Nano-analytical transmission electron microscopy of advanced semiconductor devices

Keith Thompson
Thermo Fisher Scientific, Madison, WI, USA

INTRODUCTION
The ability to develop and to mass produce complex structures on the nanometer scale is critical to the continually accelerating pace of semiconductor development [1].
Advanced gate structures form the core of nanotransistors that power modern-day electronics devices. Smaller transistors switch faster, consume less power and utilize less real estate on the silicon wafer. The result is higher-performance and lower-cost electronics. This scaling down in size is not simply a linear function of improved lithography. It is also the implementation of novel materials whose properties are carefully engineered for the specific purpose at hand. In these materials, the location and quantity of each element plays an important role in the characteristics of the device. In addition to materials development, extremely tight process control on a massive scale, as occurs in a modern-day semiconductor facility, necessitates an ability to measure and to control key device parameters of interest. Advances in analytical electron microscopy are a key part of this process.

In order to demonstrate the advances in analytical capability, a test structure for an advanced nanotransistor was analyzed in a JEM-2800 transmission electron microscope (TEM) with the NORAN System 7 microanalysis platform. The accelerating voltage of the TEM was set at 200 kV. The magnification was 4,800,000 X. The EDS detector had an active area of 100 mm² and was located close enough to the sample to achieve a solid angle of X-ray collection of 0.95 steradians. The X-ray collection rate during the experiment was 13,200 counts per second with a dead time of 22%, and the collection time was 51 minutes. At this collection rate, the energy resolution of this specific EDS system as measured at the Mn k-alpha X-ray was 120.5 eV.

NANO-ANALYSIS OF A SEMICONDUCTOR DEVICE
Figure 1a is a TEM image of the nanotransistor of interest. One can see the advanced gate structure in the center with tungsten-via interconnects on either side. Figure 1b displays a phase map of the same structure generated with COMPASS spectrum-based phase analysis software on the NS7. The corresponding energy-dispersive spectroscopy (EDS) elemental maps are shown in Figure 2.

While the element maps are instructive, they provide only a broad analysis of the sample under study. The COMPASS phase maps provide a much more detailed analysis of the actual chemical structure of the transistor. For example, the element maps show a discrete region of Ti between the W via, and a large Si, O and N region that encapsulates the device. The phase maps, on the other hand, directly discriminate the unique TiN regions in between the W vias and the SiON region. In addition, the phase maps clearly demonstrate that the TiN is missing from the bottom section of the via, likely etched away with an anisotropic (i.e. directional) plasma etch prior to the deposition of W in the etched vias [2].

An up-close analysis of the phases that compose the gate structure (Figure 3) indicates that four primary regions are evident. The largest consists of the W plug that connects the metal gate to the upper-level interconnect layers (upper-level layers not shown). The remaining three phases are nanometer-thick layers that have been carefully engineered to control the electrical characteristics of the nanotransistor being fabricated [3].

The chemical composition of the metal gate (below the W-via and above the gate dielectric) is carefully tailored to engineer the local work function of the gate in order to precisely set the threshold voltage that flips the transistor between the “on” and “off” states. With a goal of minimizing power consumption, desired operating voltages are kept at <0.5 V. Maintaining a threshold voltage variation of <0.05 V is therefore critical to consistent device operation across the several billion transistors that make up the entire microprocessor. This threshold is typically tuned to the mid-point of the Si band gap. As measured against vacuum, the conduction band = 4.85 eV, the valence band = 3.75 for a gap of 1.1 eV and a mid-point = 4.3 eV. Tuning the work function directly to this mid-point enables the nMOS and pMOS transistors to both be naturally “off” with an appropriately positive and negative threshold voltage respectively while utilizing a single metal gate structure for both transistor polarities.

The gate dielectric (final layer before the crystalline Si) is deposited as thin as possible without being so thin as to allow an excessive amount of current leakage back from the channel into the metal gate while the transistor is in the “off” state. This gate dielectric is also composed of the highest-k dielectric material possible, while maintaining a thermally stable, defect-free (or defect-minimized) structure. The end goal is...
to maximize the capacitance of the gate structure (maximum permittivity, minimum thickness) while reducing the overall surface area. Much like the metal gate region, the composition of this dielectric must be controlled as tightly as possible in order to maintain consistent electrical parameters across all gate structures. Finally, inbetween the metal gate region and the dielectric is an intermediate layer. This layer either occurs (1) by intent – with the purpose of further adjusting the threshold voltage of the gate stack; or (2) by a small amount of interdiffusion of the metal gate into the dielectric as the overall structure goes through many heating and cooling cycles.

Returning to the COMPASS phase maps, the spectra from phase 5, i.e. the metal gate region, demonstrates a region composed of an Al-Ti-silicide mix with a very small amount (~1 at%) of Ta (8.1, 9.3 kV) and Mo (17.5, 18.5 kV) present. Ta and Mo are commonly used for fine tuning the work function and are therefore expected. The presence of the three major elements of Ti, Al and Si indicates that this transistor utilizes a substituted Al metal gate [4]. The Al and Ti are incorporated because they beneficially pin the Fermi level of the metal gate near the mid-point of Si (~4.3 eV). This type of metal gate is formed by depositing Al and Ti onto a thin polysilicon layer, which subsequently sits on top of the gate dielectric. The structure is annealed at approximately 450°C until the poly-Si is consumed. The presence of the poly-Si layer is sacrificial, to consume the Al and Ti while inhibiting Al penetration through and/or Ti reduction of the thin gate dielectric. This has the effect of placing a thin Ti-Al metal gate proximate to the gate dielectric while avoiding either one of these two effects that would short the device. The Ti has the additional convenient characteristic of reducing any native oxide layer formed on the surface of the poly-Si layer if the wafers are exposed to ambient between processing steps.

The next phase down, Phase 6, has a spectrum that shows a mixture of Ti-Al-Si and Hf, Ta and O and N. The Ta, present in both this layer and the full metal-gate region, is a definite byproduct of work-function engineering. The Mo is noticeably absent in this layer. This indicates that the Mo was either added intentionally later in the process or was a contaminant that entered during either the Ti-Al deposition or the W deposition. The Hf, O and N likely diffused into this region from the gate dielectric below, as some small amount of interdiffusion between the metal gate and the underlying dielectric is expected. In addition, some of the intermixing is also due to limits of the spatial resolution of analytical techniques. Some modest blooming of the X-ray interaction...
Volume will occur through the ~50 nm depth of the sample. Additionally any tilt away from the direct access of the structure will smear the distinct layers together. This middle layer, then, is likely a function of both interdiffusion between the metal gate and the gate dielectric and some amount of overlap in the analytical acquisition.

The final phase, Phase 3, represents the thin gate dielectric. A study of the spectrum shows Hf, Si, O and N. Hf provides the high-k needed to maximize the overall capacitance of the structure. Si provides significant thermal stabilization to the gate dielectric and minimizes dangling bonds from the surface of the crystalline Si wafer. These dangling bonds alter the gate work function, either in a direction to leave the gate structure closer to the “on” state or, conversely further turned “off.” In addition, this shift in work function is highly variable, therefore being more difficult to control. While pure HfO2 will chemically reduce in the presence of Si (Si + HfO2 → SiO2 + Hf), the Si and N present significantly minimize the driving force for this reduction process. The Si is typically introduced by depositing the HfON onto an ultrapure, ultrathin (~1 nm) SiO2 layer and annealing.

These observations can be further clarified by extracting line scans within both the horizontal and vertical orientation of the gate structure (Figure 4). The line scans are overlaid on the electron microscope image for spatial reference. The amplitudes of the line scans are based on the collected X-ray count after matrix corrections and background subtraction and provides a semiquantitative analysis of the elemental distribution. The amplitudes of the line scans are scaled to assist in viewing each element relative to the other and are not scaled for relative %at or %wt. The vertical line scans show the Al and Ti intermixed in phase 5, clearly occurring at the bottom of the W. This Ti-Al layer is then followed by a distinct Ta layer (i.e. phase 6) which is subsequently followed by the Hf layer and then Si substrate below. The horizontal line scans follow the same pattern as the vertical line scans: W in the center, followed by a Ti-Al region, a Ta region and then a Hf region. As seen in the phase analysis, both vertical and horizontal line scans show that the Mo exists in the W region and, to a smaller degree, in the Ti-Al region. Because the Mo has no impact on the work function in these regions (i.e. separated away from the gate dielectric), the presence of Mo is most likely due to contamination and is not present by intent.

CONCLUSIONS
An EDS analysis of an advanced, nanotransistor structure was undertaken in an analytical transmission electron microscope. The application of spectrum-based phase analysis software enabled a consistent, high-resolution identification of the phases present along with a semiquantitative analysis of the chemistry of each phase. The extracted line scans provided an additional level of information on the relative concentration and spatial location of each element of interest through the structure.

This sort of high-quality analytical information is critical to the study and development of such advanced nanostructures and additionally supports the analytical requirements to controllably fabricate such devices on a massive scale.

REFERENCES
1. www.itrs.net/Links/2012ITRS/Home2012.htm