Design, Performance and Control of the CDF Run II Data Acquisition System
2004 IEEE Nuclear Science Symposium

William F. Badgett, Jr.

on behalf of the CDF Collaboration

Abstract--**The Run II Data Acquisition (DAQ) system of the CDF Detector at Fermilab's Tevatron accelerator has been operational since July 2001. CDF DAQ has collected over 350 inverse picobarns of proton-antiproton collision data with high efficiency. An overview of the design of the pipelined, deadtime-less trigger and data acquisition system will be presented. CDF can receive and process a maximum crossing rate of once per 132 ns, with the rate reduced in three stages to the final output of approximately 1 to 2 terabytes per day. The DAQ system is controlled and monitored via a suite of Java based control software, with connections to front end VME crate processors running VxWorks/C and back end Oracle databases. Included are a flexible and easy to use Run Control java application and associated system monitoring applications, both stand-alone and web based. The performance and operational experience of three years will be presented, including data taking efficiencies and through-put, and the role of intelligent software in tagging and solving problems. We also review future upgrades designed to increase data collection rates to cope with increased Tevatron luminosity.**

I. INTRODUCTION

THE Collider Detector at Fermilab (CDF) detects 980 GeV energy proton-antiproton collision at the Tevatron Accelerator in the Fermi National Accelerator Laboratory. CDF began collection data for Run II in July 2001, and the Tevatron has delivered over 640 pb^{-1} of collisions. The Tevatron will continue to deliver luminosity to CDF until approximately 2009. Prior to 2001, an extensive upgrade of the CDF detector took place to enable it to perform in the higher luminosity environment of the upgraded accelerator complex. With a new Main Injector for the Tevatron and the ability to recycle antiprotons, the target instantaneous luminosity of 2×10^{32} cm^{-2}sec^{-1} for Run IIa was 10 to 15 times

larger than the original Run I data collection period. To cope with these changes, the CDF trigger and data acquisition systems were completely redesigned.

Over the last two years the Tevatron performance has steadily improved, with a record instantaneous luminosity of 1.12×10^{32} cm^{-2} sec^{-1} achieved before the fall 2004 shutdown, as demonstrated in figure [1]. For the latter half of the run, Run IIb, we expect continued improvements in the luminosity, planning for peak from 3.3 to 4.0×10^{32} cm^{-2} sec^{-1}, for which various CDF upgrades are planned.
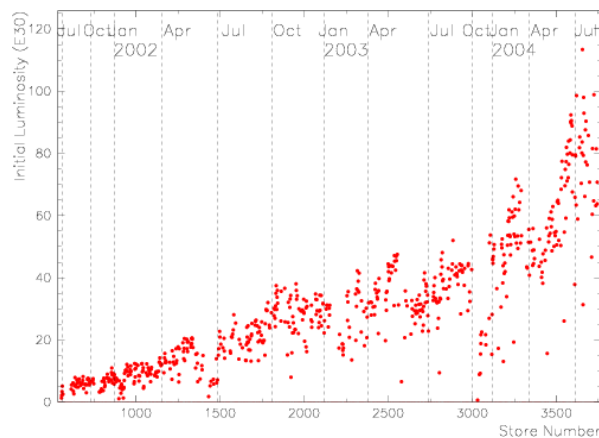

Figure 1: Peak instantaneous luminosity as a function of date and Tevatron store number

II. CDF DATA ACQUISITION SYSTEM

The new CDF data acquisition (DAQ) consists of over one hundred front end VME crates, holding hundreds of custom analog-to-digital converters, nearly five hundred time-to-digital converters, and many other specialized readout and trigger cards. The trigger and DAQ is pipelined and deadtime-less at the first level (Level 1) of triggering/DAQ at the front end crates. The bunch spacing of 132 nano-seconds is not

sufficient time in order to make a decision of whether to accept a particular bunch crossing; therefore, at the front end of all data acquisition elements there is a synchronous pipeline holding forty-two crossings' worth of data to store until a Level 1 trigger decision can be made. The Level 1 trigger decision is made from a reduced amount of data processed entirely via custom trigger hardware. The information includes the longitudinal component of central drift chamber tracks (P_T), calorimeter energy thresholds, and muon identification. As the Level 1 system is fully pipelined, it incurs no deadtime. The design rate out of Level 1 was 50 kHz.

After a Level 1 trigger accept, the pipelined data are copied into one of for Level 2 readout buffers to await a trigger decision from the second level trigger processors. The Level 2 system uses a combination of dedicated hardware and software algorithms to make a fast decision. In addition to Level 1 information, Level 2 has a list of displaced tracking vertices, calorimeter cluster, calorimeter object isolation, and tighter track to muon identification. The Level 2 decision takes approximately 20 to 60 microseconds. The design rate out of Level 2 is 300 Hz. See figure [2] for a macroscopic diagram of data flow and figure [3] for an internal diagram of data flow through the pipeline and buffers.
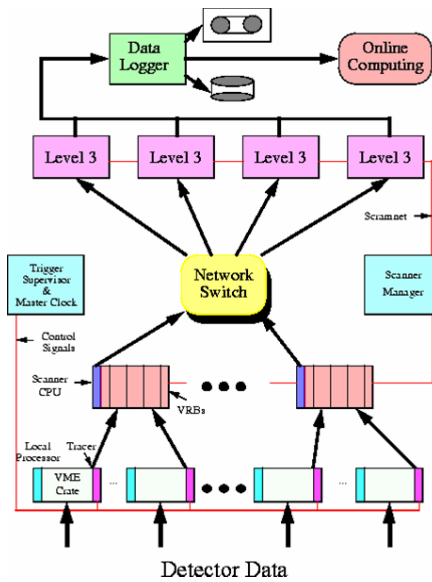

Figure 2: Macroscopic data flow through the CDF DAQ networks

At the final state triggering and DAQ, Level 3, the full event record is sent out of the Level 2 buffers into an ATM switch based event builder, and off to a farm of Linux PCs which can perform full reconstruction of the event data. Level 3 software uses offline style modules written in C++. When accepted by Level 3, the event data are shipped off for permanent storage on tape for later offline processing and analysis.
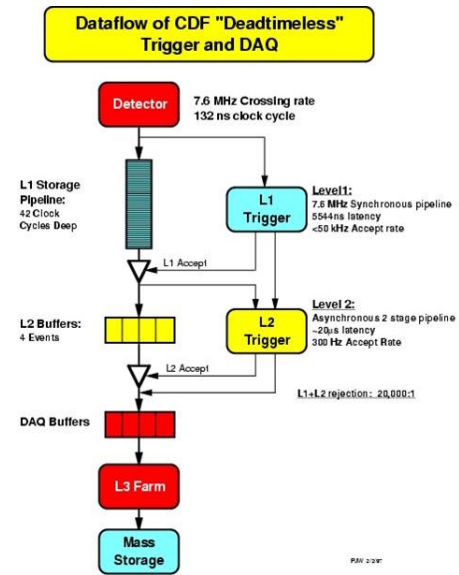

Figure 3: The CDF DAQ and Trigger Pipeline and Buffers

III. CDF PERFORMANCE

The net efficiency of the CDF experiment can be factored into different categories: Accelerator beam quality, over which the experiment has little control; Operational downtimes due to detector and electronic malfunctions, test runs, calibrations, and human error; and Intrinsic system limits, which are limited by the system throughput performance and can be adjusted via physics choices in the trigger cuts.

Since the nominal start of physics-quality data in March 2002 until August 2004, the Tevatron has delivered 642 pb^{-1} of collisions of which CDF has collected 515 pb^{-1}, live to tape. This indicates an 80% efficiency on the part of CDF as a whole, and as such is a primary indication of the success of the experimental operation. See figure [4] for a plot of integrating luminosity. For the first few months of operation, the efficiency rapidly increased from 40% to 85%, becoming asymptotic at this value, shown in figure [5]. Analyzing the causes of downtime over the data taking period indicates the leading reason is unacceptably high beam losses in the Tevatron, preventing the detectors from being powered up, accounting for 1.4% loss of luminosity. The second leading cause related to event builder malfunctions, with 1.0% loss. CDF specific (non accelerator) operational problems indicate gave about 12% total loss.
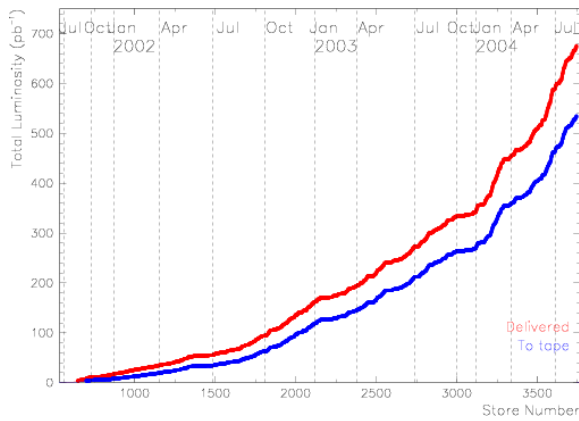
Figure 4: Integrated luminosity delivered (red) and live to tape (blue) as a function of date and Tevatron store
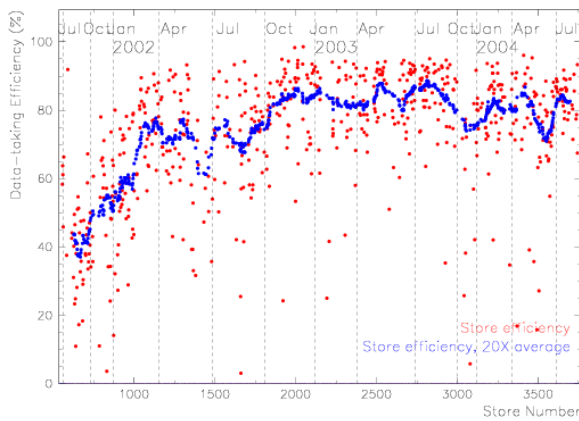

Figure 5: Efficiency of CDF data taking as a function of date and Tevatron store number

The sources of downtime over a span of two and a half years are myriad; automatic tabulation of downtime occurrences are displayed on the web, and linked to the shift-crew electronic logbook, data acquisition run summaries, and Tevatron store information.

The intrinsic inefficiency, also known as deadtime, can be divided into the main categories: Busy at 1.0%, due to the event builder, Level 3, or data logger back-pressure; Level 2 at 1.7%, due to the finite latency of the Level 2 trigger decision; and Readout at 1.1% loss, from the VME readout time after a Level 2 accept. Ambiguous Level 2 or readout deadtime comes in at 0.8% luminosity lost.

The intrinsic efficiency is limited by the throughput of the data acquisition system. During the peak luminosity run in August 2004, we were able to measure the maximum event rates at each trigger level, running at approximately 30% deadtime. The Level 1 maximum rate was 27 kHz, limited by the Level 2 decision processing time of 40 to 60 microseconds. The Level 2 peaked at 390 Hz, limited by ATM network message passing, TDC hit processing time, and Level 3 CPU power. The Level 3 rate reached 135 events per second, or 20 Mbytes per second, limited by the data logging bandwidth.

Examining the contributions of each category to the intrinsic deadtime during a long high luminosity run, we see that the

Level 2 deadtime dominates at lower luminosity, up to about 88×10^{30} cm-2 sec-1, gradually increasing with luminosity. At higher luminosities up to 100×10^{30} cm^{-2} sec^{-1}, the busy (event builder) deadtime quickly becomes dominant and rapidly rises to about 30%, effectively creating a brick wall to further throughput. See figures [6] and [7] for plots of cumulative and individual deadtime contributions from these categories, respectively.
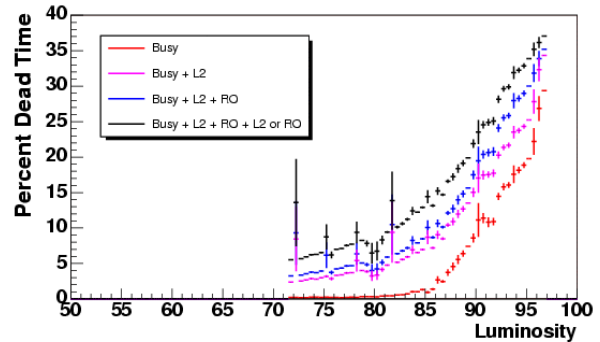

Figure 6: Cumulative deadtime contributions as a function of instantaneous luminosity
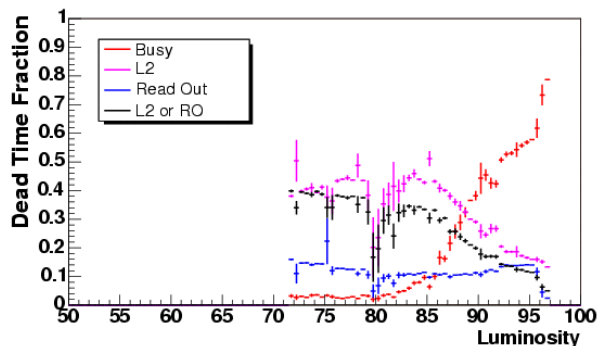

Figure 7: Individual contributions to total deadtime as a function of instantaneous luminosity

In preparation for the Run IIb running beginning in late 2005, a number of upgrades will expand the system throughput. The Level 2 global trigger will be completely replaced with a new system based on fast serial links and commercial processors; refinements in the Silicon vertex trigger will also improve the latency to about 20 microseconds, yielding about 35 to 40 kHz or more event Level 1 rate. The event builder ATM switch will be replaced with gigabit Ethernet; TDC data will be compressed and additional Level 3 processing power will yield a Level 2 rate capability of at least 1 kHz, up to 1.5 kHz. The data logger improvements will yield an eventual 410 Hz, or 60 Mbytes per second output to tape. Note that most of the upgrades rely on the proliferation of inexpensive high speed serial links.

IV. CONTROL AND MONITORING

Critical to the success of the CDF DAQ operation has been the control and monitoring applications developed in the Java language. Most control software has been written in Java 1.4.2 and runs on commercial PCs using Fermilab RedHat Linux 7.3.2. The central control program is *RunControl*, which directs, configures and synchronizes the actions of about 150 front end crates and clients across the DAQ. *RunControl* is a real-time multi-threaded Java application which provides a simple but powerful graphical user interface, extensible through object oriented inheritance to many sub-applications. *RunControl* and its clients use the *SmartSockets* commercial TCP/IP communications package it issue commands and provide status messages in a publish/subscribe paradigm. See figure [8] for a picture of the *RunControl* interface and figure [9] for the client status display and client fine grained control box. Examples of the extensible nature of *RunControl* are shown in figure [10].
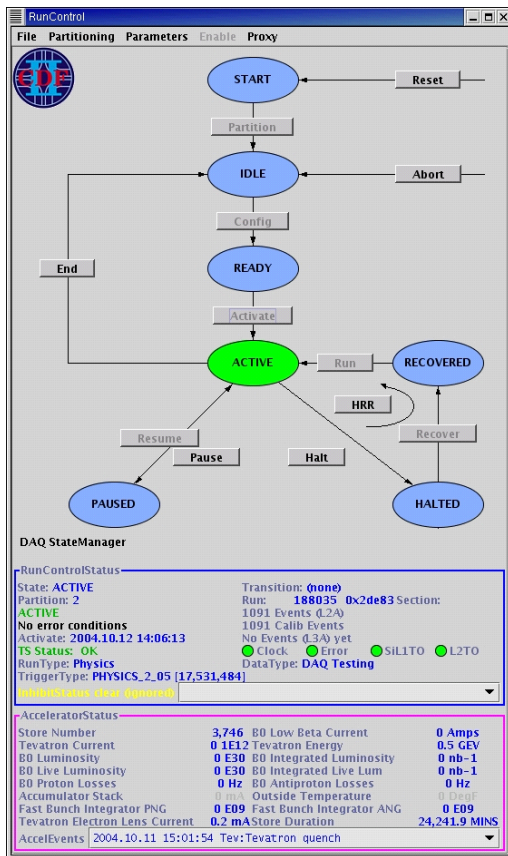

Figure 8: A view of CDF RunControl while taking data

RunControl is closely linked to the Oracle database which contains hardware descriptions, run options, calibration constants and trigger cuts. A real-time Resource Manager system allows users to partition the system into arbitrary and disjoint pieces. In conjunction with RunControl, a java ErrorHandler program monitors the status of the system and makes intelligent decisions when problems or errors are encountered; once a DAQ run has commenced, the RunControl operator has little to do.
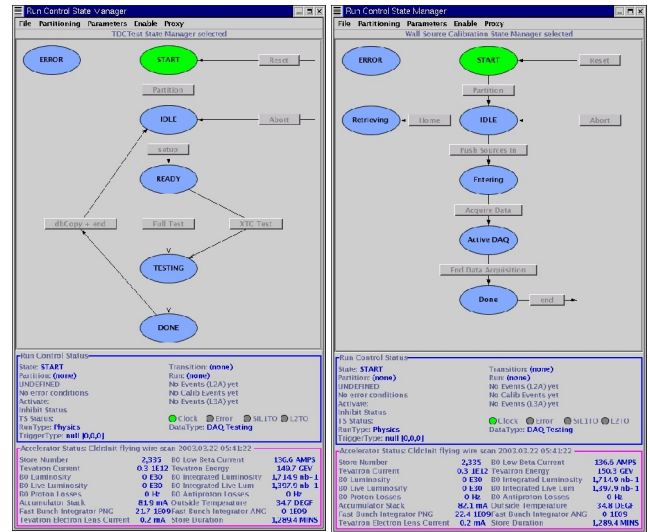

Figure 9: RunControl state managers for TDC testing and source calibrations
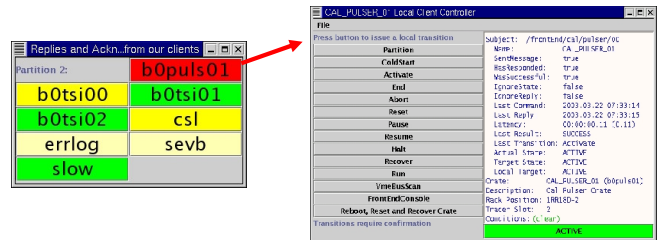

Figure 10: RunControl client status window and client fine grained control box

A suite of Java based monitors runs in conjunction with *RunControl* in the CDF control room and provides the shift-crew with fast feedback of the entire system status. The applications can be run stand-alone, or remote users can view the many web pages pertaining to the run status. For example, the *TevMon* application monitors accelerator conditions and informs the shift-crew and RunControl when the detector is in physical danger. The *ScalerMonitor* watches trigger rates and automatically maximizes bandwidth utilization when luminosity and trigger rates fall significantly below full capacity. Figure [11] demonstrates this process of trigger prescale relaxing.
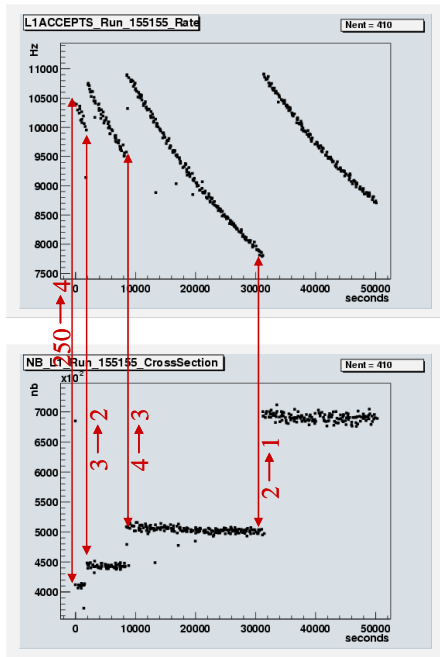
Figure 11: Maximizing bandwidth usage through trigger prescale reduction

The online monitors are fed by periodic messages from the front end crate CPUs, including status and data snapshots. Examples of monitoring displays are found in figure [12]. When data arrive at Level 3, full cross-detector synchronization and consistency checking takes place, with errors flagged in real-time and fed back to the ErrorHandler and *RunControl*, where the expert system decides on the appropriate recovery scenario. To minimize deadtime incursions due to error recovery, a fast reset sequence has been built in to the CDF DAQ, limiting the need for costly run restarts.
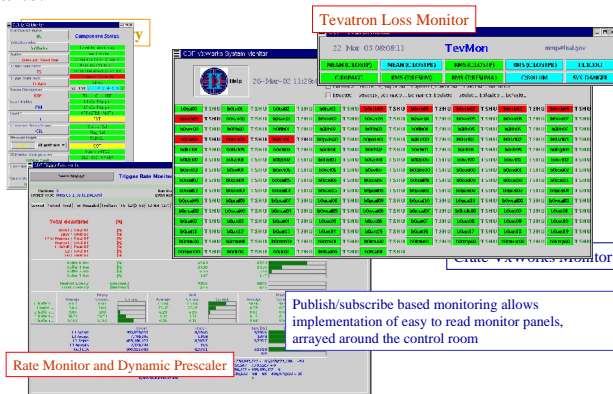

Figure 12: Examples of Online Monitoring Displays

Online monitoring takes place off-site at remote institutions around the world. A comprehensive set of run summary and run conditions web pages have been developed to provide both real-time and historical access to all characteristics of the CDF DAQ. Trigger rates, event errors, luminosities, backgrounds, and environmental quantities can are hot linked from the public run summary pages, including plots thereof, cross correlated, examples of which may be found in figure [13].
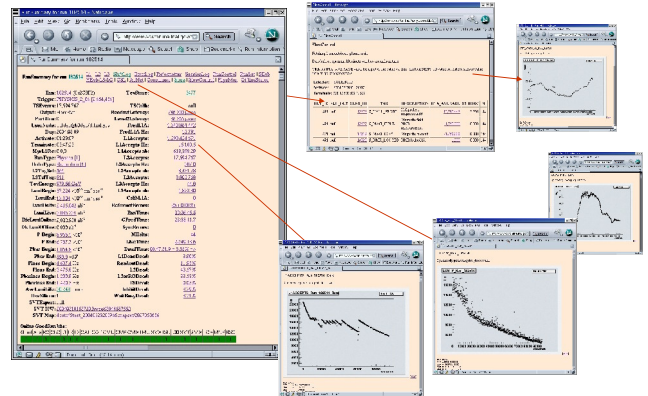

Figure 13: Examples of Run Summary web pages and plots

V. CONCLUSION

CDF has built a deadtime-less trigger and DAQ system that has been working very well over the last four years, and accumulating data over the last two and a half years. A suite of control and monitoring software, all web interfaced, has been developed and has made the shift-crew's job easier and more efficient. A smart monitoring system is in use, with fast error detection and recover, dynamic bandwidth maximization, and automated downtime calculations. The Run IIb upgrades scheduled for late 2005 will significantly increase system bandwidth.

VI. ACKNOWLEDGMENT