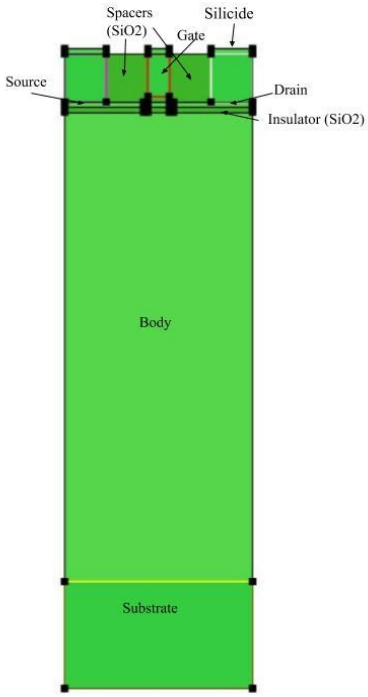
6.012 Final Project: Optimal Transistor Design

Team members: Julia Arnold, Nancy Hidalgo, and Timi Omotunde

Results Summary:



Final optimized transistor Design

Changes Made:

1. Short channel (20 nm)
2. Silicon-on-Insulator
3. LDD on both extensions
4. Higher doping on the drain/source
5. Silicide contacts
6. Lower V_{DD}

Key Metrics:

EDP: 14.59 (ns×pJ)
Energy: 3.646 pJ
Speed: 249.9 MHz

Approach:

When we approached this project, we initially brainstormed several different changes we thought we could implement within design specifications.

Some of the changes we brainstormed were:

- Changing the K dielectric
- Adding Silicides
- Raising Gates
- Shortening channel length
- Implementing Lightly Doped Drain (LDD) regions
- Using shallow junctions
- Adding an SOI region or a Halo/Super Halo regions or a Retrograde doping region in the channel
- Lowering the V_{DD} value

Some of these ideas were more effective than others in their implementation, and we will include an explanation of the ideas that were not in the final implementation and why. For example, changing the dielectric constant was not entirely necessary in Sentaurus as the list of dielectric constants was limited, and some were unable to be manufactured - ie vacuum and air. In the end, changing the dielectric constant was not necessary due to the device's already high performance. Similarly for raising the source, drain, and gates, we were unable to raise these because of consumer restrictions. For Halo/Super Halo and Retrograde doping, we will expand more upon in later section, but these additions were not the most compatible for the size of the device.

Since we had a plethora of changes to try, we decided to approach this task modularly. With each proposed change to the device, we would add only that change for the first few changes so we knew the effect of the modification in comparison to the initial transistor design. Thus, if one change drastically changed the performance, it could be more easily adjusted for and optimized. Once we optimized one feature, we would later combine changes to ensure that each change complemented the others.

Methodology:

After initially modeling the initial transistor, we observed it was far from the specs that we need to deliver to the customer as the clock frequency was below the required 200MHz, the Energy Delay Product (EDP) was almost 270 (ns×pJ), and the power consumption was a vast amount of 45.57 (pJ). Seeing how much improvement our initial device needed led us into the first few iterations of our design. Because increasing energy efficiency generally decreased performance, we decided to first increase the clock frequency until it was above the threshold, then from there increase energy consumption to lower the overall EDP.

Short channel:

In order to increase the clock frequency, the first change we made was to shorten the channel of the MOSFET (while extending the regions next to the source and the drain to maintain a constant transistor width). When the channel is shorter, the electric field in the MOSFET is bigger, resulting in higher carrier velocities and thus higher currents and higher clock frequencies.

This lead us to first change the channel length to 4 nanometers. From this change we expected to see a significant increase in the clock frequency and an increase in the power consumption. When we performed this change our expectations were met as we saw an increase in the clock frequency and a significant increase in the power consumption, which came from large increases in both the I_{ON} and the I_{OFF} as V_{TH} lowers. However, these large currents resulted in punch-through effects that we observed on the $V_{GS} - I$curve, so we tried several additional different channel lengths of 5 (nm), 10 (nm), 15 (nm), and 20 (nm) to find the shortest channel length with sufficiently low currents. We found that 20 nanometers was the optimal channel length in order to avoid device limitations.

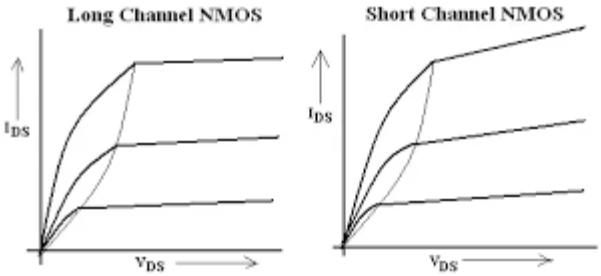


Figure 3. Effect of Channel Length Modulation on the SLOPE.

Effects of channel length modulation on long and short channel NMOS

SOI:

The next change we thought to do was to minimize the power consumption by decreasing the short-channel effects, which is the non-ideal behavior such as punch through that leads to leakage current and dielectric or avalanche breakdown that occurs when a channel is too small. Because we had shortened the channel, our MOSFET began exhibiting non-ideal behavior such as significantly higher currents, which then resulted in higher power consumption. We came up with two solutions: a super-halo doping scheme or a Buried Oxide Layer (BOX) layout, also known as an SOI scheme.

We chose the the SOI layout as it would tackle the most of our short channel effects more effectively than Halo/Super Halo or Retrograde doping for transistors of this size. Halo/Super Halo doping works best for bigger transistors (those over 100 nm), while SOI is better suited for smaller transistors, specifically those under 45 nanometers, as it mitigates punch through and leakage current some of the more salient short channel effects that we strived to reduce. SOI works because when the height of the channel decreases there is a proportional decrease in the punch through as there is less area for it to occur. In the SOI, you need only to optimize the type of insulator (dielectric constant) and the height of the insulator (thickness of the insulator region).

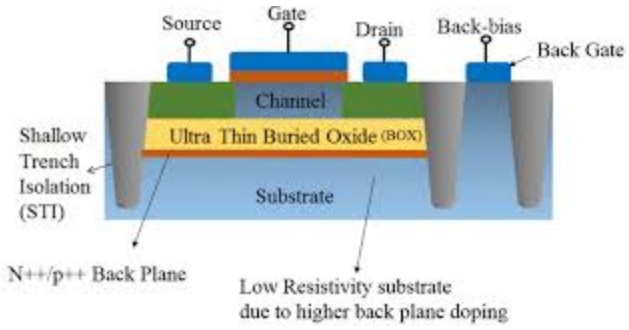


Illustration of an SOI MOSFET

This scheme lowers power consumption by replacing traditional silicon substrates with alternating layers of silicon and an insulator (typically silicon dioxide). With this change, the parasitic capacitance decreases, as do the leakage currents. Furthermore, the MOSFET can operate at lower input voltages (V_{DD}), which also decreases the total power consumption.

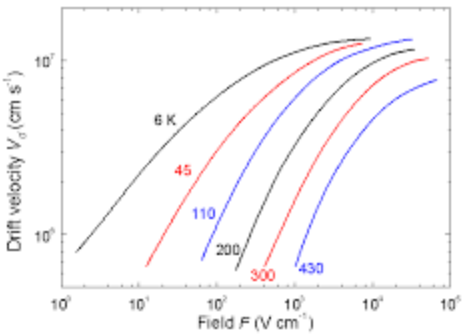
We expected this change to result in a steep decrease in power consumption with little to no effect on the clock frequency. When we modeled it we realized that using an SOI scheme by itself increased the energy output, but in combination with the 20 nanometer channel length, it lowered the energy output. We initially began with an insulator height of 15 nanometers but then modulated this height until finding the optimal length of 5 nanometers. We later learned that

typically, the thinner the SOI region the more effective the layer is, and thus our SOI layer was 5 nanometers thick so it is called an Extremely Thin SOI (ETSOI).

Drain and Source Doping:

At this point, we needed to increase our clock frequency. To do that, we thought to increase the doping of the source and the drain because this would increase the observed currents. This change would increase the observed currents because the increase in doping would result in a increase in the total charge. This in turn would increase the electric field and thus the force on the charge carriers would be larger, resulting in faster carrier speeds. Of course, larger currents also result in higher power consumption (since $P = I^2R$). Therefore, we expected this change to cause an increase in both the clock frequencies and the energy consumption, with a bigger increase in the frequency than the energy.

$$I_{Dsat} = \frac{W}{2L}\mu C_{ox}(V_{GS} - V_T)^2$$



Carrier Speed vs Electric Field

We tried a doping concentration of 1E20 centimeters^{-3} (cm^{-3}) (up from the original 1E19 cm^{-3}), with good results: our frequency increased to 382.8 MHz, although our power consumption increased to 15.22 pJ.

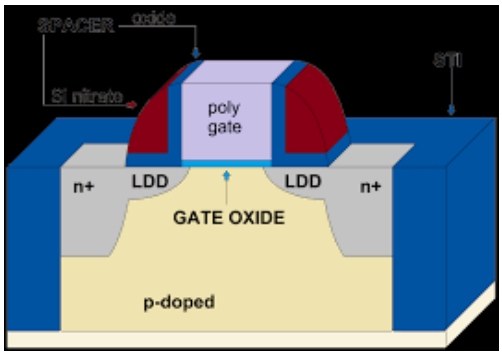
LDD:

At this point our frequency was well above the threshold, so we decided to lower the energy consumption by implementing a Lightly Doped Drain (LDD) scheme. This doping scheme would minimize the depletion width at the junction by lowering the internal potential (ϕ_B), which then reduces the electric field and the voltage. The LDD also works to stop punch through in order to decrease the leakage current by making sure there is no punch through. This

also mitigates the hot carrier effect, or the destruction of a MOSFET structure due to high carrier velocities, by lowering these velocities.

As we were discussing this change, we considered designing our source extension to be bigger than the drain or other asymmetrical schemes. However, we eventually decided against it because that would add more variability and thus make our design harder to optimize.

$$x_d = \sqrt{\frac{2\varepsilon_s(N_A+N_D)(\Phi_B-V_D)}{qN_DN_A}}$$

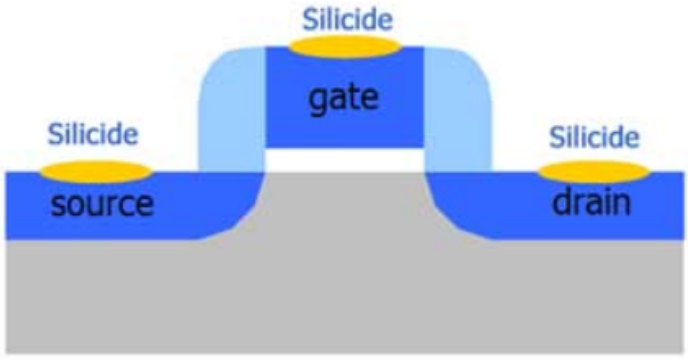


MOSFET with symmetric LDD regions

We expected this to make a significant change in the energy consumption, but not a significant change in the speed of the processor. We also expected lower doping values to result in a better EDP. We generally saw significant changes in the energy consumption with little effect on the clock frequency. However, we observed optimal results with a doping concentration of 1E19 cm^{-3}.

Silicide contacts:

The second to last change we made was adding silicide contacts to the top of our transistor. We expected this change to decrease the energy consumption by minimizing the parasitic resistance that comes from the junctions. Silicide does this because of its low resistance and little to no electromigration (the transport of material due to a lot of ion movement). Rather than the traditional aluminum, silicide contacts overall reduce the power consumption.



MOSFET illustration with silicide contacts

We expected this implementation to lower the power consumption, but when combined with all of the previous changes, it didn't. In fact, it actually increased the frequency by 1 MHz. We think that this change came from higher currents and that these higher currents come from the lowered resistances.

Lower VDD:

Now, our EDP was 15.97 (ns×pJ) with an energy consumption of 4.628 (pJ) and a clock frequency of 289 MHz. We had gone through most everything on our list, and our last change only decreased our EDP by 0.05 (ns×pJ). More importantly, we were approaching the limit to the number of changes we were allowed to make. So we decided to modify the gate voltage. We decided to try this change since it lower operating voltages result in lower power consumption (and lower frequency) since $P = IV$. We first tried the minimum of 0.40 volts, which resulted in punch-through. Then we tried several different voltages around 1 V before finding the optimal voltage of 1.6 V. This final change resulted in a decrease in the currents as expected and an overall decrease in the power consumption to a mere 3.646 (pJ). This resulted in a clock frequency of 249.9 MHz with an EDP of 14.59 (ns×pJ).

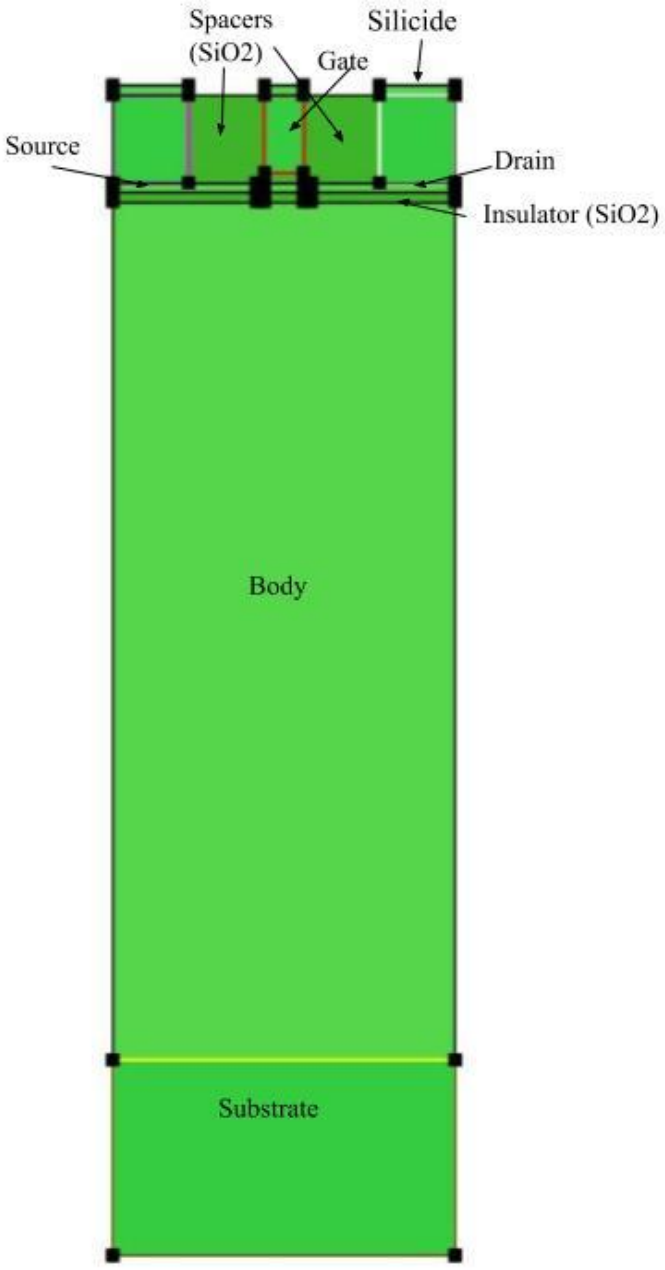
Results:

From these changes, we obtained the following results (although in reality we made more changes, which are omitted for brevity):

	initial	soi20tf	soi20tf, short	soi20tf, short	silicide	Vdd .4	vdd9	silicide4	Silicide1.6
Vth	0.558371	0.604572	0.622765	0.837772	0.836451	0.325387	0.71735	0.325411	
Gmmax	0.00047351	0.00040849	0.00045343	0.00041612	0.00041646	2.98E-007	0.0001811	3.01E-007	
Idsat	5.88E-010	4.88E-010	5.34E-010	4.01E-010	4.01E-010	2.25E-014	3.35E-011	2.27E-014	3.07E-010
Ioff	1.99E-023	1.61E-033	1.58E-031	-1.01E-033	-4.02E-034	-7.21E-029	7.02E-031	9.36E-027	
measure Ioff	4.07E-007	2.00E-007	2.76E-007	1.79E-009	1.80E-009	3.88E-011	2.02E-010	3.90E-011	1.19E-009
converted Idsal	587.941	488.403	533.794	400.584	401.46	0.0225438	33.467	0.02272	306.952
converted Ioff	407	200.203	276.266	1.79485	1.80211	0.0388164	0.202345	0.0390308	1.18833
capacitance		0.264	0.119	0.119	0.119	0.119	0.119		0.119
energy	45.57	15.95	15.22	4.628	4.628	5.318	1.287		3.646
frequency	0.1706	0.2826	0.3828	0.2888	0.2898	0.000734	0.04832		0.2499
EDP	267.1	56.44	39.76	16.02	15.97	7245	26.64		14.59

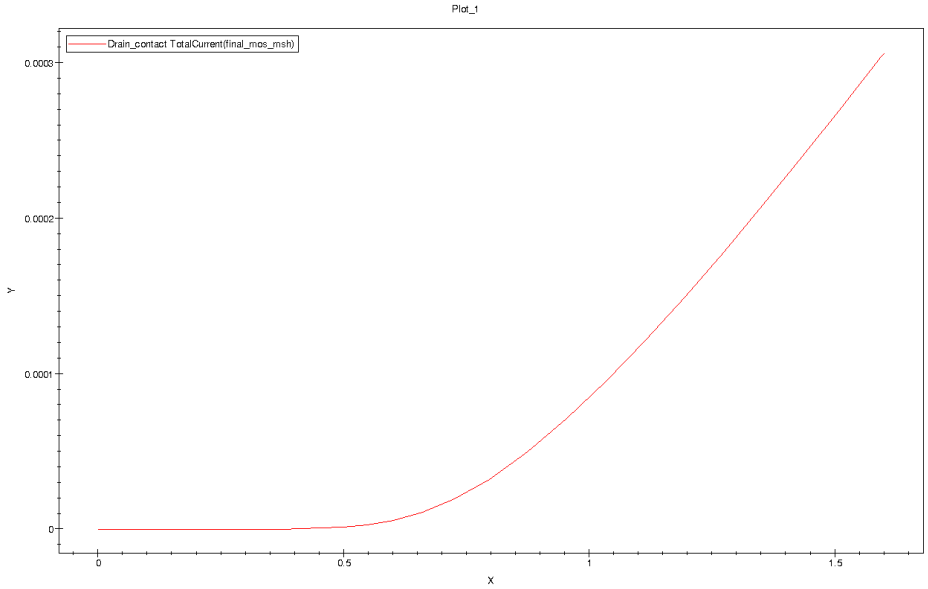
Spreadsheet detailing design changes from the inception to the end

Below, we've included our final transistor design with silicides, an SOI of thickness of 5 nm, highly doped drain and source (both with doping concentrations of 1E20 cm^{-3}, lightly doped drain and source extensions (both with doping concentrations of 1E19), shorter source and drain extensions (both at 20 nm), and a shorter channel (both at 20 nm).



Final optimized transistor Design

Finally, the below plot shows the V_{GS} - I curve of our final design as extracted from sentaurus. From it we can note that the MOSFET operates with minimal short-channel effects (since it is curvilinear).



V_{GS}-I curve of our final design

Conclusion:

We propose a transistor design more than fit for our customer's purposes using the techniques provided for in 6.012. This transistor has a threshold voltage of 0.84098 V, an I_{ON} of 306.952 (μA/μm), an I_{OFF} of 1.1883 (μA/μm), a clock frequency of 249.9 MHz (almost 50 MHz above the requires spec), and an EDP of 14.59 (ns×pJ). This design exploits the short channel effects, the material properties of insulators, the varied doping, and the material properties of silicide while also mitigating negative aspects of the short channel effects of the high electrostatic forces, the leakage current, and the punch through to meet our consumers' needs.