A versatile 0.25 micron CMOS technology

Steve Poon, Cindy Atwell, Chuck Hart, Dave Kolar, Craig Lage, Bob Yeargain
Semiconductor Products Sector, Motorola Inc.,
3501 Ed Bluestein Blvd., MD: K-31, Austin, Texas 78721
Ph: (512) 933-2450, Fax: (512) 933-2476, email: rwvc50@email.sps.mot.com

Abstract

Manufacturing issues on 0.25 micron CMOS technologies including devices for high performance processor, dense SRAM, and low power circuits are discussed. Tradeoffs between performance, cost, and statistical effects such as gate length and Vt control are presented. Requirements for future technologies are identified to improve maturity prior to implementation.

Introduction

Growth is dependent on providing increased value and development of enabling technologies. A fundamental aspect of this relies on the ability to manufacture using advanced technologies at acceptable cost for a variety of circuits. This paper describes the implementation of a 0.25 micron CMOS technology used in the manufacturing of high performance processors, dense SRAMs, and low power circuits. Features included are trench isolation, retrograde wells, cobalt salicide, in-laid tungsten local interconnect, and up to six layers of planarized metal [1]. Critical layout pitch of this technology are summarized in Table 1.

Table 1 - Critical pitch of the technology

Isolation	0.6 µm
n+/p+ space	0.6 µm
Gate poly	0.6 µm
Local interconnect	0.6 µm
Metal 1	0.7 µm
SRAM bitcell	5.3 µm2

Core Process

A 0.35 micron deep trench with sloped sidewall, thin thermal oxide liner, and filled with deposited ozone TEOS oxide is used to minimize seam and void formation in narrow isolation space. Planarization was developed with a combination of oxide RIE and CMP using a two-mask process and nitride as the polish stop layer to support arbitrary width trench. A one-mask process which replaced the two-mask process is implemented on this technology to reduce cost by eliminating a photo layer and an etch step. This is achieved by placing 'dummy active' in wide field regions to compensate for CMP dishing caused by pattern density variations. A comparison of the two approaches is shown in Figure 1.

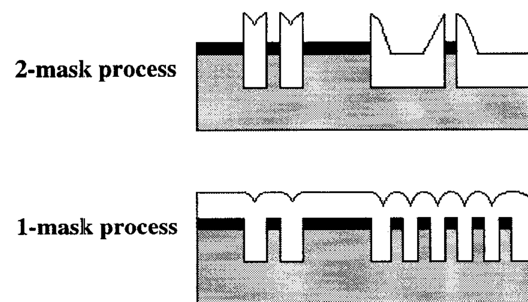

Figure 1 - Two-mask and one-mask trench

Minimum n+ to p+ spacing of 0.6 micron is achieved by optimization of resist profile [2], implant doses, and Dt control by ensuring implant straggle and lateral diffusion of dopants do not result in shorts between junctions and adjacent wells (Figure 2).

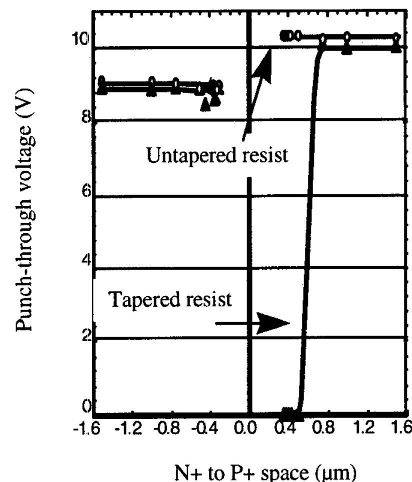

Figure 2 - Inter-well isolation

Cobalt salicide is used in place of titanium salicide to achieve low resistance at short gate length (Figure 3) and improve scaling while eliminating a need for pre-amorphorization

0–7803–4774–9/98/$10.00 © 1998 IEEE

implant (PAI) used in the titanium salicide process. Salicide formation is performed by PVD of cobalt and a capping layer, low temperature RTA, a wet strip to remove the capping layer and unreacted cobalt, and a high temperature RTA to complete the phase conversion to achieve a thermally stable, low resistance salicide.

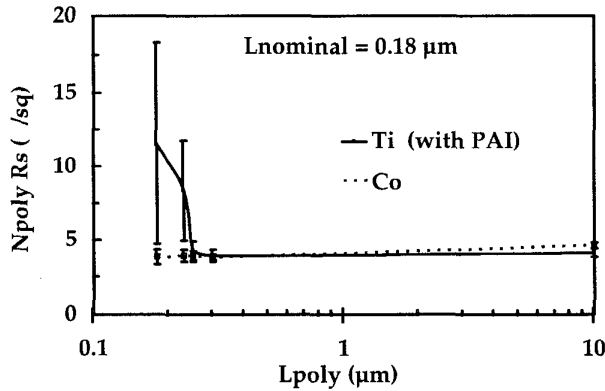

Figure 3 - Sheet resistance of Ti- and Co-salicide

In-laid tungsten local interconnect is implemented with 2-D OPC to correct for line-end pullback that enabled layout of dense SRAM bitcell and standard cells, as well as improving interconnect routing prior to aluminum metallization [3]. Planarized metal using oxide and tungsten CMP with up to six layers of aluminum-alloy complete the core process.

Modules

Multiple Vts are obtained with adjustments to implant doses and gate lengths to achieve high current drive, meet low standby current requirements, and provide static noise margin.

Assessment [4] that combined intrinsic reliability study and simulated burn-in that stressed transistor arrays as capacitors were used to determine dielectric integrity for 35Å to 70Å thermally grown oxides used in this technology to meet core and I/O reliability requirements. Vmax that meets a 1 FIT criteria for 10 years of operation for the different oxide thickness are summarized in Table 2.

Table 2 - Vmax @ 105°C

Tox	Vmax
35Å	2.4V
50Å	3.4V
70Å	4.4V

Self-aligned local interconnect (Figure 4) is developed with etch stop layers to achieve a 5.3 micron square bitcell (Figure 5) for high density embedded and stand-alone SRAM circuits.

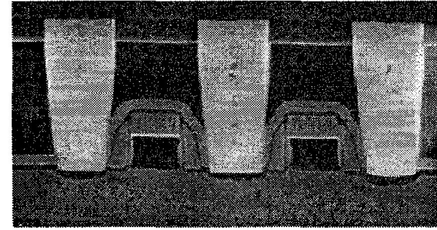

Figure 4 - SEM of self-aligned local interconnect

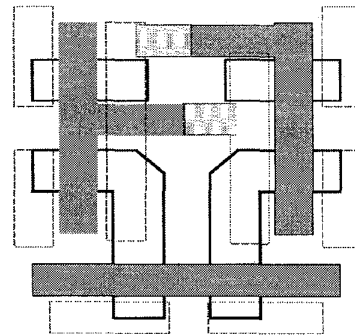

Figure 5 - SRAM bitcell

Mixed signal application is enabled (Figure 6) with a masked implant to the substrate. Other process modules such as dual gate oxide and NVM can also be integrated into the technology.

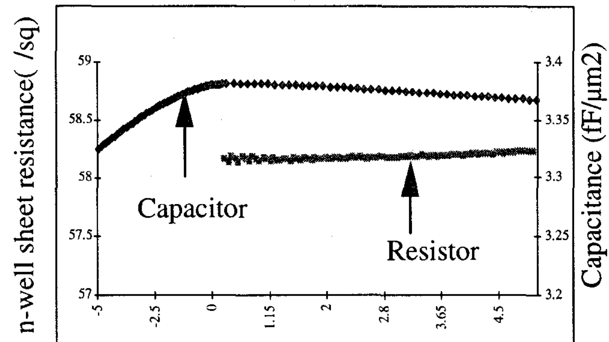

Figure 6 - Voltage characteristics of Resistor and Capacitor

Manufacturing Issues

Manufacturing issues involving tradeoffs between performance and complexity must be comprehended to achieve cost effective solutions.

Current drive and AC performance required for high performance processors is achieved by operating at short Lgate using devices with good short channel characteristics.

These devices are fabricated using steep retrograde channels, 35Å gate oxide, shallow extensions/halos, and relatively deep source/drains [5] at the expense of adding masking steps for the extension and halo implants.

50Å gate oxide needed for 2.5V Vdd operation is integrated with longer Lgate and higher Vts for circuits that require low standby current. Hard source/drains are also used in these devices to reduce junction resistance. A comparison between the high performance and low leakage devices is shown in Table 3.

Table 3 - Device Characteristics

Parameter	High Performance	Low Leakage
Vdd	1.8V	2.5V
Tox	35Å	50Å
Lgate	0.15μm	0.23μm
Vt,n	0.35V	0.60V
Ids, n	750μA/μm	550μA/μm
Ioff,n	1nA/μm	1pA/μm

Static noise margin (Fig. 7) and fast access time for SRAMs at 1.8 volt Vdd is achieved with by using different Vts for bitcell and peripheral devices.

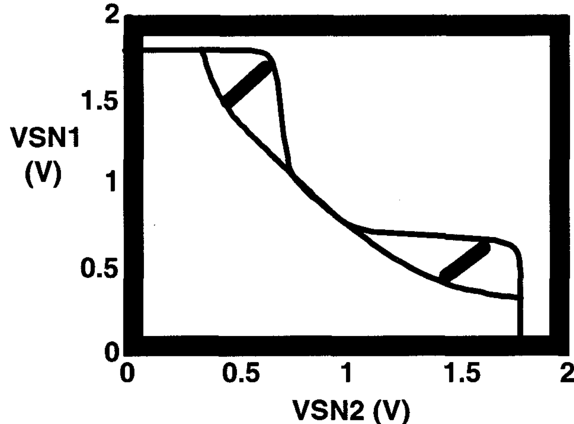

Figure 7 - Static noise margin at 1.8 volt Vdd

The Vts are also adjusted to suite applications that require higher performance or lower standby current. Vt targets for an application that requires low standby current is shown in Table 4.

Table 4 - SRAM Vt Targets

Vt,n array: 0.68V
Vt,p array: -1.05V
Vt,n peripheral: 0.59V
Vt,p peripheral: -0.57V

Gate length control is critical to high performance processors where speed requirements continue to outpace the ability of process tools to deliver robust outputs without special controls. This issue is especially critical in manufacturing where tools that are not perfectly matched must be used for volume production. A control methodology that relied heavily on feedback based on in-line measurements was developed [6] to control gate length and speed distribution on circuits where a small differences in speed can result in either revenue or scrap. A representation of the process used to limit gate length variations is shown in Figure 8. The impact of this controller on variations (one sigma) of Lpoly, Leff and circuit speed (Fmax) is shown in Table 5.

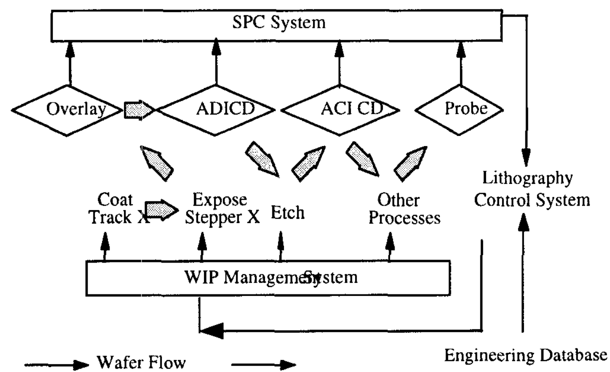

Figure 8 - Automated CD controller

Table 5 - Impact of controller (1 sigma)

	Lpoly	Leff	Fmax
Before	11.1nm	16.0nm	15.9MHz
After	7.1nm	9.4nm	8.4MHz

Statistical variations in Vt [7] on devices with small channel area due to variability in the finite number of dopant atoms and process induced effects such as misalignment and gate length variation can impact stability in SRAMs and lead to variations in drive current and delay in logic circuits. Improvements by increasing dopant atoms in the channel must be considered at the expense of bitcell area and current drive. Typical one sigma variation in Vt as a function of channel area for this technology is shown in Table 6.

Table 6 - Vt,n variation as function of channel area

WxL (μm2)	Vt,n (mV)
20x20	2.8
20x0.275	6.7
20x0.25	9.5
0.5x20	10.6
0.325x20	3.2
0.325x0.25	20.8
0.3x.225	26.1

Future Requirements

As we transition to more advanced technologies beyond 0.25 micron, a new set of requirements emerge. Device designs that can accommodate higher levels of integration with acceptable cost to provide additional functionality on the same chip such as high performance processors integrated with dense level 2 cache and analog functions must be provided to improve system performance and put more value within the chip. However, same devices must be sufficiently modular such that components not required for certain applications can be removed to avoid unnecessary cost. This is needed to keep the same libraries and processes to leverage design reuse and yield learning that can only be achieved with volume.

Cu-based metallization required to meet electromigration requirements for high current density, high temperature applications must be developed and implemented with reduced complexity and lower system cost compared to traditional aluminum-alloy based processes.

Smarter optical proximity correction and phase shift mask [8] are required to stretch and extend the capability of lithography to preserve capital investments made on 248nm tools. Significant enhancements to the quality of CMP tools and the quality and cost of CMP consumables are necessary to improve productivity.

These manufacturing technologies must be developed and implemented at a cost to satisfy shareholders, keep our jobs, and to fund technology development needed in the next millennium.

Acknowledgments

We acknowledge the dedication, ingenuity, and hard work of our colleagues in what is known as the Y-building team. Encouragement and support from our management is sincerely appreciated.

References

1. Tsui et. al., 1998 VLSI Tech Symp
2. Woo et. al., 1998 VLSI Tech Symp
3. Chuang et. al., 1997 IEDM
4. Anderson et. al., 1997 ECS, vol. 97-31
5. Venkatesan et. al., 1997 IEDM
6. Sturtevant et. al., 1998 SPIE
7. Burnett et. al., 1994 VLSI Tech Symp
8. Roman et. al., 1997 BACUS, #3236-31