Metal Gate Technology for Fully Depleted SOI CMOS

Pushkar Ranade†, Yang-Kyu Choi, Daewon Ha, Hideki Takeuchi and Tsu-Jae King

†Department of Materials Science and Engineering, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, USA

ABSTRACT

This paper reviews recent approaches in the development of a tunable work function metal gate CMOS technology and describes the application of one such approach to the fabrication of metal gate fully-depleted (FD) SOI transistor structures such as the ultra-thin body (UTB) FET and the FinFET.

INTRODUCTION

Continued scaling of CMOS technology beyond the 90 nm technology node is one of the most difficult challenges facing the microelectronics industry today. Research and development efforts in this area will most likely lead to fundamental paradigm shifts; initially in the transistor materials used and subsequently in the device architecture. The identification of compatible high-permittivity ("high-k") gate dielectrics and metal gate electrode materials dominates the former while the development of SOI-based transistor designs are expected to dominate the latter. It should be noted however that most of the promising non-classical FET structures reported to date will also rely on advanced materials and processes for optimum performance.

FD-SOI transistors offer significantly improved control of off-state leakage current, since sub-surface leakage paths between the source and the drain are effectively eliminated by the buried oxide [1]. With short-channel effects controlled effectively by the use of a very thin Si body, it is possible to keep the channel regions undoped. This offers the benefits of immunity from dopant-fluctuation-induced variations in threshold voltage V_T, improved carrier mobility, and lower gate tunneling current [2]. However, it requires that V_T be adjusted by gate work-function engineering, in the range 4.4-5.0V [3]. This implies that metal gate materials will be needed in order to optimize the performance of future FD-SOI CMOS technology. The use of metallic gate electrodes also eliminates the problems of gate depletion and dopant penetration through thin gate dielectrics, and provides lower gate resistance, for improved performance and reliability.

The integration of metallic gate materials into a CMOS process is not trivial. To achieve self-aligned source/drain regions without significantly increasing process complexity, a "gate first" fabrication approach is preferred. Thus, the gate metal should have a high melting point and should be thermodynamically stable on the gate dielectric at high temperatures (at least up to 1000°C). It also should have a low resistivity and a thermal expansion coefficient comparable to that of the Si substrate, to avoid issues with thin-film stress. Perhaps the most important requirement for a metallic gate material is that it must have a suitable work function. Since the requirements for NMOS and PMOS FETs are different, new work-function engineering approaches are needed to provide multiple gate work functions with minimal process complexity.

This paper discusses approaches to metal gate work-function engineering, and describes the application of a tunable-work-function molybdenum gate technology to the fabrication of lightly doped UTBFETs and FinFETs.

WORK FUNCTION ENGINEERING APPROACHES

There are at least three general approaches to modifying the work function of thin metal films. Structural modification can include changes to the crystalline phase (*e.g.* fcc to bcc), changes to the preferred orientation or a change in the long-range crystalline order (*e.g.* amorphization). Chemical modification can include a selective nitridation or oxidation (provided the compounds have reasonably high conductivity). Chemical modification can also be achieved through alloy formation, metal interdiffusion or metal silicidation. Alternatively, structural and chemical changes can be simultaneously incorporated through the use of ternary compounds like TaSiN or WSiN which are known to remain amorphous over a fairly wide temperature range.

Recently, Misra *et al.* have shown that the work function of binary alloys of Ru and Ta depends strongly on the composition of the alloy [4]. Alloys with less than 20 at. % Ta demonstrated work functions greater than 5 V while those with larger Ta content displayed work functions between 4.4 V and 4.2 V. This material system is thus promising for bulk-Si as well as FD-SOI CMOS applications. A possible integration approach could use bilayer stacks of Ru and Ta as the gate electrodes for NMOS and PMOS transistors. A thermal anneal can be used to intermix the films and form alloys, with the final electrode composition controlled by varying the thickness ratio of the films. Binary alloys of Ta with Pt and Ti have also been investigated for CMOS gate applications and demonstrate a wide range of work functions [5].

In some bilayer metal systems, interdiffusion is preferred over chemical alloying (at low temperatures). Bilayers of Ni (work function ~5 V) and Ti (work function ~4 V) have been used for MOSFET fabrication by Polishchuk *et al.* [6]. Ti capped with Ni was used as a PMOS gate electrode. As deposited, the gate work function of the bilayer was identical to that of Ti, which

is too low for PMOS application, but well suited for NMOS application. Upon thermal annealing (400°C, 10 min), the Ni segregated to the dielectric interface, increasing the gate work function to be equal to that of Ni, well-suited for PMOS application. Dual-work-function metal gates for CMOS application can achieved by selective removal of the Ni capping layer in the NMOS regions prior to thermal annealing, to prevent a change in work function in these regions.

Nitridation of thin molybdenum (Mo) gate films by implantation of N has been shown to be an effective way to lower the gate work function [7]. By controlling the implant dose, energy and subsequent annealing temperature, the work function of the film can be modulated between 4.4 V and 5.0 V, making it suitable for FD-SOI CMOS application. Another advantage of such a process is the relative ease of integration into conventional CMOS process flows. The application of this approach to UTBFET and FinFET fabrication is described in the next section. Another way to introduce N into thin Mo films is through the use of a sacrificial diffusion source like TiN. Lander *et al.* [8] have shown that substantial amounts of N can be introduced into thin Mo films from N-rich TiN capping layers. After high temperature annealing, the N was observed to segregate at the dielectric interface, resulting in a reduction in the gate work function. Such an approach will be relatively difficult to integrate; however it does provide immunity from potential implant induced damage to the gate dielectric.

Recent studies have also demonstrated the use of metal silicides as CMOS gate electrodes. Kedzierski *et al.* [9] used doped Ni-silicide gate electrodes to fabricate FinFETs with various values of V_T. This technique is potentially easy to integrate since it can be accomplished through a simple extension of the conventional self-aligned silicide (salicide) process.

MO GATE UTBFET FABRICATION

The UTBFET structure is considered to be a promising alternative to the conventional bulk-Si FET for gate lengths less than 15 nm [10]. This structure was first demonstrated by Choi *et al.* [11] using a highly scaled Si body thickness (<5 nm) and was shown to provide excellent control of short-channel effects with reasonable drive current. As explained earlier, a tunable-work-function gate electrode technology is highly desirable for V_T control in fully-depleted SOI transistor structures such as the UTBFET. Mo-gate PMOS UTBFETs were fabricated in order to demonstrate the efficacy of N implantation for V_T adjustment [12]. Figure 1 shows a schematic cross-section of the Mo gate UTBFET along with a cross-sectional TEM image of the gate stack. The Si body is lightly doped and thin (10 nm). The gate stack consists of thermal SiO_2 as the dielectric (2 nm), a thin layer of Mo (40 nm) capped with *in-situ* n+ doped poly-Si and a SiO_2 (LTO) hard mask. Some devices were implanted with N after Mo deposition, to adjust V_T.

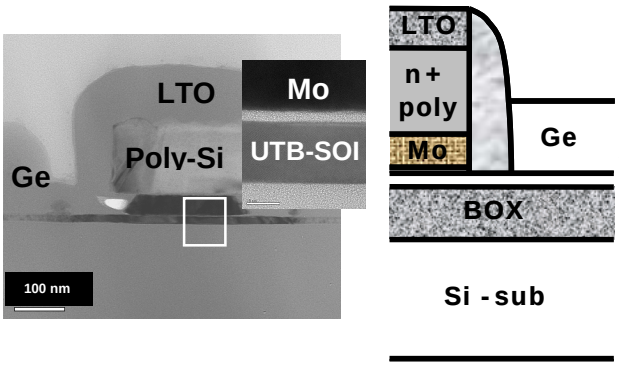

Figure 1: Schematic cross-section of Mo-gated UTBFET and corresponding high-resolution TEM of an actual device.

The purpose of the poly-Si capping layer was to provide mechanical support to the thin Mo layer and also to prevent dopant ions from penetrating the metal gate during the self-aligned S/D ion implantation step. After gate-sidewall spacer formation, Ge (60 nm) was selectively deposited to form elevated-S/D structures for reduced parasitic resistance. Afterwards, SiO_2 was deposited and the wafer was implanted with B ions to dope the Ge. A 700°C, 60 s anneal was used to activate the dopant atoms. Figure 2 shows the measured I_D-V_G characteristics of PMOSFETs fabricated using this process. It is evident that the unimplanted Mo gate device displays the desired low V_T for PMOSFETs (–0.2 V). The effect of N implantation is to lower the Mo gate work function, which results in a higher $|V_T|$, as shown in Figure 2.

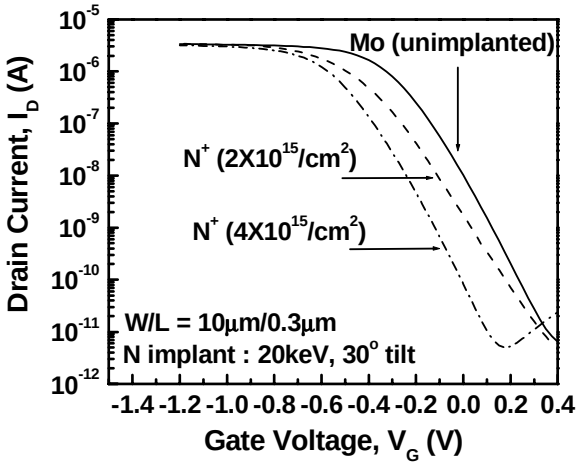

Figure 2: I_D-V_G characteristics of Mo-gated UTBFETs. The effect of N implantation on V_T is clearly seen.

It should be noted that V_T can be readily adjusted by changing the N implant dose. Depending upon the post-implant thermal annealing conditions, a shift of ~0.1 V for every 1×10^{15} cm^{-2} increment in dose can be achieved [7]. NMOS devices with heavily N implanted Mo gates are expected to display low V_T since the gate work function would be appropriate for such devices.

MO GATE FINFET FABRICATION

Another promising structure that can allow scaling of Si-based CMOS technology to sub-10 nm gate lengths is the FinFET. The FinFET is a vertical self-aligned double-gate structure wherein a gate electrode straddles a narrow Si fin on either side. In the first demonstration of a tunable-work-function metal gate FinFET technology, Mo gate PMOS FinFETs were fabricated using aggressively scaled fin widths (10 nm) with light doping [13]. Figure 3 is a schematic view of the FinFET gate stack showing the large-angle tilted ion implantation that was used to introduce N into the Mo gate along the vertical sidewalls of the Si fin. As shown in Figure 4, the V_T of an unimplanted Mo gate device is –0.2 V, as desired. N implantation controllably increases the $|V_T|$. Again it should be noted that a low gate work function achieved by high-dose N implantation would be ideal for achieving low V_T in NMOS FinFETs. In addition, multiple values of V_T can be achieved on a single chip by using different implant doses in different regions. Such a CMOS technology is desirable for applications where high performance (high drive current, low V_T) and low power (low off-state leakage, high V_T) transistors are needed on a single chip.

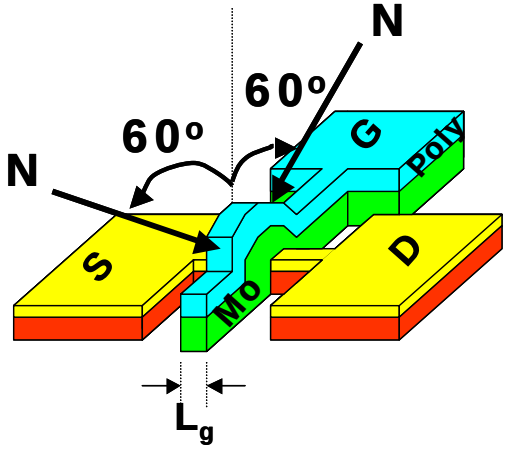

Figure 3: Schematic view of a FinFET with the poly-Si/Mo gate stack used. N implantation was performed twice to introduce N into both the front and back gates.

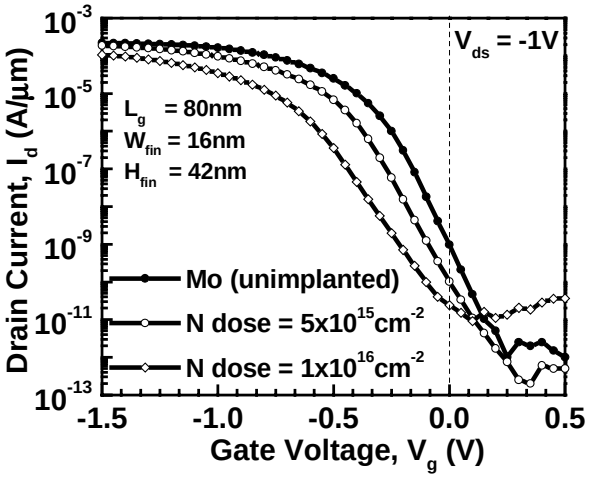

Figure 4: I_D-V_G characteristics of Mo-gated FinFETs. Nitrogen implantation into the Mo gate can controllably alter the device V_T.

SUMMARY

Advanced transistor structures such as FD-SOI single-gate or double-gate MOSFETs will require tunable work function gate technology for optimal performance with minimal process complexity. This paper discussed various approaches to gate work function engineering. A tunable-work-function Mo gate technology is shown to be effective for adjusting V_T in UTBFETs and FinFETs.

ACKNOWLEDGEMENTS

The authors would like to acknowledge financial support from SRC-MARCO under contract 2001-MT-887. Devices were fabricated in the UC Berkeley Microfabrication Laboratory. Ion implantation support from Axcelis Technologies Inc. is gratefully acknowledged.

REFERENCES

1. B. Yu *et al.*, *1997 ISDRS*, p.623, December 1997
2. L. Chang *et al.*, *IEDM Tech. Digest*, p.99, 2001
3. L. Chang *et al.*, *IEDM Tech. Digest*, p.719, 2000
4. V. Misra *et al.*, *IEEE Elect. Dev. Lett.*, **23**, p.354, 2002
5. C.-F. Huang *et al.*, *Ext. Abstracts 2002 SSDM*, p.184
6. I. Polishchuk *et al.*, *IEEE Elect. Dev. Lett.*, **23**, 200, 2002
7. P. Ranade *et al.*, *IEDM Technical Digest*, 2002
8. R.J.P. Lander *et al.*, *Proc. ESSDERC*, p.103, 2002
9. J. Kedzierski *et al.*, *IEDM Technical Digest*, 2002
10. R. Chau *et al.*, *IEDM Tech. Dig.*, p.621, 2001
11. Y.-K. Choi *et al.*, *IEEE Elect. Dev. Lett.*, **22**, p.447, 2002
12. D. Ha *et al.*, *Jpn. Journal of Appl Sci.*, April 2003
13. Y.-K. Choi *et al.*, *IEDM Technical Digest*, 2002