Gate Stack Technology for Next 25 Years

Hiroshi Iwai
Frontier Collaborative Research Center, Tokyo Institute of Technology, iwai.h.aa@m.titech.ac.jp

Introduction

Among the three main components of the transistors, – gate stack, source/drain, and channel – the gate stack has been the most sophisticated and sensitive part for their performance, yield and reliability, and huge R&D efforts have been focused on the gate stack technologies on every generation of MOSFET development, since their beginning as a single Si-MOSFET in 1960.

In this paper, historical development for gate stack technologies for the past 40 years are reviewed and, then, current issues for the gate stack are described. Finally, gate stack technologies for next 25 years are predicted.

Past overview

Looking back over the history of Si MOSFETs, there have been so many issues to be solved. In 1960's, threshold voltage instability was a critical issue for the MOSFETs to be used as commercial devices. The cause was found to be the mobile ions such as sodium involved in the gate SiO_2 in the middle of 1960's. At the beginning of 1970's, gate electrode was changed from aluminum to polysilicon so that the self alignment of source/drain to the gate stack was established. In 1970's and 80's, yield degradation caused by defect in SiO_2 gate films were major concern for thinning the gate oxides for every generation. The defects problems were considerably reduced by the control of dusts in the clean room and purity of the oxygen gas. From the middle of 1970's to the end of 1990's until the supply voltage approaches 1V, hot carrier reliability was a big concern and many different gate-sidewall scheme were introduced to provide LDD (Lightly Doped Drain) or GDD (Graduated Doped Drain) structures in order to decrease the lateral field and hence to suppress the hot carrier injection to the gate oxides. Until the middle of 1990's when the beautiful transistor operation of 1.5 nm thick gate SiO_2 MOSFETs were demonstrated 1990's, most of the engineers had not believed that gate oxide less than 3 nm works as gate insulator. The success of thinning the gate oxides less than 3 nm opened the way to reduce the gate length less than 50 nm, and the roadmap for the downswing had been quite accelerated.

However, there had been bitter arguments for the time-dependent dielectric breakdown (TDDB) for such thin gate oxide whether the TDDB of gate oxides less than 2 nm would meet the the long term reliability required for commercial devices. The argument tended to cease when 1.2 nm thick gate oxides were used in the product of leading semiconductor companies.

Until the middle of 1990's, the gate insulator, the gate insulator had been SiO_2 for 35 years since the beginning of MOSFETs, and the gate electrode has been N^+-poly Si for both N-and P-MOS transistors for 25 years. However, P^+-poly Si gate electrode became necessary for small gate length P-MOS around 0.25 μm to reduce the short channel effects, and gate insulator material had to be changed from SiO_2 to SiO_xN_y in order to suppress the boron penetration from the heavily boron-doped P^+-poly Si gate electrode to the silicon substrate though SiO_2 gate insulator during high-temperature heat process for the LSI manufacturing. So, after 35 years since the 1st MOSFET had been produced, SiO_2 was slightly modified with smaller amount of nitrogen.

Silicide on the poly-Si gate electrode has been used to reduce the gate electrode resistance. The materials changed from $MoSi_2$, WSi_2, $TiSi_2$, $CoSi_2$, NiSi, depending on the adaptability the silicide material to the process temperature, gate electrode line width (or gate length) and source/drain junction depth.

High-k gate dielectrics were thought to be introduced relatively easily, when good transistor operation of ZrO_2 and HfO_2 gate insulator MOSFETs were demonstrated at the end of 1990's. However, there have been so many severe problems – although the majority of the problems had been the ones already anticipated before --, and it took almost 10 years until major semiconductor companies

announced the introduction of the HfO_2 based high-k gate dielectrics in this and next years.

Gate poly-Si depletion is a big problem for thinning the effective or equivalent oxide thickness, and now poly-Si gate electrodes are being changed to metal or fully-silicided gate electrode with combination to the high-k gate dielectrics.

Current issues and next 25 years

So far HfO_2 based gate dielectrics are being introduced to advanced logic integrated circuits from the 45 nm commercial technology node, there are many issues for the high-k and metal gate stack for use for the next generation.

First of all, the current HfO_2 based oxides need intentionally grown interfacial layer in order to arrange a good interface between the silicon surface and the gate insulator, and hence to suppress the mobility degradation of the channel carrier of MOSFETs. This interfacial layer is typically made of SiO_xN_y and its typical EOT is 0.7 nm. Maybe the interfacial layer EOT could be reduced to around 0.5 nm. However, the SiO_xN_y interfacial layer will certainly prevent the total gate dielectric EOT reduction less than 0.7 nm almost impossible, and thus, direct contact of the high-k gate dielectrics with silicon would be inevitable. For the rare earth oxide material such as La_2O_3, we have already demonstrated very low interface state density of 6 X 10^{10} cm^{-2}, and nice peak mobility value of more than 300 $cm^2V^{-1}s^{-1}$ for channel electrons. However, this is supposed to be the hydrogen or maybe even OH termination effects of the dangling bonds at the Si surface. The hydrogen and OH are supposed to be introduced by moisture absorption of water with the La_2O_3 exposure to air before the gate electrode deposition. The problem for this termination is that at least the interfacial concentration of the hydrogen decreases during the anneal at higher temperature above 300^{o}C. Thus, the interface state density increases and the mobility degrades after higher temperature. Another problem is that higher temperature anneal around at 500^{o}C make an interfacial silicate layer grow and this silicate layer increases the EOT of the gate oxide and interface state density. It is known that higher temperature annealing in hydrogen ambient decreases the interface state density, however the interfacial silicate layer formation is a problem.

Another way to improve the interfacial layer is to add other 3rd elements, such as Al. It seems that the Al repairs the oxygen vacancy and probably annihilates the interface states. We did confirm these effects but they are not sufficiently large at this moment and process optimization or maybe quest of more effective another 3rd element is necessary.

Further solution is to insert higher k interfacial layer than that of SiO_xN_y. Using Sc_2O_3 as the interfacial layer, we have achieved good suppression of the silicate interfacial layer growth at 500^{o}C. However, interface state density between the Sc_2O_3 and Si is still necessary to be improved at this moment.

Fermi level pinning is another important factor to be considered. Eventually, the process temperature for the high-k will goes down to 400 ~ 500^{o}C, as the gate stack process moves from ‘gate first’ to ‘gate last’ process and the Fermi level pinning effect will become less significant. However, it cannot be ignored. Addition of the 3rd element will be again a solution.

When, the EOT value of the gate insulator becomes less than 1 nm, the channel carrier mobility degradation due to the remote Coulomb and remote roughness scattering caused by the charge and roughness at the interface of high-k gate dielectrics and gate electrode becomes very significant. Thus, good mobility cannot be obtained without improving that interface.

For the middle and long term issues for the gate stack, we have to consider that the MOSFETs structure will change from planar to 3-dimensional structure, such as FinFET. Eventually, the FinFET will changes to nanowire MOSFETs, and maybe to CNT MOSFETs. Or maybe other semiconductor substrate such as GaAs or Ge could be used. Source and drain will be changed from semiconductor to metal. In the case of metal source drain, maybe process temperature could be decreased to 300~400oC, which would solve some critical issues of interfacial problems of high-k/semiconductor and high-k/metal gate. Some more details of gate stack for next 25 years will be explained at the symposium.