NOVEMBER 2010

Solid State TECHNOLOGY®

THE INTERNATIONAL MAGAZINE FOR SEMICONDUCTOR MANUFACTURING

CMP for Metal-gate Integration p. 14

www.solid-state.com

EST. 1970 PennWell 100TH ANNIVERSARY

CMP for metal-gate integration in advanced CMOS transistors

EXECUTIVE OVERVIEW New materials complicate the process integration in high-volume manufacturing of high-*k* metal-gate (HKMG) CMOS transistors. The gate-last HKMG process requires two new CMP processes, both requiring extreme control over final gate height and topography. Because the gate stack is at the heart of the active device, it is far more sensitive to dimensions compared to passive interconnect and isolation structures. The poly-open CMP process has been tuned to handle new selectivity challenges. A new family of aluminum CMP slurries has been developed to provide <10nm dishing regardless of the layout. The methodical tuning of slurry additives provides for the successful integration of both poly-open and Al CMP into HKMG process flows.

For over twenty years of IC manufacturing, the creation of planar device structures has required the use of technologies to reduce topographic variation. Chemical-mechanical planarization (CMP)—pressing wafers into rotating pads in the presence of special slurry blends to produce removal through chemically amplified nano-scale abrasion— has become a critical part of modern IC fabrication.

The initial application of CMP technology was to planarize silica dielectrics for interconnects. This early use of CMP was driven by the need for basic planarity. It reduced the depth-of-focus requirement for the microlithography used to pattern the dielectric and metal layers. However, in part due to the empirical nature of original CMP process development, the technology had historically suffered from poor process control. A technology ecosystem of users, OEM's, specialty materials suppliers, and academics recognized the challenge and worked together to ensure that CMP can meet the evolving needs of the state-of-the-art in IC fabrication. This led to the adoption of CMP for a wider variety of uses.

The first metal CMP application in manufacturing was for tungsten (W) plug formation prior to the 0.35µm node. CMP processes for polysilicon (poly), shallow trench isolation (STI), and copper (Cu) were developed and deployed to solve problems in the quest for continued scaling. Also, as the requirements for CMP have become more difficult, it is now typical for the overall material removal in a CMP process to be broken into multiple specialized steps, often with unique consumables.

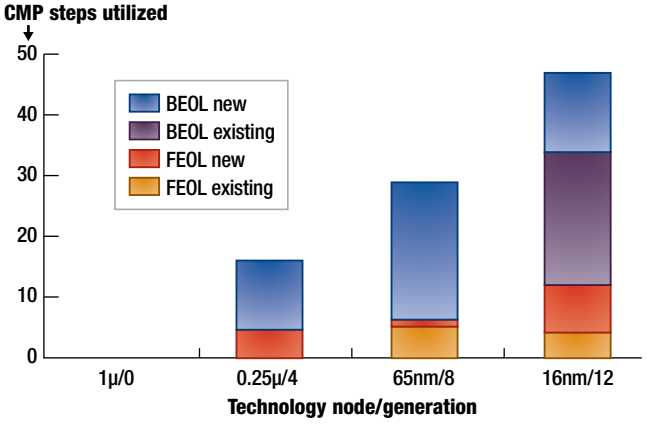

Figure 1. Starting from the 0.25µm node, about four new CMP steps have been added with each new CMOS technology generation.

Scaling and planarization of materials

CMOS scaling has led to transistor devices with critical structures approaching a few atoms in size. New materials and new structures of existing materials have been found to optimize device size and performance.

Many of these changes in IC device structures and materials have driven new planarization requirements. **Figure 1** shows that since the era of 1 micron minimum features, a net average of four CMP steps have been, or are being added, with each new CMOS manufacturing technology node [1].

The replacement of traditional materials for the gate dielectric and gate conductor has recently been a major focus for our industry. Traditional CMOS transistors are made with silicon-oxide/nitride (SiON) gate dielectrics and polysilicon gate conductors. The dielectric begins to suffer worse leakage as the dielectric thickness gets into the single-digit range of atoms. More exotic dielectrics are needed to resist leakage while managing other fundamental device parameters.

High-*k* metal-gate (HKMG) transistors were first implemented with 45nm technology. The dielectrics are based on oxides of hafnium, which can maintain low electrical leakage levels. The polysilicon conductors also need to be replaced because they do not work well with the new dielectrics. The NMOS and PMOS transistors require independently optimized complex stacks of thin work-function metals topped by a bulk conductor layer.

There are different flows under development for 32nm node processing, depending upon whether the metal-gate is formed before or after the source and drain regions. The source and drain formation includes high-temperature steps such as implants and anneals, so forming the gate first restricts the choices of conductors that can be successfully integrated. In contrast, forming the gate last—also known as the replacement metal gate (RMG) process

Paul Feeney, CMP Fellow, Cabot Microelectronics Corp., Aurora, Illinois, USA

Reprinted with revisions to format, from the November 2010 edition of **SOLID STATE TECHNOLOGY**

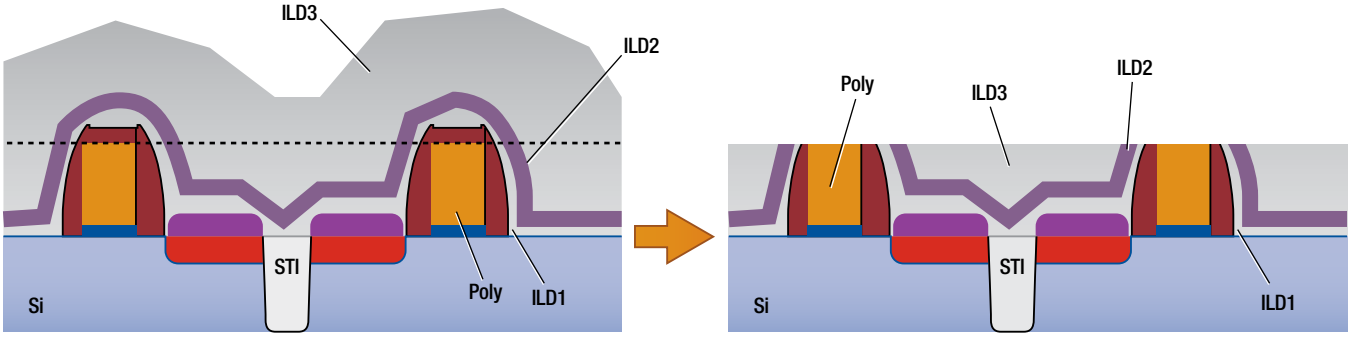

Figure 2. Cross-sectional schematics of "gate-last" CMOS transistors before (left) and after (right) the "poly-open planarization" (POP) CMP process, showing that this step controls the height of the sacrificial-poly gate.

flow—allows for the use of aluminum (Al) as the bulk conductor and has gained momentum. There is also the possibility of a hybrid approach where NMOS transistors are formed gate-first while PMOS transistors are formed gate-last.

As the device structures have become more complex, so have the manufacturing processes needed to form them. The gate-last (RMG) HKMG process flow is initially almost identical to that used to form traditional SiON/poly gates. Only after all of the high-temperature process steps are complete are the poly gates etched out and replaced by metal. The essential flow is as follows [2]:

1. STI, implants for wells and V_T control,
2. ALD of high-*k* gate dielectric and poly deposition,
3. Lithography and gate etch,
4. S/D extensions, spacer, Si recess and SiGe deposition,
5. S/D formation, Ni salicidation, ILD_0 deposition,
6. Poly open CMP, poly etch,
7. PMOS work-function metal deposition,
8. Metal gate lithography and etch,
9. NMOS work-function metal deposition, and
10. Al metal gate fill and CMP.

CMP for gate-last HKMG

Since the gate is essentially at the heart of the transistor, extreme control is needed over all gate processing steps to ensure proper device function. Control is made even more challenging by the atomic-scale dimensions in advanced devices. Variation in gate height of just a few atomic layers now leads to measurable transistor performance variability [3].

CMP is central to the above integrated gate-last process flow, being used in two challenging processes to form the active device. The new dielectric process—referred to as poly open planarization (POP) CMP —has several additional challenges compared to the oxide polish utilized for making standard devices. **Figure 2** shows that the transistor structure includes a combination of oxide, nitride and polysilicon films instead of just oxide.

A preferred way to address the total removal is to first polish the oxide back to the nitride in a step that is very similar to what is done in STI CMP today. Then, a final step is used to expose the tops of the polysilicon features. The process in this final step must remove the correct amount of each film while preventing local topography from being generated due to the film differences. Compared to a traditional ILD0 CMP step, even tighter thickness control is required in order to manage the height, and thus resistivity, of the gate conductor.

After the STI-like step, the nitride will be raised versus the oxide, so it is desirable for planarization to have the nitride removal rate above the oxide rate. Typical Silica-based slurries for the planarization of oxides have nitride to oxide selectivity well below 1 and have material removal rates too high for easy control.

New ways were found to produce a moderate nitride removal rate. The nitride removal mechanism is dependent on hydrolysis of the Si_3N_4 to Si-OH and NH_3 [4]. This reaction is pH dependent, with low pH being faster. At very low pH, silica is less anionic and is not as attracted to the nitride surface, which leads to a lower removal rate.

A new slurry platform, iDiel N3100, was produced. The silica particles in this slurry are uniquely charged to boost attraction to the to the dielectric. In this system, the oxide removal rate is controlled by pH (**Fig.3**). The oxide removal rate is affected by the presence of the nitride, so the removal rate on patterned wafers starts higher and then slows as the nitride being removed decreases. This results in the prevention of local step heights. The polysilicon removal rate is driven down by additives (**Fig.3**). Having a low polysilicon rate and a slowing oxide rate improves the ability of the process to stop at the desired thickness target.

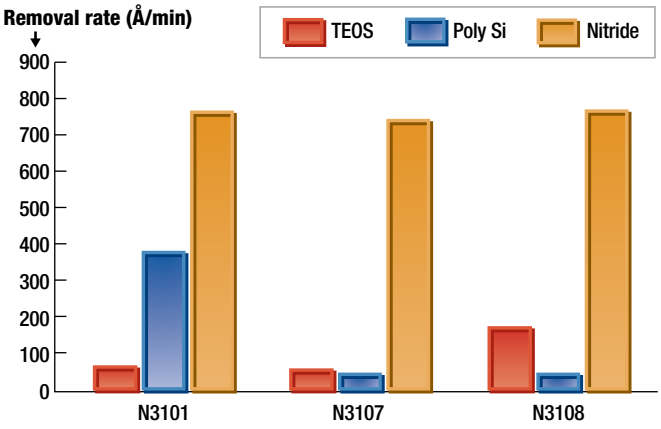

Figure 3. A range of blanket removal rates are possible through the use of additives to suppress poly removal and pH to accelerate oxide removal in the iDiel N3100 slurry family.

RMG Al CMP

The new metal CMP process—referred to as the replacement metal gate (RMG) aluminum CMP step for this gate-last flow—also has significant challenges compared to that used for W contacts. The process must planarize Al and the complex stack of work-function metals. It also must do so while stopping well on oxide and minimizing recess, both of which contribute to the gate conductor final height.

The Novus A7100 series of slurries was designed with alumina particles that provide the ability to remove Al and the work-function metals selective to the oxide underneath. As in the polishing of other metals, oxidizers and chelators play a role in creating passivation and ion-complex formation, but the mechanisms employed are unique in this system.

There are significant ways in which Al CMP is different than other metals. In W CMP, oxides of W are formed that are softer than the bulk metal and are removed more easily. Oxidation of Al creates a surface that is harder than the bulk film. This aluminum oxide surface is critical in slowing removal after clearing the bulk film in order to minimize recess. CMP using the series slurry noted above roughly follows Preston's Law—removal rate is a linear result of pressure and velocity—demonstrating the mechanically limited nature of removal.

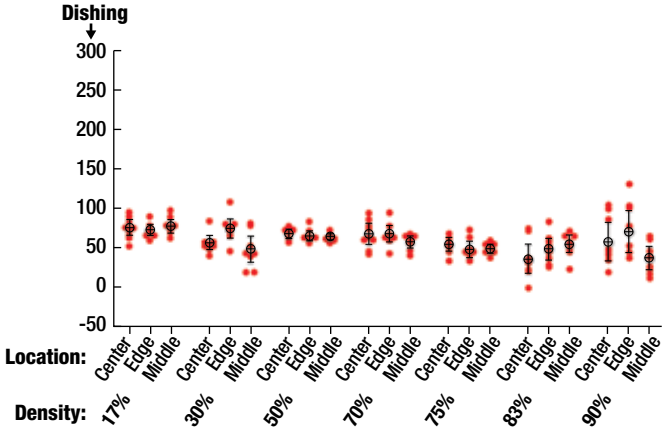

Figure 4. Within-die non-uniformity (WIDNU) data from Al CMP using A7100 slurry on a D100 pad, showing good control across a wide range of feature densities (95% CI for the mean).

This passivation and removal mechanism leads to other differences. In CMP of Cu, a technique called a soft landing—where pressure is decreased during clearing to slow removal—is a common approach for minimizing recess at the expense of process time. In the Al system, higher pressure actually improves recess.

Too much mechanical energy has, however, a unique downside here. As the nano-abrasion from the particles increases, it can overcome the kinetics of passivation. The result is abrasion of the soft bulk Al that leads to a buildup of "black debris" on the polishing pad [4]. This debris is comprised of small Al particles that remain dark in color for a short amount of time until they become fully oxidized. At the onset of black debris, removal rate and defectivity both change significantly.

The optimal Al CMP process balances rate and recess (**Fig.4**) versus debris and defects. Recess levels are below 10nm across a wide range of feature sizes and densities after the aluminum CMP step. This recess can be further reduced by an optional buff step that removes a very controlled amount of dielectric.

The advent of replacement metal gate (RMG or gate-last) HKMG process flows for 45nm node and below CMOS manufacturing has led to a significant amount of development going into the new dielectric and metal CMP steps. The needs are now being met with processes using consumables designed specifically for these steps.

Conclusion

As CMP technology has matured, it has become a more attractive option for wafer processing. In addition, CMOS device scaling has led to a steady increase in the number of CMP steps, due to both an increase in the number of layers, as well as a need to create novel structures with exotic materials.

The advent of replacement metal gate (RMG or gate-last) HKMG process flows for 45nm node and below CMOS manufacturing has led to a significant amount of development going into the new dielectric and metal CMP steps. The needs are now being met with processes using consumables designed specifically for these steps.

Acknowledgments

Many thanks to scores of people across Cabot Microelectronics and our industry for their contributions to these technologies. Novus, iDiel, and Epic are trademarks of Cabot Microelectronics Corp. ■

References

1. P. Feeney, et al., "The Increasing Needs and New Solutions in CMP," ICPT 2008.
2. K. Mistry et al., "A 45nm Logic Technology with High-*k* + Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging, IEDM Tech. Dig., p.247 (2007).
3. J. Steigerwald, "CMP Technologies for 45nm and Beyond," ICPT 2009,
4. J. Dysard, et al., "CMP Solutions for the Integration of High-*k* Devices," ECS 2010.

Biography

Paul Feeney received degrees in engineering from the U. of Illinois and from Rensselaer Polytechnic Institute. He is CMP Fellow at Cabot Microelectronics Corp., 870 North Commons Drive, Aurora, IL 60504 USA; ph.: 630-375-5592; email Paul_Feeney@cabotcmp.com.