RC23900 (W0603-032) March 3, 2006
Electrical Engineering

IBM Research Report

Simulation Study of a Metal-gate / High-k Gate Stack for Low-Power Applications

Arvind Kumar, Paul M. Solomon
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. Ithas been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distributionoutside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g, payment of royalties). Copies may be requested from IBM T. J. Watson Research Center , P. O. Box 218, Yorktown Heights, NY 10598 USA (email: reports@us.ibm.com). Some reports are available on the internet at http://domino.watson.ibm.com/library/CyberDig.nsf/home.

Simulation Study of a Metal / High-κ Gate Stack for Low-Power Applications

Arvind Kumar and Paul M. Solomon
IBM Semiconductor Research and Development Center (SRDC), T.J. Watson Research Center,
P.O. Box 218, Yorktown Heights, NY 10598; email: arvkumar@us.ibm.com, phone 914-945-3786, FAX: 914-945-2141

I. INTRODUCTION

Conventional polysilicon gates (PG) suffer from carrier depletion which limits the effective scaling of the gate insulator stack. As a result, metal gates (MG) are currently of great interest as a means to continue device scaling through reduced effective gate insulator thickness [1-6]. Their competitiveness relative to PG, however, depends strongly on the MG work function as well as the off-state leakage target (I_{soff}) of the application desired [7,8]. If a near-midgap gate is used for a high-performance application, the required reduction in channel doping to attain the same I_{soff} leads to a buried channel device due to the weak confining field at the surface [7-9]. On the other hand, the intrinsically high V_t of near-midgap gates makes them well-suited for low-power applications [8]. Fig. 1 illustrates this basic concept by showing how a low-power application (I_{soff} =300 pA/μm) with near-midgap MG requires channel doping comparable to that of a high-performance one (I_{soff} =300 nA/μm) with PG, thereby restoring the confinement needed to avoid a buried-channel device.

The purpose of this work is to study the competitiveness of a near-midgap metal/high-κ gate stack for a low-power application using mixed-mode simulations of inverter delay chains. We address key design issues including (1) choice of gate stack based on leakage requirements; (2) matching of rolloff characteristics between PG and MG; (3) effect of potential mobility degradation due the high-κ gate stack; and (4) junction leakage for ultralow power requirements.

II. SIMULATION RESULTS

Table 1 shows the three gate stacks on bulk Si studied in this work. We consider a low-power technology with an off-state leakage target of I_{soff}=300 pA/μm for both nFET and pFET, at V_{dd}=1 V and 27C. The gate stacks have been chosen to each have a gate leakage of ~12% of I_{soff}, or ~0.1 A/cm^2 based on Ref. [6]. Cases M1,M2 have workfunctions shifted 0.2 eV from midgap towards the conduction (valence) band edge for nFET (pFET). The gate lengths of cases P (L_{nom}=47 nm) and M1 (L_{nom}=42 nm) were chosen using the method in Ref. [8]; the 5 nm difference reflects the thinner effective gate insulator of case M1. Case M2 is intended to study the effect of increasing the gate length to match the rolloff of case P.

To compare performance, 5-stage inverter delay chains (β-ratio of 2) were simulated using mixed-mode FIELDAY with quantum-mechanical corrections [10] to accurately model carrier confinement. Fig. 2 shows the main result of this work, plotting quiescent leakage current (I_{ddq}) as a function of unloaded delay per stage, for the three cases. Case M1 offers a delay reduction of about 25% for gate lengths between -3σ and nominal.

Although case M1 offers substantial speedup compared to case P, it does so with markedly poorer rolloff which arises from the lower halo doping. Case M2 offers one possible solution, in which the gate length is increased by 13 nm, until its rolloff in I_{ddq} matches that of case P. *Somewhat surprisingly, there is only a slight (~1%) performance penalty for this increase in gate length.*

Table 2 shows the effective inverter capacitance (C_{eff}) and effective inverter drive current (I_{eff}) extracted from unloaded vs. loaded delay. The speedup is primarily due to the high increase in drive current. Case M1 has a C_{eff} only slightly higher than case P because of its higher V_t as well as its lower junction capacitance which offsets the gate capacitance increase [8]. The transfer curves of the nFET shown in Fig. 3, with key parameters in Table 2, confirm the drive current gains seen for MG. Furthermore, the reduced DIBL of case M2 compared to M1 leads to a higher FET on-current (I_{dsat}), consistent with the load capacitance analysis which finds higher I_{eff} (offset by higher C_{eff}) for case M2.

Mobility degradation is a possible concern for high-κ gate stacks [11]. Fig. 4 shows the effect of mobility degradation for cases P and M2 (both phonon mobility and saturation velocity are reduced). Mobility reductions of 10% and 20% result in delay increases of 8% and 19%, respectively. However, significant speedup relative to case P is still observed.

A final consideration for ultralow power applications is the junction leakage current (I_j). To understand its impact, we repeated our study using I_{soff}=10 pA/μm (and 0.1 nm thicker oxide). The lower doping level of the MG cases results acceptably in I_j<4 pA/μm while case P has a junction leakage of 22 pA/μm, exceeding the I_{soff} target.

III. CONCLUSION

We have shown that a near-midgap metal/high-κ gate stack demonstrates strong advantages for low-power applications. Moreover, the poorer rolloff arising from lower halo doping in the metal gate case can be overcome by increasing the gate length with only minor performance penalty. The advantage is sustained even when mobility degradation is accounted for. Lowering the channel doping also benefits junction leakage, which should become a more severe limitation with increased scaling as band-to-band tunneling begins to dominate.

Acknowledgments: The authors thank X. Wang, P. Oldiges. S. Fischer, and S. Furkay for FIELDAY support; M. Khare, T. Hook, and W. Haensch for discussions; and S. Fang for process conditions.

References:

[1] B. Tavel *et al.*, *IEDM Tech. Dig.* 2001, p. 825.
[2] W.P. Maszara *et al.*, *IEDM Tech. Dig.* 2002, p. 367.
[3] Z. Krivokapic *et al.*, *2003 Symp. VLSI Tech. Dig. Papers*, p. 131.
[4] J. Kedzierski *et al.*, *IEDM Tech. Dig.* 2003, p. 315.
[5] K.G. Anil *et al.*, 2004 *Symp. VLSI Tech. Dig. Papers*, p. 190.
[6] E. Gusev *et al.*, *IEDM Tech. Dig.* 2004, p. 79.
[7] Y. Abe *et al.*, *IEEE Elec. Dev. Lett.*, vol. 20, p. 632, 1999.
[8] A. Kumar and P.M. Solomon, to appear in *IEEE Trans. Elec. Dev.*
[9] A. Kumar and R.H. Dennard, submitted to *IEEE Elec. Dev. Lett.*
[10] M. Ieong *et al.*, *Proc. SISPAD* 1998, p. 129.
[11] M.V. Fischetti *et al.*, *J. Appl. Phys.*, vol. 90, p. 4587, 2001.

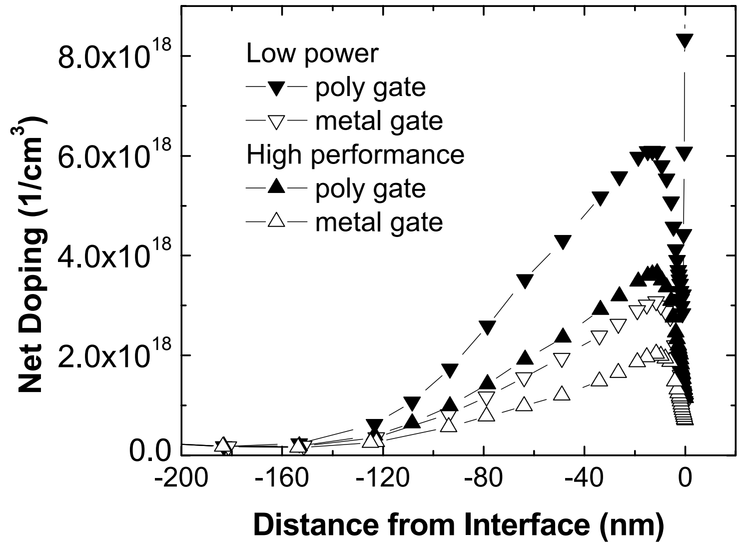

Fig. 1: Comparison of channel doping for polysilicon and near-midgap metal gates required to meet high-performance and low-power off-state leakage targets. Note that the doping for a metal gate for a low-power application is comparable to that for polysilicon gate for a high-performance application so that the confinement is restored at the lower off-current target.

case	Gate electrode	Gate dielectric	Equiv. SiO_2 thickness (nm)	L(-6σ) (nm)	L(-3σ) (nm)	L(nom) (nm)	L(+3σ) (nm)
P	poly	2.3nm SiON	1.6	37	42	47	52
M1	metal	0.6 nm SiO_2/2.5 nm HfO_2	1.1	32	37	42	47
M2	metal	0.6 nm SiO_2/2.5 nm HfO_2	1.1	36	50	55	60

Table 1: Gate dielectric stacks used in this work. The stacks are chosen to have gate leakage ~0.1 A/cm^2, about 12% of the off-current leakage target. Also shown are the gate lengths simulated for each stack, chosen using method in Ref. [8] for P and M1. Gate length for case M2 is chosen to match rolloff of case P.

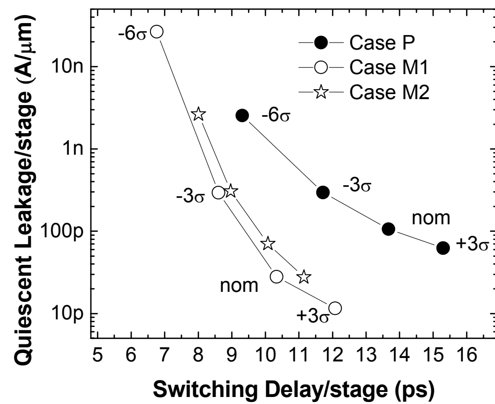

Fig. 2: Comparsion of quiescent leakage current vs. delay characteristic, showing significant advantage for metal gate cases. When the metal gate length is increased, only minor degradation in performance is observed.

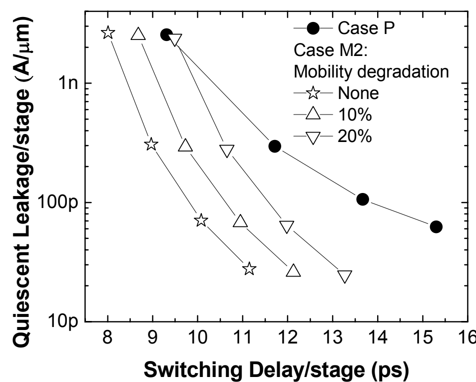

Fig. 3: Effect of mobility degradation on quiescent leakage current vs. delay characteristic. Even with mobility degradation, near-midgap metal gate is seen to outperform polysilicon gate.

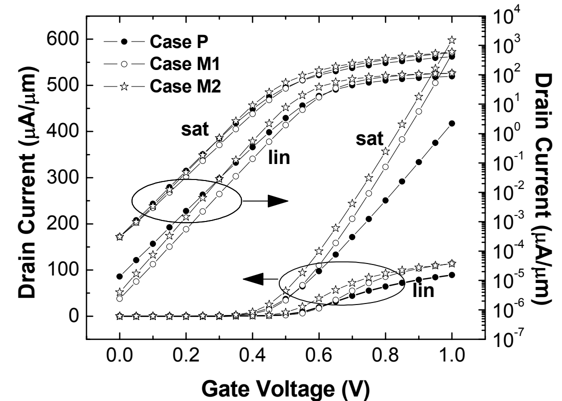

Fig. 4: Comparison of nFET transfer characteristics using devices with gate length $L_{-3\sigma}$. Key parameters are summarized in Table 2.

case	Eff. cap. C_{eff} (fF/μm)	Eff. drive current Ieff (μA/μm)	nFET Vtlin (V)	nFET DIBL (mV/V)	nFET SSsat (mV/dec)	nFET Gmlin (μS/μm)	nFET Gmsat (μS/μm)	nFET Idsat (μA/μm)
P	1.71	146	0.538	142	89	248	841	417
M1	1.78	207	0.561	144	92	374	1240	568
M2	1.98	220	0.490	113	85	531	1230	598

Table 2: 5-stage inverter delay chains with nFET and pFET widths of 45 nm and 90 nm, respectively, were simulated. Delay chains are either unloaded (τ_U) or loaded (τ_L) with capacitance C_L=1.2 fF/μm. Effective inverter capacitance is obtained from $C_{eff}=C_L\tau_U/(\tau_L-\tau_U)$ and effective inverter drive current is $I_{eff}=C_{eff}V_{dd}/\tau_U$. Basic device properties of the nFETs at gate length $L_{-3\sigma}$ are also compared.