Associative Memory Design for the FastTrack Processor (FTK) at ATLAS

A. Annovi, R. Beccherle, M. Beretta, E. Bossini, F. Crescioli, M. Dell'Orso, P. Giannetti, J. Hoff, T. Liu,
V. Liberali, I. Sacco, A. Schoening, H. K. Soltveit, A. Stabile, R. Tripiccione, G. Volpi

Abstract—We propose a new generation of VLSI processors for pattern recognition, based on associative memory architecture, optimized for online track finding in high-energy physics experiments. We describe the architecture, the technology studies and the prototype design of a new associative memory project: it maximizes the pattern density on the ASIC, minimizes the power consumption and improves the functionality for the fast tracker processor proposed to upgrade the ATLAS trigger at LHC.

I. INTRODUCTION

THE track reconstruction in high-energy physics experiments requires large online computing power. The Fast Tracker for ATLAS triggers [1] is an evolution of the Silicon Vertex Tracker (SVT) in CDF [2], [3]. The Fast Tracker is an online processor that tackles and solves the full track reconstruction problem at a hadron collider.

The SVT track fitting system approaches the offline tracking precision with a processing time of the order of tens of microseconds, compatible with 30 kHz input event rates. This task is performed with negligible time delay by a Content Addressable Memory (CAM), also called Associative Memory (AM), i.e., a device that compares the event hits in parallel with all the stored pre-calculated low resolution track candidates (patterns), and returns the addresses of the matching patterns. A second processor receives the matching patterns and their related full-resolution hits to perform the final track fitting (Track Fitter, TF).

A critical figure of merit for the AM-based track reconstruction system is the number of patterns that can be stored in the bank. For the SVT upgrade [2], [4], we developed a version of the AM chip (AMchip03) [5], using a 180 nm

CMOS technology and strictly standard-cell based VLSI design approach. The AM chip upgrade increased the number of patterns stored in chip from 128 to 5 000 and it could work at 50 MHz frequency.

The FTK processor proposed for the ATLAS experiment is much more ambitious than SVT. In fact, a very high efficiency and high quality track reconstruction, already shown possible by SVT, must be achieved in a much more complex detector. Moreover, the higher luminosity (10^{34} cm^{-2} s^{-1}) will increase the complexity of events. As a consequence a very large bank is necessary: candidate tracks have to cover with more than 95% efficiency the whole tracking detector ($|\eta| < 2.5$), with high efficiency on transverse momentum down to 1 GeV, and the pattern recognition has to be extended to 12 silicon detector layers (4 pixel layers and 8 SCT layers) with a reasonable resolution to drastically reduce both the number of fakes and the track fitting processing time. The FTK operations during Phase I LHC condition set a goal of 80 000 patterns per chip for the AMchip04. In order to meet this requirement a big improvement of all AMchip parameters with respect to the existing AMchip03 is required. This led to a redesign of most aspects of the device that are described below.

II. NEW ASSOCIATIVE MEMORY

A full custom cell is the most important goal of the R&D devoted to a new ASIC associative memory device. The new chip, with extremely high pattern density, will need also to be enriched of new functional elements and will need to be faster (at least a factor 2) than the previous version. Today, a 65 nm technology is a good choice to design a chip characterized by the speed, density and flexibility required by the results of the FTK performance studies. We started with a 90 nm technology, available as "mini-ASIC" (for small-volume production) already in 2009, to test early the full custom cell advantages [6]. Early in the design phase, we switched to a 65 nm technology, which in 2010 became cost-effective for a MPW prototype. The new full custom cell includes all the hardware necessary for the elementary functions of a single pattern layer: SRAM bits, comparators, and a final latch to store the match. In the previous AM chip these functions were implemented putting together a set of standard cells, unavoidably more expensive in terms of silicon area. We can get a reduction of a factor more than two in the area required for a single pattern layer.

Since the bank patterns occupies roughly 70% of the full AM chip (50% for the 14 mm^2 prototype), it is reasonable to
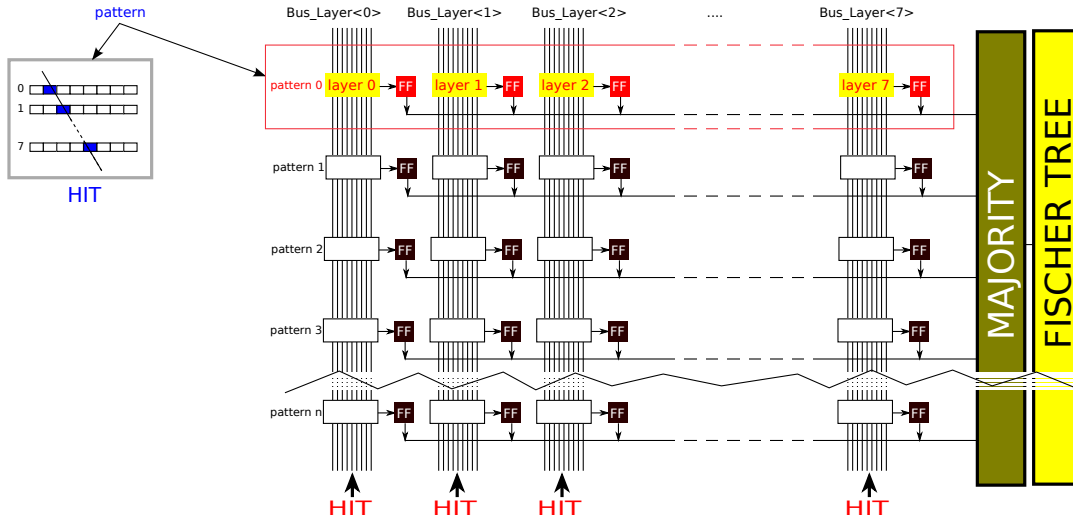
Fig. 1. AMchip array.

think that the new layout will produce a factor 2 more patterns per chip. The global readout and control logic is implemented with standard cells to optimize the development time. The density gain of the new full custom cell combined with the gain due to the technology scaling from 180 nm to 65 nm produces an estimate global increment of a factor 15 for the number of patterns. We can also expand the used silicon area, since the currently used TQFP208 package can house a 16 mm × 16 mm chip, while the AMchip03 is just 1 cm^2. Considering that for ATLAS we will need 8 layer patterns, while in CDF we used 6 layer patterns, the final gain factor can be as large as 28. For the goal of 80 × 103 patterns per chip a die of 12 mm × 12 mm would be sufficient.

The full custom cell offers also the possibility to implement important new strategies to reduce the power consumption of the chip. This is a very important issue because the pattern density growth will be eventually limited by power consumption. The clock cycle is limited to at least 10 ns by the board complexity: each hit found in the detector has to be distributed to 128 AM chips per board, with a very high fan-out. However inside the 65 nm chip the 10 ns clock cycle is conservative. We use this clock period to ease the distribution of the data through the chip. In addition we can reduce the speed of the match operation in exchange for a reduced power consumption. To achieve this goal, we perform the pattern comparison with the "pre-match" technique that first compares the 4 least significant bits of each layer word, then after successful pre-match the remaining bits are compared. The "pre-match" technique can save up to 80% power consumption. The overall power consumption is only slightly increased, despite the big jump in pattern density and the higher operating frequency. The power consumption is another reason to limit the maximum operating frequency to 100 MHz.

III. THE AMCHIP04 MPW PROTOTYPE

The features of the new AMchip04, and in particular the performances of the full custom cell, must be evaluated in

TABLE I
AMCHIP PARAMETERS

	AMchip03	AMchip04	Effect
Technology	180 nm	65 nm	×8 pattern density
Clock freq.	50 MHz	100 MHz	faster, higher power cons.
Die size	10 mm × 10 mm	12 mm × 12 mm	× 1.5 patterns
Core voltage	1.8 V	1.2 V	lower power consumption
Core power	1.3 W	0.7 W	at 40 MHz and 100 MHz
Selec. Prech.	No	Yes	80% power saving
Full custom	No	Yes	× 2 pattern density
Layers	6 (or 12)	8	$\frac{3}{4}$ pattern density
Patterns/chip	5 k	80 k	
Bits/layer	up to 18	up to 15	
Ternary/layer	N/A	3 to 6	better S/N (see text)

order to validate our expectations and to gain experience for further improvement of the associative memory technology. For these purposes we are designing a MPW (Multi Project Wafer) prototype of the new AM chip with 65 nm technology. The main goal is to verify that the full custom associative memory cell works properly and to verify the expected gain in terms of pattern density and power saving. The AMchip04 MPW prototype uses a reduced silicon area of 14 mm^2 and it is designed to store 8192 patterns, with an estimated core power consumption of 70 mW. The power consumption is the most difficult parameter to predict. It will be important to measure it with the real device. The remaining logic of the AMchip04 mini-ASIC is implemented with standard cells and is very similar to the AMchip03 logic. The main change is that the new chip will perform pattern recognition with 8 layers instead of 6 or 12.

Table I compares the performance of AMchip03 with the expected performance of a full AMchip04 extrapolated from the current MPW design.

IV. AMCHIP WORKING PRINCIPLE

As most memories, the AMchip is based on an array (Fig. 1). Columns are used to distribute the hit information
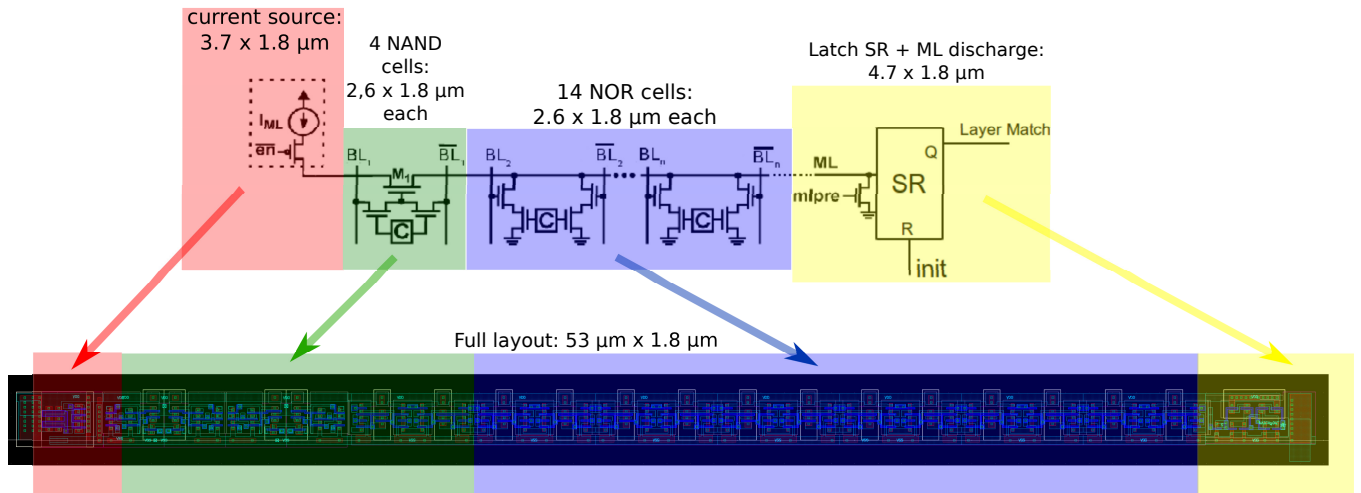
Fig. 2.   Schematic and layout of a single layer.

over vertical buses called search bit lines (or just bitlines), whereas rows are used for the write lines (the signals to enable the write operation) and for the match lines (the signal that identifies a match). Each bit line bus is made of 36 lines: the 18 bits and the 18 corresponding inverted bits. Each row in the array corresponds to one pattern. A row is organized in sub-blocks of 18 CAM cells that we call a layer block. Each layer block stores the position of the intersection between a pre-calculated track trajectory and a real detector layer. A pattern is composed of 8 layer blocks, so it can identify a track crossing up to 8 detector layers, while additional layers will be ignored. Each pattern is provided with the logic necessary to compare the stored position with actual hit position for one event. A pattern matches when all, or almost all, the stored positions correspond to the input data for one event.

The event hit positions are received over 8 input buses of 15 bits each. This limits the maximum number of positions to 32 000 positions for each layer. This might seem a strict limit, however since different AM chips (or different groups of AM chips) can independently process data of different parts of the detector it is enough to reach a granularity below 10 micro-strips or an area of 12×36 pixels respectively along the r-ϕ and z directions that will be sufficient for the FTK project and drove the choice of 15 bits.

The positions stored in each layer block are encoded in this way. Of the 18 CAM cells each storing one bit, 12 are used to store the 12 MSBs of the word, while the other 6 are combined in 3 pairs and used as ternary cells storing either 0, 1 or X values. The X value means don't care, so the hit present on the hit bus will match the stored word regardless of the value for the bits set to X. The use of the "don't care" feature, as in ternary CAMs, allows us to have variable size patterns. Normally the least significant bit of each layer corresponds to a fixed area on the detector (e.g., 20 consecutive micro-strips). This area is the width of the matching window implemented by one pattern. When a bit is set to the "don't care" value, the effective pattern size for that bit is doubled because it will match two numbers. With 3 "don't care" bits, we can enlarge

the coincidence window up to a factor 8 for each pattern and for each layer independently. In this way patterns can be tailored to maximize the acceptance for valid tracks, while reducing the probability to match spurious hit combinations. In other words each pattern has a better signal to noise ratio. Therefore, AM patterns are employed in a smarter and more efficient way. Preliminary estimates show that patterns using only two ternary cells per layer are as effective as in increase of a factor 3 to 5 in number of patterns without ternary cells [7]. This gain required an increase in chip area of just 3 CAM cells per layer that correspond to 17% of the layer or 1 mm^2 total area, this is a very important improvement of the associative memory technology.

Input data in the columns are compared in parallel with the data stored inside the layer blocks. If a layer block matches all the 18 bits (accounting for "don't care" bits in ternary cells), a Set-Reset Flip-Flop (SR-FF) is set to the high logic value. As we can see in Fig. 1, the majority block counts the number of SR-FF set to 1. If this number is equal to 6, 7 or 8, the data are transferred to the AM readout block that is able to generate the address of the matched layer by using a priority list. The latter block is based on a modified Fischer tree [8]. It is worth noting that the majority block identifies not only the match of all layers, but also a partial matching where one or two layers for a given pattern are not hit.

V. THE SINGLE LAYER

In more detail, the layer block shown in Fig. 2 is composed by a current source, 4 NAND CAM cells, 14 NOR CAM cells, and a SR-FF. The current source charges the match line with a constant current value only if the data stored inside the cells match with the input data on the bit lines. On the contrary, if at least one of the NAND cells do not match the input data it interrupts the match-line preventing it to charge. If a NOR cell does not match, a path from the match line to ground is created, and consequently, the match line is discharged. In terms of node voltages, if all the cells match the data input, the match line is charged up to a high logic value (from

TABLE II
LAYER BLOCK SIZES

Current source	2.6 μm × 1.8 μm
NAND cell	2.6 μm × 1.8 μm
NOR cell	2.6 μm × 1.8 μm
SR-FF	3.6 μm × 1.8 μm
Total	53 μm × 1.8 μm

simulations we found about 1 V); on the contrary, if at least one of cells does not match, the match line is not charged and the voltage is about 0 V. To save power we have used two different match line driving schemes: (1) current race scheme, and (2) selective pre-charge scheme. For this reason we have used also two different type of cells. Both schemes are based on standard CAM designs [9].

Obviously, our goal consists in increasing the cells density and therefore we need to save area. To this purpose, we have designed the layer block with a full custom approach. In this way we have minimized the area by placing staggered transistors with a gate width equal to the minimum value allowed by the technology rules. To guarantee the compatibility and easy integration with the standard cell library, we have designed all the cells using a VDD-to-VSS wire pitch equal to 1.8 μm and using VDD/VSS wire widths equal to 0.33 μm. The layer block size and the size of each component are given in Table II. The transistors of the current source have been carefully designed to ensure that they are identical (same W/L ratio, same gate orientation, same number of contacts and distances, and same diffusion areas) to the transistors of the reference block, in order to have a good current matching. A single reference is used for 64 current sources to save area and power. In the SR-FF block, we have also used a reset transistor controlled by the initialization signal.

VI. FULL CUSTOM ARRAY

To design the 64-half-pattern block (Fig. 3) we have used an array of 64 × 2 layers called TOP. In the TOP block we have also placed: (1) one dummy layer with a programmable delay; (2) the voltage reference; (3) and the resistance ties. After that, we have duplicated this block. In this way, the 64-half-pattern block (64 × 4 layers) has been designed. We have called TOP2 this 64-half-pattern block. TOP2 is the larger full custom block that we have designed for this project. TOP2 has been designed using Cadence Virtuoso and the TSMC design kit at 65 nm. Overall in the Fig. 3 (upper part) we can see the floor plan of the 64 × 2 layers (TOP), and in the lower part two TOP2 blocks with the majority and Fischer tree logic in the middle. Actually, the majority and the Fischer tree blocks have been designed with a standard cell approach.

VII. STANDARD CELL AND FULL CUSTOM INTEGRATION

The entire chip has been designed with a hybrid approach. More repetitive regions have been designed with a full custom approach. On the contrary, the more complex logics have been designed with a standard cell approach. To place and route standard cells, we have used Cadence Encounter. Fig. 4 shows the floorplan of the entire chip.
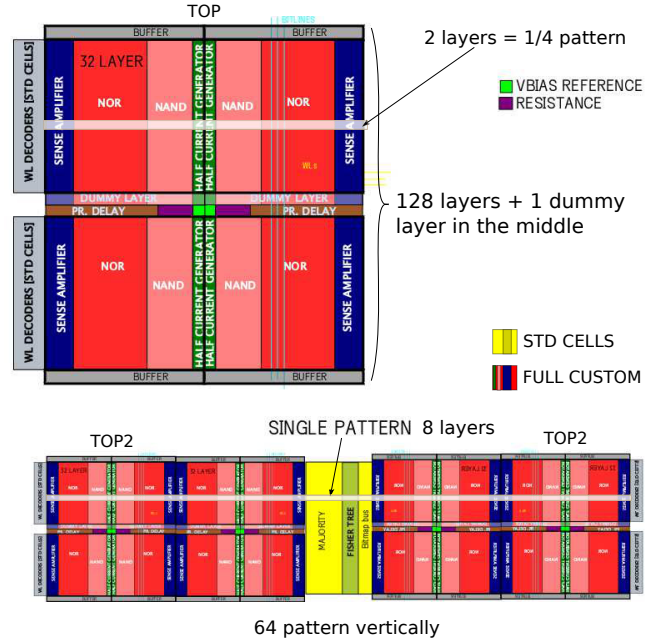

Fig. 3. Full custom floorplans. The block in the upper left corner (called TOP) is an array of 64×2 single layers.

The AMchip has an area of 14 mm² (3510 μm × 3985 μm). The memory is organized as an array 22 columns × 12 rows of full custom macro blocks (TOP2). The majority logic and the Fischer tree have been placed between two TOP2 blocks. The blocks not designed with a full custom approach, have been automatically placed by Encounter. This placement has been done using a flat description of the logic. In addition, to decrease the routing congestion, we have designed fence areas that contain the majority, the Fischer2001 tree, and 4 TOP2 macro blocks. We placed the left-lower TOP2 without rotation, the right-lower TOP2 with a horizontal mirroring, the left-top TOP2 with a vertical mirroring, and the right-top TOP2 with a rotation of 180°. To prevent routing congestion issues, we have also designed partial placing blockages (13.2% of cells are allowed) in correspondence of the narrow routing channels at the boundary between two TOP2 blocks. In this way only bit line buffers are placed in these narrow channels. The 13.2% value has been carefully calculated for this aim. In the middle of the chip, we have left a free area for the control logic and the JTAG machine.

Overall, we have placed a ring of 208 pads, and we have left a space of 30 μm between the ring pads and the standard cells. This space has been used to place power rings for VDD and VSS just inside the pad frame. The width of each power ring is 10 μm. The power ring is connected to horizontal strips distributing power inside the chip using metal1 (width = 330 nm) and vertical strips using metal6 (width = 1300 nm). The connection between the power ring and strips running on metal6 is made of a 2D grid of 6 × 39 vias on both ends of each strip. The maximum current allowed per via is 0.3 mA meaning that up to 70.2 mA can flow from the power ring to each power strip in metal6. The maximum current allowed on a metal6 strip is of 4.416 mA/μm × (1.3 μm − 0.02 μm) = 5.65 mA. Whereas, the maximum current allowed on a
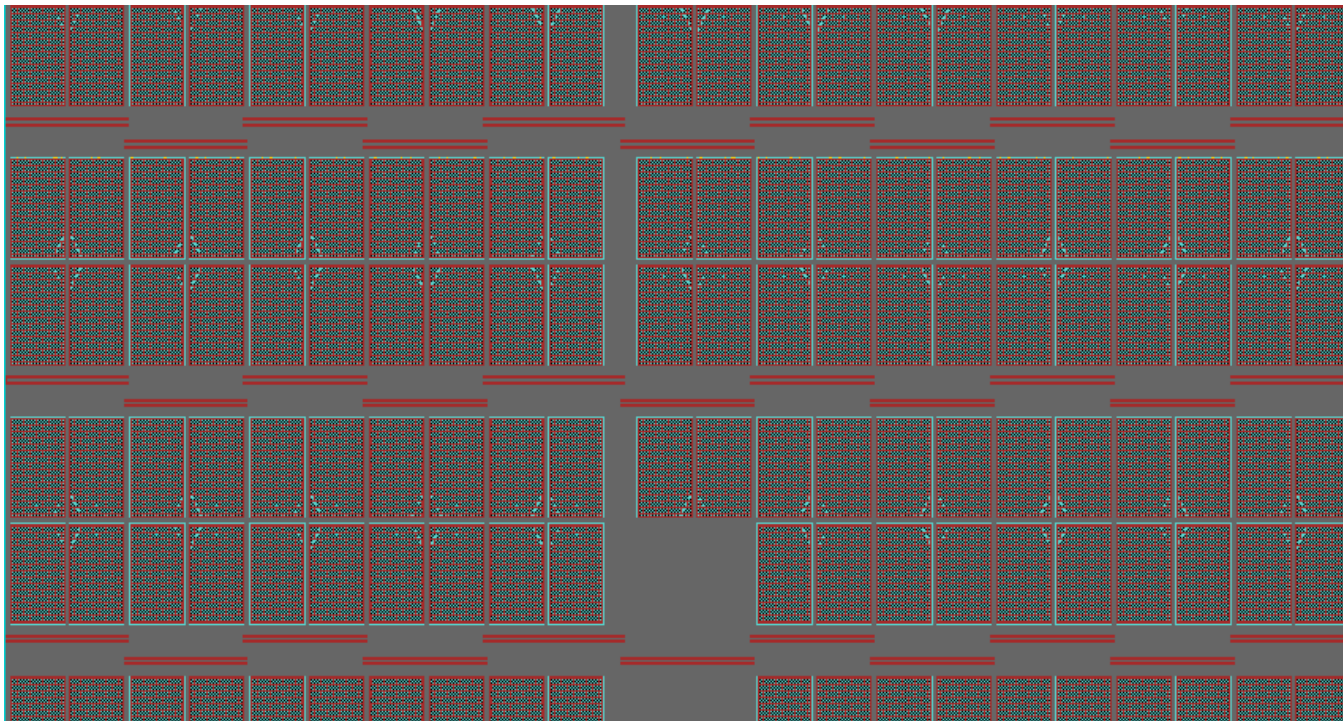
Fig. 5. Horizontal power strips in metal5.

metal1 strip is of 1.509 mA/μm $\times$ (0.33 μm – 0.016 μm) = 0.47 mA. Overall, the maximum current allowed that can flow on the sum of the one metal6 strip and one metal1 strip is 6.12 mA. By considering that we have placed about 1800 strips in the entire chip, the maximum current that can enter the power strips on metal6 from the power ring is 6 mA $\times$ 1800 strip equals about 10 A. Here all strips are counted because half strip brings VDD and half brings VSS, but each strip is connected on both sides. Hence, the maximum current that can flow vertically in the power ring that is 10 μm wide on metal2 is 1.877 mA/μm $\times$ (10 μm – 0.016 μm) = 18.7 mA. We have estimated that our chip will have a current consumption of about 70 mW for the core 1.2 V. For this reason, metal1 and metal6 strips are not sufficient to guarantee a good power distribution. Hence, we have placed horizontal power strips in metal5 (width = 3600 nm). These strips have been placed with a staggered approach to prevent routing congestion (Fig. 5). We have also placed a large number of power supply and ground pads. This is done, to guarantee a correct power supply of the core (Fig. 6).

We have chosen bidirectional pads with a driving current of 2 mA or 4 mA. Even with only 2 mA the estimated pad to output time (including line capacity on PCB for up to 3000 mil) is less than 4 ns. As the power consumption of this chip will have a significant value (about 70 mW), we have filled the empty spaces with decoupling capacitors in this way the power supplies are filtered. The expected power consumption of the I/O pad ring is 100 mW at 3.3 V.

To perform place and route, we have used the Foundation Flow by Cadence Encounter. This flow contains some important optimization steps: pre-CTS (Clock Tree Synthesis) optimization, post-CTS optimization and post-route optimiza-

TABLE III
LAYER BLOCK SIZE

Area name	Size	Rel. size
big vertical channels (maj, readout)	2.018 mm^2	14.34%
horizontal channels	0.178 mm^2	1.26%
small distribiuted channels	0.880 mm^2	6.25%
TOP2	7.050 mm^2	50.13%
central region	0.415 mm^2	2.95%
pads (including boundary scan)	3.522 mm^2	25.04%

tion. These steps have been performed to enhance the timing performance of the chip. The clock tree has been generated and results confirm that the clock distribution is good. The maximum clock skew is equal to 400 ps. We have described the timing constraints in a .sdc file which contains: (1) Setup time for inputs clocked register ranges from 0.1 ns to 2.5 ns; (2) A output time after clock for all outputs ranges from 0.1 ns to 2.5 ns; and (3) minimum period for clock equal to 10 ns. All these constraints are fulfilled in all optimization steps. Table III shows a summary of the area occupied by each part of the part of the chip. The largest fractions of area correspond to the full custom TOP2 blocks, to the pad ring and to the big vertical channels.

VIII. CONCLUSION

We are designing a new associative memory device (AMchip04) that exploits full custom CAM cells along with 65 nm technology to increase the patterns available per chip by a factor 25 with respect to the AMchip03. At the same time we use low-power techniques to achieve these goals while keeping the power consumption at about 2 W per chip. The AMchip04 design is well underway and a MPW prototype is ready to be
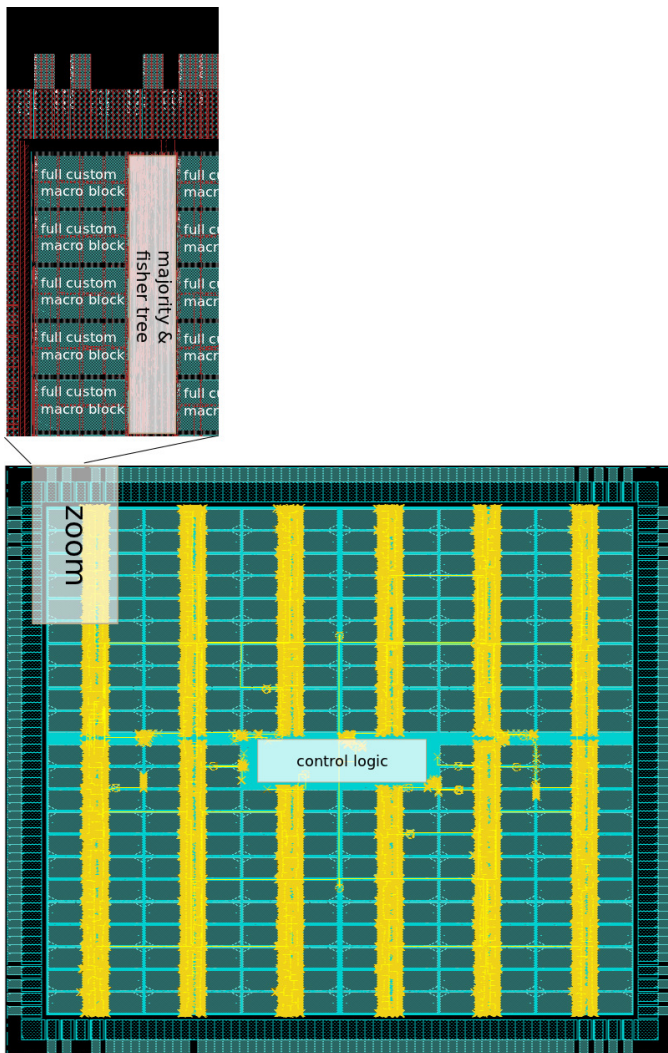
Fig. 4. Entire chip floorplan.

Fig. 6. Minimum number of power pads to ensure a good power supply.

submitted to the silicon foundry. This is the first prototype of the AMchip being developed for the Fast Tracker processor.

ACKNOWLEDGMENT

REFERENCES

[1] A. Annovi, A. Bardi, M. Campanelli, R. Carosi, P. Catastini, V. Cavasinni, A. Cerri, A. Clark, M. Dell'Orso, T. Del Prete, A. Dotti, G. Ferri, S. Giagu, P. Giannetti, G. Iannaccone, M. La Malfa, F. Morsani, G. Punzi, M. Rescigno, C. Roda, M. Shochet, F. Spinella, S. Torre, G. Usai, L. Vacavant, I. Vivarelli, X. Wu, and L. Zanello, "Hadron collider triggers with high-quality tracking at very high event rates," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 3, pp. 391–400, Jun. 2004.
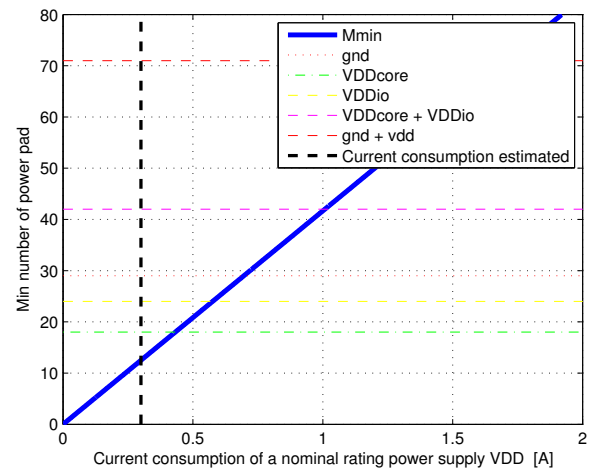
[2] J. Adelman, A. Annovi, M. Aoki, A. Bardi, M. Bari, J. Bellinger, M. Bitossi, M. Bogdan, R. Carosi, P. Catastini, A. Cerri, S. Chappa, M. Dell'Orso, B. D. Ruzza, I. K. Furić, P. Gianetti, P. Giovacchini, T. Liu, T. Maruyama, I. Pedron, M. Piendibene, M. Pitkanen, B. Riesert, M. Rescigno, L. Ristori, H. Sanders, L. Sartori, M. Shochet, B. Simoni, F. Spinella, S. Torre, R. Tripiccione, F. Tang, U. Yang, and A. Zanetti, "The Silicon Vertex Trigger upgrade at CDF," *Nucl. Instr. and Meth. in Phys. Res. – Sect. A*, vol. 572, no. 1, pp. 361–364, Mar. 2007.

[3] J. Adelman, A. Annovi, M. Aoki, A. Bardi, J. Bellinger, M. Bitossi, M. Bogdan, R. Carosi, P. Catastini, A. Cerri, S. Chappa, M. Dell'Orso, B. D. Ruzza, I. Furic, P. Giannetti, P. Giovacchini, T. Liu, T. Maruyama, I. Pedron, M. Piendibene, M. Pitkanen, B. Reisert, M. Rescigno, L. Ristori, H. Sanders, L. Sartori, M. Shochet, B. Simoni, F. Spinella, S. Torre, R. Tripiccione, F. Tang, U. Yang, and A. Zanetti, "Real time secondary vertexing at CDF," *Nucl. Instr. and Meth. in Phys. Res. – Sect. A*, vol. 569, no. 1, pp. 111–114, Dec. 2006.

[4] J. Adelman, A. Annovi, M. Aoki, J. Bellinger, E. Berry, M. Bitossi, M. Bogdan, R. Carosi, P. Catastini, A. Cerri, S. Chappa, F. Crescioli, M. Dell'Orso, B. D. Ruzza, S. Donati, I. Furic, P. Giannetti, C. Ginsburg, T. Liu, T. Maruyama, F. Palla, I. Pedron, M. Piendibene, M. Pitkanen, G. Punzi, B. Reisert, M. Rescigno, L. Ristori, H. Sanders, L. Sartori, F. Schifano, F. Sforza, M. Shochet, F. Spinella, F. Tang, S. Torre, R. Tripiccione, G. Volpi, U. Yang, and A. Zanetti, "On-line tracking processors at hadron colliders: The SVT experience at CDF II and beyond," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 581, no. 1–2, pp. 473–475, Oct. 2007.

[5] A. Annovi, A. Bardi, M. Bitossi, S. Chiozzi, C. Damiani, M. Dell'Orso, P. Giannetti, P. Giovacchini, G. Marchiori, I. Pedron, M. Piendibene, L. Sartori, F. Schifano, F. Spinella, S. Torre, and R. Tripiccione, "A VLSI processor for fast track finding based on content addressable memories," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 4, pp. 2428–2433, Aug. 2006.

[6] A. Annovi, M. Beretta, E. Bossini, F. Crescioli, M. Dell'Orso, P. Giannetti, M. Piendibene, I. Sacco, L. Sartori, and R. Tripiccione, "Associative memory design for the FastTrack processor (FTK) at ATLAS," in *Proc. IEEE-NPSS Real Time Conference (RT)*, Lisbon, Portugal, May 2010, pp. 1–3.

[7] A. Annovi, S. Amerio, M. Beretta, E. Bossini, F. Crescioli, M. Dell'Orso, P. Giannetti, J. Hoff, T. Liu, D. Magalotti, M. Piendibene, I. Sacco, A. Schoening, H.-K. Soltveit, A. Stabile, R. Tripiccione, V. Liberali, R. Vitillo, and G. Volpi, "A new variable-resolution associative memory for high energy physics," in *Proc. IEEE Int. Conf. on Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA)*, Ghent, Belgium, Jun. 2011.

[8] P. Fischer, "First implementation of the MEPHISTO binary readout architecture for strip detectors," *Nucl. Instr. and Meth. in Phys. Res. – Sect. A*, vol. 461, no. 1–3, pp. 499–504, Apr. 2001.

[9] K. Pagiamtzis and A. Sheikholeslami, "Content-addressable memory (CAM) circuits and architectures: A tutorial and survey," *IEEE J. Solid-State Circ.*, vol. 41, no. 3, pp. 712–727, Mar. 2006.