

# The GigaFitter: Performance at CDF and Perspectives for Future Applications

**S Amerio** <sup>a,b</sup>, **A Annovi** <sup>c</sup>, **M Bettini** <sup>b</sup>, **M Bucciantonio** <sup>d,e</sup>, **P Catastini** <sup>f,e</sup>, **F Crescioli** <sup>d,e</sup>, **M Dell’Orso** <sup>d,e</sup>, **P Giannetti** <sup>e</sup>, **D Lucchesi** <sup>a,b</sup>, **M Nicoletto** <sup>b</sup>, **M Piendibene** <sup>d,e</sup> and **G Volpi** <sup>d,e</sup>

<sup>a</sup> University of Padova, via 8 Febbraio, 2 - 35122 Padova, Italy

<sup>b</sup> INFN Padova, via Marzolo, 8 - 35131 Padova, Italy

<sup>c</sup> INFN Laboratori Nazionali di Frascati, via E.Fermi, 40 - 00044 Frascati (Rome), Italy

<sup>d</sup> University of Pisa, Lungarno Pacinotti, 43 - 56126 Pisa, Italy

<sup>e</sup> INFN Pisa, Largo B.Pontecorvo, 3 - 56127 Pisa, Italy

<sup>f</sup> University of Siena, via Banchi di Sotto, 55 - 53100 Siena, Italy

E-mail: [silvia.amerio@pd.infn.it](mailto:silvia.amerio@pd.infn.it)

## Abstract.

The Silicon Vertex Trigger (SVT) is a processor developed at CDF experiment to perform online fast and precise track reconstruction. SVT is made of two pipelined processors, the Associative Memory, finding low precision tracks, and the Track Fitter, refining the track quality with high precision fits. We will describe the architecture and the performances of a next generation track fitter, the GigaFitter, developed to reduce the degradation of the SVT efficiency due to the increasing instantaneous luminosity. The GigaFitter reduces the track parameter reconstruction to a few clock cycles and can perform many fits in parallel, thus allowing high resolution tracking at very high rate. The core of the GigaFitter is implemented in a modern Xilinx Virtex-5 FPGA chip, rich in powerful DSP arrays. The FPGA is housed on a mezzanine board which receives the data from a subset of the tracking detector and transfers the fitted tracks to a Pulsar motherboard for the final corrections. Instead of the current 12 boards, one per sector of the detector, the final system will be much more compact, consisting of a single GigaFitter Pulsar board equipped with four mezzanine cards receiving the data from the entire tracking detector. Moreover, the GigaFitter modular structure is adequate to scale for much better performances and is general enough to be easily adapted to future High Energy Physics (HEP) experiments and applications outside HEP.

## 1. Introduction

Real time event reconstruction plays a fundamental role in High Energy Physics experiments. Reducing the rate of events to be saved on tape from millions to hundreds per second is critical. In order to increase the purity of the collected samples, rate reduction has to be coupled with the capability to simultaneously perform a first selection of the most interesting events.

The CDF II detector, located at Fermilab (Batavia, USA) can perform a high-resolution track reconstruction at the trigger level due to the Silicon Vertex Trigger (SVT). This tracking processor was originally built for B-physics event reconstruction, and had an extremely significant impact on the CDF II physics program as a whole. The SVT allows the selection

at trigger level of events with displaced tracks; it increased by several orders of magnitude the efficiency for the identification of hadronic decay of B mesons.

Tevatron is improving its performance (initial instantaneous luminosity  $> 3 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  and  $\sim 50 \text{ pb}^{-1}$  of collected data per week) providing a greater amount of data for analysis. Nevertheless this improvement has the drawback of an exponential increase in trigger rates and a degradation of the reconstruction efficiency of physical objects due to the higher occupancy of the detector.

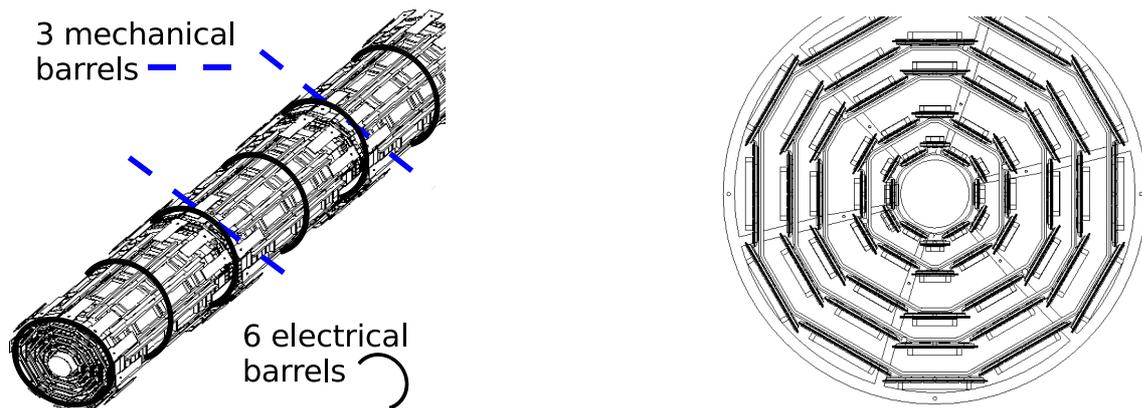
The GigaFitter (GF) is a next generation track fitter designed as a possible upgrade for SVT system, in order to enhance its performances in a very high luminosity environment. The GF is based on a modern Xilinx Virtex-5 FPGA chip, rich in powerful DSP arrays and features high speed operation, modularity, flexibility and reduced size with respect to the current system. In this paper we will describe in detail its architecture and the preliminary tests made at CDF II.

## 2. The CDF II Detector

CDF II is a general-purpose, azimuthally and forward-backward symmetric detector located at the Tevatron  $p\bar{p}$  collider at Fermilab. It consists of a charged-particle tracking system immersed in a 1.4 T magnetic field followed by calorimeters which are in turn, surrounded by muon detectors. A detailed description can be found elsewhere [1].

The CDF II coordinate system uses  $\theta$  and  $\phi$  as the polar and azimuthal angles respectively, defined with respect to the proton beam axis direction,  $z$ .  $x$ -axis points towards the centre of the accelerator while  $y$  axis upwards from the beam. The pseudo-rapidity  $\eta$  is defined as  $\eta \equiv -\ln[\tan(\theta/2)]$ . The transverse momentum of a particle is  $p_T = p \sin \theta$ .

The part of the detector relevant for this work is the tracking system, in particular the Silicon Vertex Detector (SVXII) and the Central Outer Chamber (COT) whose data are used by SVT track reconstruction. SVXII is a 5-layer double sided silicon micro-strip detector, extending from  $r = 2.1 \text{ cm}$  to  $r = 17.3 \text{ cm}$  outside the beampipe and providing  $|\eta|$  coverage up to 2. Along  $z$  direction SVXII is divided into 6 electrical and 3 mechanical barrels, while in the transverse plane is divided in 12 sectors (wedges) each  $30^\circ$  wide in azimuthal angle (fig.1)



**Figure 1.** SVXII barrel (left) and wedge (right) structure.

The drift chamber consists of cells divided into 8 super-layers, each containing 12 layers of sense wires. The odd superlayers have wires parallel to the beam axis (*axial* layers) while the even ones have wires tilted by  $2^\circ$  (*stereo* layers) in order to provide stereo information. It provides  $|\eta|$  coverage up to 1.

### 3. CDF II Trigger System

The CDF II trigger system has a three level architecture designed to reduce the event rate from 2.53 MHz, the bunch crossing rate, to approximately 150 Hz to be written on tape. At Level-1 (L1) raw muons, tracks and calorimeter information are processed to produce a L1 decision. L1 is a synchronous 40 stage pipeline based on custom-designed hardware which can provide a trigger decision in  $5.5 \mu s$  with a rate typically below 30 kHz. When an event is accepted at L1, subsets of detector information are sent to the Level-2 (L2) system, where some limited event reconstruction is performed and a decision is taken. The L2 is an asynchronous pipeline and it is based on a combination of custom-designed hardware and commodity processors. Its average latency is  $20 \mu s$  and its maximum output rate is 1 kHz. Upon L2 accept, the full detector data is readout and sent to Level-3 (L3) processors farm for further processing. Events accepted at L3 are sent to mass storage.

Tracks are reconstructed at trigger level by the eXtremely Fast Tracker (XFT) at L1 and by the Silicon Vertex Trigger (SVT) at L2.

#### 3.1. XFT

XFT [2, 3] has been developed to reconstruct tracks in the plane of the drift chamber transverse to the beam axis in time for the L1 decision using hit data from the 4 *axial* superlayers of the chamber. Track identification is performed searching for and combining track segments in the 4 axial superlayers of the drift chamber. The XFT reconstructs transverse momentum,  $p_T$ , and azimuthal angle,  $\phi$ , of tracks with  $p_T > 1.5 \text{ GeV}/c$  with an efficiency greater than 96% and a resolution  $\sigma_{p_T}/p_T^2 \sim 2\%$  ( $\text{GeV}^{-1}$ ) and  $\sigma_\phi \sim 6 \text{ mrad}$ .

Track segments are also found in the outer stereo layers of the chamber. This feature allows the rejection at L1 of fake axial tracks by requiring the association with *stereo* segments. Stereo segments are also sent to L2 and matched to the axial tracks for 3D-track reconstruction which provides a good resolution on  $\cot \theta$  ( $\sigma_{\cot \theta} = 0.11$ ) and  $z_0$  ( $\sigma_{z_0} = 11 \text{ cm}$ ).

#### 3.2. SVT

SVT [4, 5] is a L2 trigger processor dedicated to the reconstruction of charged particle trajectories in the plane transverse to the beam line. Its inputs are the  $p_T$  and the azimuthal angle  $\phi$  of the tracks found by XFT and the data from SVXII. The SVT proceeds through three main steps: hit finding, pattern recognition and track fitting. First SVXII hits are found by a hit finder algorithm and stored in hit buffers; then association between XFT tracks and SVXII hits is performed by the *Associative Memory* (AM), a massive parallel mechanism based on the search for hits from tracks within low resolution (*roads*) among the list of SVXII hits and XFT tracks; a road is a coincidence between hits on four of the silicon layers and a XFT track and it is precalculated and stored in large memories. Upon receiving a list of hits and tracks, each AM chip checks if all the components of one of its roads are present in the list of hits and XFT tracks. When it has determined that a road might contain a track, the road's hits are retrieved from the hit buffer and passed to the track fitter (TF). The TF calculates the track parameters by a simple scalar product, using a linear approximation in each SVXII wedge. It provides precise measurement of track impact parameter  $d_0$ , curvature and azimuthal angle for all tracks with  $p_T > 2 \text{ GeV}$ , as well as the  $\chi^2$  of the fit. The impact parameter is measured with a resolution of  $35 \mu m$  for  $2 \text{ GeV}/c$  tracks, which is comparable to the resolution obtained for offline reconstruction. Overall the SVT tracking efficiency is about 80% with respect to offline tracks with  $p_T > 2 \text{ GeV}$  and at least 4 hits in the silicon detector.

*3.2.1. SVT limitations* The possible roads are precomputed offline and their number is limited by the memory available to store them. The current version of SVT is provided with 512,000 roads per SVXII wedge. This number doesn't account for all the possible tracks which can be

left by charged particles traversing the detector. For example, tracks with  $p_T < 2$  GeV are not considered, neither are tracks crossing more than one of the three SVXII mechanical barrels.

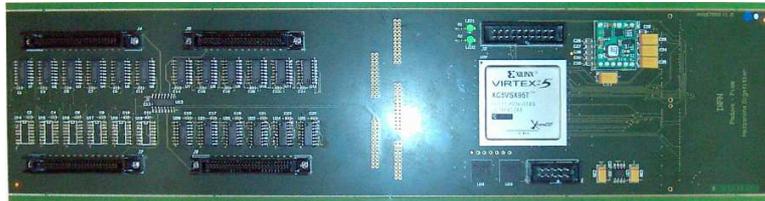
The size of the road is  $200 \mu\text{m}$  in  $r - \phi$ . This value is the best compromise achievable with the current hardware between track reconstruction efficiency and processing time. A larger value would provide a greater efficiency but also higher processing time, because many tracks could be associated to the same road. On the other hand, a smaller size would reduce the processing time, at the price of a lower efficiency, because the number of possible roads is limited.

The track coordinates are expressed by 18 bits, while the TF board is provided with  $8 \times 8$  bit multipliers. The scalar product performed by the TF has to be split into two terms: one is a function of the road boundary and can be precalculated; the other, function of the distance between the road boundary and the hit, is calculated online. The precalculated terms need a large memory to be stored.

Finally, TF memory limits to 32 the maximum number of tracks which can be fitted for each road.

#### 4. The GigaFitter

The GigaFitter (GF) has been designed as a possible upgrade of the current TF, in order to overcome SVT limits and enhance its performances at high luminosity [6]. It is based on a modern FPGA, the Xilinx Virtex 5 XC5VSX95T [7]. This device can run up to 550 MHz and is provided with 640 DSPs, each consisting of a  $18 \times 25$  bit multiplier and a 48 bit adder. The Virtex 5 is mounted on a mezzanine card (fig.2), which can receive the data from 4 SVXII wedges.

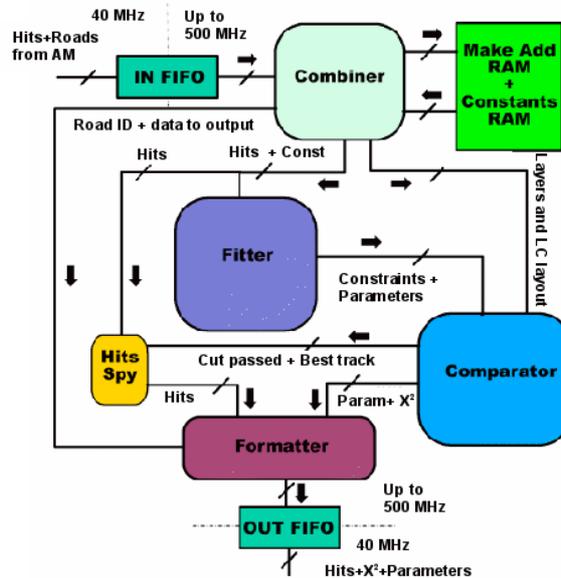


**Figure 2.** Mezzanine card hosting the Xilinx Virtex-5 FPGA. On the left the 4 connectors used to receive data from 4 SVXII wedges.

##### 4.1. The basic fitting line

The basic fitting algorithm implemented inside the Virtex 5 is shown in fig. 3. It is naturally divided in five different parts, the *Combiner*, a large RAM, the *Fitter*, the *Comparator* and the *Formatter*. Hits from SVXII, XFT tracks and road information are first received by the *Combiner*. As a road can have more than one hit per SVXII layer and every hit can be associated to a track, the *Combiner*, for each road, forms all possible combinations of hits. Then it serially sends all the track candidates to the *Fitter*, together with the appropriate fit constants retrieved from the RAM. The *Fitter* receives the hits and the constants and, exploiting the large number of DSP inside the chip, simultaneously calculates the track and quality ( $\chi^2$  components) parameters. The scalar products are calculated by a serial pipeline composed of a  $18 \times 25$  bit multiplier and an adder/accumulator inside the DSP processor. Once the results are ready, the  $\chi^2$  components are sent to the *Comparator* while the track parameters are stored in a FIFO waiting for the *Comparator* decision. The *Comparator* also receives information on the hit combination layout (layer used and quality of the hits) in order to fine-tune the decision. The *Comparator* calculates the  $\chi^2$  and compares it with a threshold. If a track candidate doesn't

pass the  $\chi^2$  cut, it is discarded, while, in the opposite case, it is accepted and its parameters and  $\chi^2$  value are stored in small FIFOs. The scalar product in the *Fitter* is made using 4 hits from SVXII and the  $p_T$  and  $\phi$  from the XFT track. In the case of a track with hits in all 5 SVXII layers (*full of hits track*) the current TF uses only 4 out of 5 hits, discarding one hit on the basis of the layers used and the quality of the hits. In a high luminosity environment, with increasing probability of fake hits due to noise, this choice can reduce track reconstruction efficiency: if a real hit of low quality is discarded in favour of a fake hit, the corresponding track does not pass the  $\chi^2$  cut and is rejected. The GF, instead, in the case of a track with hits in all SVXII layers, fits all possible combinations of 4 hits out of 5, deleting one particular layer in each fit. For each fit, a quality measure  $g$  (as in goodness) is calculated in the *Comparator*, based on the  $\chi^2$ , the layer configuration and the quality of hits. Among the five track candidates coming from the same *full of hits track*, only the one having the best  $g$  value is accepted. Finally the *Formatter* receives the parameters and the  $\chi^2$  of the accepted tracks, and merges them with the hits, the road number and some status data in accordance with the SVT output protocol.



**Figure 3.** GigaFitter basic fitting pathway.

#### 4.2. GF speed optimization

The basic fitting pathway is able to perform a fit every 6 clock cycles using 7 DSPs. In order to increase the SVT speed we can exploit the high number of DSPs inside the Virtex 5 by parallelising the fitting tasks. A first step is using two or more parallel *Combiners*. The *Combiner* takes at least 7 clock cycles to read an event, and at least three to send it in output, so continuous read-write mode can be obtained using at least two parallel *Combiners*. Secondly, we can reduce the fit time from 6 to 1 clock cycle using 6 *Fitters* in parallel, for a total of 42 DSPs per fitting pathway. In this way for each track the 6 output parameters can be calculated in parallel with a reduction of the overall fitting time per track.

A Virtex 5 is provided with enough DSP and RAM resources to host three of such parallelized fitting pathways for each mezzanine input. As each mezzanine can receive the data from 4 wedges, the parallelized system will use  $42 \times 3 \times 4 = 504$  DSPs and will be able to perform 12 fits per clock cycle.

## 5. GF final system

The three mezzanines necessary to read all the SVXII data are mounted on a Pulsar board [8], a general purpose 9U VME interface board originally designed for the CDF II L2 trigger system upgrade. The Pulsar receives the fitted tracks from the three mezzanines, corrects them for small non-linearity's effects and merges the corrected tracks into a single output for the L2 trigger decision. The correction factors can be stored in a fourth mezzanine (a Pulsar can host up to 4 mezzanines) to be used as a RAM.



**Figure 4.** GF final system, consisting of a Pulsar board and 4 mezzanines.

The final system will consist of a single Pulsar board hosting 4 mezzanines (fig.4), to be compared with the current track fitter system consisting of 16 boards (12 TF boards and 4 boards to merge the TF outputs).

## 6. Advantages over the current TF

The GF system will be much more compact than the current TF system. It will also be much faster: three mezzanines, each receiving the data from 4 SVXII wedges, each wedge processed by three fit pathways result in a total of 36 fits/clock cycle. Supposing a clock frequency of 100 MHz, the GF can perform three fits/ns.

Thanks to the 25x18 bit multipliers inside the Virtex-5, in the scalar product we can use the full hit resolution. The memory released by the precalculated terms can be used to store more patterns, for example for tracks crossing the mechanical barrels or to increase SVT  $\eta$  coverage using only SVXII data for track reconstruction. The latter improvement would be useful to trigger on forward leptons and thus to increase the acceptance for detecting Higgs decay products. The possibility to store more patterns will also allow an extension of SVT acceptance for track  $p_T$ , reconstructing tracks with  $p_T > 1.5$  GeV/c instead of 2 GeV/c. This will significantly improve the b-tagging capability. The SVT acceptance for impact parameter can also be increased: currently SVT reconstructs only tracks with impact parameter smaller than 1.5 mm but the upgraded system could be sensitive for impact parameter up to 2-3 mm, improving the lifetime measurements.

## 7. Preliminary tests

A GF prototype implementing the basic fitting algorithm and running at 120 MHz was tested in early 2009 at CDF II. The GF was inserted parasitically in the SVT system, receiving a copy of the events input to the TF. The processing time of both tracking processors was measured as the time difference between the input and the output of each event. Preliminary results are shown in fig. 5. Even though the fitting algorithm is not optimized yet, the GF is faster than the current system: it can fit different types of events in a limited time window, while the TF processing time increases on events rich in tracks. It has also to be noted that the GF, in case of a *full of hits track*, performs 5 fits, while the TF fits only once.

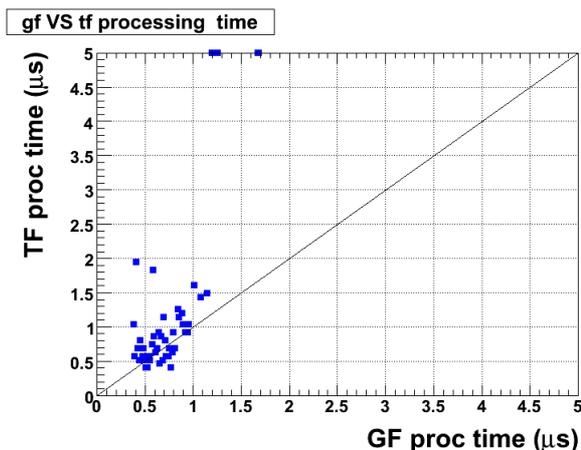


Figure 5. Comparison between GF and TF processing time.

## 8. Future applications

A GF type processor can be used in future online tracking processors where high quality track finding at very high rates is needed. Its architecture is general enough to be adapted to different experimental environments. For example, Fast TracK (FTK) [9] is a proposal for a SVT-like tracking processor to be used in the ATLAS Level-2 trigger. It is intended to provide offline quality tracks for all charged particles with  $p_T > 1$  GeV/c up to the Level-1 rate of 50-100 MHz.

## 9. Conclusions

The GF is an online tracking processor built as a possible upgrade of the CDF II trigger system, following the successful strategy implemented by the SVT tracking processor. Based on a modern FPGA, full of DSP arrays, it features high computation power, performing up to 3 fits/ns, and high flexibility and modularity. At CDF II, it will allow the reduction of trigger dead time and rates at high instantaneous luminosity and improve the physics reach of the experiment, thanks to a greater track reconstruction efficiency.

## 10. References

- [1] CDF II Collaboration, “Technical Design Report”, FERMILAB-PUB-96-390-E (1996)
- [2] A. Abulencia *et al.*, “The CDF II extremely fast tracker upgrade”, FERMILAB-CONF-06-280-E, 2006
- [3] D.Cox, “CDF L2 Track Trigger Upgrade”, in *Proc. 19<sup>th</sup> Hadron Collider Physics Symposium*, Galena, May 27-31 2008
- [4] W. Ashmanskas *et al.*, “The CDF Silicon Vertex Tracker”, *Nucl. Instrum. Methods A*, vol. 518, pp.532-536, Feb. 2004.

- [5] A. Annovi *et al.*, “A VLSI Processor for Fast Track Finding Based on Content Addressable Memories” *IEEE Transactions on Nuclear Science*, vol. 53, issue 4, part. 2, pp. 2428-2433, Aug. 2006
- [6] S. Amerio *et al.*, “The GigaFitter for fast track fitting based on FPGA DSP arrays”, Nuclear Science Symposium Conference Record 2007, NSS 2007, IEEE Vol.3, 2007, pp. 2115-2117.
- [7] “Virtex-5 family overview”, [http://www.xilinx.com/support/documentation/data\\_sheets/ds100.pdf](http://www.xilinx.com/support/documentation/data_sheets/ds100.pdf)
- [8] “The Pulsar Board Project”, <http://hep.uchicago.edu/~thliu/projects/Pulsar>
- [9] E. Brubaker *et al.*, “Performances of the Proposed Fast Track Processor for Rare Decays at the ATLAS experiment”, *IEEE Transactions on Nuclear Science* 55 (2008) 145-150, doi:10.1109/TNS.2007.913476