CDF Level 2 Trigger Upgrade – The Pulsar Project

M. Bogdan[1], R. Demaat[2], W. Fedorko[1], H. Frisch[1], K. Hahn[3], M. Hakala[2], P. Keener[3], Y. Kim[1], J. Kroll[3], S. Kwang[1], J. Lewis[2], C. Lin[2], T. Liu[2], F. Marjamaa[2], T. Mansikkala[2], C. Neu[3], M. Pitkanen[2], B. Reisert[2], V. Rusu[1], H. Sanders[1], M. Shochet[1], H. Stabenau[3], R. Van Berg[3], P. Wilson[2], D. Whiteson[3], P. Wittich[3]

[1]University of Chicago (UofC), Chicago (USA)
[2]Fermi National Accelerator Laboratory (FNAL), Batavia (USA)
[3]University of Pennsylvania (UPenn), Philadelphia (USA)
thliu@fnal.gov, reisert@fnal.gov

Abstract

The CDF data acquisition and trigger system is being upgraded to significantly increase the bandwidth for the upcoming high luminosity running of the Tevatron Collider (Run IIb). This paper focus on the upgrade for the Level 2 (L2) Trigger Decision Crate. This crate is at the heart of the L2 trigger system and has to interface with many different subsystems both upstream and downstream. The challenge of this upgrade is to have an uniform design to be able to interface with many different data paths upstream, merge and process the data at high speed for fast L2 trigger decision making, and minimize the impact on the running CDF experiment during the commissioning phase. In order to meet this challenge, the design philosophy of the upgrade is to use one type of general purpose motherboard, with a few powerful modern FPGAs and SRAMs, to interface any user data with any industrial standard link through the use of mezzanine cards. This general purpose motherboard, named "Pulsar" (PULSer And Recorder), is fully self-testable at board level as well as at system level. CERN S-LINK is chosen to allow Pulsar to communicate with commodity processors via high bandwidth, low latency S-LINK-to-PCI cards. Knowledge gained by using S-LINK at CDF will be transferable to and from the LHC community.

I. Overview of CDF Trigger System

The CDF Run II trigger is a three level hierarchical system. The first two levels, Level 1 (L1) and Level 2 (L2) shown in Figure 1, use custom-designed hardware to find physics objects based on subsets of the detector information. Level 3 uses the full detector resolution to reconstruct complete events in a processor farm. The goal of each stage in the trigger is to reject a sufficient fraction of the events to allow processing at the next stage with acceptable dead time.

The L1 system is a synchronous 40 stage pipeline. When an event is accepted by the L1 trigger, all data is moved to one of four L2 buffers in the front end electronics, and trigger data is sent to the asynchronous L2 system. Here, some limited event reconstruction is performed and a final L2 decision is evaluated at "Global Level 2". The final L2 trigger decision is based on primitives from upstream, as shown in Figure 1. L2 has at its disposal all trigger objects used in L1, such as tracks from the extremely fast track trigger (XFT/XTRP), muon information, global energy information, as well as the full L1 trigger decision information. In addition, the Shower-Max (CES/XCES) information for electron/photon identification, is available. Moreover, other higher level objects found in two dedicated L2 sub-systems, the Silicon Vertex Trigger (SVT) and the L2 Calorimeter (L2CAL), are used.

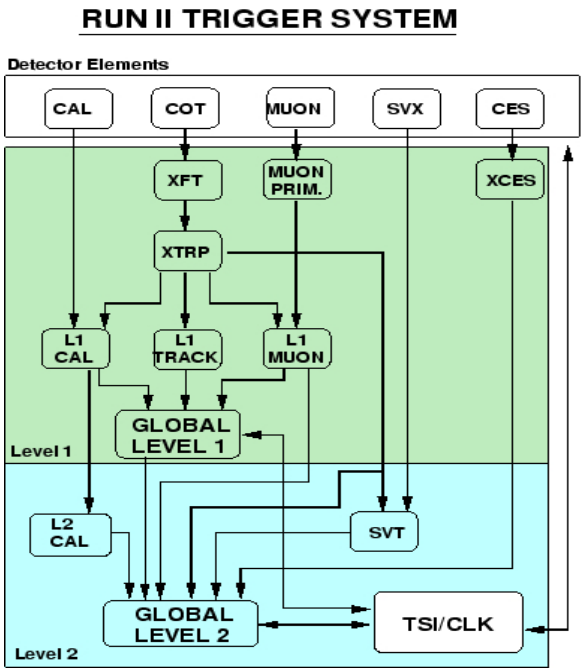

Figure 1: CDF Run II Trigger architecture. Note that only the first two levels are shown.

Dead time arises when an event is accepted by the first level trigger while all L2 buffers are occupied. There are only four L2 buffers available in CDF, this relatively shallow system requires that events be processed quickly in order to reduce deadtime. The Run IIa L2 trigger was designed to be able to handle up to 40kHz L1 Accept rate with L2 Accept rate around 300Hz. Based on these requirements, the job of the L2 trigger system can be split into two parts: *loading* data and *processing* data. *Loading* describes how long it takes from a L1 Accept until the data is available to the L2 decision making processors. *Processing* describes how long it takes to unpack the data, form objects and make a L2 decision based on simple kinematic cuts on objects or correlated sets of objects.

The current Run IIa L2 decision crate was designed and built in the mid to late 1990's based on technology available at that time. The design is relies on a custom bus (Magic Bus), several custom processors (DEC Alpha) and many different custom interface boards. The strategy we choose for the upgrade is to convert and pre-process all trigger fragments from upstream into a self-describing data format by a universal interface board. The common data stream is then merged and transferred via a standard link into a commodity processor, where the trigger decisions are made. In this approach, the only custom element involved is the universal interface board. The details of this upgrade is presented in this paper.

II. Challenges for CDF Trigger Upgrade

The CDF Run IIb trigger environment will be very challenging. The Run IIb baseline bunch spacing is 396 ns rather than the previously expected 132 ns. At the expected peak Run IIb luminosity of $3x10^{32}$ $cm^{-2}s^{-1}$, we will see ten interactions per crossing. This implies that the average data size will increase substantially, and the combinatorics will grow in processing multi-object triggers. For example, as the occupancy in the detectors increases with luminosity, the time it takes the L2 SVT to find trigger primitives increases, as does the number of primitives found. As a consequence, CDF will need to improve both the loading stage and the processing stage of the L2 trigger system. The overall goal of the upgrade is to improve the Level 2 trigger system performance so that it can handle a Level 1 accept rate above 30 kHz with a Level 2 accept rate around 750 Hz for high luminosity data taking.

One other challenge of this upgrade is to have a uniform design to be able to interface with many different data paths upstream. CDF uses different types of LVDS cable and optical fiber links for data interfaces at Level 2. For example, different types of LVDS cables and data protocols are used to transmit Level 1 trigger decision, global energy sums, tracking information from XFT/XTRP and SVT. The interface with Muon System, Level 2 calorimeter triggers (cluster and isolated cluster) and ShowerMax system (CES/XCES) are implemented via various types of optical fiber links (Hotlink [3] and Taxi [4]). The diversity of the interfaces involved (hardware as well as protocol) already makes the design of an universal interface board very challenge. Moreover, an universal board design has to be able to interface with the Trigger Supervisor (TSI) as well as Level 2 decision processors.

In addition to the requirements in speed and universal design, there is one challenge unique for any upgrade project, that is how to minimize the impact on the running experiment during the commissioning phase.

III. Pulsar Board Design

In order to meet the challenges, the design philosophy of the upgrade is to use one type of general purpose motherboard, with a few powerful modern FPGAs and SRAMs, to interface any user data with any industrial standard link through the use of mezzanine cards. This general purpose motherboard can be used either as a data sink or data source, hence its name “Pulsar” [1] (PULSer And Recorder), and is fully self-testable at board level as well as at system level. More information about Pulsar can be found at the Pulsar web page [2].

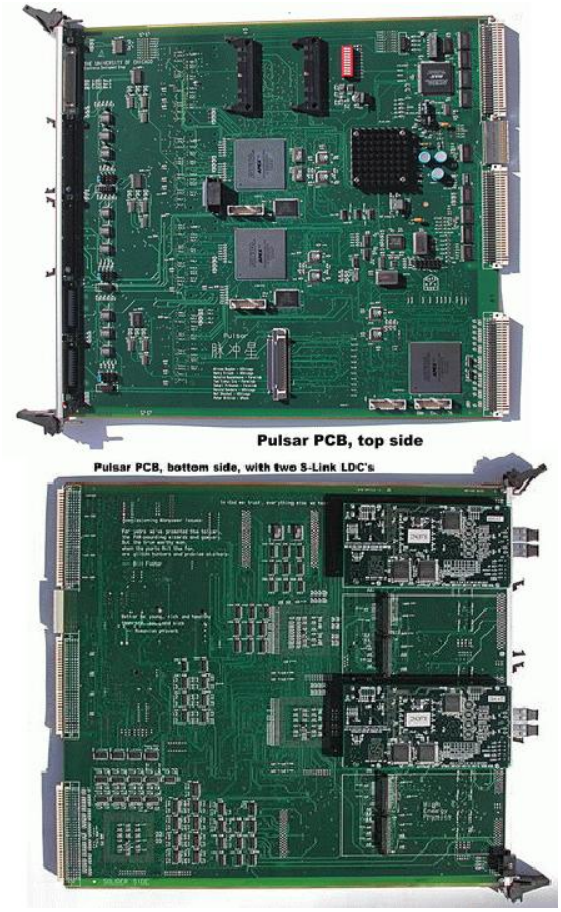

Figure 2: Top and bottom view of the Pulsar board [2].

Pulsar is a general purpose 9U VME board. Figure 2 shows the actual board, while Figure 3 shows the design of the board. Key devices on the Pulsar board are three large FPGAs: two DataIO FPGAs and one control FPGA. Each DataIO FPGA provides the interface to two mezzanine cards. The mezzanine card connections are all bi-directional (i.e. one can plug either transmitter or receiver cards). The implementation is similar to the CMC standard (Common Mezzanine Card) and the actual design followed S-LINK64 specification [5]. The four mezzanine card slots at the front of the board (on the bottom side), each has up to 83 user defined signals directly visible to motherboard FPGAs. Pulsar has user defined interface to P3 connector and this interface has up to 117 signals directly interfacing with the Control FPGA on board. This allows users to define which standard (or custom) link to interface with on the transition module on the back of the crate. The board also has user defined interface to P2 connector with up to 50 signals visible to all three main FPGAs on board via buffer chips. The user defined interfaces to both P3 and P2 are all bi-directional. In addition, there are three different types of LVDS connections at the front of the board which are specific to CDF application. However, they can be also used for applications outside CDF as well.

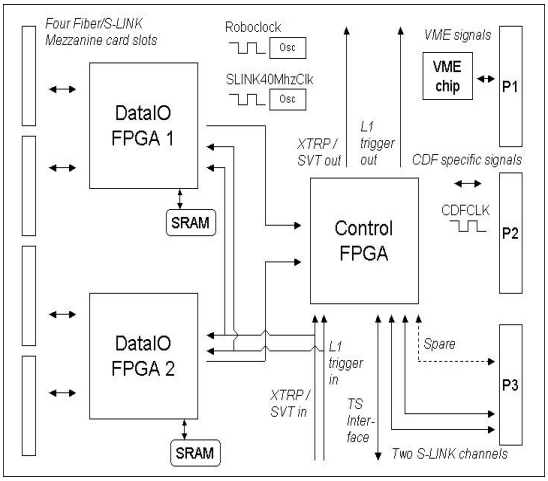

Figure 3: Schematic view of the pulsar design.

Pulsar board has all the interfaces to the L2 decision crate. The dedicated LVDS connections are for L1 trigger information, global energy sums, track information from XFT as well as from SVT, and the interface with CDF Trigger Supervisor. The rest of the interfaces are optical fibers and can be absorbed via two types of custom mezzanine cards (hotlink and taxi). For CDF Level 2 trigger application, CERN S-LINK [5] (stands for "Simple LINK") is chosen to be the standard link to allow Pulsar to communicate with commodity processors via commercially available, high bandwidth, S-LINK to PCI/PMC interface cards. This is done by using a simple transition module on the back of the crate to interface with SLINK mezzanine cards. The four mezzanine card slots on Pulsar board are also compatible with S-LINK interface mezzanine cards. In this application, Pulsar is used as an universal interface board to convert and merge many different trigger data paths into S-LINK standard. In addition, both the Level 1 trigger and track trigger information are made available to each Pulsar, allowing Pulsar to act as pre-processors to pass only Region-of-Interest trigger data downstream. This design feature is driven by physics requirements, providing flexibility in performance.

IV. Pulsar Design Methodology

A significant fraction of the design effort was dedicated to extensive design verifications by using state-of-the-art Computer Aided Design tools. The tools used for FPGA firmware development and gate level simulation areLeonardo Spectrum for VHDL synthesis and Quartus II for place and routing for logic arrays. Mentor Graphics QuickSimII using Smart Models together with netlist files created by QuartusII is used for board and multi-board level simulation. In addition, Interconnect Synthesis tool is used for trace and cross talk analysis to check signal integrity, and IS MultiBoard tool is used for signal integrity checks between the motherboard and mezzanine cards. The sophisticated tools significantly helped streamline the design process The prototype boards were also tested with on board clock speed up to 100 MHz and no problems were found. No layout or fabrication errors were found on the prototype boards, allowing them to serve as the production version.

V. PCI Interface and Processing Node

The communication between Pulsar board and commodity processors is done via commercially available, high bandwidth and low latency, S-LINK to PCI interface cards (S32PCI64) [6]. The S32PCI64 is designed to have low PCI-bus utilization and needs minimal host processor control. The S32PCI64 cards are used both in receiver mode (to send data into processors) and in transmitter mode (to send L2 decision back to a Pulsar).

The decision processors chosen are commodity dual-processor x86 type. Historically, to guarantee performance, real-time operating systems have been required. The standard 2.4 Linux kernel, however, provides a mechanism to schedule processes with real time priority. In addition, it provides the means to bind peripheral interrupts to specific CPUs in a multiprocessor environment; we use this feature to leave the second CPU free to process data and make trigger decisions. Together, these features allow operational performance which approaches that of a real time system.

The Run IIa system uses 500MHz DEC α processor on a custom designed processor board to reconstruct events and make trigger decisions. More recent and powerful processors are expected to reduce the processing time dramatically. We compared the performance of the α processor to a Intel XEON 2.4 GHz processor and AMD Opteron 2.4 GHz processor running the trigger algorithms on real events. To isolate the CPU requirements, we removed data transmission delays from the timing measurements. Figure 4 shows a marked decrease in the mean processing time and a sharp reduction in the long tail.

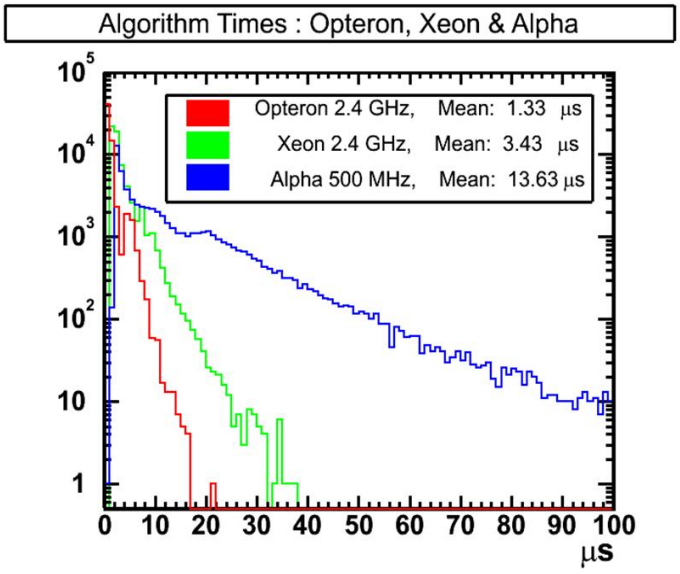

Figure 4: Comparison of the event processing time of the Run IIa processor, DEC α, and two possible choices for the Run IIb processor, Intel Xeon and AMD Opteron.

The differences between the XEON and Opteron processors, which nominally run at the same clock frequency, are to a large extent due to differences in the PCI bus architecture of both processors. The round trip latency, for Pulsar to send data into CPU memory and for CPU to send decision back to Pulsar, has been measured to be around 10 μs when the S32PCI64 cards are used. This meets our Level 2 trigger upgrade specification.

VI. Initial System Configuration for CDF Level 2 upgrade

The Pulsar system configuration for the initial phase of CDF Level~2 trigger upgrade is shown in Figure 5. A few Pulsar boards act as preprocessor boards to interface with all data paths upstream. Among them, three are preprocessor boards for the ShowerMax system, one is to receive L1 muon, tracks from XTRP and L1 trigger information, another one is for interface with the calorimeter information. Two Pulsars are used as S-LINK merger, one is to merge the output of three ShowerMax Pulsars, while the other one is for the final S-LINK merging to deliver the final S-LINK package into the processor. The SVT data, the path with the longest latency, is delivered on a separate PCI path via a separate Pulsar board. In addition, one Pulsar (L2toTS) is dedicated to receive the decision from the processor via S-LINK and communicate the decision to the CDF Trigger Supervisor.

With this system configuration, there are total of six different types of Pulsar used, all with the same motherboard but different mezzanine cards and FPGA firmware design. For all Pulsar boards used in the system, diagnostic DAQ buffers

have been implemented allowing us to readout the intermediate information (data as well as timing information) throughout the Pulsar system into the data stream. This design feature is essential for commissioning, optimizing, as well as long term maintenance of the system.

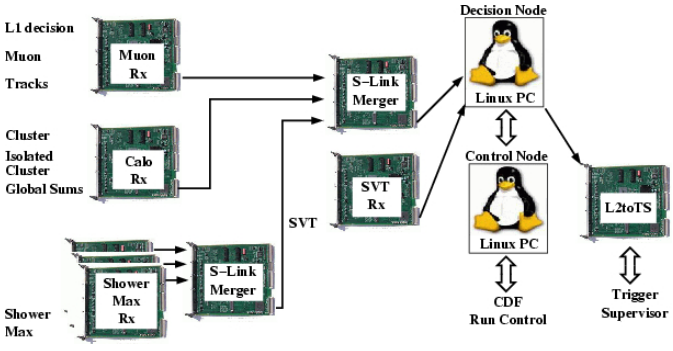

Figure 5: CDF Level2 Trigger upgrade system configuration.

VII. Commissioning strategy and initial experience

The self-testing capability of Pulsar design allows us to test each data path, hardware as well as firmware, in a test-stand using additional Pulsars configured in transmitter mode. As described above, there are six different types of Pulsar boards in the final system, and there are six different types of transmitter Pulsar board used in the teststand configuration. All hardware and firmware have been tested extensively in this controlled environment before integrating into the trigger system.

In order to minimize the impact on the operation of the CDF experiment during the system commissioning phase of the Level~2 upgrade, all input data paths have been split so that a copy of the input data is made available to the new upgrade system. The initial system commissioning work has been done using cosmic and other non-beam trigger configurations, as well as with beam in pure parasitic mode. The system has been tested extensively using this methodology before we requested for dedicated beam time to allow the new Level 2 to drive CDF. This approach has been very successful for the commissioning this summer. We only used a few hours of dedicated beam time for testing the system while delivering the triggers. In fact, the Pulsar Level 2 system worked on the first attempt in the initial com-missioning test run with dedicated beam. For all Level 2 trigger algorithms implemented sofar the trigger decision from the upgrade system perfectly matches that one expected from the legathy Level 2 system.

The goal of the initial phase of commissioning during summer of 2004 is to test the system robustness, and in particular, to demonstrate (at the proof-of-principle level) that we can deliver the Level 2 trigger decisions to drive CDF with beam. For this initial commissioning the emphasis was not on system performance optimization, rather to collect enough beam data and timing information to guide system optimization during the shutdown period this fall. The initial system performance was measured in a dedicated test run using a subset of Level 1 triggers and Level 2 trigger algorithms, including tracks, muon and SVT algorithms at Level 2. A comparison of the overall Level 2 latency, the time from Level 1 accept to the broadcasting of the Level 2 decision, is shown in Figure 6 for the legacy Level~2 system (top) and the upgrade system (bottom). Overall, the new system is already performing as good as the old system, if not better.

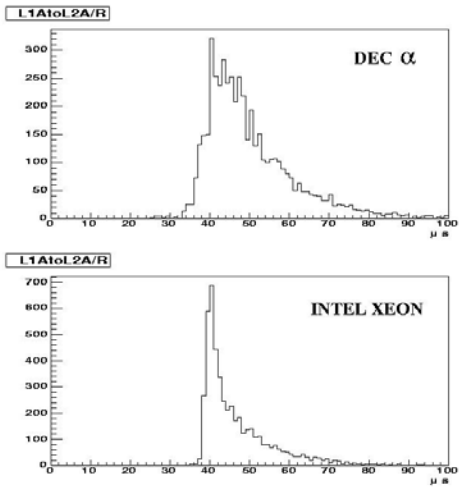

Figure 6: Global L2 latency, from Level 1 accept to broadcast of Level 2 decision for alpha system(top) and new Pulsar-based system (bottom).

For the new system, one un-optimized data path (the ShowerMax) dominates the latency. We expect that there is much room for improvement when the data handling in the Pulsar board for this data path is optimized during the shut-down. The tail to the right is dominated by the late arrival of the SVT data, this will be improved significantly with the upcoming SVT upgrade.

VIII. Future Improvements

The system performance may be optimized in various ways. For example, at board level, the data volume can be suppressed further for some data paths, and timing of the firmware can be improved.

At system level, one promising upgrade option is to use the CERN FILAR [7] instead of S32PCI64. By using FILAR, we can eliminate the need for the two Pulsar S-LINK mergers, thus allowing all data fragments to be sent directly to the CPU memory via PCI bus. Using FILAR also allows one to run the S-LINK mezzanine cards at higher speed. In addition, FILAR has less PCI overhead than that of S32PCI64. Figure 7 shows one possible system configuration using FILAR.

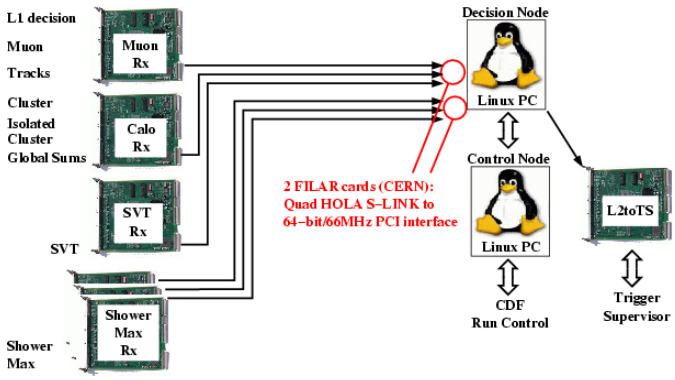

Figure 7: Possible future system configuration of CDF Level~2 Trigger upgrade using Slink Filar cards.

Further improvements may be achieved by using four Level 2 decision nodes each dedicated to a given Level 2 buffer event. This is possible since Pulsar has two S-LINK channels over P3, the potential gain here is that the processing of a given event does not have to wait for the previous event processing to be finished.

IX. Other Applications of Pulsar

Although Pulsar is designed primarily as an upgrade path for the CDF Level 2 trigger decision crate, the design is general enough that it can be potentially used in many other applications, within CDF or outside CDF. It can be used as a general purpose interface board, as a standalone DAQ system (such as a test beam environment) or software based trigger system when combined with modern CPUs, or even as a general purpose diagnostic test tool. Pulsar design is powerful, modular, universal and self-testable. It is capable to interface any user data with any industrial standard link (for example, CERN S-LINK or Gigabit Ethernet) through the use of custom mezzanine cards. The design is such that users can choose which standard link to interface with via simple custom transition module or mezzanine card.

Within CDF, Pulsar board is also used for the SVT upgrade, to replace three major components in the current SVT system with the goal to improve the overall Level 2 performance (i.e. latency). In addition, it will be used for the XFT upgrade, to interface with the new stereo segment finders and make the information available to Level 2 trigger.

X. Summary

Pulsar is a general purpose 9U VME interface board designed for HEP applications. In this paper, we have presented its application for the CDF Level 2 trigger decision crate upgrade. The Pulsar-based design departs significantly from the previous implementation of the existing Level 2 decision crate. This new system is designed to have sufficient safety margin and flexibility in performance to meet the Run IIb trigger challenges, to have built-in self-test capabilities to speed up the commissioning process and to ease the long term maintenance effort all the way through the end of Run IIb.

The board and system design of the new Level~2 decision crate, as well as the commissioning experience this summer and initial system performance, are described. This is a project where the S-LINK technology developed at CERN for the LHC experiments is used for the first time at high rate hadron collider environment. Knowledge to be gained by using S-LINK at CDF will be transferable to and from the LHC community.

References

[1] T.Liu et al., “Pulsar Design and Testing Methodology for CDF Level 2 Trigger Upgrade” presented at IEEE NSS conference, Portland, Oregon, Oct. 22, 2003.

[2] http://hep.uchicago.edu/~thliu/projects/Pulsar/

[3] Cypress Semiconductor Corporation, “HOTLink™ Transmitter/Receiver (CY7B923/CY7B933)” Data Sheet, 1999.

[4] TAXIchip™ Integrated Circuits “Transparent Asynchronous Transmitter/Receiver Interface Am7968/Am7969-125 Am7968/Am7969-175” Data Sheet and Technical Manual, 1994.

[5] H. C. van der Bij *et al*, “S-Link, a Data Link Interface Specification for the LHC Era”, 1997. Published in the Proceedings of the Beaune 97 Xth IEEE Real Time Conference. See also http://hsi.web.cern.ch/HSI/s-link/

[6] W. Iwanski *et al*, “Designing an S-LINK to PCI Interface using an IP core. Article presented at the 12th IEEE-NPSS Real Time Conference, June 4-8, Valencia (Spain). See also http://hsi.web.cern.ch/HSI/s-link/devices/s32pci64/}.

[7] H. C. van der Bij *et al*, “FILAR, Quad HOLA S-LINK to 64-bit/66 MHz PCI Interface, Users Guide” https://edms.cern.ch/file/337904/1/userguide.PDF .