Integration of a Double Polysilicon, Fully Self-Aligned Bipolar Transistor into a 0.5μm BiCMOS Technology for Fast 4MBit SRAMs

J. D. Hayden, J. D. Burnett, A. H. Perera, T. C. Mele, F. W. Walczyk, V. Kaushik, C. S. Lage, Y.-C. See‡

Advanced Products Research and Development Laboratory, Motorola Inc.,
3501 Ed Bluestein Blvd., Austin, TX 78721 (512) 928-5026
‡Advanced Technolgy Center, Motorola Inc., 2200 W. Broadway Road, Mesa, AZ 85210.

Abstract

The single polysilicon, non-self-aligned bipolar transistor in a 0.5μm BiCMOS technology has been converted to a double polysilicon, fully-self-aligned bipolar with little increase in process complexity. Improved bipolar performance in the form of smaller base resistance, larger knee current, higher peak cut-off frequency, and shorter ECL gate delay has been demonstrated. This technology will prove useful in meeting the requirements for higher performance in fast, high density, SRAM circuits.

Introduction

BiCMOS is the process of choice for fast, high density, SRAM circuits [1]. As circuit speeds continue to increase and ECL circuit design techniques are more extensively employed, improved bipolar performance is required. In particular, a double polysilicon, emitter-base or 'fully' self-aligned (FSA) bipolar transistor offers reduced collector-base capacitance and base resistance and improved ECL gate delay when compared to a single polysilicon, non-self-aligned device (NSA).

In many 4Mb generation SRAM technologies, an additional polysilicon layer has been introduced in order to form self-aligned contact landing pads which allow the bit cell area to be reduced to below 20μm^2 [2-3]. In such a structure, the second polysilicon layer makes a contact with single crystal silicon between a pair of first polysilicon word lines. In the present work, this self-aligned landing pad structure has been used to form an FSA bipolar transistor.

Process

The starting point for this work is a 0.5μm BiCMOS technology for fast 4Mb SRAM circuits with a single polysilicon NSA bipolar transistor [3]. The process assumes a 5V power supply. This technology employs self-aligned complementary buried layers and twin wells. Framed-Mask Poly Buffered LOCOS (FMPBL) isolation is used to achieve an active pitch of 1.3μm [4]. The process features three layers of polysilicon: the first is used to form the MOSFET gates; the second the bipolar emitter, a self-aligned contact landing pad and a global interconnect; and the third a teraohm load resistor in the bit cell. The second polysilicon layer is optionally strapped with titanium salicide or tungsten polycide. Surface channel NMOS and buried channel PMOS transistors are fabricated with a 150Å gate oxide thickness and display good MOSFET characteristics to minimum effective channel lengths of 0.35 and 0.50μm respectively. Both n-channel and p-channel transistors incorporate moderately doped drain (MDD) extensions to increase their resistance to short channel effects and hot carrier degradation. A variety of MDD spacer schemes are available in this process, including reverse sequence, disposable polysilicon spacers [5], and permanant dielectric spacers comprised of silicon nitride or deposited silicon dioxide. The original process features a single polysilicon NSA vertical npn bipolar transistor, where the second poly layer is used for the emitter. The extrinsic base is formed with an implant into single crystal silicon. This implant is offset lithographically from and, hence, not self aligned relative to the emitter. A rapid thermal anneal step serves to break up the interfacial oxide between emitter poly and single crystal silicon, activate the emitter dopant, and drive arsenic from the poly emitter into single crystal silicon, as well as to provide planarization prior to deposition of the first metal layer. Two levels of metallization are provided in this technology.

The single polysilicon NSA bipolar transistor was converted to a double polysilicon FSA structure by using the first polysilicon layer to form the extrinsic base contact, which now completely surrounds the emitter. The second polysilicon continues to be used for the emitter. The emitter and extrinsic base are separated by a dielectric spacer, and are thus self-aligned. The base resistance, R_B, and collector-base capacitance, C_{CB}, are then reduced relative to the NSA transistor where emitter and extrinsic base are offset lithographically. The spacer consists of a silicon dioxide portion formed during the emitter window etch as well as whatever spacer remains after MOSFET MDD formation. Formation of the FSA bipolar is thus identical to that of the self-aligned contact landing pad except that the gate oxide is removed from the bipolar area prior to deposition of the first polysilicon layer, the extrinsic base poly is doped p-type, and the source/drain implants are blocked from the emitter area. The optimization of the FSA bipolar transistor and the landing pad can be performed simultaneously. Interfacial oxide and dopant loss during polycide formation are found to affect both structures similarly. Figure 1 is a schematic cross-section of the technology.

The extrinsic base region was experimentally optimized by varying the extrinsic base implant dose while maintaining the remainder of the process unchanged. FSA and NSA transistors were fabricated on the same wafers, providing for excellent comparisons.

Experimental Results and Discussion

Excellent bipolar characteristics are demonstrated for FSA transistors with emitter widths as small as 0.4μm. Figure 2 and 3 are SEM and TEM cross-sections of an FSA bipolar. A Gummel plot of an FSA device with a patterned emitter dimension of 0.4X1.2μm exhibits ideal bipolar characteristics to below 0.1pA as shown in Figure 4. Arrays of 1000 FSA transistors do not exhibit any higher incidence of shorts than comparable arrays of NSA transistors, indicating no significant etch damage during extrinsic base formation.

Figure 5 compares Gummel plots and Table 1 presents transistor parameters for FSA and NSA bipolar transistors fabricated on the same wafer. The FSA transistor exhibits higher base current and lower collector current at intermediate emitter-base voltages. Lateral encroachment of the extrinsic base reduces the emitter efficiency at the emitter edge as well as decreasing the effective emitter area. The FSA transistor displays lower base resistance at high current levels and some evidence of forward excess tunneling [6] at very low currents for the highest extrinsic base implant doses.

The FSA transistor provides an improvement in base resistance, R_B, over the NSA device because of the reduced spacing between emitter and extrinsic base regions. This improvement comes, however, at the expense of some degradation in emitter-base breakdown voltage BV_{EBO}. Figure 6 illustrates that base resistance can be traded off against emitter-base breakdown voltage by varying the extrinsic base implant dose. Additional process modifications, in the form of reduced back-end heat and wider dielectric spacers would further improve this tradeoff. Much of the reduction in R_B is due to the double base configuration. Figure 7 compares results from NSA single and double base transistors with FSA transistors where the extrinsic base wraps around the entire emitter region. The FSA structure simultaneously reduces R_B and device area, which in turn decreases collector-base capacitance, C_{CB}. The roll-off of base resistance with increasing base current is reduced with the FSA structure because lateral encroachment of the extrinsic base dopant inhibits charge modulation of the intrinsic base around the periphery of the emitter.

The effect of hot carrier stressing on bipolar characteristics is studied in Figure 8 for FSA and NSA transistors. At low extrinsic base

CH3020-5/91/0000(0017)$1.00 © 1991 IEEE

implant doses, there is little difference between the FSA and NSA devices. Higher extrinsic base doses result in a more rapid increase in the non-ideal base current component, ΔI_B, under reverse-bias emitter-base stress. The higher doping around the periphery of the emitter-base junction results in increased electric field levels during emitter-base avalanche breakdown and a more rapid generation of interface states in the oxide overlying the junction edge.

The wrap-around base contact in the FSA transistor produces a more uniform electron injection around the emitter edge when compared to a single base, NSA bipolar. By reducing current crowding at the emitter edge, the onset of the Kirk effect or base push-out is delayed and a higher knee current results [7], as shown in Figure 9.

The geometry dependence of transistor current gain manifests itself differently between FSA and NSA transistors. This is demonstrated in Figure 10. The collector current density increases with smaller periphery-to-area ratio for both type devices. Diffusion of emitter arsenic into single crystal silicon increases the effective area for electron injection. For NSA transistors, the base current density is relatively constant and the current gain or ßeta increases. The base current is dominated by hole recombination at the emitter polysilicon/silicon interface and thus scales with patterned emitter area. For FSA transistors, encroachment of the extrinsic base dopant causes the base current density to increase more rapidly than collector current, forcing ßeta to decrease. Lateral encroachment of boron from the extrinsic base poly pinches the corners of the emitter, decreasing the emitter Gummel number and increasing hole injection at the emitter edge.

The base transit time is not expected to vary significantly between FSA and NSA transistors on the same wafer. Figure 11 shows that the peak cut-off frequency, f_T, on the other hand, is higher for the FSA device. This is due to a delay in the onset of base push-out and a reduction in parasitic capacitances (C_{CB}), compensating for any narrow emitter effects [8]. Higher cut-off frequencies could be obtained with a selectively implanted collector (SIC) [9] and/or a double diffused emitter/base [10] approach.

The FSA structure provides a significant improvement in ECL gate delay relative to the NSA transistor, as illustrated in Figure 12. The reduction in collector-base capacitance and base resistance are responsible for the improvement. Converting from an NSA to an FSA bipolar reduces ECL gate delay from 140 to 70ps/gate and 110 to 60ps/gate for conservative and aggressive layout rules respectively. Loaded BiCMOS gate delay is also reduced 10-20% with the FSA structure, again because of the reductions in R_B and C_{CB}. ECL gate delay can be further improved by eliminating the p^+ buried layer which decreases collector-substrate capacitance, C_{CS}, and by reducing the epitaxial layer thickness, which increases the knee current and lowers the collector resistance. These trends are illustrated in Figure 13.

Conclusions

A double polysilicon, fully-self-aligned bipolar transistor has been successfully substituted for a single polysilicon, non-self-aligned bipolar transistor in a 0.5µm BiCMOS technology with little increase in process complexity. Greatly improved bipolar performance has been demonstrated.

References

[1] M. Takada, et. al., "A 5-ns 1-Mb ECL BICMOS SRAM," *IEEE J. Solid-State Circuits*, vol. SC-25, pp. 1057-1062, 1990.

[2] T. Hirose et. al., "A 20ns 4Mb CMOS SRAM with Hierarchical Work Decoding Architecture," *ISSCC Dig. Tech. Papers*, , pp. 132-133, 1990

[3] T. C. Mele et. al., "A High Performance 0.5µm BiCMOS Triple Polysilicon Technology for 4Mb Fast SRAMs," *IEDM Tech. Digest*, pp. 481-484, 1990.

[4] J. D. Hayden et. al., "A High-Performance Half-Micron Generation CMOS Technology for Fast SRAM's," *IEEE Trans. Electron Dev.*, vol. ED-38, pp. 876-886, 1991.

[5] L. C. Parrillo et. al., "An Advanced 0.5µm CMOS Disposable LDD Spacer Technolgy," *Proc. Symp. on VLSI Technology*, pp. 31-32, 1989.

[6] G. P. Li, E. Hackbarth, and T.-C. Chen, "Identification and Implication of a Perimeter Tunneling Current Component in Advanced Self-Aligned Bipolar Transistors," *IEEE Trans. Elec. Devices*, vol. ED-35, pp. 89-95, 1988.

[7] A. Watanabe, T. Nagano, S. Shukuri, and T. Ikeda "Future BiCMOS Technology for Scaled Supply Voltages," *IEDM Tech. Digest*, pp. 429-432, 1989.

[8] D. D. Tang, et. al "Design Considerations of High-Performance Narrow-Emitter Bipolar Transistors," *IEEE Electron Devices Letters*, vol. EDL-8, pp. 174-175, 1987.

[9] S. Konaka et. al., "A 20-ps Si Bipolar IC Using Advanced Super Self-Aligned Process Technology with Collector Ion Implantation," *IEEE Trans. Electron Dev.*, vol. ED-36, pp. 1370-1380, 1989.

[10] B. van Schravendijk, et. al., "Thin Base Formation by Double Diffused Polysilicon Technology," *IEEE Bipolar Circuits and Technology Meeting*, pp. 132-134, 1988.

Table 1 NPN Transistor Parameters

Parameter	Units	NSA	FSA
BV_{CB0}	(V)	18	18
BV_{CE0}	(V)	6.5	6.5
C_{EB}	(fF)	8	8
C_{CB}	(fF)	13	8
C_{CS}	(fF)	48	41
R_E	(Ω)	13	12

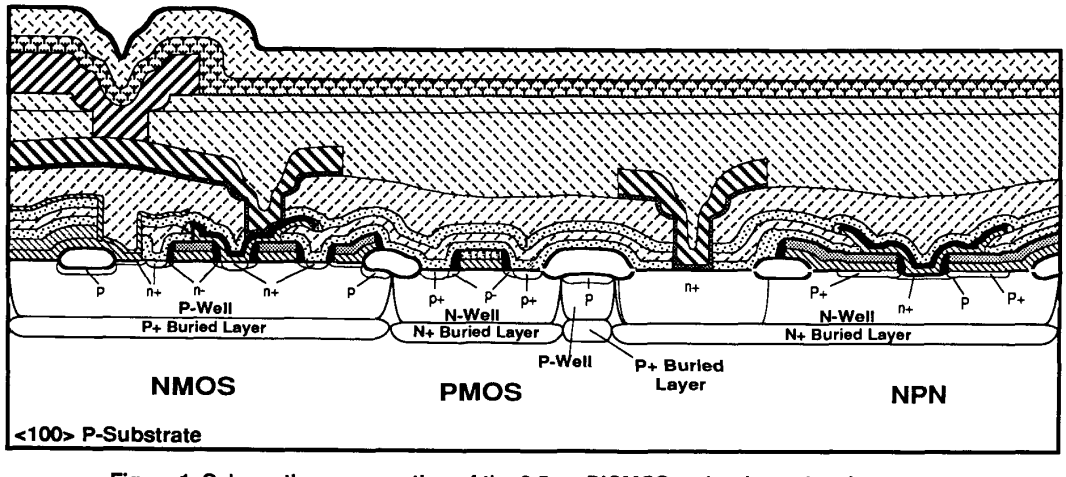

Figure 1 Schematic cross-section of the 0.5µm BiCMOS technology showing the double polysilicon, fully self-aligned bipolar transistor; NMOS transistor with self-aligned contact landing pad; and PMOS transistor.

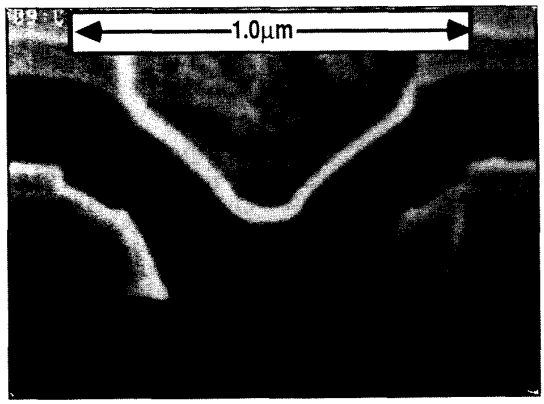

Figure 2 SEM Cross-section of double polysilicon, fully self-aligned bipolar transistor with a 0.4X1.2μm emitter.

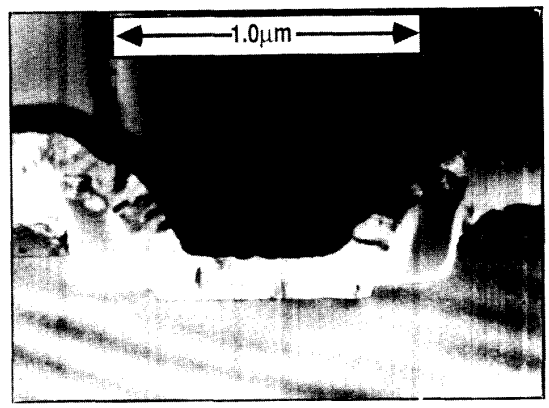

Figure 3 TEM Cross-section of double polysilicon, fully self-aligned bipolar transistor with a 0.8X2.4μm emitter.

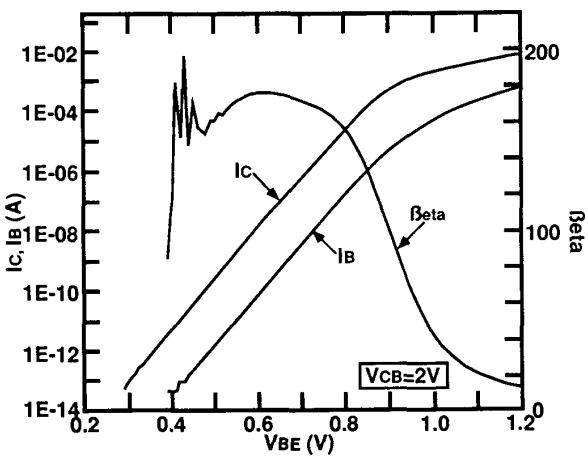

Figure 4 Gummel Plot of double polysilicon, fully self-aligned bipolar transistor with 0.4X1.2μm emitter.

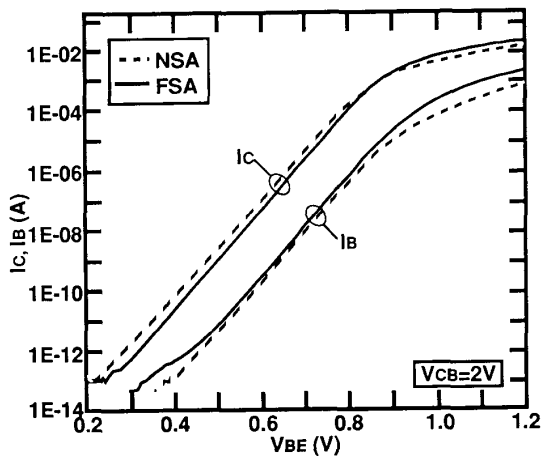

Figure 5 Gummel Plots for fully self-aligned (FSA) and non self-aligned (NSA) bipolar transistors with 0.8X2.4μm emitters.

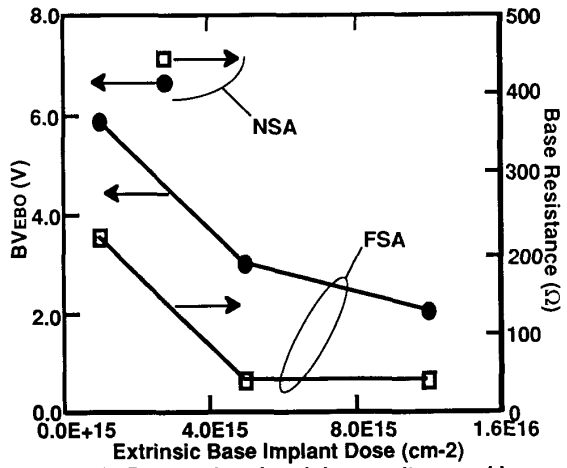

Figure 6 Base-emitter breakdown voltage and base resistance as a function of extrinsic base implant dose for bipolar transistors with 0.8X2.4μm emitters.

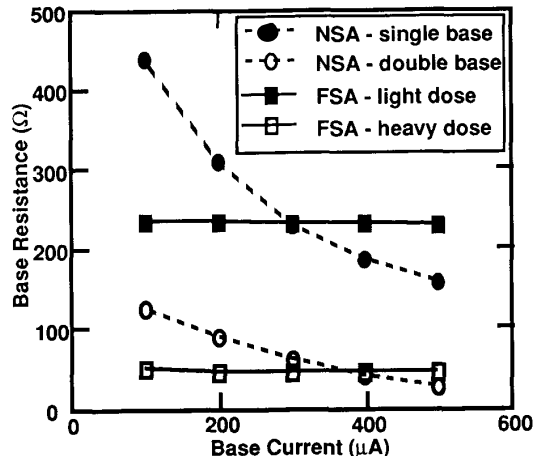

Figure 7 Base resistance as a function of base current for NSA bipolars with single and double base contacts and FSA bipolars with light and heavy extrinsic base implant doses.

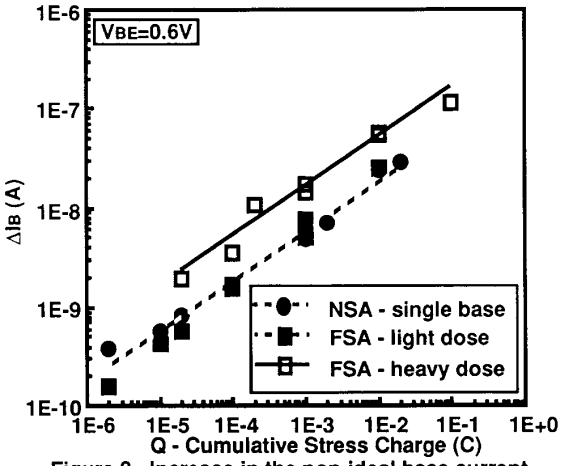

Figure 8 Increase in the non-ideal base current component as a function of cumulative stress charge during emitter-base reverse bias stress for FSA and NSA transistors with 0.8X2.4μm emitters.

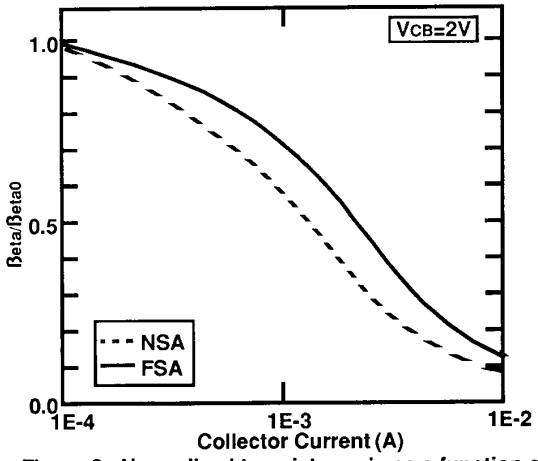

Figure 9 Normalized transistor gain as a function of collector current for FSA and NSA bipolar transistors with 0.8X2.4μm emitters.

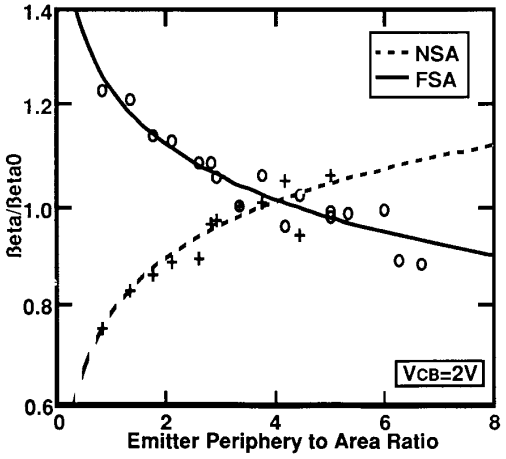

Figure 10 Normalized transistor gain as a function of emitter periphery to area ratio for FSA and NSA bipolar transistors.

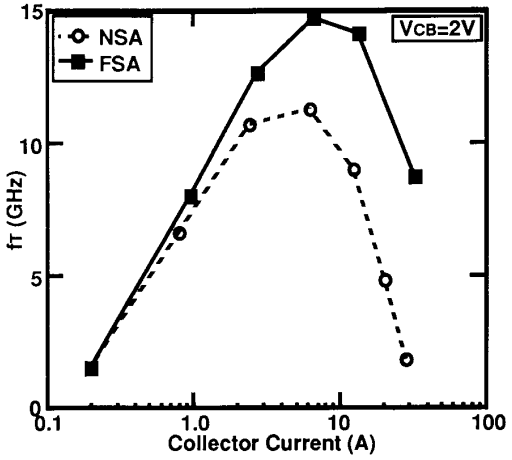

Figure 11 Cut-off frequency as a function of collector current for FSA and NSA bipolar transistors with 0.8X20μm emitters.

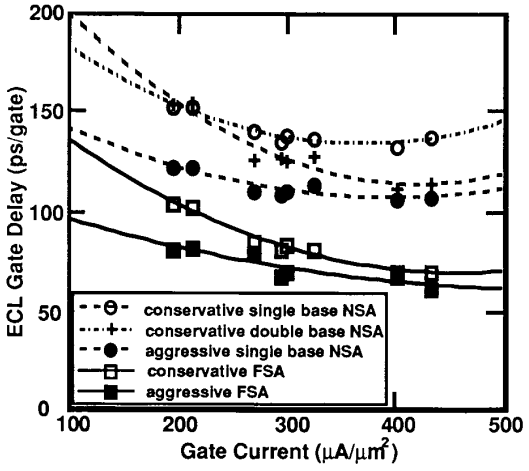

Figure 12 ECL gate delay as a function of gate current for NSA and FSA bipolar transistors with conservative and aggressive design rules. The FSA bipolar has a 2E15 extrinsic base dose.

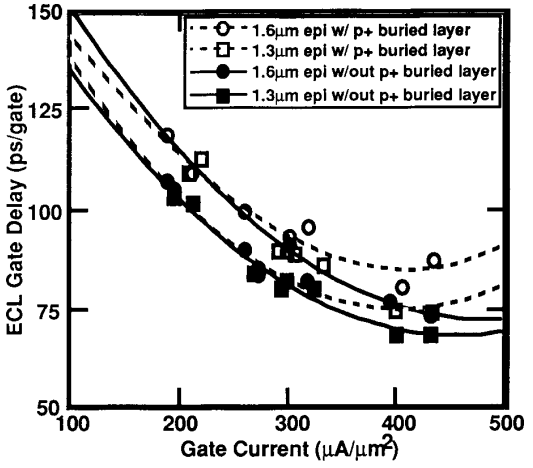

Figure 13 ECL gate delay as a function of gate current for fully self-aligned bipolar transistors with 1.3 versus 1.6μm epitaxial layer thickness and with and without p+ buried layer. The FSA bipolar has a 2E15 extrinsic base dose.