

Advancing RIT to Submicron Technology: Design and Fabrication of 0.5 μ m N-channel MOS Transistors

Michael Aquilino, *Student Member, IEEE*

Abstract—The design and fabrication of N-channel MOS transistors with effective gate lengths of 0.5 μ m or smaller have been completed at the Semiconductor and Microsystems Fabrication Laboratory at the Rochester Institute of Technology. An NMOS device with $L_{\text{mask}}=0.7\mu\text{m}$ results in an $L_{\text{eff}}=0.5\mu\text{m}$. The drive current for this device with supply voltage of 3.5V is 108 $\mu\text{A}/\mu\text{m}$. The sub-threshold slope is 100mV/decade and a DIBL parameter of 29mV/V is reported. An NMOS device with $L_{\text{mask}}=0.6\mu\text{m}$ results in an $L_{\text{eff}}=0.4\mu\text{m}$. The drive current for this device with supply voltage of 3.5V is 140 $\mu\text{A}/\mu\text{m}$. The sub-threshold slope is 103mV/decade and a DIBL parameter of 110mV/V is reported. These are RIT's first sub-0.5micron MOS transistors.

Index Terms—CMOS, NMOS, Process Development, Silvaco, Athena, Atlas, Terada-Muta

I. INTRODUCTION

The goal of this project is to design and fabricate n-channel MOS transistors with gate lengths down to 0.5 μm . To date, the smallest transistor fabricated in the Semiconductor and Microsystems Fabrication Laboratory (SMFL) at RIT has an effective channel length, $L_{\text{eff}}=0.75\mu\text{m}$. The motivation for fabricating 0.5 μm transistors is to allow RIT to continue the microelectronics industry trend of scaling towards smaller, faster and cost effective devices. The integration of a new manufacturing process flow has been implemented. The device technology employed includes: n⁺ poly gates, LOCOS isolation, aluminum metallization, shallow ion implanted n⁺ sources/drains, blanket p-type well, p⁺ channel stop and is designed to operate at a supply voltage of 3.5V. In addition, this project has provided hands-on experience in transistor level design and layout, mask design and fabrication, process development and integration, and electrical testing.

This work is part of capstone design project for a B.S. degree in Microelectronic Engineering at the Rochester Institute of Technology (RIT), Rochester, NY. The results of the project were first presented as part of the 22nd Annual Microelectronic Engineering Conference, May 2004 at RIT.

M. Aquilino is with the Microelectronic Engineering Department, RIT, (e-mail mva6237@rit.edu)

Design and layout of the test chip has been done using Mentor Graphics IC Station. The test chip includes transistors scaled down to 0.5 μm as well as test structures, capacitors, and alignment verniers. Four lithography levels: Active, Poly, Contact Cut, and Metal are exposed on the Canon FPA 2000i1, a 365 nm i-line exposure tool. The mask set is fabricated using a MEBESIII electron beam mask writer. Control wafers along with the device wafers are processed for in-line process verification and characterization.

Theoretical calculations have been made to determine process parameters such as doping profiles, junction depths, material thicknesses, and threshold voltages. A process flow is developed to achieve the required device parameters. Simulation of the process is carried out using Silvaco Athena to modify theoretical values and obtain more accurate process parameters. Electrical simulations are in progress using Silvaco Atlas for electrical performance analysis. At the completion of fabrication, the devices will be electrically tested for on and off-state performance. Parameter extraction will then be done to create a SPICE model for these devices that can be used in more complex circuit designs.

II. CHIP DESIGN

A new test chip was designed that can be used for either NMOS or PMOS processes. There are a total of 122 transistors on the design with various Length/Width ratios. Three unique sets of transistors were designed that include fully-scaled NMOS based on MOSIS 2.0 λ design rules, and two flavors of non-scaled devices. The fully scaled devices are designed with mask defined gate lengths of 0.5 μm to 10 μm where the gate length is two times the minimum feature size, λ . Mask defined poly gate lengths range from 0.35 μm to 100 μm and widths of 10 μm to 20 μm for non-scaled devices.

The results of the device discussed in this paper are based on a non-scaled design with contact cut dimensions of 5 $\mu\text{m}\times 5\mu\text{m}$ and 12 μm spacing between contact and poly edge. Other design rules have been relaxed in terms of metal and poly overlap. A second set of non-scaled devices feature 2 $\mu\text{m}\times 2\mu\text{m}$ contact cuts and a 4 μm spacing between contact and poly edge. The fully scaled devices cannot be tested due to isotropic wet etching of the Aluminum metal.

Also included on the chip are capacitors with areas that range from $40\text{k}\mu\text{m}^2$ to $250\text{k}\mu\text{m}^2$. These can be used for Capacitance-Voltage analysis to determine Accumulation, Depletion, and Inversion capacitance, oxide charge effects, and threshold voltage.

Various other test structures such as Van-der-Pauw structures for sheet resistance measurements of n⁺, poly, and metal, and Cross-Bridge Kelvin Resistors, for metal contact resistance to n⁺ and contact resistance to poly are included. There are also contact cut test structures with various contact cut dimensions ranging from $0.5\mu\text{m}\times 0.5\mu\text{m}$ to $5\mu\text{m}\times 5\mu\text{m}$. A simple test for conduction between a contact pad and a reference pad will yield if a conductive current path is made and whether it is ohmic or non-ohmic.

Transistors with field oxide as a gate oxide are included to determine field threshold voltage to be sure it is high enough to stop on parasitic conduction channels from source to drain or device to device.

Included in Figure 1 is the layout of the test chip designed in Mentor Graphics IC Station. It uses 5 design layers: active, poly, metal, contact to active, contact to poly. The two contact layers are combined to form one contact mask and individual masks for active, poly, and metal were fabricated using a MEBESIII electron beam mask writer.

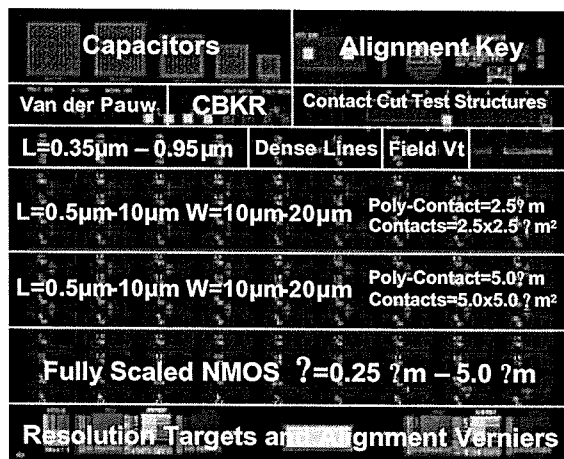


Figure 1: Layout of NMOS Test Chip

III. SIMULATION

A new process flow was developed at RIT to fabricate the NMOS transistors. This included shallow source/drain junctions, gate lengths approaching the limit of the exposure tool, and thermal processing with temperatures low enough with Transient Enhanced Diffusion and lateral diffusion into the channel designed for. The process simulation was done in Silvaco Athena (SUPREM) and the electrical simulation in Silvaco Atlas. The on and off-state performance of the devices can be compared to results from fabricated devices.

A. Process Simulation

The process used to fabricate the NMOS transistors has been simulated in Silvaco Athena to verify the correct process parameters such as implant dose and energy, thermal steps, and film deposition, result in the desired doping profiles and device structure. Figure 2 is the final simulated device structure.

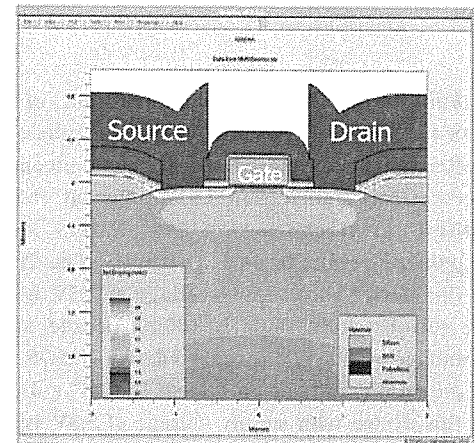


Figure 2: Final Device Structure with 0.5 μm Physical Gate length

The final simulated doping profiles achieved are shown in Figure 3. They include shallow S/D junctions with a depth of approximately $0.15\mu\text{m}$. The junctions must be scaled deeper to give the gate more control over turning the device on and off. The p-well is doped moderately high to reduce the depletion region extending from the S/D into the well. This will enable deep submicron transistors to perform as long channel devices by reducing channel length modulation and preventing lateral punch-through.[1]

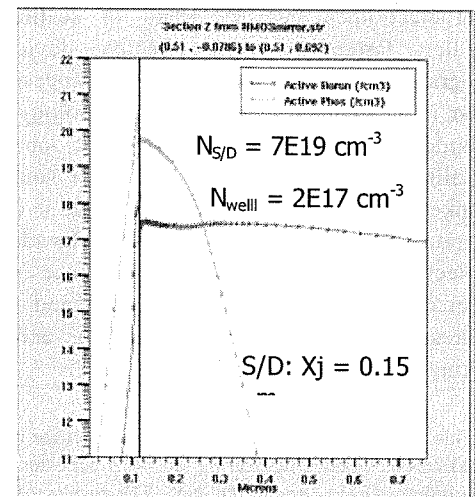


Figure 3: Simulation of Final Doping Profiles

B. Electrical Simulation

The final device structure was imported into Silvaco Atlas and simulated for on and off-state performance. For the 0.5 μ m physical gate length device, an on-state saturation drive current with $V_G=V_D=3.5$ V is simulated to be 300 μ A/ μ m. The threshold voltage is targeted to be 0.85V. In the off-state, a sub-threshold slope of 112mV/decade and sub-threshold swing of 9.5 decades is obtained. Figure 4 is a simulated ID-VD plot with the transistor biased in the linear and saturation mode. Figure 5 is a ID-VG plot with ID on a linear scale for threshold voltage extraction and on a log scale for sub-threshold characteristic extraction. I_{off} is simulated to be 10fA/ μ m with $V_G=0$ V.

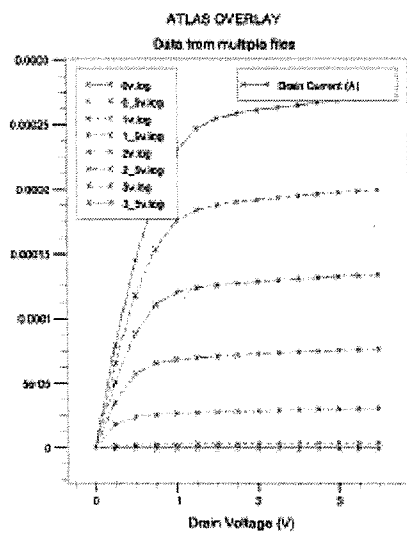


Figure 4: ID-VD for NMOS Transistor

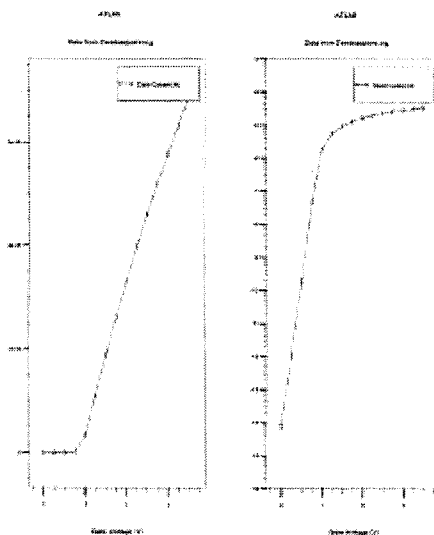


Figure 5: ID-VG in Linear and Sub-threshold Mode

IV. FABRICATION

All fabrication for these devices was done in the Semiconductor and Microsystems Fabrication Laboratory at RIT. The entire process takes approximately 4-5 weeks to complete with an average of 5-6 hours of processing time per day. A total of 32 process steps are required to bring the devices from bare silicon to test. A cross-section is shown in Figure 6 with the final device topology.

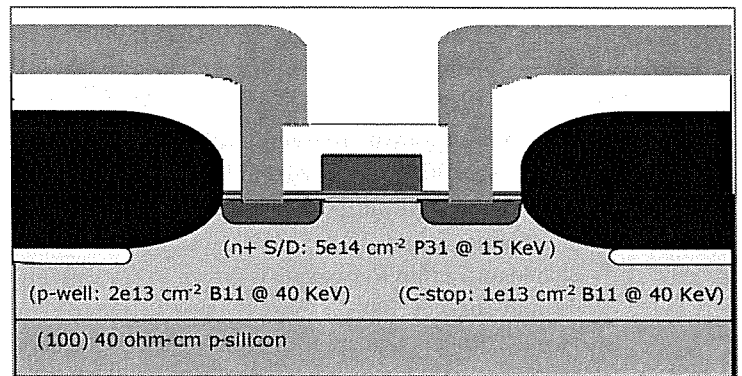


Figure 6: Cross-section of NMOS Transistor

A. P-Well Formation

- Grow 500Å Pad Oxide
- Ion Implant 2e13cm⁻² B11 @ 40 KeV
- Drive in for 60 min @ 1025°C in N₂
- NA = 2e17cm⁻³
- Xj = 2 μ m

A pad oxide is used as a screening layer for the implant so that channeling of ions along certain crystal planes in the silicon lattice does not occur. Also, the pad oxide will be used as a stress relief layer between the subsequent Si₃N₄ deposition. The doping of the p-well is chosen to be moderately highly doped so that depletion regions extending from the source/drain do not touch when the device is biased; leading to a short channel effect known as punch-through.

B. Creation of Active Area

- Deposit 1500Å of Silicon Nitride
- Level 1 Lithography: Active Area Define
- Channel Stop Implant: 1e13cm⁻² B11 @ 40 KeV
- Grow 3500Å Field Oxide

The silicon nitride is patterned so that the active area of the device is masked. A channel stop implant is done through the pad oxide to place additional boron in the field region so that parasitic conduction channels do not form between source and drain or from transistor to

transistor. A moderately thin field oxide is grown to reduce the birds beak effect inherent in a LOCOS process and to reduce the amount of boron out-diffusion into the field oxide.

C. Gate Formation

- Strip Nitride and Etch Pad Oxide
- Grow 250Å Kooi Oxide & Etch
- Grow 150Å Gate Oxide
- Deposit 3500Å Polysilicon
- Dope Poly with Solid Source Phosphorous
- Level 2 Lithography: Poly Gate Define
- Poly Etch in SF_6 Plasma

A thin oxynitride layer is formed after LOCOS field growth so a 30 second Buffered HF etch is performed. Next the silicon nitride is stripped off in a wet hot phosphoric acid bath at 175°C. The pad oxide is etched and a thin Kooi oxide is grown in dry O_2 with a 99 min soak at 900°C. This is done to remove any nitride stringers that are formed along the edges of the active region. This effect is called "white ribbon" and is created by a reaction between NH_3 and Silicon at the LOCOS edge.[1] Next a 150Å gate oxide is grown with a 45 min soak at 900°C in dry O_2 . A chlorine pre-growth clean is done to neutralize sodium ion contamination in the tube and in the quartz boat. The poly is deposited via LPCVD and doped with a solid source dopant. A 10-minute soak at 1000°C in N_2 is done to drive-in the n+ poly dopant. Level 2 lithography is performed to pattern the gate and is etched in an SF_6 gas in a LAM490 plasma etcher.

D. Source/Drain Formation

- Ion Implant $5 \times 10^{14} \text{cm}^{-2}$ P31 @ 15 KeV
- $\text{ND} = 7 \times 10^{19} \text{cm}^{-3}$
- $X_j = 0.15 \mu\text{m}$

Shallow source/drain implants are done with a P31 ion to introduce phosphorous into the source/drain regions. The field oxide is being used as a hard mask for the implant into the field region. A $5 \times 10^{14} \text{cm}^{-2}$ dose at 15 KeV is performed with the use of a Varian Ion-Implanter. The resulting junction depth after anneal is targeted to be $0.15 \mu\text{m}$ with a surface concentration of $7 \times 10^{19} \text{cm}^{-3}$.

E. Back End Processing

- Deposit 3000Å PECVD TEOS
- Anneal for 30min at 900°C in N_2
- Level 3 Lithography: Contact Cut
- RIE Oxide in SF_6/CHF_3 mixture
- Sputter Deposit 5000Å Aluminum
- Level 4 Lithography: Metal
- Wet Etch Al wiring

- Sinter

An inter-level dielectric is deposited to a thickness of 3000Å using an Applied Materials PECVD tool with TEOS chemistry to form an insulating oxide. The source/drain anneal step is done after the oxide deposition to densify the TEOS so that its insulating properties are enhanced. Level 3 lithography patterns the contact cuts, which are then etched in an SF_6/CHF_3 mixture in a Drytek Quad RIE plasma etcher. The resist is stripped, the wafers cleaned, and 5000Å Aluminum is deposited via a CVC601 sputtering tool. Level 4 lithography patterns the metal features and a wet etch in a 50°C Phosphoric/Nitric/Acetic acid is performed. Finally, the wafers are sintered in a 5 slpm H_2/N_2 forming gas mixture for 10 minutes at 420°C.

V. RESULTS

The results of the first lot of devices did not yield characteristic transistor IV. The ID-VD plot of a $10 \mu\text{m}$ NMOS device is shown in Figure 7.

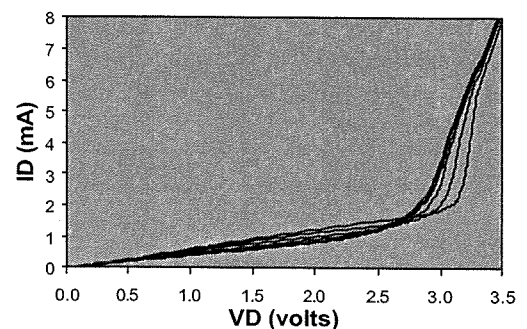


Figure 7: L/W=10 μm /20 μm Failed NMOS Device

It was determined that the contact cuts to the source/drain were etched through the n+ silicon and into the p-well due to aggressive RIE plasma etch. This causes a parasitic conductive path from source to drain through the well, bypassing the inversion channel as the primary current conductor. It was also found that the channel stop implant was too low in dose and energy. To fix the problem, 2 wafers that were left behind in the process after gate oxide growth were processed. Additional boron was implanted through the field oxide with a $1.5 \times 10^{13} \text{cm}^{-2}$ dose at 75 KeV. The etch time in the RIE tool was cut nearly in half to ensure no silicon would be etched.

The results of the second lot were successful in producing transistor characteristic curves, which exhibited long-channel behavior. Figure 7 is a cross-sectional SEM micrograph of the final device fabricated.

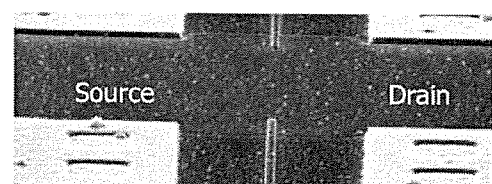
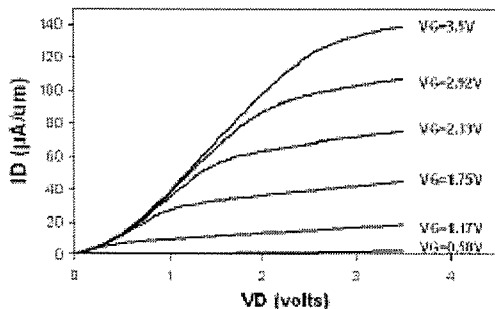


Figure 7: SEM Micrograph of $L_{\text{mask}}=0.6\mu\text{m}$ NMOS

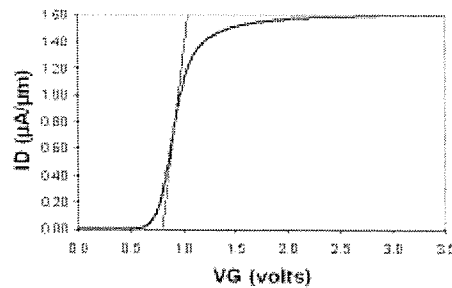
An ID-VD plot for a device with $L_{\text{mask}}/W_{\text{mask}}=0.6\mu\text{m}/20\mu\text{m}$ is shown in Figure 8. Through the Terada-Muta Method of L_{eff} extraction, an L_{eff} of $0.4\mu\text{m}$ is reported. This L_{eff} takes into account process biases such as isotropic polysilicon etch and source/drain lateral diffusion under the poly gate which makes the channel length smaller. This is RIT's first sub- $0.5\mu\text{m}$ MOS transistor.

Figure 8: NMOS ID-VD for $L_{\text{eff}}=0.4\mu\text{m}$

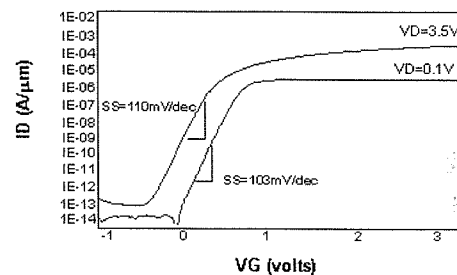
There is some non-ohmic behavior, which has been attributed to a residual interfacial oxide layer that exists between the Aluminum and silicon. A more aggressive sinter recipe is needed to correct this problem. It is also noticed that there is high series resistance, which is causing the ID-VD curve to slope over and not deliver the entire 3.5V that is applied to the drain to actually appear across the transistor. There is some slight upward slope in the saturation region of operation, which is due to channel length modulation, another short channel effect common to submicron transistors. Preliminary mobility parameter extraction yields values of around $200\text{cm}^2/\text{v}\cdot\text{sec}$, which is very low. This can be attributed to velocity saturation effects since a high voltage is being placed across a small gate length. Low doped drains can be implemented in the future to add additional resistance to the transistor, thereby decreasing the effective voltage across the S/D terminals.

An ID-VG plot is shown in Figure 9 on a linear current scale to extract the threshold voltage. The threshold voltage for this device is 0.75V . It is also noticed that

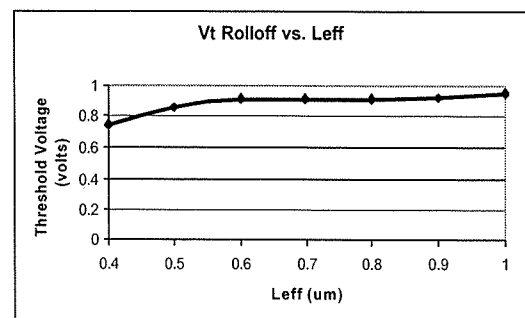
the current slopes over dramatically as V_G is increased due to high series resistance.

Figure 9: NMOS ID-VG for $L_{\text{eff}}=0.4\mu\text{m}$

A plot of ID-VG on a Log ID scale is shown in Figure 10. The sub-threshold slope is $103\text{mV}/\text{decade}$ at 0.1V drain bias and increases to $110\text{mV}/\text{decade}$ at 3.5V drain bias. I_{off} is approximately $10\text{fA}/\mu\text{m}$ at 0.1V drain bias, which matches exactly with simulation. There is about 7.5 decades difference between I_{on} and I_{off} .

Figure 10: Sub-threshold ID-VG for $L_{\text{eff}}=0.4\mu\text{m}$

Figures 11, 12, and 13 are off-state performance plots as the devices are scaled submicron. V_t roll-off is a common phenomenon as devices are scaled down in gate length. It can be seen in Figure 11 that the V_t doesn't roll-off to any significant amount until around $0.4\mu\text{m}$. At 0.75V , this is only a 12% decrease compared to the $0.5\mu\text{m}$ device with a V_t of 0.85V .

Figure 11: V_t Rolloff vs L_{eff} for NMOS Transistors

The sub-threshold slope is plotted vs L_{eff} in Figure 12, for drain biases of 0.1V and 3.5V . It is seen that the two curves do not diverge until around $0.4\mu\text{m}$.

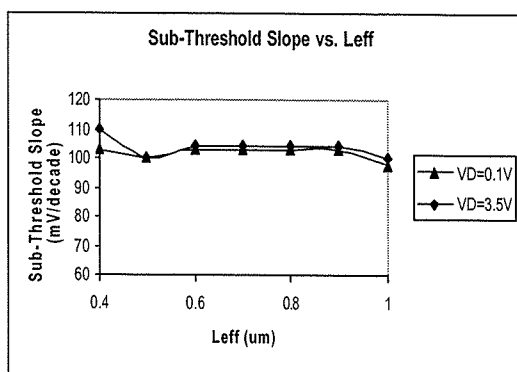
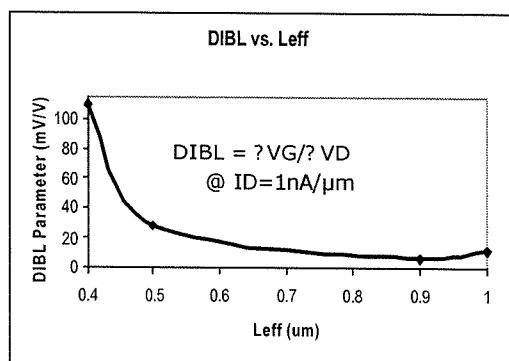
Figure 12: Sub-threshold Slope vs. L_{eff}

Figure 13 is a plot of DIBL vs. L_{eff} to characterize the short channel effect of Drain Induced Barrier Lowering. This effect causes the gate to lose control over turning the device off and increased off-state current results even when sub-threshold slope does not increase dramatically. The DIBL parameter is calculated in units of mV/V and corresponds to a change in gate voltage per change in drain voltage at a fixed current level.[1] A DIBL value of around 25 mV/V is good. It is seen that the DIBL does not start to dramatically increase until the device is scaled to an L_{eff} of 0.4μm.

Figure 13: DIBL vs. L_{eff}

A summary of the performance between an L_{eff} of 0.4μm and 0.5μm is shown in Table I.

A summary of additional electrical parameters is shown in Table II. It is noted that the sheet resistance of the n+ region is very high, 360 Ω/?; this is due to very shallow junctions.

TABLE I

SUBMICRON PERFORMANCE PARAMETERS

L_{eff} (μm)	V_t (V)	I_D @ $V_G=V_D=3.5V$ (μA/μm)	SS (mV/dec)	DIBL (mV/V)
0.4	0.75	140	103	110
0.5	0.85	108	100	29

A mask drawn gate length of 0.7μm yielded a device with an effective channel length of 0.5μm. The threshold voltage for that device is 0.85V, which came out exactly

as simulated. The drive current in the saturation mode

TABLE II

ELECTRICAL DEVICE PARAMETERS

	R_{sheet} (Ω/?)	$G_{contact}$ (mmho/μm ²)
N+	360	1.2
Poly	35	0.72
Aluminum	0.227	-

of operation with $V_G=V_D=3.5V$ is 108 μA/μm. The sub-threshold slope is 100 mV/decade at V_D of 0.1V and 3.5V. The DIBL parameter is 29 mV/V, which is very good. It can be seen that the 0.5μm L_{eff} device has well controlled short channel effects. The smallest device tested on the die was a mask drawn gate length of 0.6μm, which yielded an effective channel length of 0.4μm. The threshold voltage for this device begins to roll-off to about 0.75V with a sub-threshold slope of 103mV/decade. The DIBL parameter quickly increases to 110mV/V, which is much higher compared to the 0.5μm device, but depending on off-state current requirements, this may be acceptable. As an advantage, the 0.4μm device features a 33% increase in drive current at $V_G=V_D=3.5V$ at a value of 140 μA/μm.

VI. CONCLUSION

The goal of this project was to design and fabricate NMOS transistors with gate lengths of 0.5μm or smaller. It has been demonstrated that 0.5μm and sub-0.5μm N-channel MOS transistors are capable of being fabricated completely in-house at the SMFL at RIT using standard CMOS processing techniques. Future work will include integrating low doped drains using sidewall spacer technology and silicided source/drains and gates. This will allow for a second, deeper, implant to be done, which will lower the sheet resistance in the active area.

ACKNOWLEDGMENTS

The author thanks Dr. L. Fuller for guidance in this project. Also acknowledged are Dr. S. Rommel, Dr. K. Hirschman, Dr. S. Kurinec, Scott Blondell, Bruce Tolleson, Reinaldo Vega, Nate Westcott, Eric Woodard, Jeff Steinfeldt, Matt Holland, and Lisa Camp. The author would also like to acknowledge the RIT SMFL staff for technical assistance and equipment support.

REFERENCES

- [1] S. Wolf, *Silicon Processing for the VLSI Era – Volume III*, Lattice Press (1995), p. 219, 232, 364.

Michael Aquilino, originally from Liverpool, NY, received a B.S. degree, *Summa Cum Laude*, in Microelectronic Engineering from the Rochester Institute of Technology in 2004. He obtained co-op work experience at Atmel Corporation in Columbia, MD and Integrated Nano-Technologies in Henrietta, NY. He is currently pursuing an M.S. degree in Microelectronic Engineering at RIT.