$e=mc^3$ a review of copper annealing processes and equipment

Ken Sautter, KMS Enterprises, **Stuart Allen** & **Bill Moffat**, Yield Engineering Systems Inc., San Jose, CA, USA

ABSTRACT

In any robust microcircuit manufacturing environment, processes and equipment have to be decided on quickly, often too quickly for all the factors influencing the decision to be properly analyzed. One such process and piece of equipment is in the copper annealing area. This paper is an attempt to cover the process, physical and economic parameters that govern the decision to purchase a copper annealer. The traditional R&D unit, the hot plate, quickly moves to the vacuum hot plate. Then, as the process parameters for circuits with high aspect ratio trenches and problems with electromigration become apparent, a move to some batch process is indicated [1–4]. Further, problems with copper hillocks drive the annealing processes to longer times and possibly lower temperatures. As it was common to fight aluminum hillocks when we started aluminum multilayer metal, similar problems with copper processing should have been expected [5–7]. Problems with copper oxidation push for as low an oxygen concentration as possible [8, 9]. Problems with fluorine contamination determine the need for a totally moisture free environment [10]. Polymer voiding problems also point to a need for longer processing times and possibly some more forceful means of reducing or removing voids created during copper annealing [11, 12].

Introduction

Among other problems that may need to be addressed is the correction of unwanted copper diffusion into the low-*k* dielectric during anneal [13]. Now the latest requirement comes along. With the extreme aspect ratios of today's processes, the relatively thick metallic copper-capping layer needs to be replaced with a thin controllable silane copper-capping layer. With all these possible needs to be addressed, the search for the optimum piece of equipment can start. A couple of simple rules were established first: (1) the equipment had to fit good fab area design criteria. These include as low a particle count as possible and as versatile utilization as possible, because no fab design is cast in concrete. (2) Within reason, the equipment should be as economical as possible and incorporate a low-price nonautomatic R&D unit. It was also considered important that the transition from R&D to production would be seamless and there would be no requirement for changes in process to accommodate changes in equipment.

After a quick review of the available units, we quickly narrowed it down to two types of equipment. One, a vertical diffusion furnace where the load would be in the correct zone before the temperature is ramped up and the cool down would take place in that zone before lowering and unloading. Two, a vacuum, super low oxygen, low-*k*/or copper annealing oven with the same criteria. The one concession to higher throughput would be a load and unload temperature of 50°C with the assumption that no copper oxidation would take place at or below these temperatures [8].

Summary of equipment requirements

One of the most critical aspects of annealing copper is the removal of oxygen from the processing environment. The critical gas flow to obtain zero oxygen between the wafers was studied with the following conclusions:

Solvent or gas removal for a horizontal wafer

The process of solvent or gas removal involves the evaporation of the solvent from the solid and the diffusion of gases

TABLE 1.

Requirements	Vertical diffusion furnace	Vacuum oven
Good temp. uniformity	Excellent ±0.1%	Acceptable ±1% (optional) ±0.25%
Good temp. control from 50–400°C	Good at high temp. but poor below 200°C	Good below 200°C acceptable above 2 °C
Control of oxygen level	Below 10 p.p.m.	Below 1 p.p.m.
Good throughput	20 wafers per hour.	40 wafers per hour.
Low per unit cost	Approx. $2.3 million.	Approx. $1.0 million.
Reasonable footprint	2 to 4 square meters.	9 square meters.
Low particulates	Less than 10, 0.2 micron particles added.	0 added. Net reduction in particles.
Polymer volatile removal	Gas flow not conducive to solvent or gas removal.	Partial vacuum laminar flow gives excellent solvent and gas removal.
Versatility	1) A copper annealer and or a diffusion furnace.	1) A copper annealer. 2) A copper capper. 3) A low-k baker. 4) A low-k sealer. 5) A low-k repairer. 6) A copper adhesion unit. 7) An ammonia high-temperature copper oxide reducer. 8) Any or all of the above in combination.

through the boundary layer. The boundary layer diffusion rate is influenced by the boundary layer thickness, the diffusion rate, and the free-stream concentration of the solvent vapor.

Evaporation rate

During solvent evaporation, a microlayer of gas will be maintained at the surface of the solid in equilibrium; vapor partial pressure remains equal to the solvent vapor pressure at the surface temperature. The rate of evaporation of the solvent from the surface depends on the rate of diffusion of solvent through the solid and the rate that vapor is transported out of the surface gas microlayer.

For copper annealing, it must be remembered that the surrounding low-*k* dielectric is being cured also. For this cure process, diffusion through the solid must be limited to avoid voids caused by f bonds if a fluorine-containing low-*k* is used. This is especially true if moisture is present [8]. In order to achieve low-*k* film shrinkage uniformity, the vapor transport rate out of the surface gas microlayer must be uniform over the entire surface of the wafer.

Boundary layer thickness

For laminar flow over a flat plate, the thickness of the boundary layer δ is given by

$$\frac{\delta}{x} = \frac{5.0}{\sqrt{\mathrm{Re}}}$$

where x is the distance from the leading edge of the plate and Re is the Reynolds number. The value of the Reynolds number is given by

$$\mathrm{Re} = \frac{ux\rho}{v}$$

where u is the free stream velocity, ρ is the gas density, and v is the gas viscosity. Since, for an ideal gas, density is proportional to pressure, the above equations imply that boundary layer thickness increases as pressure decreases as follows

$$\delta \propto \frac{1}{\sqrt{\mathrm{P}}}$$

Diffusion rate

For small distances x, the diffusion rate Q is described by

$$Q = -D_{12}\frac{\Delta c}{\Delta x}$$

where D is the coefficient of diffusion and c is the concentration. For low vapor concentrations

$$D_{12} = \frac{2}{3\xi^2 P}\sqrt{\frac{k^3T^3}{\pi^3 m}}$$

which implies that

$$D_{12} \propto \frac{1}{\mathrm{P}}$$

and, finally, the combined effect of the process gas pressure on boundary layer thickness and diffusion coefficient implies a pressure dependence for the net diffusion rate as follows:

$$\mathrm{Q} \propto \frac{D_{12}}{\delta} \Rightarrow \mathrm{Q} \propto \frac{1}{\sqrt{P}}$$

Free-stream solvent or gas/vapor concentration

Figure 1 compares the gas flow during processing for a vertical laminar flow vacuum oven and a vertical diffusion furnace. For the vertical diffusion flow case, pure nitrogen flows over the outside edges of the wafers. Solvent vapor and gas is transported from the outside surface of the boundary layer to the bulk nitrogen stream through a toroidal recirculation zone. Since the recirculation zone must have a higher vapor/gas concentration than the nitrogen stream in order to transport solvent/gas into the stream, the vapor/gas concentration gradient across the boundary layer is reduced and the solvent evaporation rate is therefore also reduced. In addition, the toroidal recirculation has a stagnation zone in the center of the wafer where solvent transport away from the boundary layer surface is greatly reduced. The case of vertical laminar flow has much more uniform gas flow across the surface of each wafer and there is no stagnation zone.

With vertical diffusion, solvent vapor removed from the wafers accumulates in the nitrogen stream as the gas moves down the wafer stack. For the last wafer in a 100-wafer batch, the free-stream vapor concentration is the accumulation 99 wafers worth of solvent. The evaporation environment for the top and bottom wafers in the batch can be significantly different. The variations in gas flow and concentrations of gas above each wafer can give rise to variations in void formations in each wafer. For example, the top wafer should have the cleanest atmosphere above it and should have consistent void formation. The lower wafers will have variations in gas pressure and therefore variations in concentrations of oxygen and liberated solvents. If the void formation is a function of gas generated during bake, the variation in pressure above the wafer during bake will result in variations in void formation. If the voids are caused by the trapping of fluorine during plasma operations, a low-pressure atmosphere with no moisture will be required to stop void formation.

Particle removal and generation will also be influenced by the horizontal wafer configuration. Wafers that have particles on their backsides can (and will) drop particles onto the wafer below. Because the horizontal wafer configuration has no strong gas flow across its surface, there is no inducement to shed particles. The design and gas flow of a vertical diffusion furnace has to give rise to an increase in particles on the front of the wafers. Vertical laminar flow, on the other hand, provides a buffer of rapidly

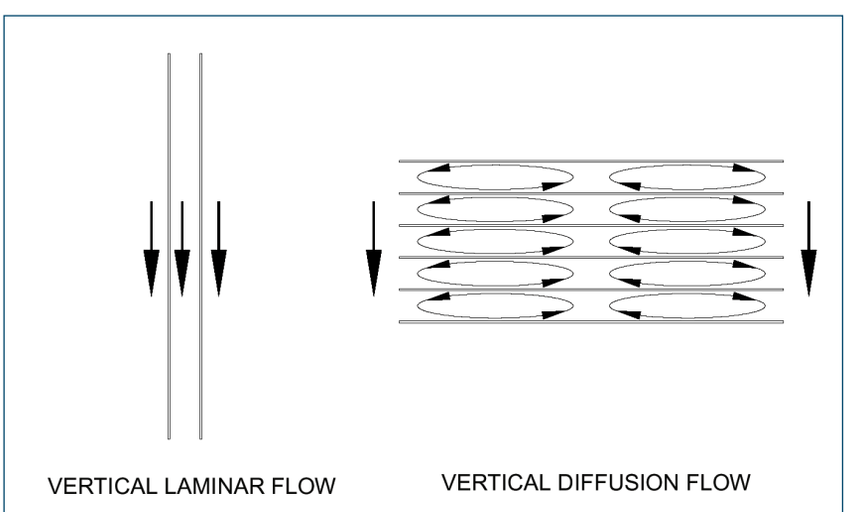

Figure 1.

moving nitrogen between wafers that results in any shed particles moving down and away from nearby wafers. In the case of the vacuum copper annealer, the wafers are 5 degrees away from vertical and the process takes place in a controlled vacuum.

Control of oxygen concentration and gas stoichiometry

A vacuum oven with sequential vacuum/hot nitrogen purges meets all requirements for low oxygen concentration. With each nitrogen purge, the oxygen level is reduced and processing levels of below 1 part per million are common.

Copper oxide removal has been reported in vacuum low-*k* material processing ovens in production conditions, possibly as a result of low oxygen levels and carbohydrate fume production from the low-*k* dielectric. One user has reported using a tarnished copper washer as a process verification tool for a cure process. If the washer came out shiny indicating that the Cu_2O had been removed, the process was OK.

Adequate ventilation to remove volatile constituents or gases

Removal of volatile constituents from a polymer in the vacuum oven begins with uniform volatile partial pressure reduction at low temperatures during the dehydration cycle purges that begin the process. After the purge cycles, the bulk of the process is carried out at a pressure of 200–300 torr of N_2. Adequate ventilation to prevent build-up of the volatile constituents is achieved by balancing vacuum draw with gas flow.

There are other factors that help to advance the case of the vacuum oven. First is the ability to provide an HMDS vapor at any temperature to assist in the reconstitution of plasma damaged low-*k* dielectric [14, 15]. Also one has the ability to provide a silane treatment of copper to promote adhesion of low-*k* dielectrics utilizing the same unit.

Trying to evaluate where we should be in 2 years time we made the following assumptions and assessed the position of 5 mythical companies.

Assumptions:

1) 1,000 wafers per day, 20-hour day utilization or 50 w/p/h.
2) 1,000 Angstrom trench width.
3) Vertical diffusion furnace, $2,000,000.00, 20 w/p/h. output.
4) Stand alone silane unit, $250,000.00, 25 w/p/h. out.
5) Combination silane copper annealer, $1,250,000.00, 40 w/p/h out.
6) Combination silane low-*k* bake, $1,250,000.00, 40 w/p/h. out.
7) 5 copper layers and 6 dielectric layers.
8) 5 layer, 50 w/p/h. needs 13 vertical diffusion furnaces, cost $26 million.
9) 6 layer, 50 w/p/h. needs 15 vert. diff. furnaces, cost $30 million.
10) 5 layer, 50 w/p/h. needs 7 combo units, cost $8.75 million.
11) 6 layer, 50 w/p/h. needs 8 combo units, cost $10.0 million.

Company A) Spends $56 million on copper anneal and low-*k* dielectric baking equipment, has 1,000 Angstrom trench width but metallic copper capping. This sacrifices 400 Angstroms of the trench space.

Company B) Spends $56 million on copper anneal and low-*k* baking equipment, has 1,000 Angstroms trench width but uses wet silane for copper capping, this sacrifices 40 Angstroms of the trench space giving nearly 40% more copper for faster speed, better power consumption. This will give significant equipment saving for metal deposition equipment and metal usage. The result is a more desirable device to the customer and savings of approximately $6.00 per wafer in chemical costs and lower equipment pricing. In short, B is a better copper capper than company A.

Company C) Spends $56 million on copper annealing and low-*k* baking equipment, has 1,000 Angstroms trench width but uses vapor silane deposition. This saves a minimum of $5.00 per wafer in chemical costs per wafer and allows multiple controlled depositions. For example 5 Angstroms followed by 5, followed by 5, followed by 5. Multiple depositions will remove any possible isolated sites of low silane adhesion. This vapor silane delivery system can also function as an HMDS delivery system to repair plasma damaged low-*k* dielectric. In short, C is a superior copper capper company, compared to company B.

Company D) Spends $30 million on low-*k* baking but spends $8.75 million on combination silane copper cappers and copper annealers. This allows silane before, during and or after copper anneal for superior copper capping and low oxygen copper annealing. This can also function as an HMDS delivery system for the repair of plasma damaged low-*k* dielectric. A superior copper capping company compared to company C.

Company E) Spends $8.75 million on a combo copper capper, copper anneal units, and $10.00 million on combination silane sealers and low-*k* dielectric bake units. This allows silane deposition before during and/or after low-*k* baking and, because the silane is a gas during deposition, all orifices in the low-*k* are silane treated. In short, a far superior copper capping company compared to company D.

Company E has superior low-*k* baking, silane sealing, superior low Oxygen copper annealing, and better more controlled copper capping. In addition, E enjoys the following savings:

1) Capitol equipment savings of $37.25 million.
2) Use of silane instead of metallic copper capping, $6.00 per operation or with 5 layers of copper $30.00 per wafer savings. At 1,000 wafer starts per day a savings of $30,000.00 per day.
3) Use of vapor silane instead of wet silane gives a saving of $5.00 per wafer operation with 5 layers of copper $25.00 per wafer, 1,000 wafer starts per day gives an extra $25,000.00 per day savings.
4) No company is using silane sealing for low-*k* dielectric at the moment so a saving is not computed for this operation.

In the final analysis company E is a Most Careful Copper Capper.

$$E=MC^3$$

REFERENCES

[1] "The recrystallization dynamics within a high aspect ratio trench can be much different than in blanket copper films due to the physical constraints of the trench sidewalls." "Trade off between reliability and post CMP defects during recrystallization anneal for copper damascene interconnects." G.B. Alers, Novellus.

[2] "As the aspect ratio of the trench sidewalls increases the kinetics of recrystallization slow down due to the increased constraints of the sidewalls." "Trade off between

reliability and post CMP defects during recrystallization anneal for copper damascene interconnects." G.B. Alers, Novellus.

[3] "The longest lifetimes (electromigration) were obtained with a 60-min 315°C anneals. This time was necessary to fully recrystallize the narrow 3:1 aspect ratio trench." "Trade off between reliability and post CMP defects during recrystallization anneal for copper damascene interconnects." G.B. Alers, Novellus.

[4] "The best anneal procedure to obtain improved electro migration performance without a high density of defects would be a long time anneal (greater than 1 h) at temperatures less than 200°C." "Trade off between reliability and post CMP defects during recrystallization anneal for copper damascene interconnects." G.B. Alers, Novellus.

[5] "The key to managing copper hillock creation is a longer time, lower temperature anneal." "Techniques to improve Cu/low-*k* Integration", Solid State Technology paper by Victoria Shannon, Novellus.

[6] "Each Copper layer must be annealed prior to each CMP. Currently this is performed at approximately 400°C in a very tight oxygen free environment." Jean-Francois Daviet, director of RTP group at Matson in a report for Semiconductor International 100 nm: The Undiscovered Country.

[7] "The choice of post-plating anneal can have an effect on the incidence of defects, presumably because it affects grain growth." "Understanding and Reducing Copper Defects", Sesha Varadarajan Novellus. A report for Semiconductor International.

Oxygen and water-vapor exposure

[8] "Copper will undergo some oxidation at environments above 50-60°C." "Fabrication Technologies for Three Dimensional Integrated Circuits" Rafael Reif M.I.T.

[9] "Minimal Copper Oxidation, sometimes no oxidation at all with 400 degrees C Oxygen concentration below 3% by weight." "Fabrication Technologies for Three Dimensional Integrated Circuits" Rafael Reif, M.I.T.

[10] "Copper with fluorine containing SiO_2-based low k material has special considerations. The F bonds can react with water and produce volatile by products. A moisture-free annealing atmosphere may help." "Techniques to improve Cu/low-*k* Integration" Solid State Technology paper by Victoria Shannon, Novellus.

Polymer voiding

[11] "Wider low k spaces exhibit less voiding." "Overview of Dual Damascene Cu/Low K Interconnect. P. Josh Wolf, Intel. International Sematech paper

[12] "Our results suggest a low temperature long time (greater than 30 minutes) anneal is needed to avoid anneal-induced voiding." "Understanding and Reducing Copper Defects", Sesha Varadarajan Novellus. A report for Semiconductor International.

Copper diffusion barrier capping

[13] "Amino-silane pre-treatment has been demonstrated to effectively limit Cu diffusion and has been subsequently qualified and ramped in production for a MCM application with no impact to electrical or reliability performance." "Evaluations of Cu Capping alternatives for Polyimide-Cu MCM-D.", Dale McHerron and Claudius Feger of I.B.M, Proceedings from the 10th Dupont Symposium on Polymers for microelectronics.

[14] "HMDS can repair the damaged structure and reduce both the leakage current density and dielectric constant of the POSG film." "Effective repair to ultra-low-*k* dielectric material (k ~ 2.0) by hexamethyldisilazane treatment". Y.S. Mor Institute of Electronics National Chiao Tung University Republic of China. Journal of Vacuum Science B20(4) Jul/Aug 2002.

[15] "Using HMDS to replace surface OH groups followed by a thermal treatment to remove hydrogen bonded OH groups, the k value of the etched/ashed film was decreased to 2.43 and the surface hydrophobicity of the film was restored". "Cleaning and restoring k value of Porous MSQ Films" Phillip G Clark FSI international. Semiconductor International 8/1/200.

ABOUT THE AUTHORS

Ken Sautter is a consultant with KMS Enterprises and specializes in semiconductor process development, as well as fab and equipment design. He has over a dozen patents in both equipment automation and wafer processing. He can be reached at kmsautter@yahoo.com, Tel: +1 (408) 720-1770.

Stuart Allen is principal engineer at Yield Engineering Systems and is responsible for the design of this and a newer copper annealing unit. He can be reached at sallen@yieldengineering.com, Tel: +1 (408) 954-8353.

Bill Moffat is C.E.O. and founder of Yield Engineering Systems with 20 years experience in photolithography and a further 20 years in equipment for the semiconductor industry. He can be reached at bmoffat@yieldengineering.com. Tel: +1 (408) 954-8353.

ENQUIRIES

Yield Engineering Systems, Inc.
2185 Oakland Road,
San Jose
CA, 95131
USA

Toll Free: (888) YES-3637
Tel: +1 (408) 954-8353
Fax: +1 (408) 954-8369