A Versatile 50ps BICMOS Technology for Mixed CMOS/ECL ICs

H. Klose, B. Hoffmann, M. Kerber, K. Ziemann, I. Kerner, R. Schreiter, R. Koepppl
Siemens AG, Corporate Research and Development, Munich, FRG

H. Kaiser, K. Mueller
Siemens AG, Components Group, Munich, FRG

A high performance BICMOS process denoted B5C for the realization of high speed/high density integrated circuits is presented. The main features of this process are 1.0 micron minimal feature size and a selfaligned bipolar transistor of 10 GHz cut-off frequency and 50 ps minimal CML stage-delay time. Moreover bipolar transistors with 70V Early-voltage are available using one additional masking step. As a more complex performance-demonstrator a high speed 16k SRAM with 3.5ns address access time is presented. Both delay times show, that B5C is one of the fastest BICMOS processes worldwide.

1. Introduction

In the past the basic differences between MOS- and BIPOLAR devices lead to two different processing worlds with separate equipment and own factories. By the implementation of CMOS-specific process modules like selfalignment into BIPOLAR processes or by using epitaxy within CMOS flowcharts, an increased synergy arose between both process-areas. This was the precondition for merging these technologies to form both CMOS as well as BIPOLAR devices on the same chip. Thus the circuit-designer is served on the same chip with devices to realize high speed signal processing paths by using BIPOLAR and high density/low power circuit parts like SRAM-cores by using CMOS (e.g. [1]).

This paper presents a 1.0 micron BICMOS process with CMOS devices of 0.8 micron drawn gate length and a high performance selfaligned BIPOLAR transistor. The key electrical data is given and the performance features are demonstrated by respective circuits and demonstrators. Focus is put onto the trade-off between analog and digital performance of the bipolar devices.

2. Technology

In Fig.1 the schematic flowchart of B5C is sketched. The additional process-modules which had to be added to the underlying CMOS process are the buried layer formation, the epi-deposition, the collector plug drive in and the realization of the emitter/base structure. The main technological features of the isolation process block are selfaligned n⁺/p buried layers and wells for a low resistive collector wiring and a tight buried layer pitch at simultaneously low mask-counts and thus low costs. The epi is of a lightly doped n-type material.

![Fig.1 Schematic B5C flowchart](image-url)
For digital applications this epi-layer is swapped by either the n- or the p-well [2]. For the use of the BIPOLAR transistor in analog circuits the n-well implant is blocked by one additional but noncritical masking steps. Thus the sub-collector region is formed by the lightly doped n-epi.

In Fig.2 a comparison is given between the Early-voltages of the "digital" and "analog" BIPOLAR transistors. Despite common models, as they are implemented for instance in network-analysis programs like SPICE, the Early-voltage vs. VCE dependencies are not constant but convex with a maximal Early-voltage at around 3V VCE. The highest Early-voltages achieved are 32V and 85V respectively for the two well types. As the Early-voltage VA reads as

\[ VA = QB \cdot \left( dQB / dVCB \right) = QB / CBC \]

with QB being the neutral base charge, VCB the base to collector voltage and CBC the base-collector junction capacitance, with increasing VCB the Early-voltage rolls off beyond 3V VCE as QB decreases due to pinching of the neutral base. CBC on the other hand remains constant within this voltage regime due to the buried layer punch-on of the base-collector depletion layer. In most analog circuits the VCE value of the bipolar transistor is larger than or equal to 1V. Thus the respective worst case data for the Early-voltage are 20V for the "digital" and 75V for the "analog" BJT (Fig.2).

![Fig.2 Early voltage vs. VCE for the digital and analog BIPOLAR transistor](image)

Fig.2 sketches the trade-off between Early-voltage, drive capability \( j_k \) and gate-delay time \( t_{del} \).

Fig.3 Tradeoff between Early-voltage, drive capability \( j_k \) and gate-delay time \( t_{del} \):

<table>
<thead>
<tr>
<th>device</th>
<th>quantity</th>
<th>unit</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJT</td>
<td>( A_{em} )</td>
<td>( \mu m )</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>( B )</td>
<td></td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>( N_B )</td>
<td></td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>( N_C )</td>
<td></td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>( f_T )</td>
<td>GHz</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>( C_{IE} )</td>
<td>( \mu F )</td>
<td>17.5</td>
</tr>
<tr>
<td></td>
<td>( C_{IC} )</td>
<td>( \mu F )</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td>( C_{IS} )</td>
<td>( \mu F )</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>( V_{EBB} )</td>
<td>V</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>( V_{EBO} )</td>
<td>V</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>( V_{CEO} )</td>
<td>V</td>
<td>6.5</td>
</tr>
<tr>
<td>NMOS</td>
<td>L</td>
<td>( \mu m )</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>( \delta (V_{GS} + V_{DD}) )</td>
<td>( \mu A/\mu m )</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>( V_T )</td>
<td>V</td>
<td>0.74</td>
</tr>
<tr>
<td>PMOS</td>
<td>L</td>
<td>( \mu m )</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>( \delta (V_{GS} + V_{DD}) )</td>
<td>( mA/\mu m )</td>
<td>190</td>
</tr>
</tbody>
</table>

Fig.4 Summary of the NMOS, PMOS and BJT device parameters

654
Thus rising the Early-voltage to a value of 75 V means a speed degradation by a factor of 1.7 and a reduction in drive-capability by a factor of 2.5 (see Fig.3).

Next to the well formation the MOS devices are formed with a drawn gate length of 0.8 microns. For improved hot-electron hardness an LDD implantation scheme is applied for the NMOS transistor. Then the base-emitter structure is realized using a double-polysilicon selfalignment-scheme [3,4,5]. Finally a two layers of metal process module is applied. A third metal interconnection scheme is optional available particularly for complex CML and ECL functions. The respective global device data is summarized in Fig.4 for the BJT, the NMOS and the PMOS transistor. The cross-section of the active BSC devices obtained by SEM analysis is given in Fig.5.

3. Basic Circuit-Evaluation

The basic performance data were determined via ringoscillators with CMOS, BICMOS (totempole) and pure BIPOLAR stages. The gate delay times of the latter stages are given in Fig.6. For a 1.0*3.0 um² drawn emitter size the minimal gate delay time is 65 ps at 600 uA/gate. The low power performance for the same transistor layout is 200 ps at 30 uA/gate. For higher speeds larger transistors have to be used e.g. with 0.8*20 um² emitter-size. Due to the higher drive-capability and the reduction of base-resistance due to the narrower emitter stripe of 0.8 micron - resulting in an effective emitter width of 0.4 micron - a further improvement of the gate delay time as low as 50 ps is achieved. The delay-time of mixed 3-input NAND/NOR gates realized in pure CMOS is 360 ps.

In Fig.7 the load lines of equally sized CMOS, BICMOS totempoles and pure BIPOLAR CML driving stages are shown. The break-even point for the use of BICMOS totempoles instead of single stage CMOS drivers is approximately 0.5 pF. With respect to the load-dependence the technologies CMOS:BICMOS:BIPOLAR relates as 6:3:1. This means that for high capacitive loading:

the BICMOS totempole is a factor of 2 faster than the pure CMOS stage

and that the BIPOLAR stage is a factor of 6 faster than CMOS and a factor of 3 faster than a BICMOS totempole.
In the case of embedded CML within a CMOS environment of course the delay-times to convert CML levels to CMOS voltage swings has to be taken into account. This approximately takes 1ns.

\[ t_d = \frac{C_L}{I} \]

Fig.7 Loadlines of CMOS, BICMOS and BIPOLAR driving stages

4. Demonstrators

To demonstrate the performance of BSC, in addition to ringoscillator data - which gives only a first impression of the performance of a technology - larger demonstrators have been designed, realized and evaluated. One example is a 16k SRAM [6] with CMOS memory-core and ECL peripherals. The respective chip-microfotograph of the SRAM is given in Fig.8.

Fig.8 Chip - micrograph of the 16k SRAM

Fig.9 Input and output wave-forms of the 16k SRAM

The measured input/output waveform for a read operation is depicted in Fig.9, where a 3.8ns readout could be obtained. Thus, due to the combination of the fast BSC devices together with a well-balanced CMOS/ECL circuit-mix, by far the fastest address access time for a 1.0 micron technology was obtained.

5. Conclusions

The presentation described a versatile BICMOS process for high speed/high density applications. Gate delay times of 50ps for CML stages and address access times of 3.8ns for a 16k SRAM show that by using BSC the whole performance regime both of high speed BIPOLAR and high density CMOS can be covered on one single chip. It was demonstrated experimentally that by using one extra masking step good analog performance is achievable even in a high performance BICMOS environment.

References:


