
White Paper On Silicon On Insulator (SOI) Implementation

June 2009

**Author: Narayana Murty Kodeti
Infotech Enterprises Ltd.,**



Table of Contents

1. INTRODUCTION TO SOI TECHNOLOGY	3
1.1 SOI BASICS	3
1.2 SOI FLOATING BODY	4
1.3 FULLY DEPLETED SOI	4
1.4 PARTIALLY DEPLETED SOI	4
1.5 HISTORY EFFECT	5
1.6 Stack Height Effect.....	5
1.7 SUBSTRATE NOISE	6
1.8 LATCHUP ELIMINATION	7
1.9 BI-POLAR CURRENTS	7
1.10 SELF-HEATING.....	8
1.11 TEMPERATURE SENSITIVITY.....	8
1.12 BODY CONTACTS	8
2. ASIC IMPLEMENTATION WITH SOI	8
2.1 MODELING SOI EFFECTS FOR THE DIGITAL IMPLEMENTATION	8
2.2 ASIC TIMING ANALYSIS AND CLOSURE	9
2.3 ASIC IMPLEMENTATION STRATEGY	11
3. SOI SUMMARY	12

Table of Figures

Figure 1 Bulk NMOS Transistor vs SOI NMOS Transistor	3
Figure 2 History Effect	5
Figure 3 Substrate Noise in Bulk	6
Figure 4 Substrate Noise in SOI	6
Figure 5 Bipolar current of the SOI FET	7
Figure 6 Simple Timing Path	9
Figure 7 Complex Clock Timing Path	10

1. Introduction to SOI Technology

Increased demand for High Performance, Low Power and Low Area among microelectronic devices is continuously pushing the fabrication process to go beyond ultra deep sub-micron (UDSM) technologies such as 45nm, 32nm and so on. Currently, chips are being designed in 55nm, 45nm and 32nm process nodes. The performance and power goals for certain applications in these advanced nodes couldn't be achieved with conventional silicon (bulk CMOS) process leading to an alternative, Silicon On Insulator (SOI) process.

Silicon On Insulator fabrication process helps in achieving greater performance and offers less power consumption compared to the Bulk Process.

1.1 SOI Basics

In a Silicon On Insulator (SOI) Fabrication technology Transistors are built on a silicon layer resting on an Insulating Layer of Silicon dioxide (SiO_2). The insulating layer is created by flowing oxygen onto a plain silicon wafer and then heating the wafer to oxidize the silicon, thereby creating a uniform buried layer of silicon dioxide. Transistors are encapsulated in SiO_2 on all sides. The blow figure shows a typical NMOS Transistor with Bulk CMOS Process and with SOI Process.

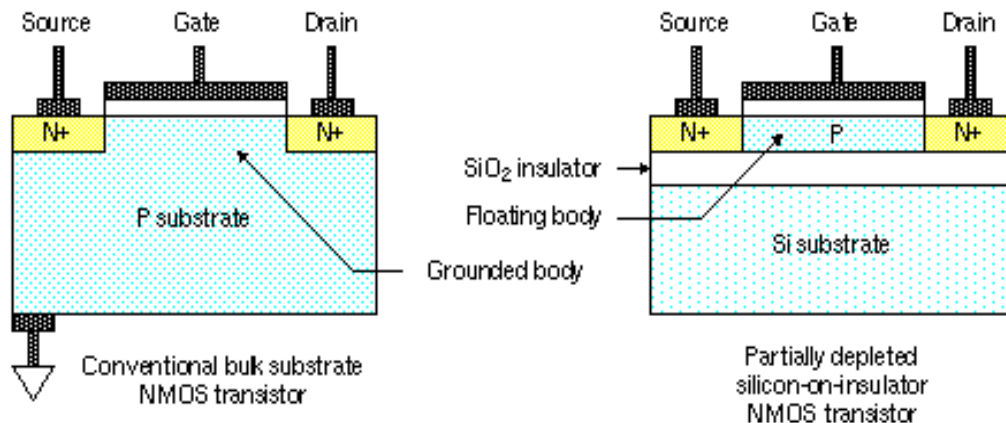


Figure 1 Bulk NMOS Transistor vs SOI NMOS Transistor

The insulating layer increases device performance by reducing junction capacitance as the junction is isolated from bulk silicon. The decrease in junction capacitance also reduces overall power consumption.

1.2 SOI Floating Body

In a standard Bulk CMOS process technology, the P-type body of the NMOS Transistor is held at the ground voltage, while in a PMOS Transistor in the Bulk CMOS process technology is fabricated in an N-well, with the Transistor body held at the VDD supply voltage by means of a metal contact to the N-well.

In Silicon-On-Insulator process technology, the source, body, and drain regions of transistors are insulated from the substrate. The body of each transistor is typically left unconnected and that results in floating body. The floating body can get freely charged/discharged due to the transients (Switching) and this condition affects threshold voltage (V_t) and many other device characteristics.

The transistor area in SOI process is less because there is no need for metal contacts to Wells used for making MOS transistors.

1.3 Fully depleted SOI

In an NMOS transistor, applying a positive voltage to the gate depletes the body of P-type carriers and induces an N-type inversion channel on the surface of the body. If the insulated layer of silicon is made very thin, the layer fills the full depth of the body. A technology designed to operate this way is called a "fully depleted" SOI technology. The thin body avoids a floating voltage. In 45nm and below CMOS, the V_t can be tuned by a midgap metal gate leaving the fully-depleted body undoped. Higher channel mobility and hence higher performance are achieved, as well as lower variability from one device to another. Fully Depleted SOI enables a CMOS LP technology with un-doped body. It gives the best performance - low leakage couple, a perfect choice for Low Power Applications.

1.4 Partially Depleted SOI

On the other hand, if the insulated layer of silicon is made thicker, the inversion region does not extend the full depth of the body. A technology designed to operate this way is called a "partially depleted" SOI technology. The undepleted portion of the body is not connected to anything. The exact voltage depends on the history of source, gate, and drain voltages leading up to the current time (the "history effect"). However, the voltage can be expected to fall within a known range.

The body voltage affects the conduction of the channel and therefore the switching speed and parasitic capacitance of the circuit. In an NMOS transistor, a lower initial body voltage results in a thinner inversion layer, lower conductivity, and slower switching. Conversely, a higher initial body voltage results in faster switching. In a PMOS transistor, the opposite is true; a lower initial body voltage results in faster switching. Charging of the transistor body (floating body) leads to a change in the

threshold voltage, if this is properly taken into account in the IC conception then the result is a faster switching, thus higher performance at same Vdd!

1.5 History Effect

The body of the NMOS or PMOS Transistors in the SOI is floating instead of tie to Ground (NMOS) or VDD (PMOS) as in bulk CMOS. This floating body can change the MOS Transistor Threshold voltage due to differences in the Body voltages. This could cause variation in the circuit delay and mismatch between two identical devices. As the SOI circuit switches, the Body Voltages of the switching transistors will change from their previous steady state condition. This is called the History Effect.

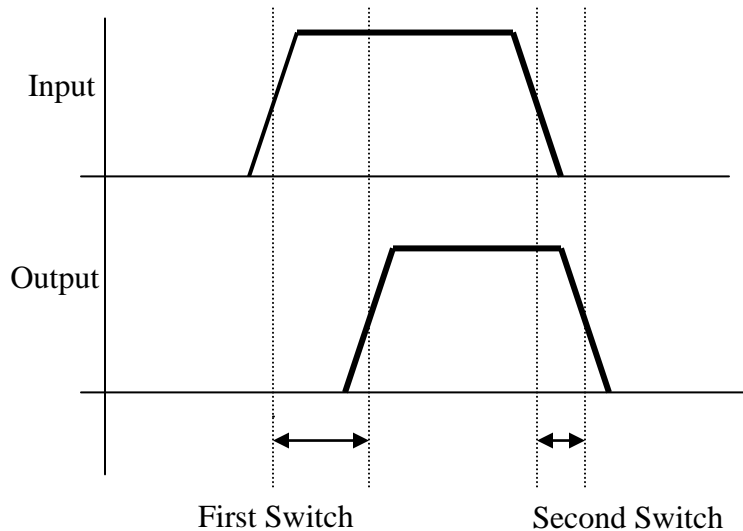


Figure 2 History Effect

A SOI logic circuit can have different (shorter) delay if switching regularly versus a circuit that has been inactive for a long time and then switches. If a circuit is not active for long enough time to be in a steady state and then switches, this switching activity is called **first switch**. If the circuit is switching more regularly, this is called **second switch**. Typically, second switch has shorter delay than first switch due to the body to source voltage of the second switch is higher than first switch, which lowers the V_t of the second switch transistor.

This is one of the most interesting circuit design issues in SOI but it is also a benefit of SOI which contributes to SOI performance advantage over bulk CMOS. Modeling Timing Parameters of the Transistor and Circuit considering the History Effect is critical for successful for Silicon Success.

1.6 Stack Height Effect

Due to floating body in PD and low substrate effect in FD, in a stack of transistors, the V_t of the upper device results smaller than in bulk, improving performances.

Then the first switch is typically 20-30 % faster than in bulk, while the second switch is even faster.

1.7 Substrate Noise

With the increased integration of Digital and Analog circuits on the same die, Substrate Noise issue is dominant in the Bulk process. Especially the digital noise can affect the sensitive analog circuits.

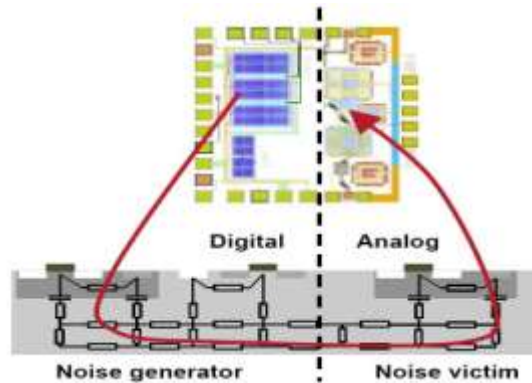


Figure 3 Substrate Noise in Bulk

In SOI technology the Buried Oxide layer acts a die-electric barrier and it helps reducing the Substrate Noise.

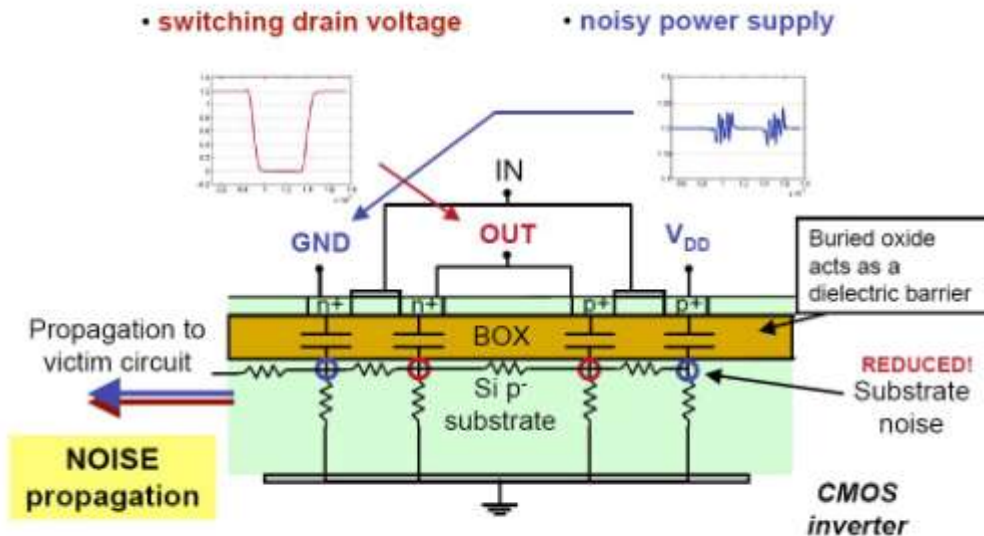


Figure 4 Substrate Noise in SOI

1.8 Latchup Elimination

Bulk CMOS relies on junction isolation between devices, while SOI uses dielectric isolation to surround the entire device sides and bottom. SOI has no wells into the substrate and therefore has no Latchup or leakage paths. It eliminates the need for guard rings, thus smaller area for same function.

1.9 Bi-polar currents

There is a low gain parasitic bipolar transistor on every floating body SOI FET transistors. This bipolar transistor is in parallel with the FET transistor and could cause false switching to the off FET transistor. In general pass gate circuit has the highest bipolar current effect.

Bipolar current of the SOI FET

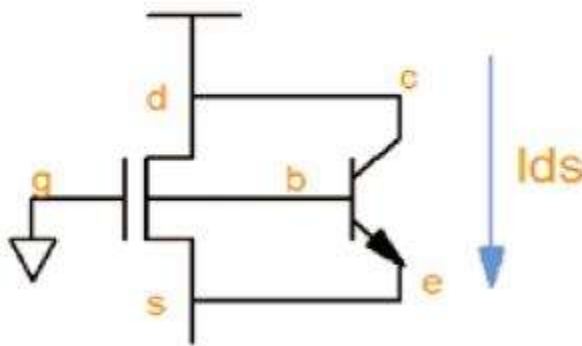


Figure 5 Bipolar current of the SOI FET

When both Source and Drain of Transistor are at High and Gate is at Low (in the case of a Pass Gate), the floating body will couple to high. When Source or Drain goes to Low then we will see a current pulse even when the gate is low.

However over the years of the technology scaling, this bipolar current effect has been pretty much eliminated due to the reduction of the operating voltage of the 90nm node and beyond. In some extreme case, where the design has many transistors connected in parallel, the designer needs to verify the bipolar current to ensure the functionality of the circuit. This is particularly important if the design requires functional burn-in at a much higher voltage than the typical operating voltage condition.

1.10 Self-Heating

The insulation layer of the SOI wafer creates a potential temperature delta between devices called local (self) heating. Self-heating is evident at the high power regions. May not have huge impact on Digital circuits, however this effect must be considered for analog type of circuits.

1.11 Temperature Sensitivity

SOI CMOS is much less sensitive to temperature than bulk. In all SOI processes, the leakage to the substrate is obviously suppressed. Further more in FD SOI, the V_t varies by about 2x less with temperature than in bulk.

1.12 Body Contacts

In the digital circuits, the transistor operates as a switch and remains in a steady state most of the times. Modeling switching characteristics with Floating Body effects are slightly complicated, but it can be modeled. Where as in the Analog/Mixed/IO design modeling the behavior of linear characteristics circuit is very difficult with varying potential of the Floating Body as it changes the output impedance of the device and its V_t -matching to the next device. A Body Contact Transistor can be used as the current source or as any matching transistors designs to eliminate the floating body effect in the SOI technology.

However Body Contact has RC delay associated with it, and shows poor transient response due to high capacitance and resistance. Also Body Contacts do not scale with the Gate Length, and requires bigger Transistor size and Low density. Body contacts are used only where needed because they increase the layout area and decrease performance.

2.0 ASIC Implementation with SOI

ASIC Digital Implementation with SOI very similar to the Bulk process, however the modeling the timing characteristics of the transistors and analyzing the circuit timing due to floating body and history effects are the challenging tasks.

2.1 Modeling SOI effects for the Digital Implementation

Modeling the history effects in the Timing Analysis is a critical step in the Digital Implementation. There are several methods available

Average body voltage method: This method calculates the average body voltage during switching with various loads and frequencies, and applied that voltage to the switching device to generate the delay used by the timing model. This method is very complicated and takes longer simulation time since the body voltage needs to be calculated ahead of time and applied to the Spice simulation during the circuit

characterization. This method can predict chip performance closer to the actual hardware.

First switch, second switch method: With this methodology, the first switch condition is used to create the timing model for the circuit to be used by the slow (max) path delay to predict the maximum frequency of the design. The second switch condition is used to create the timing model for the circuit to be used by the fast (min) path delay to predict race condition (hold issues). This method may result in over-design but the timing model is much simpler and easier to create.

2.2 ASIC Timing Analysis and Closure

For a typical vanilla design (with no clock gates, no muxes etc in clock paths) Timing analysis can be performed using First Switch (max) library model for Setup Analysis and Second Switch (min) library model can be used for Hold Analysis as clock buffers in the clock paths of the Launch and Capture flops will switch together. Either First Switch or Second Switch model is sufficient to model the clock path. Some additional timing margins are used to account the process variations and voltage drop issues.

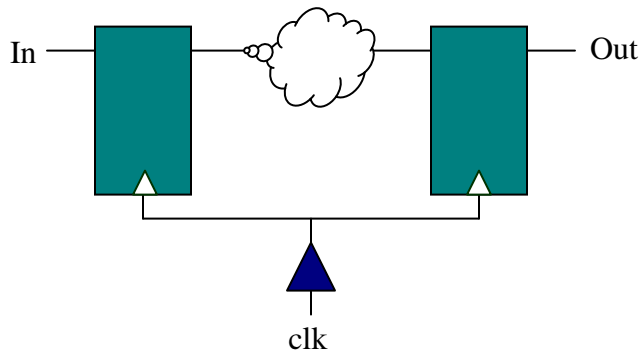


Figure 6 Simple Timing Path

But the designs today are highly complex with complex Power Management schemes, and number of Functional and Test Modes. It requires complex clocking structures that include clock gating cells at various levels of hierarchy to save the clock tree power and multiplexer logic to switch between different modes etc. As shown in the circuit below the source clock and destination clock paths can switch at different times (could be first switch or second switch) depending on the clock gating enable and mode selection signals.

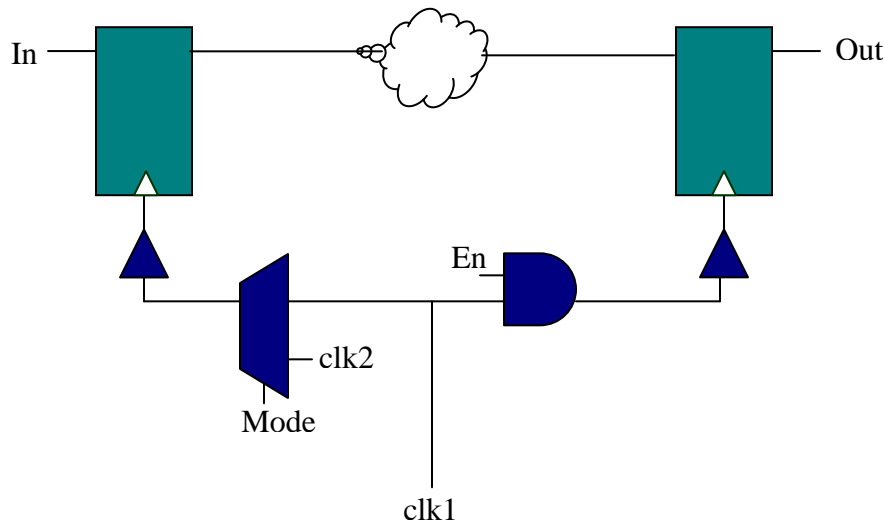


Figure 7 Complex Clock Timing Path

If we run timing analysis Either with only First Switch (max) or Second Switch (min) models, it's very likely we will encounter Setup timing violations and/or Hold timing violations due to the differences in the clock path delays due to the First switch or Second switch possibilities (depending on the modes and clock enables) of launch and capture clock paths. So, one must use both First Switch (max) and Second Switch (min) models during the timing analysis.

Typically the First Switch (max) and Second Switch (min) libraries characterized at "Slow PVT" corner used for setup timing analysis and both First Switch (max) and Second Switch (min) libraries characterized at "Fast PVT" corner used for Hold timing analysis. In multiple corner timing analysis setup and hold timing checks are performed at both PVT corners.

When both max and min libraries are used, the Timing Analysis tool uses the worst case (pessimistic) scenarios for Setup and Hold analysis. For Setup analysis, the Timer picks max library for data path, launch clock paths and picks min library for capture paths. For hold analysis the Timer picks min library for data path, launch paths and max library for the capture paths. This kind of min/max analysis makes timing very pessimistic but it models the possibility of either First or Second Switch in the paths.

SOI Transistors (also Logic Gates) are pretty fast compared to Bulk process due to reduced junction capacitance. Using min/max libraries in the hold timing analysis with additional hold timing margins to cover the process variations etc makes Hold Timing Closure very challenging and the Hold time fixing will cause significant area overhead. History effect also causes duty cycle reduction and care must be taken in the implementation of very high performance designs with both positive and negative edge flops.

2.3 ASIC Implementation Strategy

The biggest challenge in the ASIC Implementation is Hold Timing Closure due to the Timing Analysis with both First Switch (Max) and Second Switch (Min) models and extremely fast libraries. It gets worse if there are divergent clock paths to launch and destination flops from the clock root and a high number of levels in the clock paths.

The below implementation guidelines help in reducing the hold time issues.

1. Reduce the clock divergence to the launch and capture clock domains. Avoid placing the Power Management logic and dividers at the top level, place it within the sub-module hierarchy to reduce the divergent clock paths.
2. Minimize the number of level of logic in the clock paths. Use Big clock buffers or Clock spines to drive clock tree structures. Reducing the clock paths levels (delay) also help in reducing the duty-cycle distortion issues.
3. Use High-Vt flops and Gates in the short data paths, it saves leakage power and reduces the hold buffer area overhead.
4. Use High-Vt buffers to fix HOLD violations

2. SOI Summary

Advantages:

1. Higher Performance at equivalent VDD
2. Reduced Temperature sensitivity
3. Latchup Eliminated
4. Reduced Antenna issues
5. No Body or Well Taps needed
6. Small transistor saves lot of Area
7. Power savings

Disadvantages:

1. Vt Dependent Switching (History Effect)
2. Bipolar currents
3. Self Heating
4. Modeling Issues

These advantages simplify fabrication steps, improve density and reduce parasitic capacitance. The result is up to 30 percent lower power consumption, 20 percent higher performance and 15 percent higher density than traditional bulk CMOS at the same feature size.

References:

1. SOI Implementation guide – by Nghia Phan (Prominent Reference)
2. Synopsys Solvnet Article on SOI
3. D. Bol, R. Ambroise, C. Roda Neve, J.-P. Raskin and D. Flandre, "Wide-Band Simulation and Characterization of Digital Substrate Noise in SOI Technology", IEEE International SOI Conference 2007, pp.133-134, 2007.d.o.i.: 10.1109/SOI.2007.4357888
4. UCL SOI Consortium – by D. Flandre

About the Author:

Narayana Murty Kodeti is a Senior ASIC Engineer and has over Ten Years of Experience in ASIC Implementation, EDA Flow development and computerized control systems. He has extensive experience in RTL Synthesis, Physical Synthesis, DFT, Physical Design, Static Timing Analysis, Crosstalk Analysis and Timing Closure across various technology nodes and its related issues right from 0.35u to 40nm technologies.

About Infotech Enterprises Ltd,.

Infotech Enterprises Ltd (IEL) is one of the leading companies, providing ASIC design services to clients varying from networking, communication, consumer electronics and semiconductor industries. IEL brings proven methodologies and expertise to provide a competitive edge to the customers.

For more information and queries please contact hitech@infotech-enterprises.com