BICMOS Technology and Fabrication
BiCMOS Technology

Combines Bipolar and CMOS transistors in a single integrated circuit

By retaining benefits of bipolar and CMOS, BiCMOS is able to achieve VLSI circuits with speed-power-density performance previously unattainable with either technology individually.
Characteristics of CMOS Technology

Lower static power dissipation
Higher noise margins
Higher packing density – lower manufacturing cost per device
High yield with large integrated complex functions

Advantages of CMOS over bipolar

High input impedance (low drive current)
Scaleable threshold voltage
High delay sensitivity to load (fan-out limitations)
Low output drive current (issue when driving large capacitive loads)
Low transconductance, where transconductance, $g_m \alpha V_{in}$
Bi-directional capability (drain & source are interchangeable)
A near ideal switching device

Other CMOS Advantages
Characteristics of Bipolar Technology

Advantages of Bipolar over CMOS

Higher switching speed
Higher current drive per unit area, higher gain
Generally better noise performance and better high frequency characteristics
Better analogue capability
Improved I/O speed (particularly significant with the growing importance of package limitations in high speed systems).

Other Bipolar Advantages

high power dissipation
lower input impedance (high drive current)
low voltage swing logic
low packing density
low delay sensitivity to load
high $g_m$ ($g_m \alpha V_{in}$)
high unity gain band width ($f_t$) at low currents
essentially unidirectional
Combine advantages in BiCMOS Technology

- It follows that BiCMOS technology goes some way towards combining the virtues of both CMOS and Bipolar technologies.

- Design uses CMOS gates along with bipolar totem-pole stage where driving of high capacitance loads is required.

- Resulting benefits of BiCMOS technology over solely CMOS or solely bipolar:
  - Improved speed over purely-CMOS technology.
  - Lower power dissipation than purely-bipolar technology (simplifying packaging and board requirements).
  - Flexible I/Os (i.e., TTL, CMOS or ECL) – BiCMOS technology is well suited for I/O intensive applications. ECL, TTL and CMOS input and output levels can easily be generated with no speed or tracking consequences.
  - high performance analogue
  - Latchup immunity (Discussed later in course)
The *simplified* BiCMOS Inverter

Two bipolar transistors (T3 and T4), one nMOS and one pMOS transistor (both enhancement-type devices, OFF at $V_{in}=0V$)

The MOS switches perform the logic function & bipolar transistors drive output loads

**Vin = 0 :**

- T1 is off. Therefore T3 is non-conducting
- T2 ON - supplies current to base of T4
- T4 base voltage set to $V_{dd}$
- T4 conducts & acts as current source to charge load $C_L$ towards $V_{dd}$
- $V_{out}$ rises to $V_{dd} - V_{be \ (of \ T4)}$
- Note: $V_{be \ (of \ T4)}$ is base-emitter voltage of T4.

(pullup bipolar transistor turns off as the output approaches $5V - V_{be \ (of \ T4)}$)

**Vin = $V_{dd}$ :**

- T2 is off. Therefore T4 is non-conducting.
- T1 is on and supplies current to the base of T3
- T3 conducts & acts as a current sink to discharge load $C_L$ towards 0V.
- $V_{out}$ falls to $0V + V_{CE_{sat \ (of \ T3)}}$
- Note: $V_{CE_{sat \ (of \ T3)}}$ is saturation V from T3 collector to emitter
The *simplified* BiCMOS Inverter

- T3 & T4 present low impedances when turned on into saturation & load $C_L$ will be charged or discharged rapidly
- Output logic levels will be good & will be close to rail voltages since $V_{CE_{sat}}$ is quite small & $V_{BE} \approx 0.7V$. Therefore, inverter has high noise margins
- Inverter has high input impedance, i.e., MOS gate input
- Inverter has low output impedance
- Inverter has high drive capability but occupies a relatively small area
- However, this is not a good arrangement to implement since no discharge path exists for current from the base of either bipolar transistor when it is being turned off, i.e.,
  - when $V_{in} = V_{dd}$, T2 is off and no conducting path to the base of T4 exists
  - when $V_{in} = 0$, T1 is off and no conducting path to the base of T3 exists

This will slow down the action of the circuit
The *conventional* BiCMOS Inverter

Two additional enhancement-type nMOS devices have been added (T5 and T6).
These transistors provide discharge paths for transistor base currents during turn-off.
Without T5, the output low voltage cannot fall below the base to emitter voltage $V_{BE}$ of T3.

**Vin = 0:**

T1 is off. Therefore T3 is non-conducting
T2 ON - supplies current to base of T4
T4 base voltage set to $V_{dd}$.
T5 is turned on & clamps base of T3 to GND. T3 is turned off.
T4 conducts & acts as current source to charge load $C_L$ towards $V_{dd}$.
$V_{out}$ rises to $V_{dd} - V_{be}$ (of T4)

**Vin = $V_{dd}$:**

T2 is off
T1 is on and supplies current to the base of T3
T6 is turned on and clamps the base of T4 to GND. T4 is turned off.
T3 conducts & acts as a current sink to discharge load $C_L$ towards 0V
$V_{out}$ falls to $0V + V_{CE_{sat}}$ (of T3)

BiCMOS Technology
The *conventional* BiCMOS Inverter

Again, this BiCMOS gate does not swing rail to rail. Hence some finite power is dissipated when driving another CMOS or BiCMOS gate. The leakage component of power dissipation can be reduced by varying the BiCMOS device parameters.
More advanced BiCMOS structures

Various types of BiCMOS gates have been devised to overcome the shortcomings of the conventional BiCMOS gate.

BiCMOS devices are available which provide the full Vdd -> GND voltage swing.

There is a common theme underlying all BiCMOS gates:
- all have a common basic structure of a MOSFET (p or n) driving a bipolar transistor (nnp or pnp) which drives the output.

BiCMOS can provide applications with CMOS power & densities at speeds which were previously the exclusive domain of bipolar. This has been demonstrated in applications ranging from static RAMs to gate arrays to u-processors.

BiCMOS fills the market niche between
- very high speed, but power hungry bipolar ECL (Emitter Coupled Logic)
  and
- very high density, medium speed CMOS.
More advanced BiCMOS structures

When the power budget is unconstrained, a bipolar technology optimised for speed will almost always be faster than BiCMOS and will most likely be selected.

However, when a finite power budget exists, the ability to focus power where it is required usually allows BiCMOS speed performance to surpass that of bipolar

The concept of ‘system on a chip’ becomes a reality with BiCMOS.

Most gates in ROM, ALU, register subsystems etc do not have to drive large capacitive loads. Hence the use of BiCMOS technology would give no speed advantage.

To take maximum advantage of available silicon technologies, the following mix of technologies in a silicon system might be used:

- **CMOS** for logic
- **BiCMOS** for I/O and driver circuits
- **ECL** for critical high speed parts of the system.
Comparison of logic families

e.g., 74BCT have similar speeds to 74F but with greatly reduced power consumption

<table>
<thead>
<tr>
<th>Device</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>74</td>
<td>Standard TTL</td>
</tr>
<tr>
<td>74S</td>
<td>Schottky clamped TTL – transistors do not enter saturation</td>
</tr>
<tr>
<td>74LS</td>
<td>Low power Schottky – as 74S but larger resistor values</td>
</tr>
<tr>
<td>74AS</td>
<td>Advanced Schottky – same as 74S but improved processing</td>
</tr>
<tr>
<td>74ALS</td>
<td>Advanced low power Schottky – low power version of 74AS</td>
</tr>
<tr>
<td>74F</td>
<td>Fast – compromise between S and ALS</td>
</tr>
<tr>
<td>74C</td>
<td>Standard CMOS – first CMOS parts in TTL pinout</td>
</tr>
<tr>
<td>74HC</td>
<td>High speed CMOS – improved CMOS</td>
</tr>
<tr>
<td>74HCT</td>
<td>High speed CMOS with TTL i/p voltage levels</td>
</tr>
<tr>
<td>74AC</td>
<td>Advanced high speed CMOS (1.5 µm CMOS)</td>
</tr>
<tr>
<td>74ACT</td>
<td>Advanced high speed CMOS with TTL i/p voltage levels</td>
</tr>
<tr>
<td>74(A)BCT</td>
<td>High speed BiCMOS for line drivers</td>
</tr>
<tr>
<td>74LVC</td>
<td>Low voltage (2.7–3.6 V) 1 µm CMOS</td>
</tr>
<tr>
<td>74LV</td>
<td>Low voltage (2.7–3.6 V) 2 µm CMOS</td>
</tr>
<tr>
<td>74LVT</td>
<td>Low voltage BiCMOS (optional 5 V inputs, 3 V outputs)</td>
</tr>
<tr>
<td>74ALVC</td>
<td>Advanced low voltage 1 µm CMOS</td>
</tr>
<tr>
<td>4000B</td>
<td>Early CMOS, not TTL pin compatible, 5–12 V supply</td>
</tr>
<tr>
<td>F100K</td>
<td>100K ECL series – very fast but poor noise margins</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technology</th>
<th>Delay(ns)</th>
<th>Pstatic</th>
<th>Vohmin Volmax @Iomax</th>
<th>Vihmin Vilmax</th>
<th>Ihmax Ilmax</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTL</td>
<td>10</td>
<td>10 mW</td>
<td>2.4/0.4</td>
<td>2/0.8</td>
<td>40 µA/1.6 mA</td>
</tr>
<tr>
<td>TTL</td>
<td>3</td>
<td>20 mW</td>
<td>2.7/0.5</td>
<td>2/0.8</td>
<td>50 µA/2 mA</td>
</tr>
<tr>
<td>TTL</td>
<td>10</td>
<td>2 mW</td>
<td>2.7/0.5</td>
<td>2/0.8</td>
<td>20 µA/0.4 mA</td>
</tr>
<tr>
<td>TTL</td>
<td>2</td>
<td>8 mW</td>
<td>2.7/0.5</td>
<td>2/0.8</td>
<td>20 µA/0.5 mA</td>
</tr>
<tr>
<td>TTL</td>
<td>4</td>
<td>1 mW</td>
<td>2.7/0.5</td>
<td>2/0.8</td>
<td>20 µA/0.1 mA</td>
</tr>
<tr>
<td>TTL</td>
<td>3</td>
<td>4 mW</td>
<td>2.7/0.5</td>
<td>2/0.8</td>
<td>20 µA/0.6 mA</td>
</tr>
<tr>
<td>CMOS</td>
<td>30</td>
<td>50 µW</td>
<td>4.2/0.4</td>
<td>3.5/1</td>
<td>± 2 µA</td>
</tr>
<tr>
<td>CMOS</td>
<td>9</td>
<td>25 µW</td>
<td>4.3/0.33</td>
<td>3.5/1</td>
<td>± 0.1 µA</td>
</tr>
<tr>
<td>CMOS</td>
<td>10</td>
<td>25 µW</td>
<td>4.3/0.33</td>
<td>2/0.8</td>
<td>± 0.1 µA</td>
</tr>
<tr>
<td>CMOS</td>
<td>4</td>
<td>25 µW</td>
<td>4.3/0.44</td>
<td>3.5/1.5</td>
<td>± 0.1 µA</td>
</tr>
<tr>
<td>CMOS</td>
<td>6</td>
<td>25 µW</td>
<td>4.3/0.44</td>
<td>2/0.8</td>
<td>± 0.1 µA</td>
</tr>
<tr>
<td>BiCMOS</td>
<td>3.5</td>
<td>600 µW</td>
<td>2/0.55</td>
<td>2/0.8</td>
<td>0.07 mA/0.65 mA</td>
</tr>
<tr>
<td>CMOS</td>
<td>5</td>
<td>50 µW</td>
<td>2/0.55</td>
<td>2/0.8</td>
<td>± 1 µA</td>
</tr>
<tr>
<td>CMOS</td>
<td>9</td>
<td>50 µW</td>
<td>2.4/0.4</td>
<td>2/0.8</td>
<td>± 1 µA</td>
</tr>
<tr>
<td>BiCMOS</td>
<td>4</td>
<td>400 µW</td>
<td>2/0.5</td>
<td>2/0.8</td>
<td>± 1 µA</td>
</tr>
<tr>
<td>CMOS</td>
<td>3</td>
<td>50 µW</td>
<td>2/0.55</td>
<td>2/0.8</td>
<td>± 5 µA</td>
</tr>
<tr>
<td>CMOS</td>
<td>75</td>
<td>50 µW</td>
<td>2.5/0.4</td>
<td>3.5/1.5</td>
<td>± 0.1 µA</td>
</tr>
<tr>
<td>ECL</td>
<td>0.75</td>
<td>20 mW</td>
<td>-0.9/-1.7</td>
<td>-1.2/-1.4</td>
<td>240 µA, 0.5 µA</td>
</tr>
</tbody>
</table>

BiCMOS Technology
Further advantages of BiCMOS Technology

Analogue amplifier design is facilitated and improved

High impedance CMOS transistors may be used for the input circuitry while the remaining stages and output drivers are realised using bipolar transistors

In general, BiCMOS devices offer many advantages where high load current sinking and sourcing is required. The high current gain of the NPN transistor greatly improves the output drive capability of a conventional CMOS device.

MOS speed depends on device parameters such as saturation current and capacitance. These in turn depend on oxide thickness, substrate doping and channel length.

Compared to CMOS, BiCMOS’s reduced dependence on capacitive load and the multiple circuit and I/Os configurations possible greatly enhance design flexibility and can lead to reduced cycle time (i.e., faster circuits).

[Peak bipolar speed is less dependent on circuit capacitance. Device parameters $f_t$, $J_k$ and $R_b$ determine Bipolar circuit speed performance (not covered here) and depend on process parameters such as base width, epitaxial layer profile, emitter width and extrinsic base formation]
Further advantages of BiCMOS Technology.

BiCMOS is inherently robust with respect to temperature and process variations, resulting in less variability in final electrical parameters, resulting in higher yield, an important economic consideration.

Large circuits can impose severe performance penalties due to simultaneously switching noise, internal clock skews and high nodal capacitances in critical paths - BiCMOS has demonstrated superiority over CMOS in all of these factors.

BiCMOS can take advantage of any advances in CMOS and/or bipolar technology, greatly accelerating the learning curve normally associated with new technologies.
Are there disadvantages with BiCMOS technology?

Main disadvantage: greater process complexity compared to CMOS

Results in a 1.25 -> 1.4 times increase in die costs over conventional CMOS. Taking into account packaging costs, the total manufacturing costs of supplying a BiCMOS chip ranges from 1.1-> 1.3 times that of CMOS.

However, as CMOS complexity has increased, the percentage difference between CMOS and BiCMOS mask steps has decreased. Therefore, just as power dissipation constraints motivated the switch from nMOS to CMOS in the late ‘70s, performance requirements motivated a switch from CMOS to BiCMOS in the late ‘80s for VLSI products requiring the highest speed levels.

Capital costs of investing in continually smaller (<1um) CMOS technology rises exponentially, while the requirement of low power supplies for sub-0.5um CMOS results in degradation of performance. Since BiCMOS does not have to be scaled as aggressively as CMOS, existing fabs can be utilised resulting in lower capital costs. Extra costs incurred in developing a BiCMOS technology is more than offset by the fact that the enhanced chip performance obtained extends the usefulness of manufacturing equipment & clean rooms by at least one technology generation.
BiCMOS - Brief Historical Perspective

Most early BiCMOS applications were analogue; BiCMOS operational amplifiers were introduced in the mid-70s followed by BiCMOS power ICs.

Digital LSI BiCMOS devices were introduced in the mid-80s, motivated by high power dissipation of bipolar circuits, speed limitations of MOS circuits & a need for high I/O throughput.
Development of VLSI BiCMOS resulted in very high performance memories, gate arrays & micro-processors

BiCMOS follows the same scaling curve as mainstream CMOS technology resulting in explosive growth in BiCMOS product growth.

BiCMOS has been established as the technology of choice for high speed VLSI.
BiCMOS Fabrication

Theoretically there should be little difficulty in extending CMOS fab processes to include bipolar as well as MOS transistors.

In fact, a problem of p-well and n-well CMOS processing is that parasitic bipolar transistors are inadvertently formed as part of the outcome of fabrication (see section on CMOS latchup).

Production of npn bipolar transistors with good performance characteristics can be achieved, e.g., by extending the standard n-well CMOS processing to include further masks to add two additional layers; the n+ subcollector and p+ base layers.

The npn transistor is formed an n-well & the additional p+ base region is located in the well to form the p-base region of the transistor. The second additional layer, the buried n+ subcollector (BCCD) is added to reduce the n-well (collector) resistance & thus improve the quality of the bipolar transistor.
Arrangement of BiCMOS npn transistor (orbit 2um CMOS) for reference