AM06: the Associative Memory chip for the Fast TracKer in the upgraded ATLAS detector

A. Annovi, a M. M. Beretta, b G. Calderini, c F. Crescioli, c L. Frontini, d,e V. Liberali, d,e,1 S.R. Shojaii d and A. Stabile d

a INFN — Sezione di Pisa,
  Largo B. Pontecorvo 3, 56127 Pisa, Italy
b INFN — Laboratori Nazionali di Frascati,
  Via E. Fermi 40, 00044 Frascati, Italy
c Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE),
  4 place Jussieu, 75252 Paris, France
d INFN — Sezione di Milano,
  Via G. Celoria 16, 20133 Milano, Italy
e Università degli Studi di Milano,
  Via G. Celoria 16, 20133 Milano, Italy

E-mail: valentino.liberali@mi.infn.it

ABSTRACT: This paper describes the AM06 chip, which is a highly parallel processor for pattern recognition in the ATLAS high energy physics experiment. The AM06 contains memory banks that store data organized in 18 bit words; a group of 8 words is called “pattern”. Each AM06 chip can store up to 131 072 patterns. The AM06 is a large chip, designed in 65 nm CMOS, and it combines full-custom memory arrays, standard logic cells and serializer/deserializer IP blocks at 2 Gbit/s for input/output communication. The overall silicon area is 168 mm² and the chip contains about 421 million transistors. The AM06 receives the detector data for each event accepted by Level-1 trigger, up to 100 kHz, and it performs a track reconstruction based on hit information from channels of the ATLAS silicon detectors. Thanks to the design of a new associative memory cell and to the layout optimization, the AM06 consumption is only about 1 fJ/bit per comparison. The AM06 has been fabricated and successfully tested with a dedicated test system.

KEYWORDS: Digital electronic circuits; Trigger concepts and systems (hardware and software); VLSI circuits

1Corresponding author.
1 Introduction

After the long shutdown planned in 2019, the LHC will enter into Phase-I in 2021 and it will reach a luminosity up to $3 \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ [1]. The event rate will increase up to 80 interactions per bunch crossing every 25 ns. Only a small fraction of data produced can be stored offline, and therefore an effective trigger system must select the information we are most interested in. This requires massive computing power to minimize the online execution time of complex selection algorithms.

A multi-level trigger is an effective solution to meet this challenge. The Level-1 trigger will perform a first selection, to reduce the event rate to about 60 kHz (with a maximum rate of 100 kHz). The Fast TracKer (FTK) will receive the data from Level-1 trigger and from silicon detectors, and will use this information to reconstruct particle tracks. The output of the FTK will be delivered to the High-Level Trigger (HLT), which is a CPU farm made of more than 2000 PCs. The HLT will refine the information of the FTK and it will perform particle identification and selection of events for long-term storage [1].

The FTK, shortly described in section 2, is an evolution of the successful Silicon Vertex Trigger in CDF [2], and it is an upgrade to the current ATLAS trigger system [3]. The FTK operates at the full rate of the Level-1 Trigger output (100 kHz) and reconstructs the coarse tracks of charged particles with transverse momentum $p_T > 1 \text{ GeV}$ over the entire inner detector.

The FTK exploits the massive parallelism of associative memories. The Associative Memory (AM) system of the FTK stores millions of pre-calculated patterns, and compares them in parallel with the data coming from the silicon detectors. A dedicated CMOS integrated circuit has been designed for this task, aiming at reducing power consumption while maintaining a high level of
efficiency. The AM06 chip, described in section 3, has been designed and fabricated in a 65 nm CMOS technology; section 4 describes the test setup and procedure, and reports some results from the AM06 characterization.

2 The FTK architecture

Figure 1 sketches the whole architecture of the Fast TracKer system.

The FTK receives data from the ATLAS inner detector read-out drivers (RODs), which read the information from pixels and semiconductor trackers (SCT).

The Data Formatters (DFs) process the incoming data from the RODs, to find clusters and thus reducing the amount of data. To increase efficiency, the inner detector volume is divided into 64 regions, so-called “towers”, each of them being processed independently. The DFs arrange the clusters in the towers and distribute them to the corresponding Processing Units (PUs).

The FTK contains 128 PUs (2 PUs per tower). Each PU is composed of a pair of boards (AM and AUX), which perform the pattern matching and the first stage fit. Data from the DFs are received by the AUX card, and the Data Organizer (DO) organizes all clusters according to a coarse resolution position identifier. The identifiers corresponding to the clusters are sent to the AM boards, which combine the data of the 8 innermost layers to find patterns, and returns the tracks at coarse resolution (called “roads”). The DO receives the roads back from the AM, and selects the corresponding packet of data, which are sent to the Track Fitter (TF). The TF builds all combinations of clusters in a road and evaluates their $\chi^2$. Then the Hit Warrior (HW) removes possible duplicated information, and all good tracks are sent to the Second Stage Board (SSB).

The SSB improves the resolution from 8-layer track fits, by adding the data from other 4 layers and performing a 12-layer fit. The FTK to Level-2 Interface Card (FLIC) collects the reconstructed...
The FTK hardware processing chain is described in detail in [4].

### 2.1 The AM system

The Associative Memory system is the core of the FTK [5]: it can store up to 1 billion patterns for pattern recognition. The whole FTK system contains 128 AM boards, and each board stores 8 million patterns. The AM chips are assembled in groups of 16 into ‘Local’ AM Boards (LAMBs), and 4 LAMBs are mounted on each AM board. Each AM board contains 64 AM chips, and the complete FTK system will employ 8192 AM chips.

Each AM chip can store $131,072 \ (= 2^{17})$ patterns, which are sets of 8 cluster identifiers (one identifier per layer). Each cluster identifier is coded with a word of 18 bit, which will be called “layer” hereinafter.

### 3 The AM06 chip

The AM06 chip is the 6th version of the associative memory originally proposed in [6]. Table 1 summarizes the main features of the AM chips designed for particle tracking in high energy physics experiments. The first three versions were designed in different technologies and with different approaches. The versions 4, 5 and 6 of the AM chip have been designed in 65 nm CMOS technology; the AM04 [5] and the two versions of the AM05 [7] are prototypes with small area, designed and characterized to demonstrate the system functionality on silicon.

The next generation of associative memory circuits is already being designed in 28 nm CMOS, for the Phase-II ATLAS upgrade [8].

#### 3.1 AM chip description

Figure 2 shows a simplified schematic diagram of the AM06 chip.

The AM chip receives the data from the DO, organized in sets of 8 cluster identifiers (one per detector layer) in parallel. Each set of 8 layers is a “pattern”, which is compared with the contents of

---

**Table 1.** Associative Memory (AM) chips designed

| Version | Design approach               | Technology | Area         | Patterns$^a$ | Package |
|---------|-------------------------------|------------|--------------|--------------|---------|
| 1       | Full custom                   | 700 nm     | 128          | QFP         |
| 2       | FPGA                          | 350 nm     | 128          | QFP         |
| 3       | Std cells                     | 180 nm     | 100 mm$^2$   | 5 Ki        | QFP     |
| 4       | Std cells + Full custom       | 65 nm      | 14 mm$^2$    | 8 Ki        | QFP     |
| mini-5  | Std cells + Full custom       | 65 nm      | 4 mm$^2$     | 0.5 Ki      | QFP     |
| 5       | Std cells + Full custom + IP blocks | 65 nm | 12 mm$^2$ | 3 Ki        | BGA     |
| 6       | Std cells + Full custom + IP blocks | 65 nm | 168 mm$^2$ | 128 Ki      | BGA     |

$^a$1 Ki = 1024
the memory. When one of the identifiers matches the memory content, the corresponding Flip-flop (FF) is set to ‘1’ to signal that a matching occurred. A “pop-count logic” compares the number of matched layers with the required threshold, which can be set to 8, 7, or 6 layers for normal operation. If the number of matching layers is larger or equal to the threshold, the matched address is passed to the “priority encoder”, which queues all the addresses and sends them to the output.

To avoid routing congestion at board level, the input and output data are serially encoded and transmitted at 2 Gbit/s. The input deserializers (DES1 to DES8 in figure 2) convert the 2 Gbit/s serial data into parallel “words” with 18 bit and 100 MHz rate. Parallel data words are simultaneously distributed to the 64 memory blocks, each of them containing 2048 patterns. Therefore, each input identifier is simultaneously compared with the 131 072 stored data corresponding to the same layer.

The AM06 chip can also receive matching addresses from neighboring chips, interconnected in a daisy-chain configuration; this feature is useful to simplify further the interconnections at board level. Two deserializers (DES9 and DES10 in figure 2) are used to decode the input addresses, which are sent to the priority encoder.

The output of the priority encoder (i.e., the sequence of all matching addresses) is converted into the serial format by the output serializer (SER), and it is sent back to the DO, which retrieves the full track information from the address.

Since all the output data is sent over a single serial link, the AM06 chip has an output latency, which depends on the number of matching patterns. As the priority encoder orders the matching addresses in a first-in, first-out (FIFO) queue, at a first approximation the AM06 latency is given by the number of matches multiplied by the clock period (10 ns), plus the latency due to SER/DES
blocks and distribution of data inside the chip core (about 10 to 20 clock cycles). Overall, the AM06 latency is only a small fraction of the maximum latency allowed for the whole FTK system (100 µs) [1]: the major contribution to the FTK latency is due to data organization and formatting.

Finally, it is worth mentioning that the AM06 chip includes a JTAG interface, which is not shown in figure 2. The JTAG interface is used for test, initialization, and memory bank writing. Memory banks include Built-In Self-Test (BIST) features: pseudo-random test pattern generation, and signature analysis based on Cyclic Redundancy Check (CRC).

3.2 AM06 design constraints

The major issues in the design of the AM06 chip were: (i) the large silicon area required to store patterns, that could result in a low yield; (ii) the serial links operating at 2 Gbit/s, to reduce I/O signal congestion at board level; and (iii) the maximum power should not exceed 250 W per AM board, because this figure is the upper limit corresponding to the capability of the rack cooling system; since each AM board contains 64 AM chips, each chip will have a power budget of 2.5 W (160 W for the 64 AM chips on the board), while the remaining 90 W are reserved for FPGAs and dc-dc converters.

The chip was designed to meet all the above mentioned constraints.

(i) To guarantee the usability of chips with localized faults, each memory block can be set off-line, so that a defective chip can still be used with a reduced capability (e.g., with 63, or 62, or even less working blocks, instead of 64).

(ii) To convert high-speed serial data into parallel data and vice-versa, the AM06 employs serializer/deserializer (SER/DES) IP blocks provided by Silicon Creation. SER/DES blocks use low-voltage differential signaling (LVDS) and the standard 8b/10b encoding; moreover, their power supplies are separated from the supplies of the core logic.

(iii) To reduce power, a dedicated AM cell has been designed and its layout has been optimized in order to reduce the length of bit lines.

3.3 The AM cell and the 18 bit memory layer

The new AM cell, called XORAM (= XOR + RAM), is made of a conventional 6T SRAM cell, merged with an XOR gate [9]. Figure 3(a) illustrates the schematic diagram of the XORAM cell. The 6T SRAM has two CMOS inverters connected in a positive feedback loop, and two NMOS switches driven by the write signal (WL) to store input data into the cell. The XOR gate is made of two complementary switches. If the cell stores a high logic value (A = ‘1’), then the output node is connected to inverted bit line (BL); otherwise, when A = ‘0’, the output is connected to the bit line BL. Therefore, the XORAM output is ‘0’ if the cell contents matches the input bit (A = BL), and ‘1’ otherwise.

Compared with conventional NAND-type or NOR-type AM cells [10], the XORAM cell employs the same number of transistors (10) as the NOR-type cell, while the NAND-type cell is made of 9 transistors. Moreover, the XORAM cell operation is purely combinational and does not require any control circuitry. The XORAM cell has a low energy requirement: about 1 fJ/bit per comparison.
**Figure 3.** The designed full custom cell: (a) XORAM schematic diagram; (b) XORAM layout in 65 nm CMOS; (c) 18-input NOR schematic diagram.

**Figure 4.** Layout of the 18 bit AM layer.

Figure 3(b) shows the layout of a single XORAM cell in 65 nm technology. The area of the AM cell is 2.17 $\mu$m $\times$ 1.48 $\mu$m.

The AM06 uses ‘don’t-care’ bits, to implement a ternary logic (‘1’, ‘0’ and ‘X’), where the ‘X’ logic value is employed in variable resolution patterns, as described in [11]. From figure 3(a), we can see that the ‘don’t-care’ (‘X’) value can be easily obtained by assigning a logic ‘0’ to both the bit-line BL and the inverted bit-line $\overline{BL}$, because either BL or $\overline{BL}$ is connected to the output of the XORAM cell, and a ‘0’ at the output indicates that the input matches the bit stored in the cell. In the AM06, two, four, or six bits of each layer can be set to ‘X’ during the initialization procedure. This information is stored into registers, and is used to set to ‘0’ both BL and $\overline{BL}$ for the rightmost bits of the corresponding layer.

An 18-input NOR cell, shown in figure 3(c), receives the match bits of the 18 memory cells and provides a ‘1’ at the layer output (O) when all the 18 bits stored into the cells match the input data (i.e., all the 18 XORAM outputs are equal to ‘0’).

Figure 4 shows the layout of one layer, which contains 18 XORAM cells (in groups of 3) and the 18-bit NOR function which corresponds to the schematic diagram in figure 3(c). The area of the single layer is 50.8 $\mu$m $\times$ 1.48 $\mu$m. The height of the layer has been kept minimal, to reduce the parasitic capacitance of vertical bit line wires which cross the 2048 layers of each memory block.

### 3.4 AM06 layout

The AM06 size is 14.674 mm $\times$ 11.434 mm, for a total silicon area of 168 mm$^2$. Figure 5 (left) shows the chip floorplan. The SERDES blocks for the 2 Gbit/s serial links are located in the central upper
Table 2. AMchip06 supply voltages.

| Name          | Domain                   | Voltage (V) |
|---------------|--------------------------|-------------|
| \(V_{DDH}\)  | SERDES I/O               | 2.5         |
| \(V_{DDA}\)  | SERDES Core              | 1.2         |
| \(V_{DD,SERDES}\) | SERDES Std Cells      | 1.2         |
| \(V_{DD,BGREF}\) | BandGap Reference    | 2.5         |
| \(V_{DD,LVDS}\) | Clock LVDS              | 2.5         |
| \(V_{DD,IO}\) | Single-ended I/O         | 2.5         |
| \(V_{DD,CORE}\) | Core Std Cells + AM     | 1.0 to 1.2  |

Part of the chip, and they have separate power supplies. Table 2 summarizes the seven separate supply voltages of the AM06.

The JTAG signals enter into the chip from the bottom side, and the JTAG interface is distributed over the whole chip in the areas between memory blocks. The small circles in figure 5 are the 1178 bumps which connect the flipped chip to the package substrate.

The 64 memory arrays occupy the largest portion of the AM06 area. The right part of figure 5 shows the layout of a single memory array, which can store 2048 patterns.

The AM06 contains 421 million transistors, and it can store \(\approx 18.9 \text{ Mbit} = 64 \text{ blocks} \times 2048 \text{ patterns} \times 8 \text{ layers} \times 18 \text{ bits}\).

3.5 Power consumption

Although the single XORAM cell requires only 1 fJ per comparison, the simultaneous operation of a large number of cells (18.9 millions) and the comparison frequency (100 MHz) make the power consumption of the AM06 a critical issue. Indeed, the average power of the AM06 core is \(\approx 2 \text{ W} = (18.9 \text{ Mbit} \times 1 \text{ fJ/bit} \times 100 \text{ MHz})\). SER/DES blocks are responsible for an additional power consumption (between 300 mW and 400 mW). However, SER/DES have separate power supplies and their current consumption is quite constant.

On the other hand, the AM and the core cells are designed with “fully CMOS” logic and their input is synchronous with the 100 MHz clock. Therefore, when the AM06 performs parallel...
comparisons, the AM06 exhibits large current spikes at positive clock edges. To limit the voltage drop and the supply ripple due to current pulses, a special effort was spent to carefully design the package substrate and the printed circuit boards.

The AM06 has been packaged into an HFC-BGA (High-performance Flip Chip Ball Grid Array). The package has 529 balls on the external side of the substrate, arranged in a $23 \times 23$ array. Most of the package balls and of the chip bumps are used for power and ground interconnections, to reduce the values of parasitic inductances and resistances in series to ground and power supplies.

The boards are equipped with decoupling capacitors, to filter all the supply voltages. As an example, in the PCB used for the test described in section 4, the $V_{DD,CORE}$ supply voltage is filtered with three electrolytic capacitors connected in parallel, for a total capacitance of 4.5 mF, and eleven surface-mount capacitors with low series resistance and inductance soldered on the bottom side of the board, for a total capacitance of 30 $\mu$F. The AM06 contains on-chip decoupling capacitance cells, placed to fill empty space in the layout; the total on-chip capacitance amounts to 21 nF.

4 Test procedure and results

The test setup is made of an FPGA evaluation board, and a custom board equipped with a zero-insertion-force (ZIF) BGA socket.

A computer-based test procedure performs the test of one chip in about 2 min. The test procedure exploits the JTAG interface and the BIST features of the core memory blocks. The test sequence verifies: (i) the operation of the JTAG interface, (ii) the operation of SER/DES blocks, (iii) the memory arrays, and (iv) the configuration functions. It has 100% coverage against single faults in the AM core blocks, in the input/output, and in the chip configuration.

Figure 6 shows the measured eye diagram with LVDS serial data at 2 Gbit/s on one serial link. From the measurement, the rms jitter is about 25 ps. Thanks to the separate power supplies, the SER/DES blocks are not affected by the core current consumption.

Figure 7 shows the measured ripple on the $V_{DD,CORE}$ voltage (set to 1.0 V), when the AM06 operation switches between the ‘idle’ mode and the ‘compare’ mode. This is the most critical situation, as the AM core average current rises from 0.1 A to 2.2 A in about 0.1 ns, and current peaks are synchronous with the 100 MHz clock rising edge. A periodic ripple can be observed on the $V_{DD,CORE}$ voltage, with a period equal to the clock period (10 ns). The decoupling capacitors limit the ripple value to about ±60 mV. This value does not prevent the correct operation of the chip.

Figure 8 shows a shmoo plot of the AM06 functionality, for $V_{DD,CORE}$ voltages ranging from 0.8 V to 1.3 V, and for clock frequencies ranging from 50 MHz to 140 MHz. The white region
corresponds to a correct operation; the black region indicates that at least one step in the test procedure has failed.

At low supply voltages, logic signals have longer delays, and setup violations occur. At higher supply voltages, the current consumption becomes more critical, and the ripple on $V_{DD,CORE}$ increases, preventing the AM06 from correct operation. The ripple on $V_{DD,CORE}$ has high frequency components which depend on interconnection parasitics. For some clock frequencies, the ripple contributions due to consecutive current pulses are in phase, while for other frequencies the ripple contributions are in antiphase and one partially cancels out the other. This explains the non-monotonic behaviour with the clock frequency, for values $V_{DD,CORE}$ from 1.225 V and 1.275 V.

At 100 MHz, the AM06 chip is operational over the whole range of the $V_{DD,CORE}$ voltage, from 1.0 V to 1.2 V. We decided to operate the AM06 chips with 1.1 V, instead of the minimum design value of 1.0 V, because $V_{DD,CORE} = 1.1$ V lies in the center of the correct operation region and the supply voltage ripple has less influence.
Moreover, the LAMB is being redesigned including a new dc-dc converter, capable to provide up to 80 A, and with a better decoupling scheme for power supplies.

The test of the first production batch (9 wafers, with $\approx 2700$ chip in total) has demonstrated a high yield: more than 80% of chips have no defects.

## 5 Conclusion

The AM06 has been successfully designed and fabricated. The prototypes are working and no redesign is needed. Tests on the first production batch show a high yield.

The AM06 current consumption exhibits peaks when the chip performs parallel comparisons, and special care is needed in package and board design to reduce the supply voltage ripple. To cope with this issue, the LAMB design is being improved. Test results indicate that the optimal voltage supply for the AM06 chip is 1.1 V, and this value is being used in FTK.

## Acknowledgments

LPNHE gratefully acknowledges the support from the ANR project ANR-13-BS05-0011.

## References

[1] ATLAS collaboration, *Letter of intent for the Phase-I upgrade of the ATLAS experiment*, CERN-LHCC-2011-012, LHCC-I-020, (2011).

[2] CDF collaboration, B. Ashmanskas et al., *The CDF silicon vertex trigger*, Nucl. Instrum. Meth. A 518 (2004) 532 [physics/0306169].

[3] ATLAS collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, 2008 JINST 3 S08003.

[4] V. Cavaliere et al., *Design of a hardware track finder (Fast Tracker) for the ATLAS trigger*, 2016 JINST 11 C02056.

[5] A. Andreani et al., *The AMchip04 and the processing unit prototype for the FastTracker*, 2012 JINST 7 C08007.

[6] M. Dell’Orso and L. Ristori, *VLSI structures for track finding*, Nucl. Instrum. Meth. A 278 (1989) 436.

[7] A. Andreani et al., *The Associative Memory Serial Link Processor for the Fast Tracker (FTK) at ATLAS*, 2014 JINST 9 C11006.

[8] A. Annovi et al., *A XOR-based associative memory block in 28 nm CMOS for interdisciplinary applications*, in Proceedings of the International Conference on Electronics, Circuits and Systems (ICECS), Cairo, Egypt, 2015, pp. 392–395.

[9] L. Frontini et al., *A new XOR-based Content Addressable Memory architecture*, in Proceedings of the International Conference on Electronics, Circuits and Systems (ICECS), Seville, Spain, 2012, pp. 701–704.

[10] K. Pagiamtzis and A. Sheikholeslami, *Content-addressable memory (CAM) circuits and architectures: A tutorial and survey*, IEEE J. Solid-State Circ. 41 (2006) 712.

[11] A. Annovi et al., *A new variable-resolution associative memory for high energy physics*, in Proceedings of the IEEE International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA), Ghent, Belgium, 2011.