Integration of High-K Dielectrics and Metal Gates into Submicron NMOS Transistors at RIT

Daniel J. Jaeger, *Student Member, IEEE*

***Abstract*—Zirconium oxide, a high-k gate dielectric, and molybdenum, a refractory metal, were successfully integrated into an existing submicron NMOS transistor process at RIT. Submicron high-k gate dielectric metal gate transistors were produced as a result of this project, and electrical characteristics were compared to reference submicron transistors fabricated with 75 Å silicon dioxide gate dielectrics and polysilicon gates.**

***Index Terms*—High-K, Metal Gate, NMOS, CMOS, Process Integration**

I. Introduction

OVER the last 15 years device performance has been continuously improved by scaling the thickness of the silicon dioxide gate dielectric. Currently, gate oxide thicknesses on the order of 12Å are being used for the 90nm node, with plans to scale the gate dielectric down to 8Å for the 65nm node [1]. However, continued scaling of the silicon dioxide gate will not be possible due to the lack of atomic layers of SiO_2, as well as the unacceptable levels of gate leakage current due to the extremely physically thin silicon dioxide gate dielectric.

The purpose of this project was to investigate the use of zirconium oxide, a high-k material, as a replacement to the silicon dioxide gate. The investigation involved integrating such a high-k dielectric gate material into an existing NMOS transistor process at RIT. The integration also involved the characterization and processing of a metal gate electrode to replace the typical n+ or p+ polysilicon gate electrode used today.

II. Process Development

The initial phase of this project involved creating deposition processes for zirconium oxide (ZrO_2) and molybdenum (Mo). A zirconium oxide deposition process was created by reactively sputtering in a Perkin Elmer 2400 radio frequency (RF) sputtering tool. A pure zirconium (Zr) target was used while flowing 30 sccm of Argon (Ar) gas and 4 sccm of Oxygen (O_2) gas, resulting in a deposition pressure

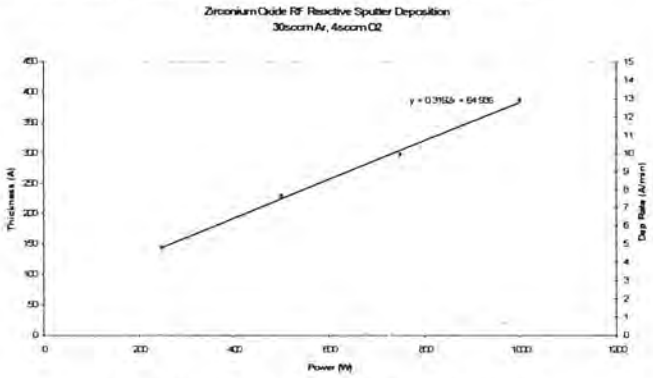

Fig. 1. Verification of the ZrO_2 Sputtering Process

of 30 mTorr. In order to verify the stability of the process, a series of four depositions were performed to compare deposition rate with sputtering power. The sputtering powers used for the zirconium oxide deposition were 250 W, 500 W, 750 W, and 1000 W, and the length of each deposition was thirty minutes with a five minute pre-sputter.

As expected, the deposition rate of the process varied linearly with sputtering power. For future work in this process, a sputtering power of 750 W was chosen because it yielded a desired film thickness of approximately 300Å. Using ellipsometry, the refractive index of the deposited ZrO_2 film at a power of 750 W was found to be 1.96 and the average thickness across the wafer of the film was found to be 297Å.

After a zirconium oxide deposition process had been

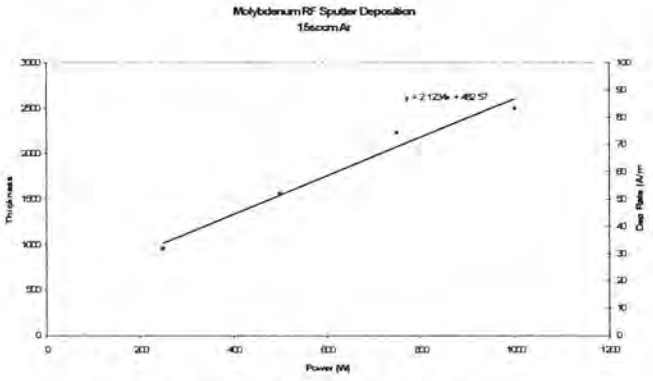

Fig. 2. Verification of the Mo Sputtering Process

This work is part of the senior design project requirement for a B.S. degree in Microelectronic Engineering at the Rochester Institute of Technology (RIT). The results of this project were presented at the 23rd Annual Microelectronic Engineering Conference on May 10, 2005 at RIT in Rochester, NY.

D. Jaeger is with the Microelectronic Engineering department at the Rochester Institute of Technology in Rochester, NY.

established, a molybdenum deposition process was created using the same RF sputtering tool. A pure molybdenum target was used while flowing 15 sccm of Argon gas, resulting in a deposition pressure of 16 mTorr. Data was collected to create a deposition rate versus sputtering power plot to verify the stability of the molybdenum process. The sputtering powers used for the molybdenum deposition were the same as for the zirconium oxide deposition and the deposition time for each was also thirty minutes with a five minute pre-sputter. Again, the deposition rate varied linearly with the sputtering power. Using an automated four-point probe, the sheet resistance of the 750 W molybdenum film was measured to be 1.72 ohm/square, while the average thickness across the wafer was found to be 2232Å using a Tencor P2 profilometer.

III. Materials Characterization

After deposition processes were created for the materials under investigation, the materials were briefly characterized

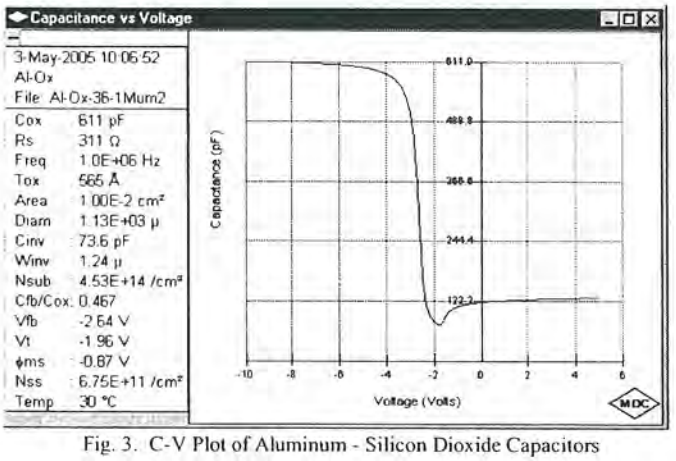

Fig. 3. C-V Plot of Aluminum - Silicon Dioxide Capacitors

using capacitance-voltage (C-V) analysis. For the purpose of comparison, a wafer with capacitors made of aluminum and 500 Å of silicon dioxide grown thermally was fabricated. Next, capacitors with molybdenum and 500 Å of silicon dioxide were fabricated. The lateral shift in the C-V

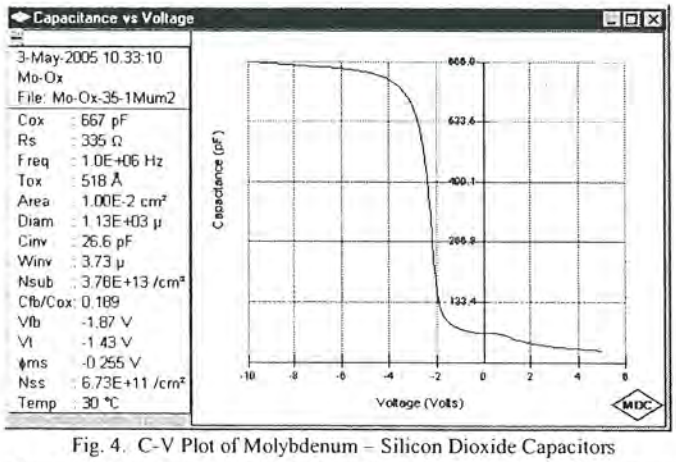

Fig. 4. C-V Plot of Molybdenum – Silicon Dioxide Capacitors

characteristic was attributed to the difference in the metal workfunction (Φ_m) for aluminum and molybdenum. Using the well established value of 4.11 eV for the metal workfunction of aluminum, the extracted metal workfunction of molybdenum was calculated to be 4.725 eV.

After the molybdenum was characterized, aluminum and zirconium oxide capacitors were fabricated in order to determine the dielectric constant of the zirconium oxide. The zirconium was deposited at 750 W yielding a thickness of approximately 300 Å. Using the difference in oxide

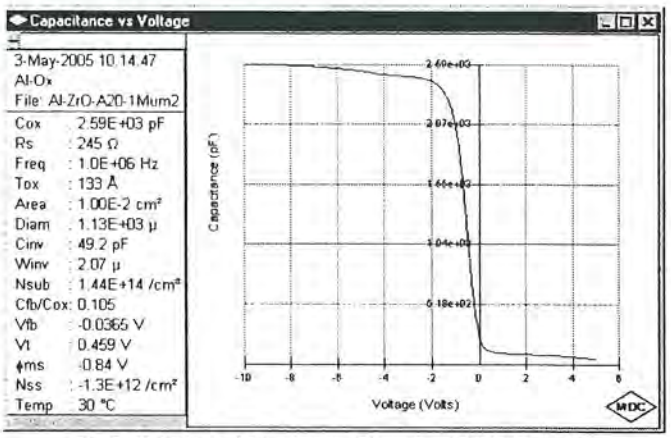

Fig. 5. C-V Plot of Aluminum – Zirconium Oxide Capacitors

capacitance obtained from each C-V plot, and also accounting for the physical thickness difference between the silicon dioxide and zirconium oxide, the dielectric constant of the zirconium oxide was determined to be 8.80. This relative permittivity is rather low for zirconium oxide as it is typically reported to have a dielectric constant between fifteen and twenty-five in literature. The most likely reason for the low dielectric constant would be the existence of a fairly thick, 50-80 Å, layer of silicon dioxide between the silicon and the zirconium oxide. Part of this interfacial layer would be due to the native thermal oxide that is grown when the silicon wafers are exposed to air, however native oxides are typically only in the range of 20-30 Å. It is possible that some oxygen diffused through the zirconium oxide during the 950° Celsius one minute anneal, forming a thicker layer of silicon dioxide at the surface. It is recommended that further work and experimentation be completed to determine the exact reason for the low dielectric constant.

IV. Transistor Fabrication

An established submicron NMOS transistor process at RIT was used as a baseline for transistor fabrication for this project. Both the baseline process and the test chip were designed by in 2004 [2]. Modifications were made to the process in hopes of further improving upon the existing process. Transistors with polysilicon gates and silicon dioxide gate dielectrics were fabricated along with the molybdenum gate and zirconium oxide gate dielectric transistors for comparison. The changes made to the

existing polysilicon and silicon dioxide process involved thinning the gate oxide from 150 Å to 75 Å, increasing the field oxide thickness from 3500 Å to 5000 Å, increasing the channel stop implant energy from 40 keV to 100 keV, and using a rapid thermal processor to activate the source and drain implants instead of a diffusion furnace. Fabricating the high-k metal gate transistors involved following the same process flow, with only a few modifications. A replacement gate process was *not* necessary for the integration of these materials into submicron NMOS transistors. The entire fabrication process is outlined below in more detail.

The process began with performing an RCA clean on the (100) p-type silicon wafers whose starting resistivity was 40 ohm-cm. Next, a 500 Å pad oxide was grown in a diffusion furnace and the p-well was implanted through this pad oxide using a dose of $2x10^{13}$ cm^{-2} of B^{11} at an energy of 40 keV. The well was then driven in at 1025° Celsius for 60 minutes in N_2. 3200 Å of silicon nitride was deposited via an LPCVD furnace for the LOCOS isolation scheme utilized in this process. The first photolithography level (Active) was completed, and the silicon nitride was plasma etched. A channel stop implant of B^{11} was implanted for a dose of 8e13 cm^{-2} at an energy of 100keV. The photoresist from the active lithography was then removed in an oxygen plasma, and 5000 Å of silicon dioxide was thermally grown in a diffusion furnace. A quick buffered oxide etch (BOE) was performed to remove any oxynitride layer that may have formed in the silicon nitride, and then hot phosphoric acid was used to remove the remaining silicon nitride, followed by another BOE step to remove the underlying pad oxide. After the pad oxide was etched, a 250 Å kooi oxide was thermally grown. Since no threshold adjustment implant was used in this process, the kooi oxide was then removed in a buffered oxide etch, followed by another RCA clean to prepare the wafers for the gate dielectric. At this point in the process, the wafers were split into two groups; one which would receive the traditional thermal oxide and polysilicon gate stack, and one that would receive the metal gate high-k gate stack. A TCL clean was performed on the furnace tube used for the gate oxidation growth, and then 75 Å of thermal oxide was grown as the gate dielectric on the reference wafers with polysilicon and silicon dioxide. These wafers then had 2500 Å of polysilicon deposited upon them via an LPCVD furnace. Meanwhile, 300 Å of zirconium oxide followed by 2300 Å of molybdenum was sputtered onto the high-k metal gate transistor wafers. It is important to note that these materials were deposited onto the wafers in the same tool without breaking vacuum. Once each set of wafers had its appropriate gate stack, the second photolithography level was completed defining the gate areas. The gate materials were then dry etched, each in different plasma etchers. The sources and drains were then formed by implanting P31 for a dose of $2x10^{15}$ cm^{-2} at 25 keV for the polysilicon silicon dioxide wafers, and for a dose of $3x10^{15}$ cm^{-2} at 33 keV for the molybdenum zirconium oxide wafers. 3000 Å of oxide deposited via PECVD using a TEOS precursor was put onto all of the wafers to serve as an inner-layer dielectric. The source and drain implants were then activated using the rapid thermal processor. The recipe used in the rapid thermal processor ramped the temperature up at 50° Celsius per second, soaked at 950° Celsius for one minute, and then ramped back down. This process was simulated to provide adequate dopant activation as well as diffusion to form the source and drain regions of the transistors. The third photolithography level (contact cut) was performed and the oxide was etched in a buffered oxide etch with surfactants. The surfactants allowed the etchant to get into the very small contact cut holes defined by the photoresist. The contact cuts on the wafers with the traditional materials etched all the way down to the bare silicon above the sources and drains, however the wafers with the zirconium oxide did not. This was due to the extremely slow etch rate of the zirconium oxide in buffered oxide etch. It was obvious that the inner-layer dielectric would etch away before the zirconium

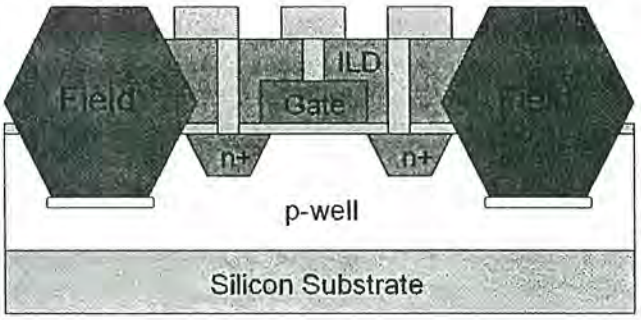

Fig. 6. Final NMOS Transistor Cross-Section

oxide was completely removed in the contact areas so a different approach had to be taken. The zirconium oxide was removed by using a dry etcher and flowing argon gas at a power of 600 W. This essentially "sputtered" the zirconium oxide off of the wafers. In the process of doing so, the photoresist was damaged and became extremely powdery. Extensive photoresist stripping had to be performed in order to remove all of the residue. Following photoresist removal, all of the wafers had 6000 Å of aluminum sputtered onto them. The metal was patterned using the fourth and final photolithography level, and was wet etched in aluminum etchant. Finally, the wafers were sintered in a diffusion furnace at a temperature of 450 ° Celsius for approximately twenty minutes.

V. Electrical Performance of Transistors

The first electrical test performed on the polysilicon gate silicon dioxide gate dielectric transistors was stepping the

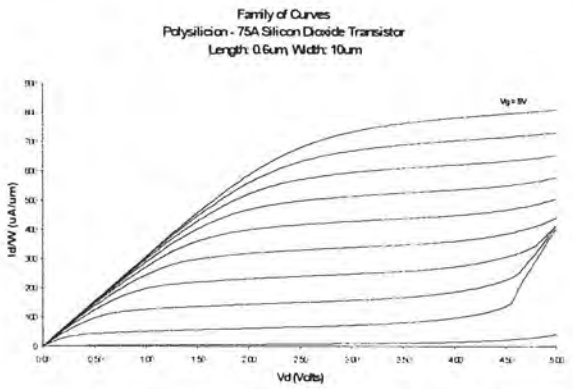

Fig. 7. Family of Curves for Poly – SiO_2 Transistor

gate voltage while sweeping the drain voltage and measuring the drain current, producing a family of curves. The

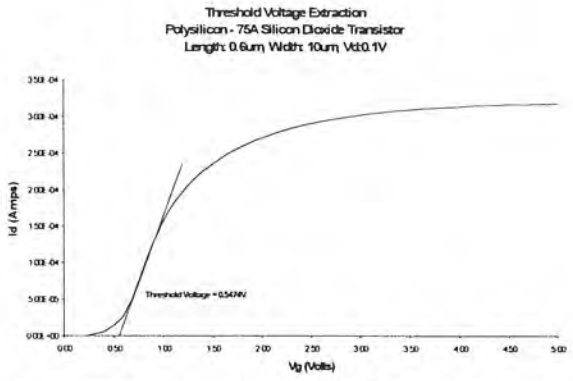

Fig. 8. Threshold Voltage Extraction for Poly– SiO_2 Transistor

transistor tested in Fig. 7 had a physical gate length of 0.6 um and showed excellent drive current, approximately 800

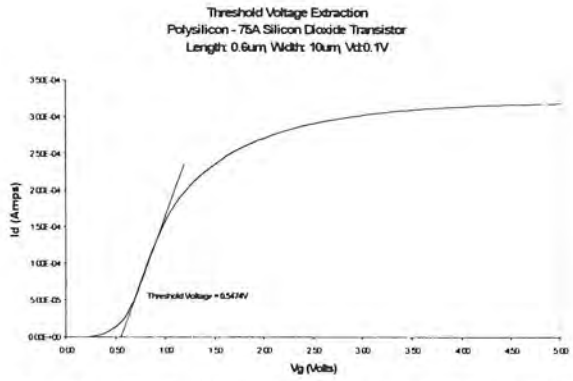

Fig. 9. Sub-Threshold Characteristic for Poly– SiO_2 Transistor

uA/um when normalized to the 10 um width of the transistor. In the next electrical test, the drain voltage was held constant at 0.1 V while the gate voltage was swept and the drain current was measured. This allowed the threshold voltage to be extracted as 0.5474 V. The same test also allowed for the sub-threshold slope to be determined by

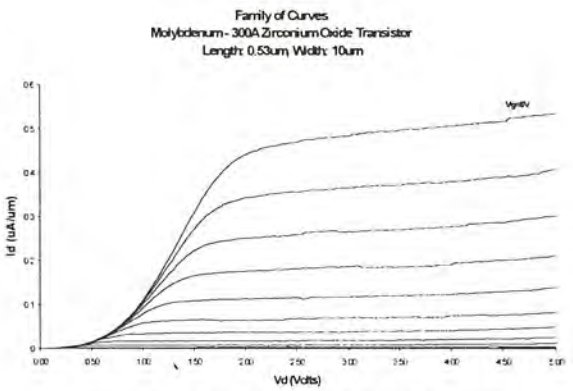

Fig. 10. Family of Curves for Moly– ZrO_2 Transistor

plotting the drain current on a log scale. The sub-threshold slope for the polysilicon silicon dioxide transistor was 96.64

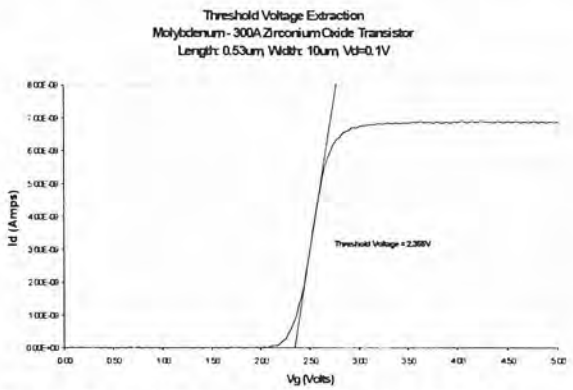

Fig. 11. Threshold Voltage Extraction for Moly– ZrO_2 Transistor

mV/decade. The same set of electrical tests were then performed on the high-k metal gate transistors. The

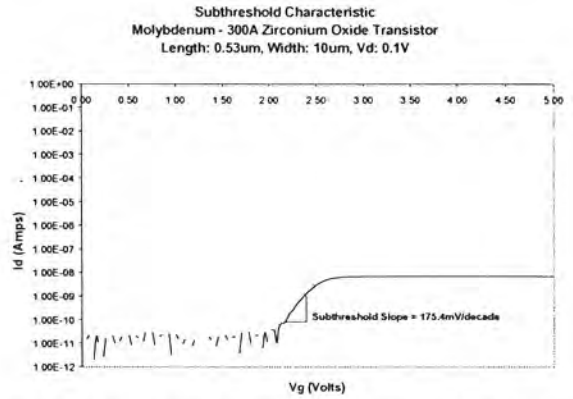

Fig. 12. Sub-Threshold Characteristic for Moly– ZrO_2 Transistor

Transistors definitely worked, as they produced a typical family of curves, however it appeared as though a large amount of series resistance was present. The drive currents for the molybdenum zirconium oxide transistors were orders of magnitude less than the reference polysilicon silicon dioxide transistors. After analyzing various Van der Pauw structures

in different layers on the test chip, it was determined that the sheet resistance of the source and drain regions in the high-k metal gate transistors was over 2900 ohm/square instead of 50 ohm/square as in the reference transistors. This suggests that the zirconium oxide blocked most of the source and drain implant, which was not anticipated. The threshold voltage was found to be 2.355 V which is somewhat high, but not unreasonable for a first run of these materials at RIT. One would expect a fairly high threshold voltage when sputtering the gate dielectric and not implementing a threshold adjust implant. The sub-threshold slope for these transistors was found to be 175.4 mV/decade, also not as good as the polysilicon gate silicon dioxide gate dielectric transistors. The molybdenum gate zirconium oxide gate dielectric electrically tested had a physical gate length of 0.53 um and a width of 10 um.

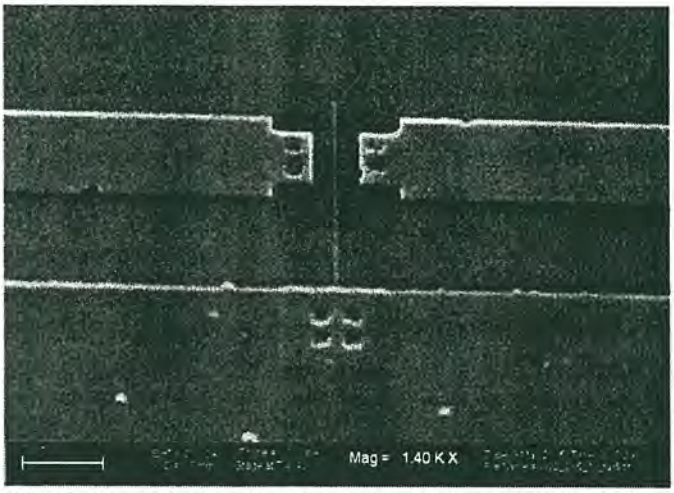

Fig. 13. Mo – ZrO_2 Transistor with Length of 0.53 um and Width of 10 um

VI. Conclusion

Over the course of this project, deposition and etching processes for both molybdenum and zirconium oxide were developed. Both of these materials went through basic characterization; however there is still much more work to be done in this area. Submicron NMOS transistors with a 75 Å silicon dioxide gate dielectric and polysilicon gate were successfully fabricated and possessed high drive current, low sub-threshold slope, and a reasonable threshold voltage of close to 0.5 V. Most importantly, RIT's first sub-micron NMOS transistors with high-k gate dielectrics and metal gates were fabricated and successfully tested. These transistors may not have had the best electrical characteristics, due to the blocked source and drain implant, but they still behaved as transistors, producing an excellent family of curves.

Acknowledgment

The author thanks Michael Aquilino, Dr. K. Hirschman, and Dr. S. Rommel for their guidance in this project. Also acknowledged are Eric Woodard, David Pawlik, Robert Manley, Robert Saxer, Lance Barron, Robert Mulfinger, Raymond Krom, Robin Joyce, Thomas Grimsley, John Nash, Scott Blondell, Sean O'Brien, Bruce Tolleson, and David Yackoff. In addition, the author specifically acknowledges the staff of the Rochester Institute of Technology Semiconductor and Microsystems Fabrication Laboratory for technical assistance and equipment support.

References

[1] R. Chau. *Advanced Metal Gate/High-K Dielectric Stacks for High-Performance CMOS Transistors.* Components Research, Logic Technology Development, Intel Corporation.

[2] M. Aquilino. *Advancing RIT to Submicron Technology: Design and Fabrication of 0.5um N-Channel MOS Transistors.* 22nd Annual Microelectronic Engineering Conference, May 2004.

[3] R. Chau, S. Datta, M. Doczy, J. Kavalieros, M. Metz. *Gate Dielectric Scaling for High-Performance CMOS: from SiO_2 to High-K.* Intel Corporation.

[4] Y. Ma, Y. Ono, L. Stecker, D. Evans, S.T. Hsu. *Zirconium Oxide Based Gate Dielectrics with Equivalent Oxide Thickness of Less Than 1.0nm and Performance of Submicron MOSFET using a Nitride Gate Replacement Process.* Sharp Laboratories of America.

[5] R. Kwasnick, E. Kaminsky, P. Frank, G. Franz, K. Polasko, R. Saia, T. Gorczya. *An Investigation of Molybdenum Gate for Submicrometer CMOS.* IEEE Transactions on Electron Devices, Vol. 35, No. 0, September 1988.

Daniel J. Jaeger, originally from Canton, MI, received a B.S. degree in Microelectronic Engineering from the Rochester Institute of Technology located in Rochester, NY in May of 2005. He obtained co-op work experience at Integrated Nano-Technologies in Henrietta, NY and is currently completing the M.S. portion of the B.S./M.S. dual degree program in Microelectronic Engineering and Materials Science at the Rochester Institute of Technology.