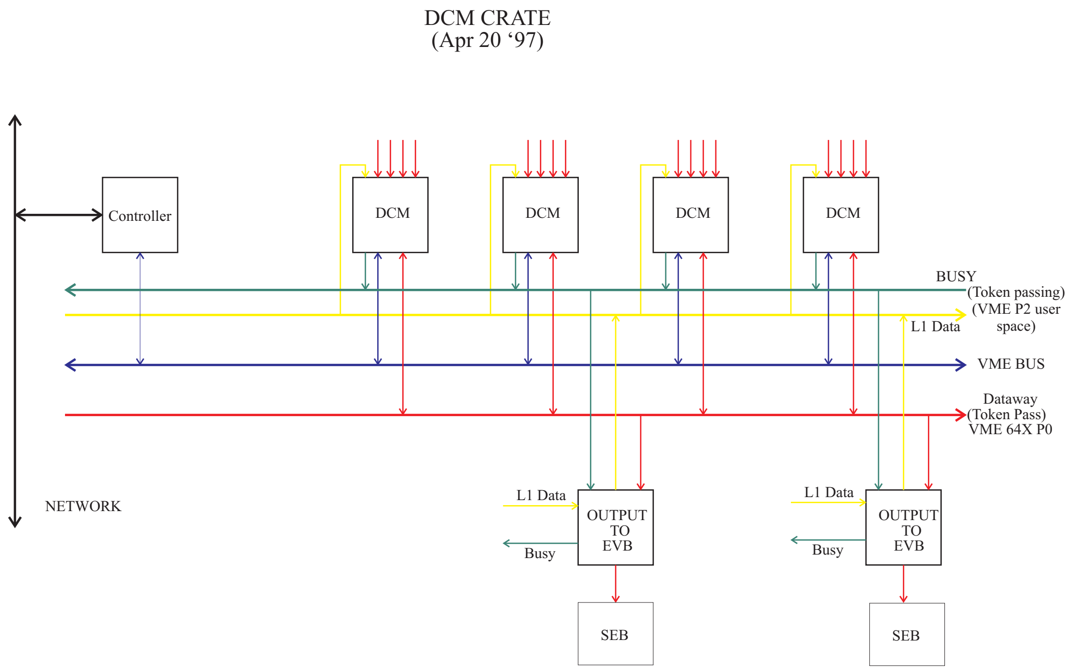



Fig. 11. Block diagram of the DCM crate. The controller is the Power PC Crate Controller and the "Output to EVB" is the block diagram for the Partition Module. The SEB's shown in the figure are outside of the DCM crate

were successfully implemented in compressor FPGA and used in the RHIC year-2 run in 2001. The performance completely met specifications.

The DSP coding was done in C with significant embedded assembler code in key data processing sections. The interrupts are currently using a standard providing a C function call. The overhead for these interrupts was reasonable for the event rates observed in the first two years, but further work on possible faster interrupts and processing will be carried out in the future.

We encountered mechanical issues with the VME64X P0 backplane where the exposed P0 pins are the first to interconnect, thus resulting in a bent pin risk. We resolved the issue with better crate mechanical specifications and a P0 connector guide on the DCM board. Overall the performance has been good and we expect the DCM's to be completely adequate for handling data rates well beyond the RHIC design luminosity.

## 7 Event Builder

The two primary functions of the PHENIX Event Builder (EvB) are to perform the final stage of event assembly in the PHENIX DAQ and to provide an environment in which LVL2 trigger processing is performed[26–28]. The EvB receives many parallel data streams from the PHENIX DCM's, assembles fragments from each data stream into complete events, performs LVL2 trigger processing on the events and transmits accepted events to the PHENIX Online Control System (ONCS) for logging and distribution to monitoring processes.

The two most important measures of EvB performance are the maximum event rate and the maximum aggregate data rate at which events can be assembled. The maximum event rate is determined primarily by per-event overheads in the receiving and distribution of events while the aggregate data rate that can be achieved is largely determined by the performance of the data transmission technology in the EvB. The ultimate performance goals of the baseline EvB were to process events at an input rate of 12.5 kHz and to handle aggregate data rates as high as 500 Mbyte/s. The event rate limit was determined by the maximum LVL1 trigger rate achievable with the $\times 2$ multiplexed readout of some of the PHENIX detectors, while the 500 Mbyte/s goal allowed for a factor of two over maximum expected data rates in early RHIC operation. Because of the large events obtained during heavy ion operation ($\geq 150$ kbyte for minimum bias Au-Au collisions) aggregate data rates in excess of 100 Mbyte/s were expected during the first few years of RHIC operation so the data rate became the most important performance goal during the initial implementation of the EvB. Because PHENIX had established a maximum average transfer rate to the RHIC Computing Facility (RCF) of 20 Mbyte/s, LVL2 trigger rejection was initially assumed to be required whenever the aggregate EvB data rate exceeded this data rate. However, a local buffering system was established for PHENIX that provided for greater instantaneous recording rates to compensate for luminosity decay during a store and for PHENIX/RHIC downtime. This system was designed to log data at rates as high as 60 MByte/s so LVL2 rejection would then only be required for rates exceeding this higher threshold.

### *7.1 EvB Design*

In the early stages of the design of the EvB it became clear that the number of input data streams that the EvB would have to accept would vary with time due to the deferred construction of some of the PHENIX detector systems and/or FEE's. It was also clear that the event rate at which the EvB and trigger systems would have to perform would increase over time with the luminosity growth of the collider. The EvB was thus designed to be scalable on both the input and output sides, allowing additional input streams to be added and trigger processing power to be provided without substantial disruption to the system. The EvB was also required to satisfy the partitioning requirements of the PHENIX DAQ system such that configurable collections of input streams (granules) could be read out independently. These considerations led to the choice of a passive switch-based EvB in which a set of "Sub-event Buffers" (SEB's) would receive and buffer data from the DCM's and transfer them

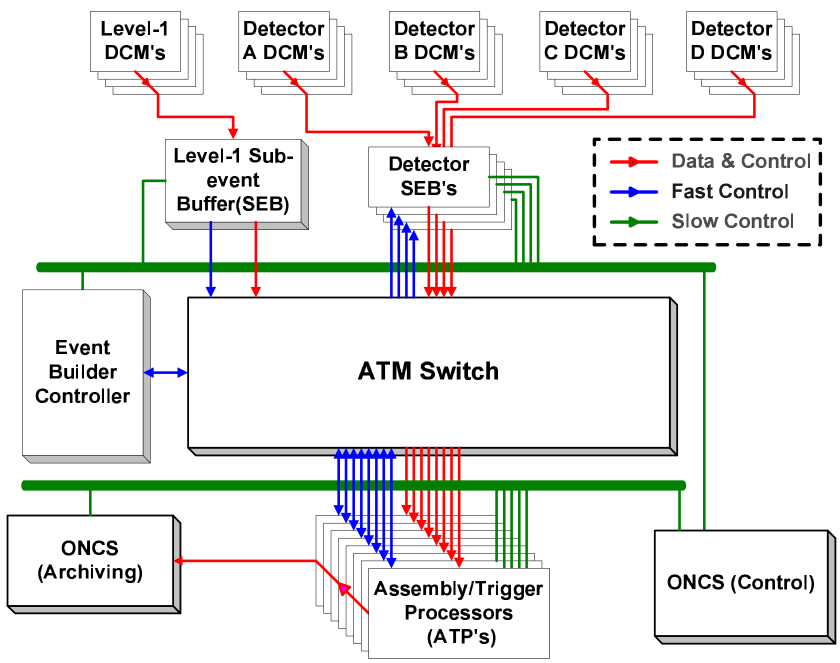



Fig. 12. Block diagram of the EvB architecture

on request to a set of "Assembly/Trigger Processors" (ATP's) under the control of the "EvB Controller" (EBC). This implementation, illustrated by the block diagram in Fig. 12, is similar to and is based on the design of the ATLAS demonstrator-C LVL2 trigger system.

A central problem in the implementation of switch-based EvB's is that data from a large number of source processors must converge on a destination processor. Assuming that both sources and destinations have the same link speed to the switch and that the sources transmit at maximum speed, the combined data rate from the $N$ source processors will overload the output link rate by a factor of $N$ resulting in either loss of data or continual attempts at retransmission. Thus, the system must be designed to either prevent two sources from sending to the same destination simultaneously or must allow the sources to send data to a particular destination at lower than the maximum link transfer rate. A variety of solutions to this problem have been tried with varying degrees of success [26,28]. However, the problem is easily solvable via the use of Asynchronous Transfer Mode (ATM) technology and its "quality-of-service" features. The ATM communication protocol splits large transfers into 48-byte (+ 5 byte header) "cells". Communication between two nodes on the ATM network occurs over a "virtual circuit", a long-term connection established through the switch between the input and output nodes. The ATM protocol allows the data transfer rate on a particular virtual circuit to be limited at the origin node by the hardware. Thus, if bandwidth limits are imposed on all virtual circuits destined for a particular ATP such that the combined data rate on those circuits does not exceed the ATP link speed, no output port overloading will occur at the switch. If the SEB simultaneously transmits to many ATP's the ATM Network Interface Card (NIC) will interleave the data streams controlling the bandwidth on each virtual circuit but allow the full bandwidth of the SEB link to be utilized.

The data flow through the switch was implemented using a "pull" architecure in which the input nodes (SEB's) transfer data to the destination nodes (ATP's) on demand. The controller allocates events to the ATP's using a load-levelling scheme, keeps track of the state of all events and notifies the SEB's when the data for a given event can be discarded or "flushed". These various steps are initiated through "control" messages passed between the various processes over the same switch network as used for the data flow. The pull architecture allows the SEB's to act primarily as intelligent buffers. They are thus responsible for parsing and checking the integrity of the data received from the DCM's and sending the data to the ATP's upon request. The architecture simplifies the ATP operation and allows the ATP's to perform zero-copy receive operations since no data arrives unless it is expected. Because the controller is responsible for initiating the processing of an event, it is necessary that the controller know when new events have been read out of the FEM's. Since the LVL1 trigger (GL1) data is read on every crossing for which a LVL1 trigger fires, the arrival of data from the GL1 sub-system at the EvB provides the appropriate signal of the presence of a new event. The GL1 SEB software contains extra code that collects the information about arriving events and transmits it to the controller periodically. In principle, multiple partitions may be triggered on the same crossing and in the current implementation of the EvB, the different partitions are allocated to the ATP's and assembled independently. The controller keeps track of the completion of the various assignments for a given crossing and only sends "flush" messages to the SEB's when assembly has been completed for all partitions.

Data is transferred from the DCM's to the SEB's via the DCM "partition module" over 32 bit parallel cables using LVDS signaling. The data is received and buffered in a custom-built PCI interface card (JSEB) implemented using the PLX 9080 PCI bridge chip. Each JSEB contains two 1 Mbyte banks of SRAM that act as local data buffers. One bank is used to receive data from the DCM's while the other is accessible by the host via the PCI bus. The functionality of the two banks is swapped under host control so that one bank is always available for writing while the other, containing the most recently read data, is accessible to the host. The partition module uses a dedicated line to raise an end-of-event signal and the JSEB explicitly keeps track of the number and locations of events in each bank. The writeable bank is considered full when either the number of events or the number of words written into the bank reach a specified limit. At that time the JSEB raises a dedicated "hold" signal that prevents the partition module from transferring additional data. Once the host swaps banks the JSEB lowers its hold and starts writing into the newly available bank. The JSEB also implements a PCI interrupt via the PLX9080 that indicates when the currently writeable bank becomes full.

We chose to use a FORE (now Marconi) ASX-1000 switch for the PHENIX EvB because of its expandibility, because it was well-matched to the bandwidth requirements of the baseline EvB and because of prior experience using the ASX-1000 in the high-energy physics data acquisition community[27,28]. The ASX-1000 can be equipped with a maximum of 64 multi-mode fiber OC-3 connections, each of which operates at a bit transfer rate of 155 Mbit/s duplex. Because of the intrinsic overheads of the ATM protocol, the maximum achievable data transfer rate on one of these connections is 130 Mbit/s or slightly greater than 16 Mbyte/s. The ASX-1000 consists of four separate 2.5 Gbit/s non-blocking switching elements which can be configured with up to 16 OC-3 ports and which inter-communicate via a 10 Gbit/s backplane. In order to fully utilize the 10 Gbit/s backplane, all ports must transmit at the maximum rate in both directions. Because the data flow in the EvB is unidirectional, the maximum data transfer bit rate is 5 Gbit/s or 512 Mbytes/s when the switch is fully equipped. Each processor in the EvB is equipped with a single FORE PCI-bus ATM NIC. The EvB was developed using the model PCA200E NIC but prior to the first operation of PHENIX a new version of the board, the model HE155, became available and we replaced all original NIC's with these higher performance cards. All ATM fiber connections were made using standard commercial 0.62 mm multi-mode fiber patch cords with SC connectors. The ATP's were connected to the ONCS system via 100 baseT point-to-point ethernet connections to a central switch.

The processing nodes in the EvB were implemented using commercial PC's of various specifications. After initial design tests, all the PC's that were used in the various stages of EvB implementation were equipped with Pentium III processors with clock speeds $\geq$ 450 MHz using a variety of Intel chipsets and PC100 or PC133 SDRAM memory. For Run-2 operation dual-processor 1 GHz PIII machines equipped with 64-bit capable, 66 MHz PCI buses were purchased for use as ATP's. All nodes were equipped with fast ethernet interfaces. Most of the PC's had the ethernet interface built into the motherboard. The SEB's were connected to the PHENIX On-Line network via fast ethernet hubs while the ATP's were provided with switched ethernet connections through which they transferred data to the logging system. All of the PC's in the EvB were connected to Master Console II KVM switches from Raritan Computer Inc. A variety of 8, 16 and 32 port switches were cascaded to provide access to all of the PC's from a single keyboard, mouse and video terminal station.

### *7.3 Software*

Because of the ready availability of vendor-supplied drivers for all commercial ATM NIC's and the PLX9080 chip, we chose to run Windows NT on all of the EvB nodes. The EvB functions were implemented in a single multi-threaded process with the run-time configuration and control of each node provided via CORBA. The applications used for the SEB's and ATP's were identical for each component except for the SEB handling data from the PHENIX GL1 trigger. An extra thread was run in the GL1 SEB to indicate to the EBC the arrival of new events. The EvB applications were written in C++ and extensive use was made of object-oriented features of C++ and the C++ Standard Template Library (STL) in developing control and bookkeeping data structures. Applications were compiled and linked using the Visual C++ development environment on a separate development PC and were downloaded to the EvB PC's via a PERL script using the PERL FTP package.

Permanent virtual circuits are used for all communication and data transfer in the EvB. Circuits are configured prior to EvB startup via a PERL script that reads configuration files to obtain the EvB topology and partitioning configuration of the DAQ, determines all of the necessary virtual circuit connections for EvB operation, configures the switch and writes the virtual circuit indices into configuration files for each node in the system. Point-to-multipoint circuits are used for multi-cast transmissions that allow the ATP's to send a single message requesting a particular event from all the SEB's and the EBC to send a single flush message to all of the SEB's. Transfers over ATM were implemented using the WINSOCK2 API which extended the original BSD sockets implementation under NT to handle modern networking technologies including ATM. Extensive use was made of the asynchronous I/O features of

Windows NT to manage the many simultaneous I/O operations taking place in the EvB nodes. Extensive use was also made of the scatter-gather interface provided by the WINSOCK2 API that allowed multiple-buffer I/O operations to be queued in a single call.

Configuration and control of the EvB was provided through a single CORBA-accessible object in each node that maintained all necessary configuration parameters and provided methods to execute state transitions under remote control. A separate monitoring object was provided for each node that kept track of processed events, data volume, error counts and other quantities relevant to the performance of the SEB's, ATP's and the EBC.

### *7.4 Implementation History and Performance*

The initial design and development of the EvB was performed from Spring 1997 through Fall 1998 using a $2 \times 2$ system consisting of 200 MHz Pentium II PC's and a single 2.5 Gbit/s ATM switch. In Summer 1999 a $4 \times 4$ system was implemented using 450 MHz PC's equipped with 128 Mbytes of PC100 SDRAM. Tests of this system using self-generated simulated data indicated that event rates as high as 750 Hz and total data rates in excess of 50 Mbyte/s could be achieved in stand-alone operation. At this time attention turned to integrating the EvB into the PHENIX DAQ. Through Spring 2000 effort was focused on implementing the CORBA control interfaces, the interface to the PHENIX data-logging system and the readout of data through the JSEB card.

For RHIC run-1 operation, the EvB was equipped with 16 SEB's and 4 ATP's. However, the JSEB boards were not completed in time for run-1 so the readout from the DCM's to the SEB's was implemented using TCP/IP over ethernet. Because the DCM's were not originally designed for high-speed readout through the crate controllers, event rates were limited to 20 Hz during run-1 operation. The corresponding maximum data rates of 20 Mbyte/s were well below any performance limitations of the EvB.

For RHIC run-2 the EvB was equipped with 32 SEB's and 28 ATP's. A combination of 733 MHz PC's and 933 MHz machines, all equipped with 128 Mbytes of PC133 SDRAM, were used for the SEB's while dual-CPU 1 GHz machines were used for the ATP's. The JSEB boards were complete and available for run-2 operation. Performance tests of the JSEB interface alone demonstrated that it could receive events at rates exceeding 5 kHz and could handle total data rates in excess of 40 MByte/s. Between run-1 and run-2, effort was also devoted to reducing the event handling overheads in the SEB's. As a result of these improvements, single-SEB event rates as high as 2.5 kHz were achieved using the 933 MHz PC's for detectors with small data volumes ($< 1$ kByte).

For detectors with larger sub-event sizes, the event rate was limited by the 17 Mbyte/s throughput of the ATM OC-3 connection. For sub-event sizes $\geq 10$ kbyte we were able to achieve 100% utilization of the ATM link throughput. However, for smaller events the SEB's were limited to event rates of 1.1-1.5 kHz depending on the detector sub-system and the speed of the PC. These limits are not understood and require further study. They appear not to be due to the ATM link speed but are also not simply due to per-event overheads in event handling at the SEB since the achieved rates are a factor of two lower than those achieved for detectors with very small data volumes

During most of run-2 operation, the luminosities produced by RHIC were such that the performance limits of the EvB were not encountered. However, by the end of run-2 LVL 1 trigger rates in excess of 1 kHz were observed in PHENIX. The maximum event rate achieved by the PHENIX DAQ system during run-2 operation was $\approx 900$ Hz corresponding to an aggregate data rate of 140 Mbyte/s in the EvB. The limitation on the DAQ event rate resulted from a combination of disabled FEM pipelining, LVL2 trigger performance, congestion in individual DCM's due to noisy channels and intrinsic SEB event rate limits.

## 8 Level-2 Trigger

The PHENIX Level-2 (LVL2) triggers are described as they existed for the second RHIC run in 2001/2002. The LVL2 system for that run was designed to operate at RHIC average design luminosity for Au-Au collisions. This was the maximum average luminosity that was expected to be achieved in run 2.

At the RHIC Au-Au average design luminosity of 2 x $10^{26}$ cm$^{-2}$s$^{-1}$, the interaction rate as seen by the BBC-LVL1 trigger is about 1400 Hz. This represents about 92% of the nuclear interaction cross section. The event size for a zero-suppressed minimum-bias Au-Au collision was about 160 kBytes, so a rate of 1400 Hz corresponded to a data rate of 224 Mbytes/s. The DAQ system as constituted for the second RHIC run, was designed to process data at that rate all the way to the ATP's. In the ATP's, the data rate had to be reduced to the rate at which data could be archived to disk, which is approximately 35 Mbytes/s. This reduction of a factor of six or seven in data rate was achieved by the use of LVL2 triggers running in the ATP's. The LVL2 triggers select events that appear to have interesting physics potential. Those events, along with a sample of minimum bias events, are written fully to disk, along with the trigger primitives calculated by the LVL2 algorithms. For events which are rejected by the LVL2 triggers, and which were not included in the sample of minimum bias events, only the trigger decisions and primitives calculated at LVL2 are written to disk.

LVL2 trigger parameters are configured via Run Control (described elsewhere in this volume [2] before a run is started. This configuration process sets the association between the LVL1 trigger bits and the LVL2 trigger algorithms. It also sets the scaledown factors for each trigger algorithm.

Monitoring of the LVL2 trigger performance was carried out using an offline monitoring code that read the LVL2 trigger primitives written to disk for each event. This was done within minutes of the data being written to disk, so that any problems with the performance of the triggers could be found quickly. LVL2 trigger performance could also be extracted from each ATP via direct CORBA calls, a feature that was useful when installing and debugging new trigger algorithms.

### *8.1 Design of the LVL2 Trigger System*

The LVL2 triggers operate in the ATP's. During the second RHIC run, 32 ATP's were in use. At the design luminosity interaction rate of 1400 Hz, the time available for one of the 32 ATP's to run all of the LVL2 trigger algorithms on a single event was about 25 ms, including the time needed to unpack the data and apply any calibrations. For this reason, the LVL2 trigger infrastructure and LVL2 algorithm code were designed to be as efficient as possible.

The LVL2 trigger code is written in C++ and the trigger classes are divided into several categories, as shown in Table 1. Aside from the infrastructure classes, all of the classes use templates to enforce a common interface, and to add common methods used by the infrastructure. The LVL2 infrastructure objects are created at the beginning of each run, and they persist for the entire run. The database accessor objects, which read calibration data and lookup tables (LUT) from the Objectivity database, are created at the beginning of the run and they also persist for the entire run. Their methods are called when they are created, and the calibration data and LUT's are stored in public data arrays. After that, the database accessor objects are read-only. Also at the beginning of a run, the LVL2 infrastructure makes a list of which trigger algorithms are enabled.

For each event the LVL2 infrastructure creates a new instance of each trigger algorithm class. Each trigger algorithm object accesses trigger primitive classes that calculate quantities needed for making the trigger decision. A trigger primitive object is created the first time a request is made for its results, the calculate methods are called once only and the results are stored in public arrays. The trigger primitive object then persists until the event is fully processed. A trigger primitive object can be accessed by any trigger algorithm or by any other trigger primitive object. Trigger primitive objects access the public data arrays of other trigger primitive objects, data accessor objects and database accessor objects. The data accessor objects are also instantiated, and their methods executed, only when they are accessed for the first time for each event, and they then persist until the event is fully processed.

No class is instantiated or its methods called unless another object specifically requests results from that class. The LVL2 infrastructure code contains methods to monitor the rejection of each trigger algorithm, and the time per event taken by the trigger, primitive and data accessor classes. For safety, all arrays are templated and all array operations are checked for bounds errors.

Following the processing of all events, the LVL2 infrastructure adds the results to the data stream for that event in the form of a decision packet and packets containing trigger primitive results. For accepted events all of the data are also written out. The LVL2 triggers are run on all events and the decision and primitive information written to the data stream, even when the event is accepted automatically, because it satisfies the conditions for a "forced accept".

The ATP's use the Windows NT operating system. However the LVL2 infrastructure is independent of the operating system and the trigger algorithm development was carried out on processors that use the Linux operating system, using a simple framework designed to make code development by a large group easier. This test framework was separate from the PHENIX On-Line system allowing identical algorithms to be tested off-line with minimum bias events. Some care was taken to test new code under Windows NT to check for any compiler compatibility problems.

### *8.2 LVL2 Trigger Algorithms*

The trigger algorithms used in the second RHIC run were based on data from the ZDC [4], BBC [3], RICH [7], EMCal [8], PC [5], TEC [5], $\mu$Id [6] and DC [5]. The coherent peripheral trigger was run only on ZDC triggers, which are sensitive to mutual Coulomb interactions as well as nuclear interactions. At the RHIC average design luminosity for Au-Au collisions, the ZDC trigger rate is about 2100 Hz. All other triggers were run on all events that satisfied the BBC-LL1 interaction trigger at LVL1. A summary of the physics triggers used in the second RHIC run is presented in Table 2. There were usually multiple versions of each trigger, with different cuts (see Table 3 for details). In most cases, trigger versions with looser cuts (for example lower threshold cuts) were prescaled so as to limit their contribution to the data stream. The trigger versions with looser cuts are used to study the behavior of the version with the tightest cuts, in the region where the cuts become effective. There

Table 1
Categories of LVL2 Trigger Classes.

| Category | Purpose | When Instantiated |
|---|---|---|
| Infrastructure Classes | Call database accessor classes<br>Loop over events and call trigger algorithms | Each run |
| Database Accessor Classes | Get calibrations and lookup tables<br>from database | Each run |
| Trigger Algorithm Classes | Call trigger primitive classes<br>Make trigger decision | Each event |
| Trigger Primitive Classes | Call data accessor classes<br>Calculate quantities needed for<br>the trigger algorithm decision | Each event |
| Data Accessor Classes | Unpack data from the data stream<br>Apply calibrations and find hits | Each event |

was no separate event centrality trigger, rather the centrality calculation was used in some cases to make centrality selections for the physics triggers.

The single electron trigger begins by finding a cluster of RICH hits that has at least two hit phototubes and a combined signal equivalent to at least 3 photoelectrons. Then a lookup table is used to identify the possible matching region of the EMCal and that region is searched for an EMCal hit cluster. If more than one cluster is found, all possible combinations are kept. Each EMCal cluster that survives an EMCal threshold cut, a PC3 track association cut and a cut on track bending angle as a function of cluster energy is regarded as an electron candidate. The electron pair trigger calculates the equivalent invariant mass for all electron pairs, and then imposes an invariant mass threshold cut. Electron pair triggers with high threshold cuts (2.0 or 2.2 GeV/c$^2$) were aimed at $J/\Psi$ decay physics. An electron pair trigger was included with an invariant mass threshold of 0.7 GeV/c$^2$ to accept $\phi$ decay candidates. Because the rejection of that trigger was poor for central events, a centrality cut was made to exclude all of the 40% most central collisions. It was then possible to secure a good sample of $\phi$ events from central collisions from minimum bias events and at the same time collect a sufficient sample of $\phi$ events from peripheral collisions. The same argument applies to the single muon trigger discussed in the next paragraph.

The single muon trigger finds roads through the 5 layers of $\mu$ID panels that end in a hit in the deepest panel. The road is allowed to have up to two missed layers. A LUT of possible hit sequences is used to test if the road is consistent with a muon from the interaction point. The muon pair trigger looks for pairs of muon candidates which satisfy a minimum opening angle cut.

The high $p_T$ EMCal trigger searches for clusters of EMCal hits for which the cluster energy exceeds a given threshold.

The high $p_T$ charged particle trigger finds possible matching hits in PC1 and PC3. It also looks for roads in the DC and calculates a bend angle for each one found. The algorithm searches for high momentum charged particle tracks by selecting tracks with bend angles less than a given threshold, and then checking for matching PC hits at the expected locations.

The coherent peripheral trigger is run only on ZDC triggers in which there were no BBLL1 interaction trigger. These triggers had to be prescaled by a factor of 30. This was done because the bandwidth available between LVL1 and LVL2 was insufficient for all ZDC triggers to be allowed through LVL1. The coherent peripheral trigger selects events that have no BBLL1 trigger, and which have 2 or more fired pads in the innermost layer of PC's (PC1) in each arm (East and West). This cut gives only a negligible contribution from random triggers, e.g. due to electronic noise.

The centrality algorithm calculates the angle of an event in the two dimensional ZDC energy and BBC charge space. A typical plot is found in Fig. 6 of the paper on the PHENIX Inner Detectors in this volume[3]. This angle has been found to be a reasonable measure of event centrality (designated the clock method) and the calibration can be obtained easily by generating a LUT of angle vs percentile of events from real data. The algorithm returns the percentage of events that are more central than the present one.

### *8.3 Results from the Au-Au Run*

The triggers used in the RHIC run 2 are listed in Table 3, along with the most important cut values, the typical prescale factors used and the typical rejection factors observed for 200 GeV Au-Au collisions.

The prescale factors for each trigger shown in Table 3 were chosen to match the data volume to the maximum disk archiving rate, which was about 35 Mbytes/s. The most important rare event triggers were either not prescaled at all, or were very lightly prescaled, while the triggers with looser cuts, used primarily for trigger studies, were frequently heavily prescaled.

The total LVL2 trigger processing time for a single event was found to be 31 ms (compare to the design value of 25 ms). This was adequate for the 2001 RHIC run with 32 ATP's active, since the maximum DAQ throughput was

Table 2
Overview of LVL2 Trigger Algorithms.

| Trigger | Method |
|---|---|
| Single Electron | Match RICH rings to EMCal clusters<br>Make EMCal energy threshold cut |
| Electron Pair | Calculate invariant mass of electron pairs<br>Make invariant mass cut |
| Single Muon | Find roads through $\mu$ID panels |
| Muon Pair | Find two muon trigger candidates |
| High $p_T$ EMCal | Find EMCal clusters<br>Make threshold cuts |
| High $p_T$ Charged | Match PC and DC hits<br>Cut on the bend angle |
| Coherent Peripheral Events | Look for ZDC trigger with no BBC trigger<br>Look for PC hits |
| Centrality Selection | Use BBC and ZDC to estimate centrality<br>Make centrality cuts on selected triggers |

800 Hz for other reasons.

In the 2001 RHIC run, the peak Au-Au minimum bias trigger rate for PHENIX at the beginning of a store was about 1400 Hz. During a typical 4 hour store, this rate would drop to 200 Hz. The minimum bias trigger rate usually dropped below 800 Hz in the first 50 minutes. This rapid variation of minimum bias trigger rate with time was accomodated by changing the prescale (which determines the fraction of events accepted unconditionally) of the minimum bias triggers at LVL1 so that the data volume to disk was approximately constant. Over the length of a store, the prescale for minimum bias interactions varied from 3 to 30.

In conclusion the PHENIX LVL2 triggers worked very well in the 2001 RHIC run, allowing PHENIX to process most minimum bias triggers at the available maximum DAQ throughput rate of 800 Hz, and to reduce the data volume sufficiently so that all events accepted by the rare event triggers could be written to disk. A few minimum bias triggers were excluded but this occurred only during the initial portion of some high-intensity fills. In addition, typically 50 Hz of minimum bias triggers were also recorded, filling up the disk archiving bandwidth not needed by the rare event triggers.

Table 3
The LVL2 Triggers Used in the RHIC Run 2. (The prescale factors shown here are the ones used at the 800 Hz event rate.)

| Trigger | Cuts | Prescale | Rejection |
|---|---|---|---|
| Single Electron | E $>$ 1.5 GeV/$c^2$, centrality 40-100% | 1 | 728 |
| | E $>$ 2.5 GeV/$c^2$ | 1 | 412 |
| | E $>$ 3.2 GeV/$c^2$ | 1 | 2475 |
| Electron Pair | M $>$ 0.7 GeV/$c^2$, centrality 40-100% | 1 | 73 |
| | M $>$ 2.0 GeV/$c^2$ | 10 | 20 |
| | M $>$ 2.2 GeV/$c^2$ | 1 | 33 |
| Single Deep Muon | centrality 40-100% | 1 | 100 |
| | | 99999 | 7 |
| Muon Pair | $\theta_{open} > 19.2^o$, centrality 40-100% | 1 | 100 |
| | $\theta_{open} > 19.2^o$ | 1 | 58 |
| High $p_T$ EMCal | E $>$ 1.5 GeV/$c^2$, centrality 40-100% | 2 | 32 |
| | E $>$ 2.5 GeV/$c^2$ | 20 | 18 |
| | E $>$ 3.5 GeV/$c^2$ | 2 | 148 |
| High $p_T$ Charged | | 5 | 151 |
| | | 2 | 250 |
| Coherent Peripheral | | 1 | 241 |

## 9 Electrical Infrastructure

The design of monitor and control systems for experimental safety, power and environmental aspects of the PHENIX detector at RHIC necessitated careful consideration of diverse requirements. For the critical safety systems, the requirement for reliability outweighed the use of advances in computerized control hardware, whereas commercial hardware and software provided a flexible solution for experimental power control and environmental monitoring. Granularity and prioritizing of safety related conditions with respect to the various PHENIX detector subsystems had to be determined by each subsystem in consultation with RHIC safety personnel. The PHENIX low voltage power supply system was required to be remotely controlled in the PHENIX IR, have many individual channels of unique low voltages for different detectors and be modular to allow rapid replacement of failed power supplies.

The monitor and control systems for the PHENIX detector are divided into two subsystems, namely the Safety Monitoring and Control System (SMCS) and the Rack Monitor and Control System (RMCS). The SMCS was designed using historically proven and reliable relay logic technology to initiate appropriate responses to critical Level One safety related occurrences. The SMCS design incorporates fail safe considerations such that wire breaks or blown fuses will cause automatic initiation of the various shutdown responses. The RMCS employs a combination of commercially available Data Acquisition and Control (DAC) modules and relays to address the requirement for electronics rack monitoring, control and Level Two safety interlocks. The Low Voltage Power System (LVPS) is controlled by the RMCS and is specially designed to meet the unique requirements of PHENIX detector FEE's.

### *9.1 Safety Monitoring and Control System*

The PHENIX SMCS is an emergency shutdown system designed to allow the PHENIX detector subsystems to operate safely. It provides protection against major equipment damage, particularly that due to the effects of fire and water leakage. The system's most important response is to automatically turn off electrical power and flammable gas supply to the experiment when hazardous conditions are detected. Each PHENIX detector subsystem contains its own local interlocks that will detect and react to internal faults. The SMCS, however, supplements those responses and provides global protection coverage.

The SMCS consists of commercially available fire detection, water detection, and flammable gas detection instruments and sensors. This equipment interfaces with custom designed relay logic circuits which, in turn, generate output signals to effect the safe shutdown of the PHENIX detector. The controls/relay logic circuits for the SMCS are contained in 19 inch equipment racks located in the CH control room and electronics rack room. The safety related sensors and controls are hard wired to these locations.

### *9.2 Fire Detection*

Over 50 individual zones are continuously monitored for smoke utilizing photoelectric smoke sensors. They are each digitally addressable and connected to an intelligent fire alarm control panel via a two wire communication cable network. The zone locations are inside electronic rack enclosures and detector chambers. In general, smoke detected in a zone will cause a SMCS response to shut off electrical power to the subsystem electronics contained within the enclosure.

Air spaces throughout areas of the PHENIX detector (outside of electronic enclosures) are monitored using a high sensitivity smoke detection (HSSD) system. Air samples are drawn through a series of perforated PVC pipes and delivered to the smoke detector. The SMCS responds to high levels of smoke by shutting off all electrical power and flammable gas flow to the experiment.

### *9.3 Water Leak Detection*

Water leak sensors are used throughout the PHENIX facility. Individual leak sensing electronic modules are coupled with various lengths of leak sensing cable. They are used in every electronic enclosure containing chilled water radiators and at floor spaces where leaking water is expected to accumulate. The leak sensors are integrated into the PHENIX SMCS. In most cases water leaks will cause the SMCS to shut off electrical power to electronics contained within the enclosure.

### *9.4 Flammable Gas Detection*

Seven zones are monitored for flammable gas leaks using catalytic bead type sensors. Three sensors monitor regions near the South $\mu$ID panels and the HVAC return air flow from the detector IR region. Four zones monitor areas in the Gas Mixing Building. Sixteen zones are monitored in various areas surrounding the PHENIX detector panels via air samples continuously being drawn to flammable gas detection panels. The sampling panels are located outside the PHENIX detector areas and draw air samples through small diameter plastic tubing.

### *9.5 PHENIX Rack Monitor and Control System*

The PHENIX electronics racks are monitored and controlled by a Supervisory Control And Data Acquisition (SCADA) system referred to as the PHENIX Rack Monitor and Control System (RMCS). This system consists of modular data acquisition and control modules in the Advantech ADAM 5000 series which provide analog and digital inputs and outputs at remote locations connected by a RS-485 network to a Human Machine Interface computer in the PHENIX Control Room (PCR). The RMC facilitates the monitoring of rack ambient air and cooling water temperatures, LVPS temperatures, voltages, currents, rack heat, water leak, smoke interlock and main rack contactor (AC power) status and other detector-specific signals such as magnet flux. The RMC controls enable signals for LVPS channels, rack AC power, alarm window box control, various Level 2 safety system interlocks and other detector-specific control functions.

## *9.6 System Architecture*

The HMI computer is a PC running the ICONICS Genesis32 suite of HMI software applications under Windows NT. These applications are Object Linking and Embedding for Process Control (OPC) clients enabling them to communicate with any OPC compliant server.[29] Genesis32 consists of the Graphworx32, Dataworx32 and Trendworx32 applications. Graphworx32 is a graphical development application providing tools for developing OPC compliant monitor/control interfaces. Dataworx32 is a data aggregation and logic handling process used to perform alarm logic for the alarm window box. Trendworx32 is a historical trending/logging program used for storing certain parameters in files for analysis.

The OPC Server is a Windows NT PC running the Fastwel ADAM OPC Server application. This computer contains an eight channel RS232 serial card. Seven of these channels are connected to RS232-to-Fiber translators for communication with major subdivisions of the PHENIX detector in the IR. The eighth channel is used to direct commands to the alarm window box in the PCR. In the IR the fiber links are converted by Fiber-to-RS485 translators into seven multi-drop, half duplex RS485 LAN's for connection to multiple DAC modules in the PHENIX racks.

## *9.7 PHENIX Low Voltage Power System*

The 11 detector subsystems in the PHENIX IR produced a list of demands on the design of the LVPS. An in-house design incorporating commercially available parts was produced to address these design requirements while allowing incremental production of power components as PHENIX grows to completion. The system is implemented as a 6U crate with a set of 340 mm ("C" size) plug-in power supply cards. A Front End converter card converts 208 VAC 3-phase power into 300 VDC distribution power for 8 to 12 configurable power supply cards which in turn produce low voltage. The low voltage outputs then proceed to 160 mm noise filter cards installed in the rear of the crate. The low voltage outputs from the filter boards are connected to feedthrough type terminal blocks on the rear panel. Control and monitoring of the LVPS is achieved through the RMCS system via ribbon cables from the rear of the crate.

A listing and short discussion of ten basic LVPS requirements are given below:

(i) Efficiency–All low voltage supplies are of the switching regulator type to meet space, weight and heat limitations presented by the IR physical design. The supply modules used have typical power efficiencies of 80% to 90%. Significant electrical energy savings are also realized over linear types when considering the estimated 100kW of power produced by the total system.

(ii) Density–The large amount of FEE's in the IR and limited rack space placed a premium on space for low voltage supplies. This drove the requirement for higher power density.

(iii) Noise–With PHENIX particle detectors designed to detect minute event signals it is essential to keep radiated and conducted noise to a minimum. The LVPS contains both passive and active noise filters to meet this requirement.

(iv) Output Voltage Adjustment–Most of the detectors incorporated in PHENIX require non-standard voltages. This is caused by voltage drops in the supply line to the detector and compliance voltage required for local regulators in the detector subsystems. The supply modules used in the LVPS are available in various nonstandard voltages required by PHENIX and are adjustable to "tune" the compliance voltage at the detector for optimum operation of the local regulators. The use of load voltage sense lines, which would have doubled the number of wires to detectors, was forgone in favor of simply adjusting compliance voltages.

(v) Power Range–The LVPS is designed to produce from 25 W to 600 W per channel of output power with as much as 1.2 kW on a single plug-in card.

(vi) Output Control–Output channels may be enabled individually or combined in groups within a card to provide simultaneous powering of channels.

(vii) Monitoring–Each channel of the LVPS has an adjustable voltage window comparator that produces a DC OK signal as long as the channel output voltage is within a preset specification. The distribution supplies used to feed the LVPS cards also produce AC OK signals to indicate status of the AC line. Channel output current and voltage is measured directly from the test jacks on the card front panel using a DVM during setup. These signals are also routed through the back panel for remote monitoring where needed. Analog voltages representing heatsink temperatures are also provided. For more on monitoring see section 9.5.

(viii) Servicing–Many of the PHENIX electronics racks in the IR are not easily accessed because of their height/location. Access to the rear of these racks is difficult at best. The LVPS was designed as a system of front panel removable cards loaded into dedicated 6U X 340 mm crates to meet this need. Power to each of the crates is switched by circuit breakers accessible in the rack.

(ix) Air Flow–The PHENIX standard electronics rack design calls for vertical airflow cooling. The LVPS system satisfies this requirement.

(x) AC Mains–Adequate power factor (0.9 or more) and balanced loading of the three phases are essential to efficient power system design. The off-line front end power converters used in the LVPS load all phases equally and have a typical power factor of 0.9.

### *9.8 LVPS Construction*

The low-voltage crate consists of 8 or 12 "C" sized (6U X 340 mm) plug-in power supply regulator cards installed in a 534 mm deep crate with a 300 VDC distribution backplane and 160 mm output noise filtering cards installed in the rear. The power supply card backplane is divided into four 1/4 backplane boards with a fifth backplane board used for the 3 kW/5 kW 3-phase front-end converter. Two types of 1/4 backplane boards can be installed in the same crate: Low Power (6HP wide, 3 cards/backplane) or High Power (9HP wide, 2 cards/backplane). The rear panels are aluminum plates containing feedthrough type terminal blocks for AC input and DC output power connections.

Several types of modules are used in LVPS construction. The Low Voltage Front End (LVFE) converter module hosts either a 3 kW or 5 kW offline converter to convert 208 VAC 3-phase power to 300 VDC distribution power to the LVLP or LVHP power supply modules. A front panel "ENERGIZED" indicator warns the user of hazardous voltage remaining in the module capacitors after removing the power and AC OK and BUS OK LED's indicate sufficient operating voltages.

The Low Voltage Low Power (LVLP) module hosts 8, 25 W or 50 W DC-DC converter modules, for converting 300 VDC distribution power to low voltage. An optically isolated control matrix enables any one of the 8 channels. Any combination of channels may be ganged together to form channel groups that "come up" simultaneously. Sockets for 8 window detector daughterboards are located along the front edge of the module. Another socket located near the rear is provided for installing a baseplate temperature monitor board.

The Low Voltage High Power (LVHP) module hosts 6, 50 W to 200 W DC-DC converter modules, for converting 300 VDC distribution voltage to low voltage. An optically isolated control matrix enables any one of the 6 channels individually or in groups. One set of jumpers allows ganging together channels in groups while several other sets of jumpers may be used to master-slave modules for up to 6 modules (1200 W) in a single output. Sockets for 6 window detector daughterboards are located along the front edge of the module. A heatsink temperature sensor provides a linear analog output of 10 mV/°C for readback by the rack ADAM 5000 modules.

Low voltage modules are constructed with different jumper settings and DC-

DC converter module types depending on the individual PHENIX detector subsystem requirements. These are denoted as variants and are given letter names. A variant is described by a block diagram sheet showing the basic DC-DC converter module groupings and ratings and a jumper settings sheet.

Vertical airflow cooling is achieved by a fan tray-water/air heat exchanger combination installed below the crate. The module front panels and backplane produce an air column for conducting vertical airflow. The LVLP module is passively cooled without added heatsinks since thermal calculations indicated sufficient cooling from baseplates alone. The LVHP module employs aluminum heatsinks mounted on conductive pads for improved heat conduction between baseplates heatsinks.

## 10 Acknowledgements

We acknowledge support from the Department of Energy (U.S.A.) from Monbusho and STA (Japan) and from the University of Lund (Sweden).

## References

[1] Article entitled "PHENIX Detector Overview" this volume.

[2] Article entitled "PHENIX On-Line and Off-Line Computing" this volume.

[3] Article entitled "PHENIX Inner Detectors" this volume.

[4] Article on ZDC this volume.

[5] Article entitled "PHENIX Central Arm Tracking Detectors" this volume.

[6] Article entitled "PHENIX Muon Arms" this volume.

[7] Article entitled "PHENIX Central Arm Particle I.D. Detectors" this volume.

[8] Article entitled "PHENIX Calorimeter" this volume.

[9] W.L. Bryan et al., IEEE Trans. Nucl. Sci. 45 (1998) 754.

[10] M.C. Smith et al., IEEE Trans. Nucl. Sci. 46 (1999) 1998.

[11] F.M. Newcomer et al., IEEE Trans. Nucl. Sci. 40 (1993) 630.

[12] Y. Arai et al., IEEE Trans. Nucl. Sci. 45 (1998) 735.

[13] M.N. Ericson et al., IEEE Trans. Nucl. Sci. 43 (1996) 1629.

[14] M.S. Emery et al., IEEE Trans. Nucl. Sci. 44 (1997) 374.

[15] M.L. Simpson et al., IEEE Trans. Nucl. Sci. 43 (1996) 1695.

[16] R.G. Jackson et al., IEEE Trans. Nucl. Sci. 44 (1997) 303.

[17] A.L. Wintenberg et al., IEEE Trans. Nucl. Sci. 45 (1998) 758.

[18] M.N. Ericson et al., IEEE Trans. Nucl. Sci. 44 (1997) 312.

[19] ARCNET Designer's Handbook, Document 61610, Datapoint Corporation, (1993).

[20] Integrator Series FPGAs 40MX and 42MX Families Preliminary Data Sheet, Actel Corp., September 1997.

[21] Virtex-E 1.8V FPGA Family: Detailed Functional Description, Xilinx Corp., April 2001.

[22] IQX Family Data Sheet, I-Cube Inc., December 1997.

[23] C.Y. Chi et al., IEEE Trans. Nucl. Sci. 45 (1998) 1913.

[24] ALTERA FPGA Databook

[25] Analog Devices SHARC User Manual

[26] J. Christiansen et al., CERN-DRDC-93-55 (1993).

[27] M. Costa et al., IEEE Trans. Nucl. Sci 43 (1996) 1814.

[28] D. Calvet et al., IEEE Trans. Nucl. Sci. 43 (1996) 90.

[29] www.opcfoundation.org