A Complementary GaAs (CGaAs™) 32-bit Multiply Accumulate Unit

Michael J. Kelley, Matthew A. Postiff, Timothy D. Strong, Richard B. Brown, and Trevor N. Mudge

Department of Electrical Engineering and Computer Science
University of Michigan
1301 Beal Ave.
Ann Arbor MI, 48109

Abstract

A high speed one-cycle, 32-bit multiply, 64-bit accumulate unit is presented in Complementary GaAs (CGaAs™) technology. A tree of 4:2 compressors is used to collect the partial products and a carry select adder is used to determine the final result. Radix-4 Booth encoding is utilized to reduce the partial product tree size. Differential cascode voltage switch logic (DCVSL) is used on critical paths. A description of CGaAs technology, including its inherent radiation hardness, is provided as a background to the discussion. Finally, a study of some of the implications of designing in CGaAs is presented, including logic styles, circuit issues, design methodology, and their effect on performance.

1. Introduction

The emerging demand for satellite communications is moving the market for radiation hardened devices out of military applications and into the mainstream. Both of these markets have need for micro- and digital-signal processors which are capable of surviving a hostile radiation environment. In addition, the satellite communications market also requires low power operation.

Motorola's CGaAs technology satisfies these requirements. CGaAs offers inherent radiation hardness, and its high electron mobility provides good device transconductance and enables high speed operation at low voltages.

The newest generation of Motorola's CGaAs has drawn gate lengths of 0.5μm and thresholds of 0.35V. The process uses refractory metal gates and self aligned source/drain implants. Ohmic contacts to the source/drain areas are formed using a refractory metal, which provides high temperature stability and the opportunity to use standard aluminum metallization for interconnect. Three levels of AlCu metallization are currently available, and a fourth level of metal is currently under development [1, 2].

The University of Michigan's PUMA research project seeks to exploit these CGaAs characteristics in a high-speed, low-power design of the PowerPC microprocessor architecture. Several technology-proving designs are also underway. We chose to design the high speed multiply accumulate unit (MAC) reported in this paper because of the importance of multiplication in microprocessor and DSP systems.

2. Technology

CGaAs is a complementary GaAs technology which provides enhancement mode n-channel and p-channel heterostructure insulated-gate field effect transistors (HFETs) on the same substrate. A fully complementary (CMOS-like) logic style, having much lower power dissipation compared to other GaAs logic families, can be implemented in CGaAs. The rail-to-rail signal swings provide better noise margins than those of direct-coupled FET logic, the most common GaAs logic family.

2.1. Device Structure

CGaAs devices are created in epitaxial layers deposited by molecular beam epitaxy (MBE). A semi-insulating GaAs substrate is used for starting material.

A cross section of a pair of nHFET and pHFET devices appears in Figure 1. Heterojunctions help confine the carriers to a conducting channel formed in the InGaAs layer, where they have high mobilities because there is no impurity scattering. The large bandgap AlGaAs layer provides gate isolation, and the top GaAs layer is used as a cap to prevent surface oxidation of the AlGaAs.

In most MESFET technologies, the gate is formed by a Schottky diode junction. In CGaAs, the large bandgap of the AlGaAs layer between the TiWN gate metal and the channel increases the n-gate diode turn-on voltage (where $I_g = 1\mu A/\mu m^2$) to 1.75V, and the p-gate diode turn on voltage to -2V, with $V_{ds} = 0V$ [1, 3]. Gate leakage

Not to scale

Figure 1. Cross section CGaAs devices

increases with higher gate voltages, larger gate areas, and higher drain-source voltage (due to drain-induced barrier lowering). This leakage must be addressed in the design phase.

The threshold voltage of the devices is set by the conduction and valence band discontinuities of the AlGaAs/InGaAs, by a silicon delta doping layer placed just below the channel area, and by a p-channel implant [1]. Earlier versions of CGaAs (without the channel implant) set the thresholds at 0.55V. The new version is being optimized for a 0.35V threshold which will provide improved speed and the possibility of using unipolar transmission gates (n-pass gates). The earlier high-threshold devices resulted in an unacceptable voltage drop across an n-transistor pass-gate when passing a logic one.

Ohmic contact between the first metal layer and the source/drain areas is provided by a refractory metal which allows subsequent high temperature processing. CGaAs provides a conventional three-level AlCu metallization with stacked vias. A fourth metal level is in development but was not used in this design. The metal feature size and spacing are coarser than those of a typical CMOS technology of comparable minimum feature size [4].

Isolation of the devices is provided by an oxygen implant that extends into the semi-insulating substrate. This technique separates the n and p active regions. An implant is used to avoid the mesa etch step common in GaAs technologies, and it results in better surface planarity. Unlike CMOS, the substrate is semi-insulating, so wells and well contacts are not necessary, and latchup is not possible.

2.2. Process Characteristics

Circuits designed in CGaAs have several important features. The low threshold voltages, full power rail transitions and high transconductance at low supply voltages make low voltage performance good. These lower supply voltages reduce both dynamic power dissipation and power lost through the gate. The MAC is designed to oper-

ate at 1.3V, which is slightly lower than the technology's power supply rating of 1.5V. This moderates performance somewhat but provides a significant reduction in gate leakage.

The gate layer has a lower resistance than its CMOS polysilicon counterpart, but it is in direct contact with the semiconductor surface, and there is a leakage current even in oxygen implanted field areas. Therefore, one must route as little in the Schottky gate metal layer as possible.

Radiation hardness of the process is excellent. It is immune to total dose radiation effects because it does not use SiO_2 as a gate or field insulator. In a CMOS process, SiO_2 accumulates charge due to radiation, which causes leakage between transistors and a shift in the threshold voltages. Recently, the single event upset (SEU) soft error rate was improved by the addition of a Low Temperature GaAs (LTG) layer to reduce charge collection from ionizing radiation (see Figure 1). The LTG layer also improves subthreshold characteristics and provides better isolation between devices. Optimization of the new structure is ongoing.

With the immunity of CGaAs to latchup, a semi-insulating GaAs substrate, and the added LTG layer, CGaAs devices are suitable for radiation intensive environments [2].

3. Multiplier Architecture

Although CGaAs is the technology chosen to implement the MAC, the high-level architecture is essentially independent of the technology. The overall design goals for the multiply accumulate chip (MAC) included high speed, low area, and low power. The main focus was to explore the design space and to determine techniques and circuit topologies for digital CGaAs designs.

With these objectives, the MAC was designed as a one-cycle, two's complement multiply-accumulate unit. The MAC is composed of two subunits, a 32-bit multiplier and a 64-bit adder (see Figure 2). Latches on the inputs, outputs and internal registers are included in a global scan chain for testing purposes.

The 32-bit, two's complement multiplier tree is implemented using 4:2 compressors. A 4:2 compressor operates on four partial product bits and compresses these to two result bits. A tree layout based of 4:2 compressors has more regular cell placement and simpler routing than a 3:2 (full-adder) implementation [5].

A full tree of 4:2 compressors takes five levels to produce the two 64-bit sum and carry results presented to the final adder. The utilization of Radix-4 Booth encoding eliminates one level of compressors and reduces area [6].
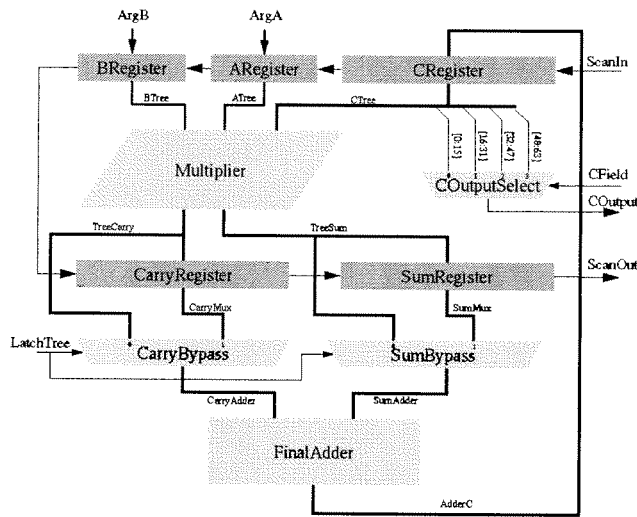
Figure 2. MAC Architecture

In addition to performing a 32-bit multiply, the 64-bit intermediate result can be used as an accumulate value; this technique is often used in signal processing circuits. The accumulate is implemented by folding the accumulate (CRegister) value into appropriate bit positions in the partial product tree of the multiplier. This is possible because some inputs to the compressors are not used by the multiplier. The value is thereby combined with the partial products from the A and B registers, and is already added to the product when the result appears at the intermediate registers.

The two 64-bit quantities output from the multiplier are added to form the final result, which is also the next accumulate value of the MAC. The 64-bit adder is composed of seven 16-bit adders so that multiplexors select the outputs of adders that have the correct carry results from lower segments. Each 16-bit adder is designed with 4-bit carry-lookahead. This architecture provides an efficient adder that requires only a few standard cells.

The adder and multiplier make up most of the MAC. To facilitate testing, the MAC incorporates several test devices. Two registers (CarryRegister and SumRegister) are included between the multiplier and the final adder. When these registers are selected, the MAC becomes a two-cycle pipelined design. A full scan-chain of all the registers is included for testing. Since the pipeline registers are included in the scan chain, the intermediate results of the multiplier tree are available for debug, and inputs can be scanned directly into the final adder to aid in its testing. The inputs to the multiplier tree and the accumulate value are latched at the beginning and end of the unit, respectively (see Figure 2).

4. Circuit Implementation

CGaAs provides a wide variety of logic families in which we could have designed the MAC unit, including full complementary, dynamic, passgate, source-coupled, and unipolar circuits. Each of these represents different points in the performance, size, power and noise immunity space [7]. The MAC was designed using primarily dynamic Differential Cascode Voltage Switch Logic (DCVSL). An example XOR gate appears in Figure 3. The gate is a standard DCVSL dual-rail domino gate with "keeper" pHFET devices to maintain voltages on the dynamic nodes. HSPICE simulations show that this "keeper" device not only allows the clock to be stopped, making the design pseudo-static, but also helps to prevent charge sharing in the evaluation logic from incorrectly influencing the output.
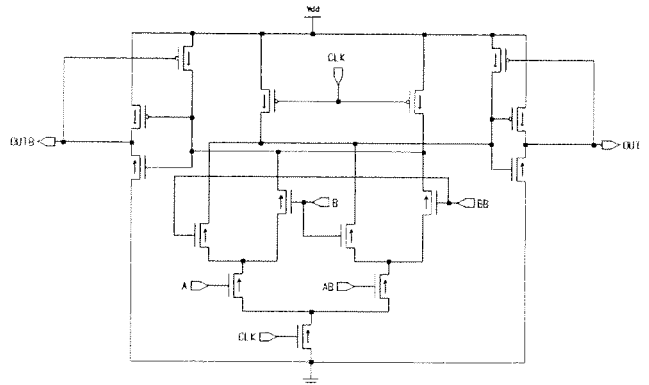

Figure 3. Dynamic DCVSL XOR gate with "keeper" transistors

The DCVSL style circuits allow higher performance than a fully complementary design because nHFET devices are 3 to 4 times faster than the corresponding pHFET devices. The gates are efficient for speed and power. The XOR gate of Figure 3 has an evaluation delay of 428ps at 189μW for a power delay product of 81fJ. While the presence of dynamic nodes might increase the SEU sensitivity of the design, the result is still expected to be much less radiation sensitive than a CMOS circuit. Evaluation of the MAC for radiation hardness will add to our understanding of dynamic CGaAs circuits.

The MAC uses full complementary circuits to perform Radix-4 Booth encoding [9, 11, 12]. The delay of these static circuits is hidden within the pre-charge time of the dynamic circuits. Once all dynamic nodes in the circuit have been pre-charged, they evaluate in domino fashion.

CGaAs also improves performance and reduces the size of the logic for Booth encoding over DCFL designs. Booth encoding is far more efficient in a technology that provides small fast multiplexors. DCFL designs must

implement the coding logic in NOR-NOR configurations which are larger [8].

4.1. General Design Considerations

Among the more obvious differences between CGaAs and CMOS is that CGaAs has no wells. This allows p- and n-channel transistor drains to be abutted, prevents latchup, and eliminates the need for well contacts.

Another difference in CGaAs is the Schottky layer. While the Schottky layer in CGaAs can be used for local routing, it leaks to the substrate, so its use as an interconnect layer should be limited especially when used with dynamic circuits. Dynamic circuits rely on the storage of charge on their internal nodes. The use of small "keeper" transistors helps maintain the charge on the dynamic nodes.

In addition, the n-transistors in CGaAs have a transconductance about 4 times better than the p-transistors. Due to the better transconductance, circuits should be designed primarily in the n-transistors. This design style lends itself very well to dynamic circuits where most of the logic is in the n-network. Complementary logic designed in CGaAs should be designed primarily of NAND structures, thus avoiding large p-stacks.

4.2. Physical Design

Although many process and circuit techniques of CGaAs may be different than in CMOS, the methodologies behind designing chips are very similar. The MAC is modeled using Verilog. Both behavioral and structural models of the MAC were simulated and verified for proper functionality by applying targeted and random test vectors.

In conjunction with the development of the MAC models, a standard cell library was created. The standard cells were custom designed using Mentor Graphics ICstation. Each standard cell was simulated and verified using HSPICE.

With a standard cell library in place, the MAC core was designed using Cascade Design Automation's EPOCH compiler. EPOCH was used for both standard cell placement and routing. The core was manually placed and routed to the padframe. Figure 4 shows a preliminary layout of the MAC unit. Further analysis on clock buffering and power dissipation needs to be completed. The MAC unit will be fabricated using Motorola's CGaAs process once the design is completed and a final layout is obtained.

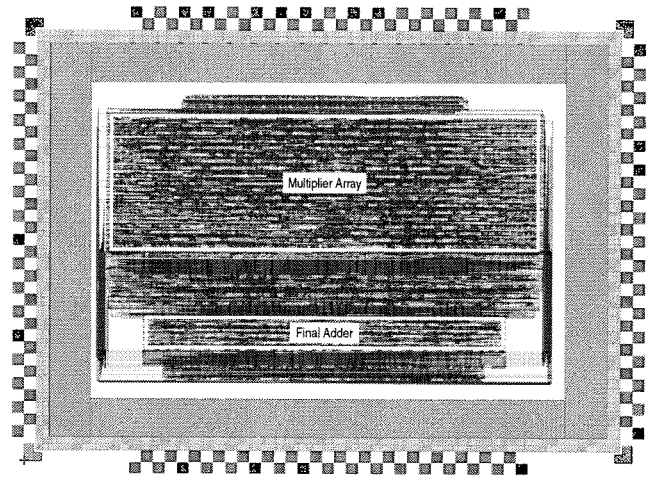The final design (without clock buffering) contains 60,904 transistors. HSPICE-estimated power dissipation

Figure 4. Preliminary MAC Layout

for the core is 239mW. The simulated critical path through the design is 13.7ns, yielding a clock speed of 73MHz in single cycle mode or about 140MHz when operated in pipelined mode.

5. Conclusion

The MAC described in this paper demonstrates several of the characteristics of CGaAs technology in a high speed design. Careful attention to layout and circuit details is necessary to avoid difficulties that can be caused by the differences between CGaAs and CMOS. However, CGaAs technology provides a high performance, low power, radiation hard alternative for digital designs.

Acknowledgment

References

[1] J. Abrokwah, J. Huang, W. Ooms, C. Shurboff, J. Hallmark, R. Lucero, J. Gilbert, B. Bernhardt, G. Hansell, "A Manufacturable Complementary GaAs Process," 1993 IEEE GaAs IC Symposium Technical Digest, pp. 127 - 130.

[2] M. LaMacchia, J. Abrokwah, B. Bernhardt, D. Foster, B. Crawforth, B. Mathes, T. McGuire, T. Weatherford, "Radiation Hardened Complementary GaAs," 1997 IEEE GaAs IC Symposium Technical Digest, pp. 59 - 61.

[3] A. Akinwande, P. Ruden, D. Grider, J, Nohava, T. Nohava, P. Joslyn, J. Breezley, "Complementary III-V Heterostructure FETs for Low Power Integrated Circuits,"

1990 IEEE International Electron Devices Meeting Technical Digest, pp. 983 - 986.

[4] P. O'Neil, B. Bernhardt, F. Nikpourian, C.Delta, Y. Abad, G. Hansell, "GaAs Integrated Circuit Fabrication at Motorola," 1993 IEEE GaAs IC Symposium Technical Digest, pp. 123 - 126.

[5] T. Huff, Architectural and Circuit Issues for a High Clock Rate Floating-Point Processor, 1995 Ph.D. Dissertation, University of Michigan.

[6] I. Koren, Computer Arithmetic Algorithms, 1993 Prentice Hall.

[7] R. Brown, B. Bernhardt, M. LaMacchia, J. Abrokwah, P. Parakh, T. Basso, S. Gold, S. Stetson, C. Gauthier, D. Foster, B. Crawforth, T. McQuire, K. Sakallah, R. Lomax, T. Mudge, "Overview of Complementary GaAs Technology for High-Speed VLSI Circuits," 1997 IEEE Transactions on VLSI Circuits.

[8] M. Riepe, T. Huff, M. Upton, T. Mudge, R. Brown, "A 7-nS 53x53-bit Parallel Array Multiplier Implemented in 0.6 μm GaAs DCFL," University of Michigan, Electrical Engineering and Computer Science, pp. 1-9.

[9] G. Bewick, Fast Multiplication: Algorithms and Implementation, 1994 Ph.D. Dissertation, Stanford University.

[10] A. Baca, J. Zolper, M. Sherwin, P. Robertson, R. Shul, A. Howard, D. Rieger, J. Klem, "Complementary GaAs Junction-Gated Heterostructure Field Effect Transistor Technology," 1994 IEEE GaAs IC Symposium Technical Digest, pp. 59 - 62.

[11] N. Burgess, "Removal of Sign-Extension Circuitry From Booth's Algorithm Multiplier-Accumulators," Electronics Letters, August 1990, Vol. 26, No. 17, pp. 1413 - 1415.

[12] M. Roorda, "Method to Reduce the Sign Bit Extension in a Multiplier That Uses the Modified Booth Algorithm," Electronics Letters, September 1986, Vol. 22, No. 20, pp. 1061 - 1062.

[13] M. Mehta, V. Parmar, E. Swartzlander, "High-Speed Multiplier Design Using Multi-Input Counter and Compressor Circuits," 1991 IEEE Symposium on Computer Arithmetic, pp. 43 - 50.

[14] J. Hallmark, C. Shurboff, B. Ooms, R. Lucero, J. Abrokwah, J. Huang, "0.9 V DSP Blocks: A 15 ns 4K SRAM and a 45 ns 16-bit Multiply/Accumulator," 1994 IEEE GaAs IC Symposium Technical Digest, pp 55 - 58.

[15] H. Ling, "High-Speed Binary Adder," 1981 IBM Journal of Research and Development, Vol. 25, No. 3, pp. 156 - 166.

[16] N. Quach, M. Flynn, "High-Speed Addition in CMOS," Stanford University Technical Report, CSL-TR-90-415, 1990.

[17] J. Huang, E. Glass, J. Abrokwah, B. Bernhardt, M. Majerus, E. Spears, J. Parsey, D. Scheitlin, R. Droopad, L. Mills, K. Hawthorne, J. Blaugh, "Device and Process Optimization for a Low Voltage Enhancement Mode Power Heterojunction FET for Portable Applications," 1997 IEEE GaAs IC Symposium Technical Digest, pp. 55 - 58.

[18] P. Song, G. DeMicheli, "Circuit and Architecture Trade-Offs for High-Speed Multiplication," 1991 IEEE Journal of Solid-State Circuits, Vol. 26, No. 9, pp. 1184 - 1198.

[19] J. Rabaey, Digital Integrated Circuits, 1996 Prentice Hall.