SiGe HBT BiCMOS FPGAs for fast reconfigurable computing

B.S.Goda, J.F.McDonald, S.R.Carlough, T.W.Krawczyk Jr. and R.P.Kraft

Abstract: Field programmable gate arrays (FPGAs) are flexible programmable devices that are used in a wide variety of applications. The flexibility of the FPGA hinders its performance due to the additional logic resources required for programmable hardware. The paper proposes a high speed SiGe heterojunction bipolar transistor (HBT) FPGA design co-integrated with CMOS in an IBM BiCMOS process. This device would be bitwise compatible with the Xilinx 6200, with operating frequencies in the 1–20 GHz range. To reduce power dissipation, the configuration bits used to define the FPGA's function will be stored in CMOS memory. Further power savings can be accomplished by integrating CMOS control into bipolar current trees and using a switchable current mirror to turn off unused current trees. The speed of bipolar combined with power savings of CMOS can now be merged to produce a new family of high speed FPGAs.

1 Introduction

A field programmable gate array (FPGA) consists of an array of reconfigurable logic blocks surrounded by segmented programmable interconnect [1]. Design entry and simulation is done in software and used to create the configuration file that defines the FPGA's logic function. Once programmed, the design is tested, validated and optimised. The flexibility of the FPGA allows it to reconfigure hardware resources to satisfy the instantaneous needs of a digital system. However, the relatively slow operating speeds of current FPGAs (30–70 MHz), prevents their use in high speed digital systems.

An example of an FPGA's programmable interconnect is depicted in Figs. 1–3, demonstrating how points 2 and 3 would be interconnected. The equivalent circuit is effectively a low-pass filter and studies have shown that the smallest delay between two configurable logic blocks (CLB) in the Xilinx 4000 FPGA family is approximately 1 ns [1, 2]. The delay in the FPGA's programmable interconnect is significantly greater than that of a simple wire, because of the significant resistance and capacitance introduced by routing signals though pass transistors. Connection delays often exceed the delay of the of the logic block and is therefore one of the fundamental limits on FPGA performance. The slow operating speed and poor bandwidth caused by the interconnect delay limits the FPGA's widespread use [1].

© IEE, 2000

IEE Proceedings online no. 20000468

DOI: 10.1049/ip-cdt:20000468

Paper first received 17th September 1999 and in revised form 7th March 2000

B.S. Goda is with the Rensselaer Polytechnic Institute, 6A Mordecai Drive, Watervliet, NY 12189, USA
E-mail: godab@rpi.edu

J.F. McDonald, S.R. Carlough, T.W. Krawczyk Jr. and R.P. Kraft are with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12181, USA
E-mail: mcdonald@unix.cie.rpi.edu
E-mail: carlough@unix.cie.rpi.edu
E-mail: krawct@rpi.edu

2 SiGe HBT BiCMOS process

The high current gain and low input capacitance of bipolar transistors make it an attractive replacement for CMOS in some high performance applications. Bipolar devices can permit low logic swings, which are essential for GHz operation at low power. Further performance increases are feasible in HBT technology where AlInAs/GaInAs devices with 500 GHz F_{MAX} have been demonstrated in research laboratories [3]. However, GaAs/AlGaAs HBT technology cannot be integrated with existing CMOS technology, thus low power memory circuits are not possible. FPGAs and other reconfigurable circuits require configuration memories to store personalisations of the fixed resources of these circuits. A 256-cell bipolar configurable array logic was developed in 1990 using 1.2 μm double-metal collector diffusion isolation technology [4]. The idea of using bipolar in programmable logic makes them attractive in such applications as ASIC emulation and signal processing.

A mature foundry capability has been developed at IBM Burlington that can produce both 0.5 micron CMOS and 0.5 micron (50 GHz f_T) HBT devices in one process. SiGe technology offers some of the speed of III-V devices at the processing costs and yields of silicon. Bipolar transistors achieve their speed through the use of a very thin base region. The speed of the collector current depends on how long it takes for charge carriers to travel through the base. By introducing Ge into the base of a Si BJT, a smaller base bandgap is created, which increases electron injection from the emitter and decreases base transit time. The decrease in transit time results in a higher f_T and higher β. A higher β permits the base doping level to be raised, lowering the base resistance. In addition, Ge at the collector/base junction gives a higher early voltage, a desirable characteristic for high-speed [5]. IBM has developed a graded Ge alloy base where the Ge content varies linearly across the base,

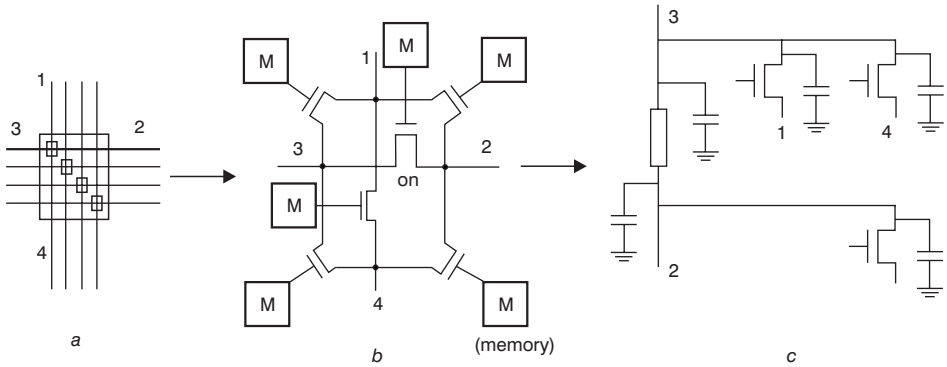

Fig. 1 *Interconnect delay*

a Interconnect
b Pass transistor
c Equivalent circuit from node 3 to node2

creating a built-in electric field to enhance speed. The emitter–base interface has 0% Ge, so the IBM device is really a graded base BJT, but the term HBT is still commonly used. Circuits can be created capable of operating up to 20 GHz with current mode logic (CML 13 ps gate delay), while incorporating CMOS (60 ps gate delay) technology for configuration memory. It is now possible to have the speed of an HBT combined with the low power consumption of CMOS.

3 Current mode logic

High speed in CML is achieved by operating all transistors out of saturation and keeping the logic signal swings relatively small (250 mV). CML differential logic uses a pair of BJT transistors as a current switch. Using this type of logic produces low switching noise, a desirable feature when operating at high-speed [6]. Another advantage of differential CML logic is that the complement signal is always available, so extra circuitry to create a signal complement is not necessary. Fig. 2 depicts a CML XOR implementation in seven transistors and three resistors. Signal A and its complement come in on level 1 (0 and −0.25 V) and Signal B and its complement comes on level 2 (−0.95 and −1.2 V). The difference between levels is slightly more than one V_{BE} (0.85 V). The tail resistor at the bottom of the current steering tree is connected to a reference current mirror that fixes the current through the circuit (0.7 mA). Current flows through only one path at a time.

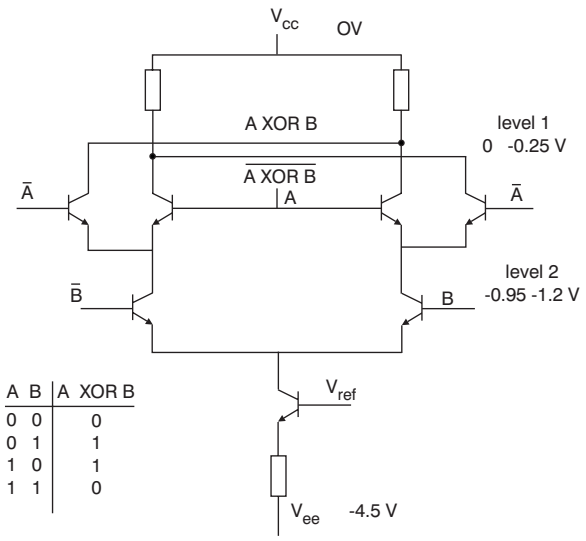

Fig. 2 *Full diffirential current mode logic XOR cell*

The voltage swing to turn on or off the opposite members of a given differential pair depends on what level the pair sits in the tree. This prevents saturation of any of the transistors, which slows down the switching of the transistor due to charge storage in this state. To avoid saturation, each transistor needs to have a bias or centering offset voltage that is different for each level. This requires a level translator circuit when the same driver drives multiple input levels.

By placing a field effect transistor (FET) in a Widlar bipolar current mirror, direct connections to memory circuits can be made to turn off unused current trees. The graphs in Figs. 3 and 4 represent the level 1 output (0 to −250 mV) of a current tree being turned off and on. A sharp on, slower off characteristic is desirable during reconfiguration operations.

Complex functions can be created in what amounts to one gate macro, with propagation delays comparable to that of a simple gate. This results in faster operation because one does not have to traverse several levels of logic to obtain such complex functions [7]. One of the difficulties of differential CML logic is the true and complementary inputs need to switch skew free. Otherwise, the differential pair will be blocked on either of its two paths, which would result in an interruption of current. Preventing this is accomplished by routing wire pairs adjacent to each other. Another problem of differential CML is the doubling of the wires in the design. Wire doubling is approximately offset by the high degree of functionality of the CML logic circuits. Differential wiring also increases the RC charging time for a given signal since twice the voltage swing is seen between conductor pairs.

Crosstalk is less of a problem in CML because the circuit family provides common mode noise rejection that will suppress noise occurring on the differential circuit inputs. Often, the coupling is difficult to visualise, and frequently the coupling can end up converting a piece of GHz logic into a feedback oscillator generating unwanted behaviour throughout the circuit. This is almost completely suppressed by the differential features of the circuit and the differential wiring.

Data-dependent signal switching occurs when neighbouring signals switch simultaneously. For single-ended logic, a fast transition occurs when signal neighbours swing in the same direction and a slow transition occurs when neighbours transition in the opposite direction. Data-dependent signal switching is a dangerous situation in system design, since digital timing simulators often do

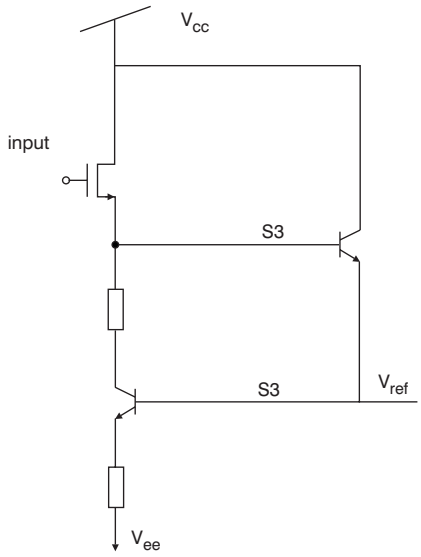

Fig. 3 *Widlar current mirror with CMOS switch*

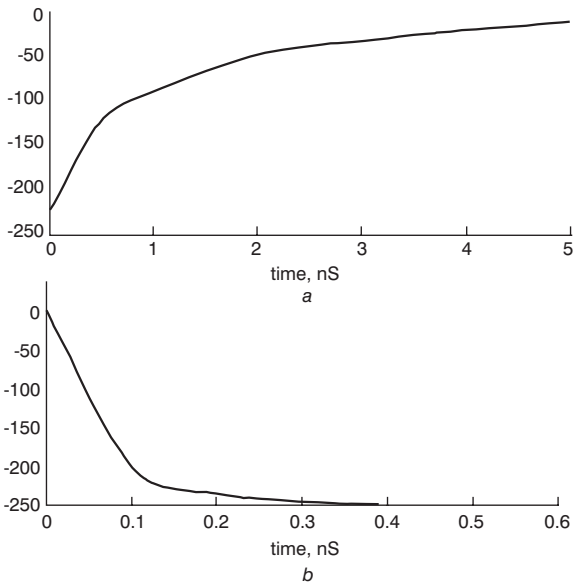

Fig. 4 *Widlar current mirror with CMOS switch*

a Current tree voltage turn-off
b Current tree voltage turn-on

not include fluctuations in local electric fields. Modelling interconnect assuming all signals undergo slow transition (designing for the worse possible case) is not always sufficient. If data signals are too fast, they may violate hold times, and, if clock signals are too fast, they generate setup time violations for the latches they clock. This effect is further complicated when signals from a low power driver are nearby signals from a high power driver.

Data-dependent signal switching still occurs in differential signals although the effect is less pronounced since the modelling already accounts for the fact that one neighbour (the complement signal) is always swinging in the opposite direction. Fig. 5 shows how data-dependent signal switching can occur in simultaneously switching routed differential pairs. Analyses of interconnect layout shows that this effect can change signal switching speeds by as much as 30%.

Often the data busses require many signals switching at the same time to occupy a dense routing region. Under these circumstances it may not be possible to separate simultaneously switching signals. One method of reducing

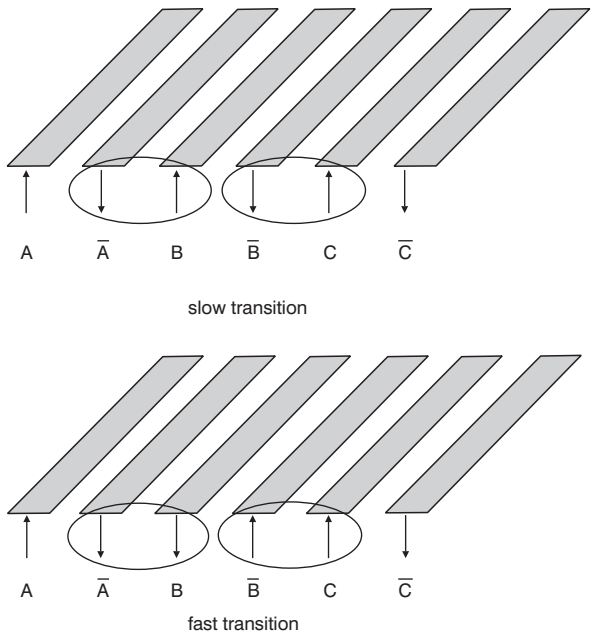

Fig. 5 *Data dependent switching*

data dependent switching that is similar to bit-line twisting used in memory cells. Dynamic capacitance is balanced between the transient signal and the wires in the differential pair, by switching which wire neighbours the transient signal.

4 Xilinx 6200 emulation

The Xilinx 6200 chip was chosen for emulation since the information about the configuration bits are in the public domain and core logic cell in the 6200 chips consists of multiplexers and flip-flops, which can easily be converted into CML logic (Fig. 6) [8]. Fig. 6 shows a 6200 configurable logic block (CLB) configuration and a redesigned bipolar version. The redesign takes advantage of the differential pair wiring and switchable current trees. A 40% savings can be realised and the redesigned CLB can

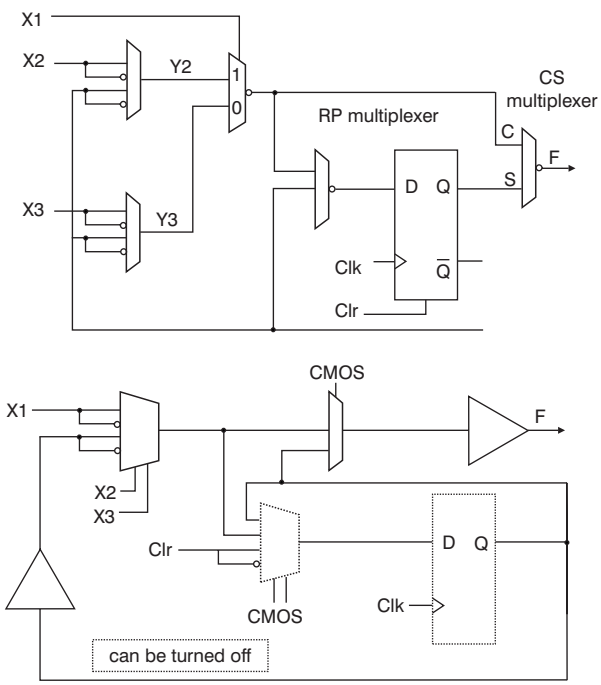

Fig. 6 *Basic logic cell (CLB) in XC6200 FPGA and redesigned cell*

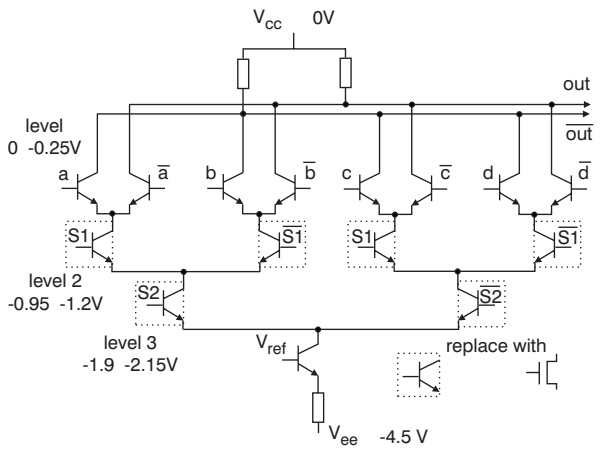

Fig. 7 *4 to 1 multiplexer in CML*

be implemented in 50 bipolar, 11 CMOS, and 20 resistors. If the memory part of the CLB is not, half of CLB can be turned off.

The XC6200 uses multiplexers to implement its logic functions. Fig. 7 is a CML implementation of a four to one multiplexer using 15 (15 bipolar or 9 bipolar/6 CMOS) transistors and 3 resistors. The inputs on levels two and three are used to select one of four level 1 inputs. The uppermost level of the CML tree is where the fastest switching capability exists, and fortunately this is also where the FPGA visible signals would enter. This means that the nearly static CMOS signals would occupy only the lower levels of the current tree where speed is not as critical [9].

The XC6200 uses multiplexers to implement its logic functions. A lookup table is used to set the inputs to the multiplexers for the desired function. Fig. 8 depicts how a logic function is implemented. The X1 input controls whether Y2 or Y3 selected. In the first example, if $a = 1$, Y2 is selected. If $b = 0$, Y2 = 1 and the output = 0. If $b = 1$, Y2 = 0 and the output = 1. If $a = 0$, Y3 = 1 and the output = 0. Thus the first example implements a 2 input AND function. In the second example, the a input controls whether Y2 or Y3 is selected ($a = 1$, Y2: $a = 0$, Y3). If $a = 1$, if b = 1, the output = 0, else if $b = 0$, the output = 1.

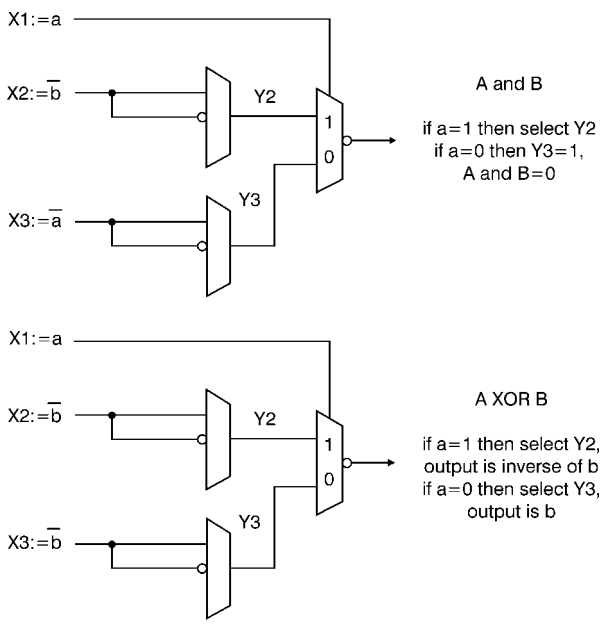

Fig. 8 *Multiplexer logic functions*

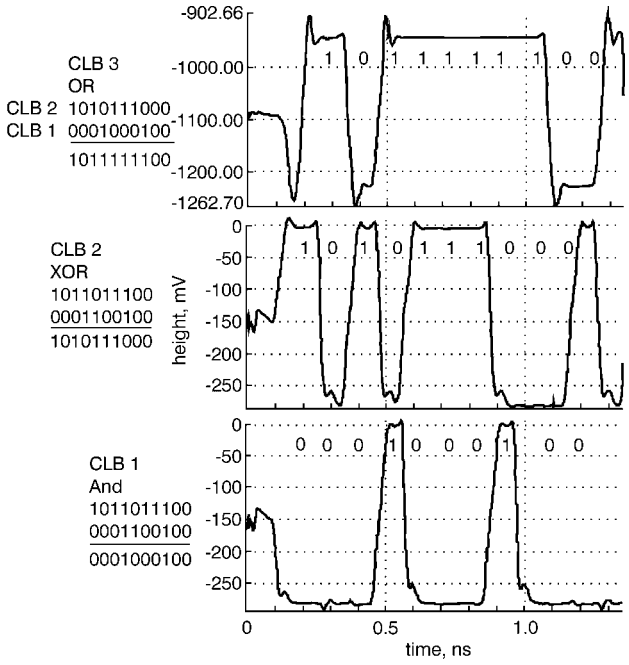

Fig. 9 *10 GHz three CLB simulation*

The opposite is true if $a = 0$. Thus the second example implements a 2 input XOR function. Similar functions are calculated in the same manner [8].

A simulation in Fig. 9 shows CLB 1 ANDing two 10 GHz signals, CLB 2 XORing the same two signals, and CLB 3 ORing the outputs of CLBs 1 and 2. The simulation demonstrates that different logic functions can be implemented at high speeds and the output can drive other CLBs. Further simulation at higher speeds result in distorted waveforms since transistors are unable to switch fast enough to maintain the design noise margin of 200 mV. The most important point is that HBT CML can be used to implement all of the logic components visible to the designer, as well as all the hidden top level routing resources. CML multiplexers pass signals with delays of about 12–14 ps in the 50 GHz SiGe HBT process. High GHz signal bandwidths are sustainable in this class of circuits, provided wire lengths are limited, or repeaters are used to reshape signal rise and fall times.

In Fig. 10 a XC6200 is surrounded with various multiplexers used for routing. Signals can come from adjacent CLBs (north, south, east, west) or from adjoining groups that are 4 CLBs away (N4, S4, E4, W4). The XC6200 uses a 'magic' wire for critical signals such as carry forward.

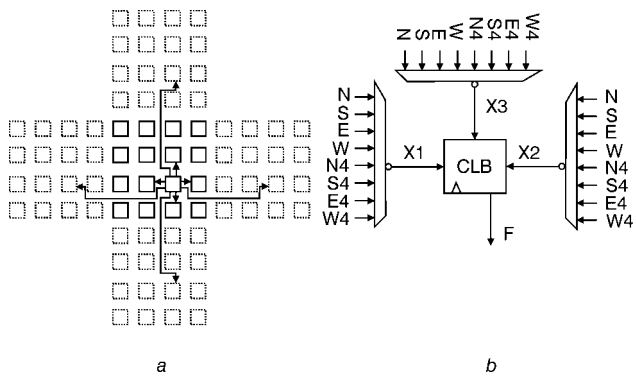

Fig. 10 *Multiplexer porting to XC6200 CLB*

a Outgoing CLB routing
b Ingoing CLB routing

The clear (clr) and clock (clk) signals travel in opposite north–south directions.

Connections from the NESW ports to other CLBs may be made *via* local connections between adjacent cells or through global connections that run vertically and horizontally on the chip. Bipolar drivers for the longer interconnections tend to encounter less RC delay and hence support higher bandwidth signals [9]. This is due to the low on-resistance of the bipolar device, especially when scaled to handle the currents appropriate for such lines. Emitter follower circuits help insure that this low device on-resistance can be used to create the low impedance drive required. This can be an important feature in creating high fan-out clock circuits for flip-flops.

5 Configuration memory

The current FPGA design includes eight memory planes for the configuration data. Each memory plane contains a different configuration for the FPGA as well as the state of the latches in each CLB. Bipolar devices are a poor choice for memory elements since they dissipate static power; however, CMOS circuits dissipate very little static power making them ideal for the configuration memory. A CMOS multiplexer is used to select between the memory planes allowing the FPGA to page between up to eight different tasks at high rates of speed. Furthermore, the memory planes may be loaded or saved to the external system while processing continues uninterrupted in the FPGA.

Fig. 11*a* shows the logic for a single CLB configuration bit if a normal CMOS to CML buffer is used. Each of the eight CMOS memory planes contain data for this configuration bit, which determines the state of the multiplexer input signals for the CLB. The 8 to 1 CMOS multiplexer is used to select the correct bit from the memory planes to pass on to the CLB. With the integration of CMOS directly into the current tree as shown in Fig. 11*b* the CML buffer is not required and signals directly from memory can be used for multiplexer selects.

6 Performance results

The results for the major parts of the SiGe FPGA have been summarised in Table 1. All but the CLB are measured in the results, the CLB numbers are from simulation. An interesting finding is that a CML gate and a CML multiplexer implementing the same function run at almost the same speed. This is due to the similar current tree structure used to implement both logic gates and multiplexers.

High power consumption is often cited as one of the biggest drawbacks to using bipolar for making large-scale logic circuitry [10]. Reduction of current trees and smaller voltage swings (250 mV), however, permits larger scale, high-speed CML logic circuits to be built in SiGe. A 1000 CLB CLB SiGe FPGA would consume ~22 watts (although most FPGAs only use about half of their resources at any one time), so a SiGe FPGA can operate at room temperature. Fig. 12 demonstrates the temperature sensitivity of a 2:1 multiplexer. Normal design rules call for at least a 200 mV noise margin, so a multiplexer can operate at 12 GHz at room temperature and 5 GHz at 125°C. Cooling the multiplexer results in a 20% improvement at −50°C. Further study at Auburn University is

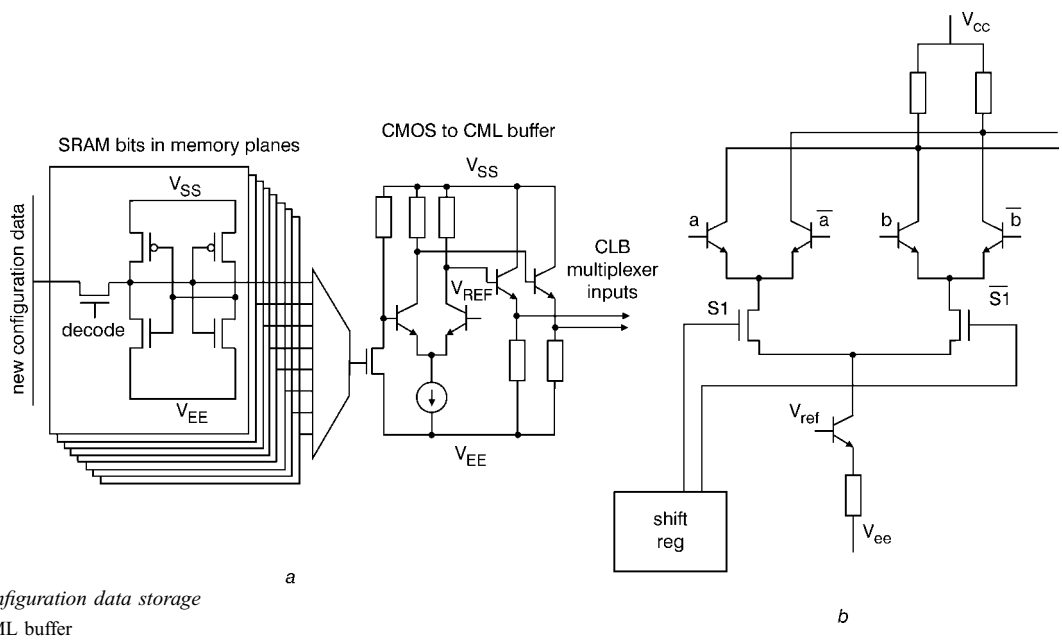

Fig. 11 *Configuration data storage*
a CMOS to CML buffer
b CMOS integrated into current tree

Table 1: SiGe performance

Circuit type	Buffer	CML XOR, AND, OR	MUX XOR, AND, OR	CLB
Propagation delay	17 ps	22–25 ps	23–26 ps	100 ps
Current used	0.7 mA	0.8 mA	0.8 mA	2.4 mA
Power used	3.15 mW	3.6 mW	3.6 mW	10.8 mW

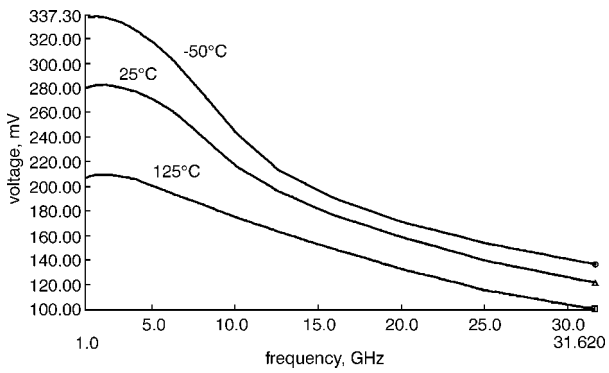

Fig. 12 *Multiplexer performance frequency, GHz*

being done for circuit operations at even lower temperatures [9].

7 Conclusions and future work

The design of an FPGA that runs in the 1–20 GHz range using CML in SiGe combined with low power CMOS in the same process has been investigated. The Xilinx XC6200 was chosen for emulation due to its public domain bit stream and implementation in multiplexer logic and flip-flops. The core of the XC6200 has been demonstrated to be feasible and the major parts work at speed at room temperature. The drawbacks of previous high power consumption have been mitigated by using CMOS to turn off current mirrors and integrating CMOS directly into the bipolar current tree.

One of the largest areas of future work will be integrating the existing XC6200 tools to work with the SiGe FPGA. Although the XC6200 is not being manufactured any more, there are many independent programmers that create tools for the XC6200. Work proposed in this paper will help continue the independent development of the XC6200.

8 References

1 SMITH, M.: 'Applications-specific integrated circuits' (Addison-Wesley, Reading, Massachusetts, 1997)

2 SINGH, S., ROSE, J., CHOW, P., and LEWIS, D.: 'The effect of logic block architecture on FPGA performance', *IEEE J. Solid-State Circ.*, 1992, **27**, (3), pp. 281–287

3 MENSA, D., PULLELA, Q., LEE, J., GUTHRIE, S., MARTIN, C., SMITH, R., JAGANATHAN, S., MATHEW, T., AGARAWAL, B., LONG, S., and RODWELL, M.: '48-GHz digital ICs and 85-GHz baseband amplifiers using transferred-substrate HBTs', *IEEE J. Solid-State Circ.*, 1999, **34**, (9), pp. 1196–1203

4 CARRUTHERS, C., and KEAN, T.: 'Bipolar CAL chip doubles speed of FPGAs'. Oxford international workshop on field programmable logic and applications, 1991, pp. 46–53

5 IBM SiGe Designer's Handbook' (IBM Inc., Burlington Vermont, 1998)

6 MCDONALD, J., PHILOWER, R., VANETTEN, J., SIMMONS, S., TSINKER, V., LOY, J., and GREUB, H.: 'F-RISC/G: AlGaAs/GaAs HBT standard cell library'. Proceedings of the IEEE international conference on computer design, 1991, Boston, MA, pp. 297–300

7 MCDONALD, J., GREUB, H., YAMAGUCHI, T., and CREEDON, T.: 'Cell library for current mode logic using an advance bipolar process', *IEEE Trans.*, 1991, **JSCC–26**, (5), pp. 749–762

8 Xilinx series 6000 user guide' (Xilinx Inc., San Jose, CA, 1997)

9 HARAME, D., CRABBE, E., CRESSLER, J., COMFORT, J., SUN, J., STIFFLER, S., KOBEDA, E., BURGHARTZ, M., GILBERT, J., MALINOWSKI, A., DALLY, S., RATHANAPHANYARAT, M., SACCAMANGO, W., COTTE, J., CHU, C., and STORK, J.: 'A high performance epitaxial SiGe-base ECL BiCMOS technology', (IEEE IEDM Tech Digest, 1992), pp. 2.1.1–2.1.4

10 SINGH, S., CHOW, P., and LEWIS, D.: 'The effect of logic block architecture on FPGA performance', *IEEE J. Solid-State Circ.*, 1992, **27**, (3), pp. 281–287

11 CARLOUGH, S., CAMPBELL, P., STEIDL, S., GARG, A., MAIER, C., GREUB, H., MCDONALD, J., and ERNEST, M.: 'Wiley encyclopedia of electrical and electronics engineering' (Wiley, New York, 1999), vol. 17

12 ROSE, J., FRANCIS, R., LEWIS, D., and CHOW, P.: 'Architecture of field-programmable gate arrays: the effect of logic block functionality on area efficiency', *IEEE J. Solid-State Circ.*, 1990, **25**, (5), pp. 1217–1225

13 TAUR, Y., and TAK, N.: 'Fundamentals of modern VLSI devices' (Cambridge University Press, New York, 1998)