

The 16x622 Mbits/s COM16M: the PRELUDE switch architecture integrated into a 6-million transistor monochip

A.Botta, A. Chemarin, J. Majos, M. Serval, D. Weil

CNET Grenoble CIT/CRB
28 Chemin du Vieux Chêne - BP 98
38243 MEYLAN Cedex
FRANCE

1. Abstract

This paper presents COM16M, a fully autonomous monochip ATM switch capable of handling 16 multiplexes carrying ATM cells at 622 Mbits/s. It is a *fully autonomous switch*, i.e. the chip includes clock adaptation, routing and cell buffering as well as header translation and control capabilities. The switch is integrated into one single chip containing 6,000,000 transistors implemented in a 0.5 μm CMOS process.

2. Introduction

Since the beginning of the 80's, CNET at Lannion has been carrying out research on asynchronous transmission techniques which were standardized by the ITU under the name of ATM (Asynchronous Transfer Mode) in 1987. The first ATM switch, named COPRIN, was part of a global ATM demonstrator - known as the PRELUDE project. The whole experiment has been finalized by mid-1987 [3].

It has been recognized that broadband switching architectures refer to one of the three following principles: space-division, shared-medium, and shared-buffer architectures [1]. COPRIN is a shared-buffer switch [5] essentially backed up by the classical time-division principle in use for synchronous switching. As early as 1986, COPRIN achieved a 16x280 Mbits/s throughput with the use of ECL technology, at the expense of a consumption reaching 400W at 5V...

Then, CMOS technology progress led to the implementation of this architecture in a single board involving 24 CMOS chips designed in CNET at Grenoble. This second version of COPRIN - known as COM16 [4] - has been finalized in 1993. It achieved a throughput of 16x155.52 Mbits/s while consumption was limited to 40W at 5V.

Ten years being past, COPRIN architecture has proved to be highly efficient for switches with a limited number of input multiplexes. Moreover, this architecture can be simply upgraded to achieve complex cell header analysis and support traffic control capabilities - which are forthcoming ATM network requirements.

Therefore, we have decided to keep up with the COPRIN architecture when sub-micron technologies have emerged. The third version of the COPRIN architecture is a monochip COM16M that is described hereafter.

3. COM16M functions

3.1. Datapath:

- handling of 16 asynchronous ATM multiplexes. Each multiplex has its own clock and can work at a speed varying from 0 to 622 Mbits/s.
- compatibility with UTOPIA [6] (as defined by the ATM Forum) and IP1 (CNET proprietary) interface.

3.2. Routing:

- 512-cell shared internal buffer and individual output queues, each of them capable of storing 128 addresses. This dimensioning guarantees that cell loss rate keeps below 10^{-10} as long as link utilization remains below 92%.
- multicasting.
- selective cell discarding with respect to the CLP bit value in case of queue congestion.
- setting the EFCI bit of the PTI field in case of congested output FIFO.
- low mean transfer delay: 1 "cell-time" in plesiochronous mode and 3 "cell-times" in totally asynchronous mode when there is no congestion (1 "cell-time" = 682 ns at 622 Mbits/s).

3.3. Header analysis:

- internal context memory of 4096 words of 72 bits allowing translation and routing according to the VPI field (Virtual Path Indicator) or alternatively according to the 12 least significant bits of the VCI field (Virtual Circuit Indicator).
- 32-bit address bus and 48-bit data bus allowing the translation of the whole ATM header through an external memory.
- self-routing in 56-byte cell format by using 3 additional bytes inserted between the 4th byte and the HEC field (Header Error Correcting code).
- special '128 bit' mode which allows the extraction of the first 128 bits of each cell on the external 32-bit wide address bus (only at 155.52 Mbits/s). This mode is used for experimentation of Fast Reservation Protocols (FRP).

4. Circuit architecture (see Fig.5)

The principle of the COPRIN architecture is described in detail in [2] and [4]. In this section we briefly recall the architecture, pointing out the functions that had to be modified for the integration of the switch in a single chip.

The three main functions of COM16M, that can also be found in a Time Division Multiplexing (TDM) switch, are:

- clock adaptation and phase alignment.
- supermultiplexing and cell buffering in a shared memory.
- header analysis and control.

One of the main differences with a TDM switch is that the read and write addresses are not given cyclically to the memory, but calculated after header analysis by a special address management block and stacked in 16 control FIFOs, one per output.

4.1. Clock adaptation and serial-to-parallel Conversion:

The 16 ATM multiplexes enter the switch in a 53 or 56 byte format and are clocked by 16 totally independent clocks having a frequency ranging from 0 to 77.75 Mhz. 16 asynchronous Receive FIFOs (Fig.1) convert the cells into the internal format of the switch (one internal ATM cell = 16 words of 28 bits) and resynchronize all of them on the internal clock. 16 Transmit FIFOs almost identical to the Receive FIFOs re-adapt the outgoing cells to the 16 external clocks and re-format the cells in 28-bit format back into a 53 byte format compatible with the UTOPIA or the IP1 interface.

4.2. Super Multiplexing, header analysis, cell buffering and demultiplexing

The *Incoming Rotation Matrix* [4] receives the 16 input multiplexes coming out of the 16 Receive FIFOs on a 28-bit format after phase alignment and transmit a *supermultiplex* of 16 28-bit wide buses, the 1st bus containing all the 1st words of the cells of each multiplex, the 2nd all the 2nd words,...., and the 16th containing the last words of the cells of the 16 multiplexes (see Fig.2).

The cell header is then extracted and directed to the *cell header processing block*. The header is analyzed, translated and outgoing directions are calculated.

The *control block* receives the output directions and computes a write address at which the ATM cell is stored in the *shared cell memory*. The address is calculated by a special *address management block* giving the first available address of

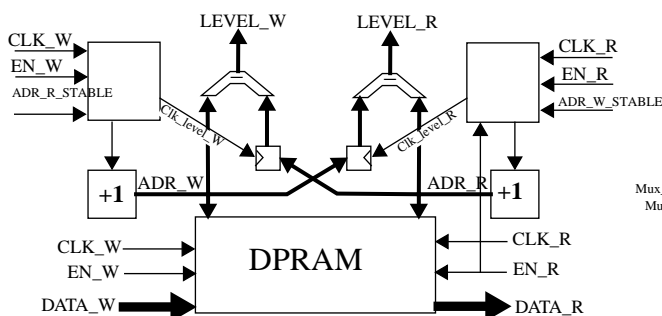


Fig. 1. Asynchronous FIFO

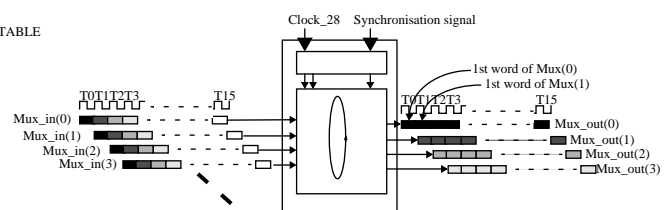


Fig. 2. Rotating matrix

the 512 addresses in the cell memory in one cycle time. While the ATM cell is stored in the shared memory, its address is stacked in one or several control FIFOs. There is 16 control FIFOs, each corresponding to an output multiplex, and the address is stacked in as many FIFOs as there are outgoing directions (more than one FIFO when multicasting is used).

The 16 control FIFOs are read sequentially, and the read addresses are extracted and directed to the cell memory (see Fig.5) where cells are extracted (§4.2). The output multiplex towards which the cell is directed depends on the number of cycles between the arrival of the read address and a periodic synchronization signal.

The ATM cell comes out of the cell memory in the supermultiplex format, it again passes through an *Outgoing Rotating Matrix* where it is *demultiplexed*, and directed to one of the 16 transmit FIFOs (§4.1).

5. Chip design methodology:

5.1. Modeling, synthesis and routing:

A VHDL model has been written for the whole chip and its environment. The circuit was designed using automatic synthesis and routing tools. RTL level simulations, using the same scenarios as the functional model, were performed after synthesis and after routing for the whole circuit (75000 RTL gates). Transistor level simulations were carried out for all critical blocks, up to 400 000 transistors.

5.2. Clock distribution:

Due to the high complexity of the chip, special attention was paid to the clock distribution. Two different approaches were used at inter-block level and intra-block level.

As the internal clock frequency was low (23.5 MHz) compared to the size of circuit and the speed of glue logic, clock distribution problems between blocks were solved by latching signals on the clock falling edge when they come out of a block, and latching them on the clock rising edge as they enter another block.

Clock distribution inside a block containing several tens of thousands of logic RTL gates is a different problem: we have to ensure that clock skew is always shorter than the shortest propagation delay between 2 temporal barriers. The worst case arises when 2 flip-flops are connected to each other; therefore, a good approach is to ensure that the clock skew is shorter than the flip-flop propagation delay, assessed to 1 ns in our case. In order to create a balanced clock tree, we handled the problem using 2 approaches depending on the size of the blocks:

- for small blocks (smaller than about 1 mm wide), the clock tree was synthesized manually with the help of the synthesis tool.

- for larger blocks containing up to several tens of thousands RTL gates, clock tree balance was carried out by the router after cell placement, taking into consideration the net capacitances.

6. Conclusion

We have presented a fully autonomous low power monochip ATM switch named COM16M. This chip is capable of handling 16 ATM multiplexes running at up to 622 Mbits/s at the UTOPIA interface, with both VP and VC switching.

COM16M offers an extremely high level of integration together with a large potential to support emerging traffic

Technology: 0.5 μm 3 Metal CMOS
Processing: HCMOS5 process of SGS-Thomson.
Number of transistors: 6 million.
Number of "standard cells": 75000.
Surface: 3.4 cm^2 .
Maximum frequency of datapath clocks: 77.75 MHz.
Maximum frequency of internal clock: 23.5 MHz.
Consumption: 13 W at 622 Mbits/s, 3.4W at 155 Mbits/s
Power supply: +3.3 V.
Package: Bull's 560 pin BGA.

Fig. 3. Chip characteristics

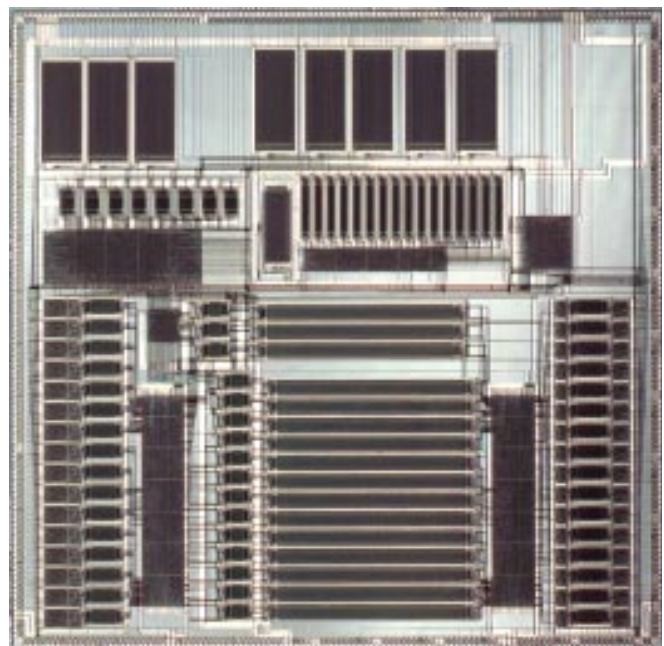


Fig. 4. Chip microphotograph

control functions: it allows analysis of the full header plus the first 8 bytes of the ATM cell payload prior to routing. Therefore, COM16M is a good candidate for ATM switching both in the wide and local area: it can be used for cross-connecting, switching, multiplexing and de-multiplexing.

COM16M can be used as a basic matrix to build larger switching systems such as routers, servers, cross-connects. In this framework, it is suggested that switch elements located at the entrance of the switch systems would run the VPI/VCI translation mode associated to a large external context memory. Besides, the internal switch elements would run the self-routing mode.

It is anticipated that network protocol processing - instead of throughput increase - will pace the evolution of switching elements. There is room for COM16M in this perspective.

7. Acknowledgments:

The authors would like to thank the CAD support team of CNET at Grenoble and SGS-Thomson for manufacturing the chip.

8. References:

- [1] H. Ahmadi, W.E. Denzel: «A survey of modern high-performance switching techniques», *IEEE, J.Selected Areas in Commun.* 7, 1989, pp 1091-1103.
- [2] A. Thomas, J.P. Coudreuse, and M. Serval; «Asynchronous time-division switching: an experimental packet network integrating videocommunications.» in *Proc. ISS'84*, Florence, Italy, May 1984, paper 32C2.
- [3] M. Devault, J.Y. Cochenec, M. Serval: «The "Prelude" ATD Experiment: Assessments and Future Prospects», *IEEE, J.Selected Areas in Commun.* Vol. 6, NO. 9, Dec 1988.
- [4] A. Chemarin, A. Andre, A. Botta, J. Majos *et al*: «A High-Speed CMOS Circuit for 1.2-Gbits/s 16x16 ATM Switching», *IEEE, J. of Selected Circuits*, Vol. 27, NO. 7, July 1992.
- [5] A. Jajszczyk, M. Roszkiewicz, J. Garcia-Haro: «Comparison of ATM Shared-Memory Switches», *ISS'95*, April 1995, Vol. 1.
- [6] ATM Forum: «Utopia, an ATM-PHY Interface Specification».

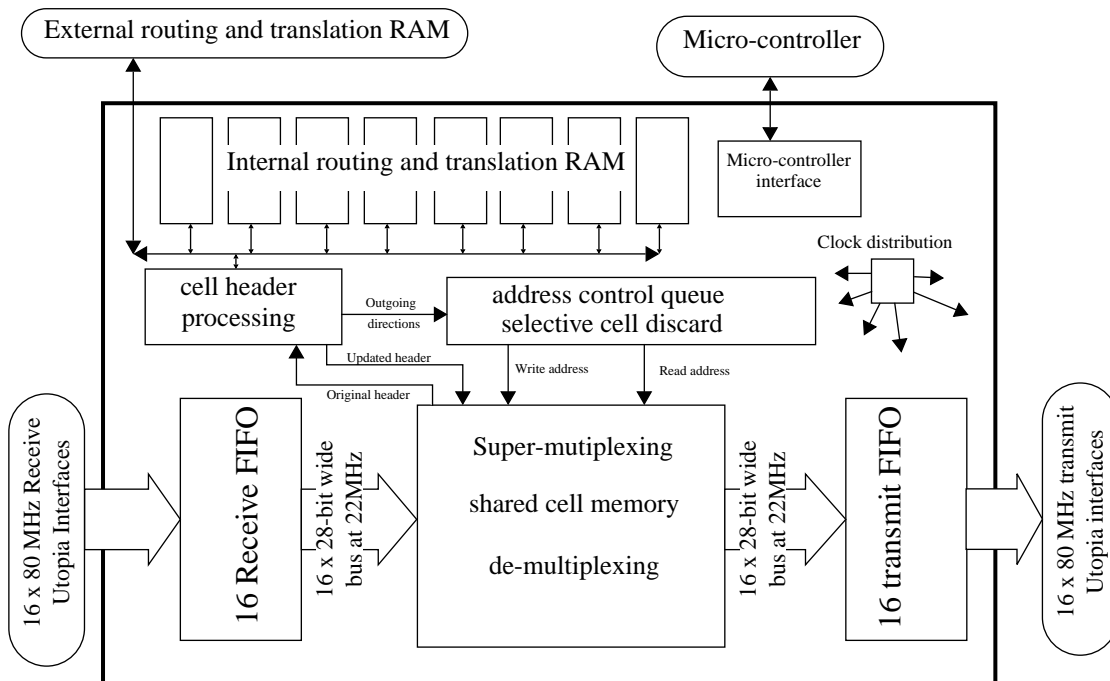


Fig. 5. COM16M architecture