A large Associative Memory system for the CDF Level 2 Trigger

A.Bardi[2], S.Belforte[2], A.Cerri[3], M.Dell'Orso[1], S.Donati[1], S.Galeotti[2], P.Giannetti[2], E.Meschi[1], F.Morsani[2], G.Punzi[3], L.Ristori[3], F.Spinella[2],

[1]Dipartimento di Fisica, Università di Pisa, Piazza Torricelli 2, 56100 Pisa, Italy

[2]INFN Pisa, Via Livornese 1291, 56010 S.Piero A Grado (PI), Italy

[3]Scuola Normale Superiore, 56100 Pisa, Italy

Abstract

A large Associative Memory system for on-line track reconstruction in a hadron collider experiment has been designed, prototyped and tested. This is the first such application of the Associative Memory concept and it is based on a full custom VLSI chip developed within this project. The Associative Memory is the heart of the Silicon Vertex Tracker, which is part of the Level 2 trigger of the CDF experiment, and is able to complete track finding in the CDF silicon vertex detector less then 1μsec after detector readout is over.

This system is a multi-board project running on a common 30 MHz clock, but critical parts multiply clock frequency to operate up to 120 MHz. The Associative Memory board architecture, design, implementation and test are described. The main characteristics of this project are the use of sophisticated clock distribution techniques and the high density of components.

I. Summary

The Associative Memory Board (AMB) is a bank of content–addressable memory for the purpose of fast pattern recognition[1]. It is part of the Silicon Vertex Tracker[2], a trigger processor dedicated to reconstruction of charged particles trajectories at the Level 2 of the CDF trigger. Every time an event is accepted by the Level 1 trigger, the digitized pulse height in the Silicon Vertex Detector are sent to the Hit Finders which calculate hit positions. The hits found by the Hit Finders and the tracks found in the Central Outer tracker by the Level 1 Fast Track finder (XFT) are then fed to both the Hit Buffer[3], where they are temporarily stored, and the Associative Memory system. The Associative Memory system is made of one control board (the Associative Memory Sequencer, AMS) and several AM boards. Its function is to perform pattern recognition, that is to select for further processing only hit combinations that represent good track candidates. This is the most computationally-intensive part of SVT data processing, and it is done by comparing the input data with a stored set of patterns in a completely parallel way, using a dedicated custom VLSI chip (AMchip[4]). This technology allows for pattern recognition to be carried out "on–the–fly" during detector readout (input phase), the results are therefore available shortly after the end of the input phase. This pattern recognition process is carried out at a coarser resolution than the full available resolution, by discretizing hit coordinates in bins called SuperStrips. The AM system outputs a list of Roads, each road is defined as a combination of SuperStrips on different detector layers that can be traversed by a single track. The Roads are sent to the Hit Buffer, that retrieves the original full–resolution hit coordinates and delivers them to the Track Fitter system for full–precision calculation of track parameters.

A. Architecture and Operation modes

The AM system is organised as a set of 9Ux400 VME boards. Up to four AMB's, each holding up to 128 AMchips, are controlled by one single AMS board through a custom P3 backplane. The Sequencer board handles all I/O and is the interface between the Associative Memory system and the rest of the SVT system and provides the proper opcode sequences to the AMB's.

In addition to the P3 bus, each AMB has its own VME interface and can be operated in two different modes. In 'VME mode' the board ignores signals coming from the P3 backplane, and operates on an internal (slow) clock; this mode is used to load the patterns into the associative memory at power–up and for diagnostic purposes. In 'running mode' the board is controlled by the P3 bus and runs with the fast clock distributed on the backplane by the AMS. The working of the AMB's and of the AMS is fully synchronous. In this mode reading and writing of the memory contents is disabled. Toggling between modes is controlled by VME.

In running mode, the board distributes the opcodes and the data received on P3 to all AMchips, collects data output by all chips and queues them to the P3 bus. This is done through two tree–like structures, an input tree and an output tree, with 128 AMchips at the bottom level. All AMboards connected to the same Sequencer are also managed as different branches hanging from a single root node, the Sequencer itself.

B. The GLUE tree

The input and output tree have four levels. Each level is implemented as a number of identical GLUE chips implemented using Field Programmable Gate Arrays (FPGA) by QuickLogic. These chips are named GLUE0, GLUE1, GLUE2 and GLUE3 following the level in the tree starting from the bottom. The GLUE0 has a fanout of 8 (it handles 8 AMchips), all the others have a fanout of 4. GLUE0, GLUE1 and GLUE2 are located on the AMB while GLUE3 is located on the AM Sequencer. At each level there are common input and output busses, and individual control lines for each chip.

To ensure fast and reliable operation, the input (downward) tree operates as a simple pipeline synchronized on the board clock, moving the data one level down at every clock cycle without any handshake. The output (upward) pipeline

works instead with a synchronous handshaked protocol and allows sequential queueing of data which may be present simultaneously on several tree nodes at the same level, one word is pushed up the tree each 3 clock cycles. The reason for two different protocols is twofold: on one side moving data up the tree from the 128 AMchips to the single AM Sequencer requires additional time to switch the common address bus from one chip to the other, on the other the typical data flow in the system is of several hundreds of data words going in and very few coming out (many hits, few tracks !). Therefore the GLUE is designed to keep the clock cycle as short as possible (in each clock cycle signals only go from one driving to one receiving chip) allocating more cycles for the more complex output function.

Data moving up the tree is handled by two finite state machines inside each GLUE chip, one for the down and one for the up side. All FSM's work on the same common clock sent by the AM sequencer. The bottom GLUE0 chip is designed to hide the AMchip complexity and interface to the higher levels of the tree through a simple interface synchronized on a single clock edge. The down FSM of the GLUE0 also controls the clock lines to the AMchips: to achieve the best timing performance the clocks can not run freely, but need to be switched on and off at the right times in order to break down the AM chip output operation into elemental steps that could be fit into the 30 ns clock cycle. The need to use FPGA chips with low current output to drive critical clock lines with a large fanout (8 chips) required special care in the PCB design.

When the board is in VME mode both pipelines are disabled and all chips are simultaneously addressed in much the same way as an ordinary RAM.

C. Clock distribution

Clock is distributed through an independent tree. The AMchip uses 4 distinct phases in each clock cycle, none of them coincident with the common signal clocking the whole pipeline. Moreover, for a proper handshake of the data exchange with the lowest lying GLUE chip (called GLUE0), an additional clock phase need to be delivered to the GLUE0 chip. This brings the number of different clock phases to be distributed over the board to 6. In order to fully exploit the speed performance obtainable from the AMchips, the clock edges must be placed with an accuracy of couple of nanosecond or so everywhere on the board. To simplify the problem of clock distribution, only one clock phase is distributed over the board (master clock), and the 5 remaining phases are locally generated where needed with definite relative timing.

Both clock distribution and multiple phase generation are accomplished by using PLL based clock buffers from Cypress Semiconductors ('Roboclocks'). Each Roboclock produces eight outputs with 4 adjustable phase shifts with respect to the input clock, also allowing zero or negative delays. The 6 clock phases are generated by 2 Roboclocks plus external logic that also provides appropriate feedback to the PLL forcing it to oscillate at double frequency than the board clock. The timing of each of the 6 phases is adjusted by controlling the individual programmable delays of each Roboclock output, in steps of about 1ns. All delays are jumper–programmable to allow fine tuning of the strobes as a function of the external driving frequency. When the board is put in 'VME mode' the PLL's are disabled using the TEST feature of the Roboclocks, that makes them act as simple fixed delays. The downstream logic however is still capable of generating 4 phases in quadrature. In this mode, the clock from the P3 connector is switched off and the clock to the AMchips is run directly under VME control. When the board is switched to P3 mode again, it is held in a wait state for about a millisecond, with all clocks to the AMchips disabled, to give time to all PLL's to lock–in.

D. PCB design

The 128 AM chips cover a very large fraction of the usable component area of the AM board, adding the GLUE and clock distribution trees left almost no place for the VME logic and routine. As a consequence the the AMboard was physically implemented as a single–width 9U PCB with 16 connectors for mezzanine cards that carry the lowest two levels of the 'tree'. The board can operate with any number of cards inserted. This design offers several advantages compared to a 'flat' design, such as: easier to change/upgrade design, extra routing space for tight optimization of the most critical parts on the mezzanine boards, saving in the number of AMchips when a smaller memory is needed, easy replacement of defective parts, possibility to test each mezzanine card on a smaller and simpler setup before mounting.

The mezzanine card PC layout required a lot of care due to the space limitation and the large fanout required. Analog simulation of the critical clock signals was extensively used to define the best layout and the chosen solution was to use series termination for most control signals, while the clock lines have been layed out in a "T" shape from the GLUE0 to the center of the AMchips array, where they split into two lines of identical length and doubled impedance to guarantee good matching and cancellation of end-of-line reflections. This design worked very well both in simulation and in the prototype.

E. Test

An AMboard prototype has been tested together with a Sequencer prototype and showed reliable operations for many days at a frequency of 30 MHz. Individual mezzanine boards have been successfully tested for several days at 37 MHz.

II. References

[1] M. Dell' Orso, L. Ristori "VLSI Structures for Track Finding", *Nuclear Instruments and Methods A, 278 (1989)* p. 436;

[2] S. Belforte et al. "The CDF Trigger SVT", *IEEE Trans. on Nucl. Sci., 42 (1995)* p. 860.

[3] S. Belforte et al. "The SVT Hit Buffer", *IEEE Trans. on Nucl. Sci., 43 (1996)* p. 1810.

[4] S.R. Amendolia et al., "The AMchip: a Full-custom CMOS VLSI Associative Memory for Pattern Recognition", *IEEE Trans. on Nucl. Sci., vol. 39, 1992* pp. 795-797.