Pseudo Epi—Cost Reduction Approach and a Paradigm Shift in Substrate Material

Mehran Aminzadeh, K. V. Ravi, George E. Sery, *Member, IEEE*, Amiad Conley, and Matt Casperson

Abstract—**Historically, P/P+ epitaxial wafers have been utilized for CMOS products for more than two decades. The epitaxial wafers have several key characteristics such as latch up immunity, oxygen-free active areas, superior oxide quality, and gettering capability compared to bulk nonepi wafers. The epi wafers, however, are costly. Pseudo epi is an alternative to epi wafers with equivalent device performance and material cost savings of 20%–25%. Pseudo epi or high-temperature hydrogen anneal is expected to save a significant percentage of start material costs over the epi substrate for logic family products where epi wafers have dominated the market.**

Index Terms—**Annealing, CMOS digital integrated circuits, costs, epitaxial layers, high temperature, silicon.**

I. Introduction

DUE TO the continual price reduction of personal computers and electronic consumer products there has been significant pressure to reduce the manufacturing cost of microprocessors and Flash memory products. Fig. 1 shows relative product cost structure for a state of the art 300-mm device factory [1]. The primary cost contributors to a 300-mm factory are tool depreciation or tool capital cost. Substrate material or silicon is the second highest cost contributor. Therefore, there is continual pressure to reduce the manufacturing cost of the starting material. This work addresses a solution in reduction of start material costs for microprocessors and chipset products without any performance degradation. Historically, many semiconductor device manufacturers have relied on utilizing epitaxial wafers as a starting material to build their products due to superior material characteristics relative to bulk (nonepi) wafers. Typically P/P+ epitaxial wafers, a lightly doped epitaxial layer over a heavily doped Si substrate, are used as the starting material since they have several key device and material advantages over bulk nonepitaxial wafers.

From a device standpoint P/P+ epi wafers provide latch up immunity, superior oxide quality, and a good resistivity control for active device region compared to bulk nonepi wafers. From a materials standpoint, epitaxial wafers provide an oxygen-free epi layer, a high level of bulk defect density such as oxygen precipitates and stacking faults for intrinsic gettering, heavy boron-doped substrates for solubility enhanced gettering, and an active area free of Si vacancies. The epi wafers, however, are large materials cost contributors. Typically, a bulk 200-mm polished wafer (test wafer) costs 40%–50% lower than an epi wafer. However, bulk wafers do not have the device or the material advantages of epi wafers so they are not suitable for products. Pseudo epi (PE) uses a high-temperature hydrogen anneal process to achieve a low surface boron concentration and low defect density. High-temperature hydrogen anneal is not a new concept. The DRAM industry has used both hydrogen and argon annealed wafers for several years. However, except for some isolated cases H_2 annealing has not been able to successfully replace the epi wafer for microprocessors (CPU), Flash, or CMOS applications. The objective of this work was to substitute a replacement to epitaxial wafer by substantially reducing starting materials cost by 20% to 25% with no device or performance compromise to logic technology. Special attention was given to silicon specification selection and manufacturability of a process which is inherently more cost effective than epitaxial process. This alternative to the epitaxial wafer is equivalent in device performance and meets the cost goals.

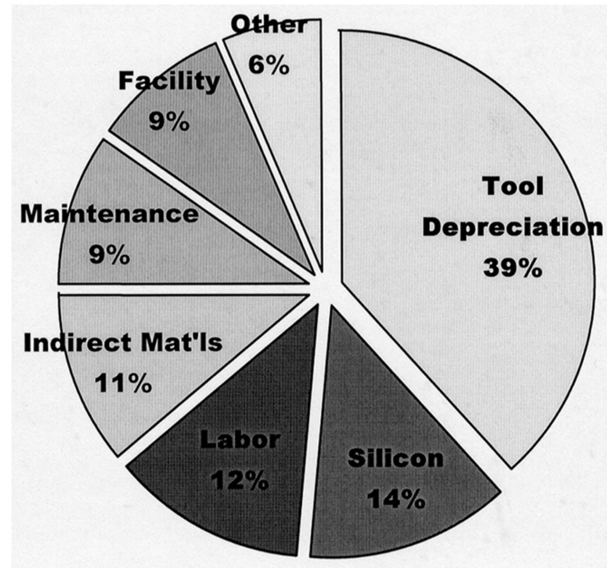


Fig. 1. Silicon costs in a 300-mm device factory.

II. Pseudo Epi and Its Benefits

PE is a process for creating a low boron doping concentration near the surface (analogous to the lightly doped P epitaxial layer in P/P+ epitaxial wafers) achieved by heat-treating a homogeneously boron doped wafer in a hydrogen atmosphere at a temperature close to 1200 °C for approximately 1 h.

The high-temperature anneal is done in a vertical diffusion furnace with typical load size of 75 to 100 wafers. The throughput is typically about 12–18 wafers per hour. Hydrogen annealing achieves four major factors: 1) creation of an oxygen

Manuscript received May 8, 2002; revised July 15, 2002.
The authors are with Intel Corporation, Santa Clara, CA 95070 USA.
Digital Object Identifier 10.1109/TSM.2002.804891

0894-6507/02$17.00 © 2002 IEEE

Authorized licensed use limited to: Teknillinen Korkeakoulu. Downloaded on October 21, 2009 at 06:00 from IEEE Xplore. Restrictions apply.

denuded zone; 2) boron out diffusion from the near surface regions of the wafer; 3) annealing of silicon vacancies [2]; and 4) improved nanotopography and surface roughness. Additionally lower junction (diode) leakage has been reported [3].

A. Denuded Zone/Intrinsic Gettering

Interstitial oxygen is not desirable in the active area since it can form oxide precipitates with thermal treatment and cause degradation in device performance. PE creates a greater than 50 μm oxygen-free surface, a denuded zone which is key in terms of low defect density. The high-temperature process also helps bulk interstitial oxygen form oxide precipitates that act as intrinsic gettering (IG) sites for metallic impurities. Some CMOS technologies have relied on the available thermal budget, during device processing, to enable the formation of bulk oxygen precipitates and stacking faults. However, as the thermal budgets for advanced processes is reduced the effectiveness of oxide precipitate formation is also reduced. The high-temperature hydrogen anneal step associated with PE can be tailored to provide an efficient means of generating bulk micro defects for IG independent of the CMOS process thermal budget. IG sites help getter metallic contamination away from the active area. The IG is typically not required in today's fab clean room environments, but is an added insurance policy.

B. Boron Out Diffusion

The combination of high temperature anneal ($\sim$1200 °C) and hydrogen ambient leads to the out diffusion of boron from the wafer resulting in about one order of magnitude lower boron concentration near the surface. Fig. 2 shows a typical boron concentration gradient after H_2 anneal extending more than 1.0 μm from the surface. This doping gradient is analogous to P/P+ epi with lower doping near the active surface layer on a more heavily doped bulk region which helps prevent latch-up. The actual boron profile, in reality, is quite different than the epitaxial wafer. At elevated temperatures hydrogen reacts with the boron at the surface of the wafer to form volatile Diborane (B_2H_6). Boron depletion at the surface creates a concentration gradient causing boron from within the bulk to diffuse to the surface and be removed by the hydrogen. This process establishes a boron concentration profile [4] and can be advantageous in some device applications. This doping gradient can help prevent latch-up in CMOS circuits. Note, other ambients such as argon do not achieve boron out diffusion and result in a flat boron profile, but have all the other benefits of high-temperature annealing. For a flat boron profile the annealing is done in argon.

C. Si Vacancy Annealing

One of the main limitations of bulk wafers is silicon vacancies [also known as crystal originated pits (COPs)] that result in poor gate oxide breakdown [5] and inferior charge to break down. High-temperature annealing removes the silicon vacancies to a depth of greater than 5 μm from the surface, which is sufficient to achieve equivalent oxide quality as epi wafers for shallow CMOS active areas of 0.18 μm and smaller technologies. The size and density of Si vacancies are a function of Czochralski (CZ) crystal pull rate, the thermal history of the ingot, and wafer diameter. Fig. 3 shows density versus size of silicon vacancies as a function of crystal pull rate. High-temperature annealing is more efficient in annealing the smaller size and higher density Si vacancies, COPs. Therefore, the crystal pull rate during the CZ process can be increased for PE to achieve a productivity increase of about 15% on crystal pullers by faster pull rate of 200 mm crystals. These COPs are later annealed in the high-temperature hydrogen anneal step. This is also true for 300-mm ingots which have a high density of COPs relative to 200-mm ingots due to differences in crystal thermal process.

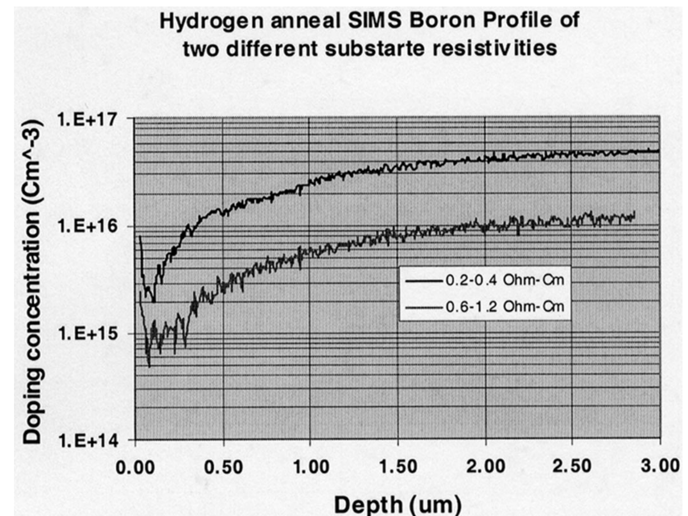


Fig. 2. Boron out diffusion profile after high temperature H_2 anneal.

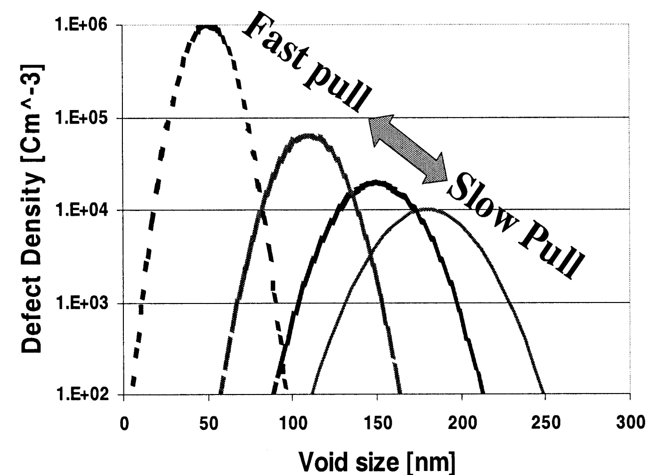


Fig. 3. Density and size of silicon vacancies as a function of Czchrolski crystal pull rate for 200-mm crystal.

D. Nanotopography/Surface Roughness Improvement

PE improves nanotopography [6] and surface roughness of the wafers. This improvement is a result of material flow at elevated temperatures in the hydrogen ambient. The silicon surface reconstructs into a 2×1 structure with step height of about 0.1 to 0.2 nm. Fig. 4 shows an Atomic Force Microscope (AFM) picture of a polished wafer in contrast to PE. The improved nanotopography results in reduced variability in the remaining film thickness post CMP and improves litho variability for CD control.

Authorized licensed use limited to: Teknillinen Korkeakoulu. Downloaded on October 21, 2009 at 06:00 from IEEE Xplore. Restrictions apply.

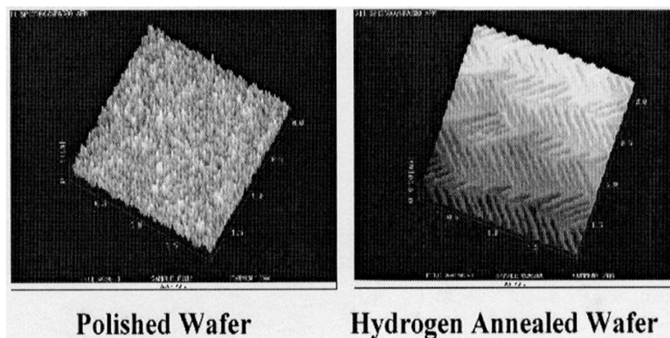


Fig. 4. AFM picture of a polished wafer and how a wafer surface reconstructs improving nanotopography after H_2 anneal.

III. Performance/Cost Optimization

Switching the substrate material from the epitaxial wafer is a paradigm shift for any of the advanced products such as chipsets or microprocessors. In order to achieve this transition, a significant amount of work was done to solve and overcome the stringent product qualification requirements. These included materials technologydevelopment, cost of ownership (CoO), modeling of furnace anneal versus epi process, CAD simulation for device modeling and optimization of PE silicon parameters, process development, and high volume factory support for engineering and qualification evaluations on microprocessors and chipset products. Design design changes were made to products to enable PE material. Reliability groups were involved in device and product reliability characterizations. Also, experimental work was done with silicon suppliers and furnace suppliers to develop a cost-effective anneal process.

A. Furnace Anneal Considerations

PE or hydrogen annealing is achieved in high-temperature vertical furnaces with very good temperature control in a batch process. In order to avoid thermomechanically-induced dislocations (slip) caused by either the thermomechanical stresses exceeding the shear strength of Si or the gravitational stresses causing wafer sagging at elevated temperatures, the temperature ramp rate must be well controlled [7]. For a cost effective annealing process, the ramp rate needs to be carefully balanced to achieve a fast ramp without creating dislocations or slip. A simple fixed ramp rate is not cost effective since it would increase cycle time for a COP free-active area. The gravitational stress strongly depends on boat design, type of wafer support, and wafer spacing in the furnace. For a 200-mm wafer, edge support is sufficient. To avoid dislocations, wafer spacing and ramp rates at a given anneal temperature must be optimized. To achieve efficient annealing of Si vacancies, a high annealing temperature around 1200 °C is required. This was experimentally verified. Lower anneal temperatures do not achieve the same efficiency of boron out-diffusion and would require substantially longer anneal times which are not as cost effective. Note lower anneal temperatures are not sufficient for effective COP annealing and increase in anneal times increases the cycle time and manufacturing costs.

Other considerations for an effective high temperature anneal process are selection of appropriate tube, boat, and wafer support materials. Quartz is not a good candidate for high temperature annealing near 1200 °C. SiC has a much longer life but can be a source of metallic contamination. Often a CVD-coated SiC film is added to a graphite boat and wafer support and subsequently the graphite is burned off at high temperatures to eliminate any metallic contamination. However, this is very costly. Alternatively, boat and wafer support material made out of silicon is also used in the industry. In this case, the boat and wafer support can be made free of metallic contamination. Machining an Si boat is also difficult and relatively expensive. Surface roughness of Si boat or support material needs to be controlled to avoid wafer bonding to the support material at high temperatures.

B. CAD Process Simulations

Computer-aided design simulations were done to evaluate the impact of different anneal time and temperatures on boron out diffusion and its impact on latch-up, junction capacitance, and transistor performance. Different N- and P-well implants and compensation implants were evaluated to closely match the epi process. At first, a reduction of anneal temperature for instance from 1200 °C to 1150 °C is desirable to reduce PE cycle time. However, this would lower the efficiency to anneal out the COPs, and in fact, would need a significantly longer anneal time to properly eliminate COPs in the active area. The lower anneal temperature also reduces the Boron out diffusion and alters the well resistance. To achieve the same boron out diffusion, the anneal time must increase from 1 h at 1200 °C to $\sim$4 h at 1150 °C. This increase in anneal time is not cost effective. Therefore, the device performance dictated the 1200 °C temperature for lowest CoO. Simulation work also defined a window for wafer resistivity of the substrate wafer. A $2\times$ resistivity range was designed in to allow the silicon manufacturing process to utilize the entire CZ ingot length for low manufacturing costs. Also note the substrate resistivity of the PE is about 3 orders of magnitude lower than the P+ epi substrate.

C. Cost of Ownership Modeling

The driving force for switching to PE is lower material cost. There are two factors contributing to this. The first factor is the technical driving force of lower CoO for annealed wafers and the second is mainly business and market driven.

1) Technical Driving Force: Silicon wafer manufacturing processes consist of three main steps: 1) CZ crystal growth; 2) wafering steps which include slicing, grinding, lapping, etching, and polishing to achieve the desired mechanical and geometrical characteristics; and 3) the epitaxial growth of a few microns of Si layer on top of the bulk wafer. Except for higher substrate resistivity, the CZ crystal growth and wafering steps are the same for PE and epi. From the cost perspective, the main difference is that the epi step is replaced with high-temperature annealing in a furnace. A detailed CoO model was developed to compare the added manufacturing cost of epi to the added manufacturing cost of PE. The PE CoO model was utilized to calculate tradeoff parameters such as anneal temperature, anneal time, wafer spacing, wafer support type and material, wafer boat life, and other parameters in the cost model. For a 200-mm wafer, the CoO model indicates that PE has about 50%

Authorized licensed use limited to: Teknillinen Korkeakoulu. Downloaded on October 21, 2009 at 06:00 from IEEE Xplore. Restrictions apply.

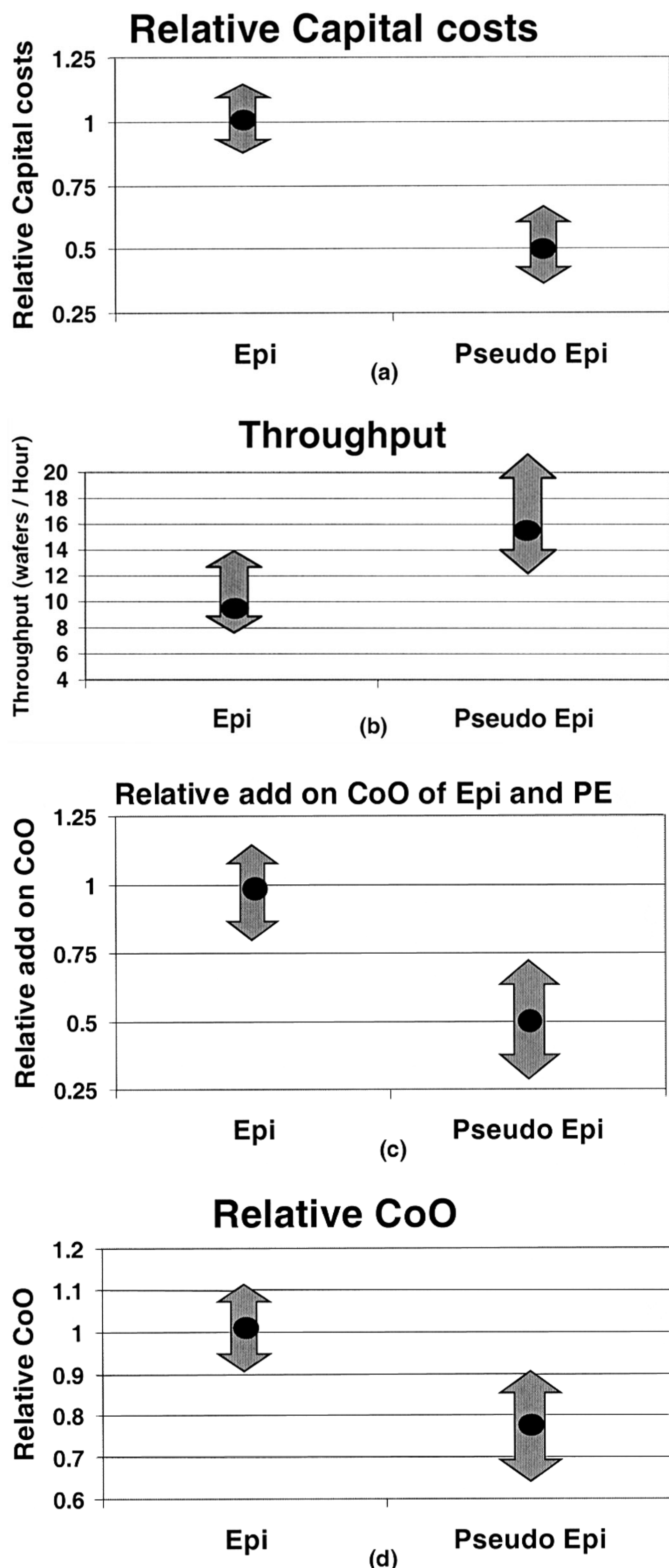


Fig. 5. (a) Relative capital cost comparison of epi and PE. (b) Relative through put of epi and PE. (c) Relative added CoO comparison of epi and PE. (d) Relative CoO comparison of epi and PE.

lower capital cost relative to epi; see Fig. 5(a). The reported throughput numbers from equipment makers varies from the manufacturing numbers reported by silicon suppliers. Fig. 5(b) shows that PE, which is a batch process, has 25% to 30% higher throughput relative to epi for 200-mm wafers. These two factors are primary reasons for lower PE CoO. Therefore, the added cost of PE beyond crystal and wafering is roughly 50% of the added cost of epi as shown in Fig. 5(c). For 200-mm wafers, the yield, metrology, consumable, and operating expenses of epi and hydrogen annealing are nearly the same. The overall impact of the final wafer is about 20% to 25% lower than manufacturing cost for PE relative to epi shown in Fig. 5(d).

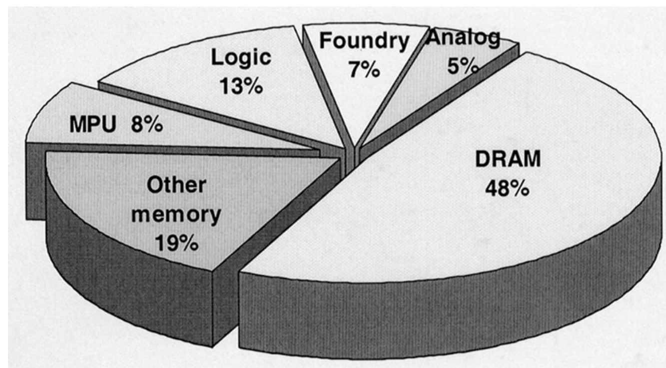


Fig. 6. Silicon consumption by product.

2) Market Driving Force: Fig. 6 shows silicon consumption by product for 2002 [8]. DRAMS are responsible for nearly 50% of the world wide silicon consumption. Logic and microprocessors, and other memory products collectively consume about 40% of the worldwide silicon. The majority of the DRAM products utilize annealed wafers rather than epi because of lower cost. Historically, most logic and other products have used epi for starting material, but recently the trend is mixed in terms of epi or annealed. Since the majority of silicon consumption in the world utilizes annealed wafers, the market driving force for lower prices tends to favor annealed wafers.

D. Robust ESD Design

Earlier evaluations of PE revealed sensitivity to electro static discharge (ESD), a reliability issue. Historical designs relied on high conductivity of P/P+ epi substrate for ESD robustness. The higher resistivity of the bulk of PE substrate resulted in ESD failures on initial attempts. A new design was incorporated to include lateral diodes for I/O structures to increase ESD robustness independent of the substrate resistivity. Starting with 0.18 μm chipset products, as well as all subsequent products, a robust ESD design has been incorporated to ensure the devices are independent of substrate resistivity. This also helps any future transition to silicon on insulator (SOI) where the active area is isolated from the substrate. Subsequent reliability tests verified that product ESD levels for PE are equivalent to the standard epi process.

IV. Reduced Device Sensitivity to Start Material

Fig. 7 shows a schematic cross section of a CMOS circuit using an epitaxial wafer with trends of reduced sensitivity to start material as device scaling continues [9]. Continual device scaling of minimum CDs, lower operating voltage, and thinner gate material has made state-of-the-art devices with 0.18 μm or smaller significantly less sensitive to some of the historically tight start material parameters.

Authorized licensed use limited to: Teknillinen Korkeakoulu. Downloaded on October 21, 2009 at 06:00 from IEEE Xplore. Restrictions apply.

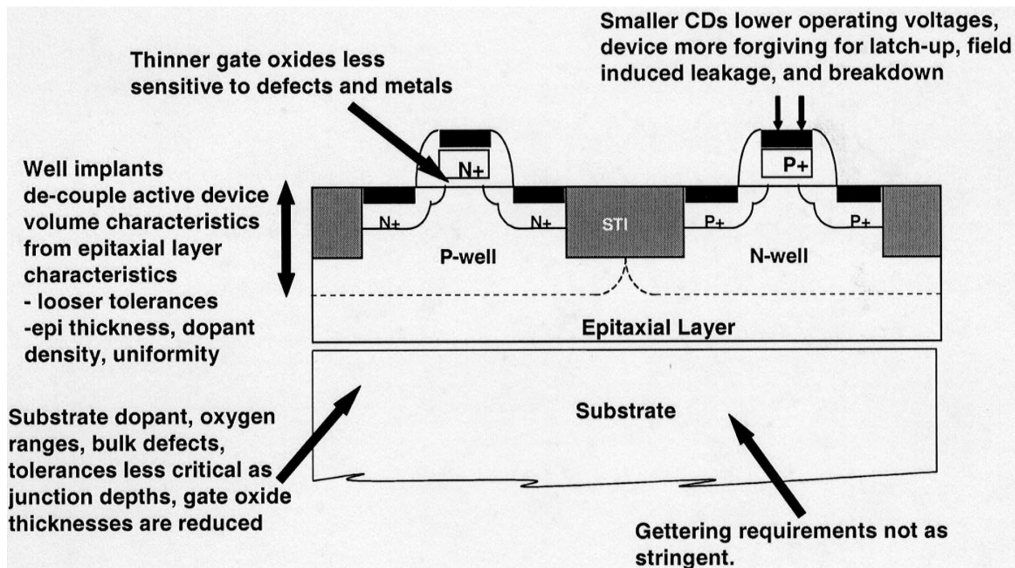


Fig. 7. Advanced CMOS sensitivities to substrate material.

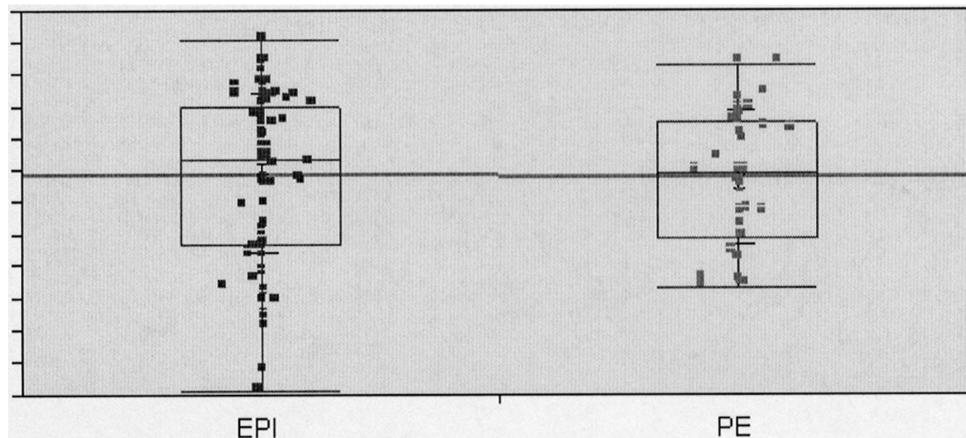


Fig. 8. Relative yield data for epi and PE on eight lots for a 0.18-μm chipset product.

The extremely thin gates have a high level of leakage due to direct tunneling, so gates are less sensitive to metallic impurities from the substrate. Intrinsic gettering requirements are not as stringent as previous technologies since the manufacturing lines are free of metallic contamination and there is less need for intrinsic gettering. Due to lower operating voltage, latch-up requirements for CMOS circuits are less of an issue. As junction depths are reduced and implants are increased, there is a higher tolerance for substrate or epi dopant control, bulk oxygen concentration, bulk defects, epi thickness, and uniformity tolerance of the start material. This makes annealed wafers good candidates for future technologies. Of course, the mechanical and geometrical parameters such as flatness and nanotopography scaling do continue with device scaling. Other major trends are shift from bulk to SOI in the industry.

V. Pseudo Epi Results

We have targeted 0.18-μm chipset products to switch to PE. Engineering evaluations of PE on this process resulted in equivalent yield and E-test results. Fig. 8 shows equivalent relative yields on epi and PE on several lots on a 0.18-μm chipset product. Note that there was no PE specific process optimization and the epi and PE wafers had exactly the same process steps. A significant amount of engineering activity was incorporated to ensure the impact of PE was understood both from an E-test and reliability perspective. PE passed all qualification requirements for E-test parameters. PE had equivalent E-test results as epi on gate oxide leakage, oxide breakdown voltage, P–N junction leakage, transistor threshold voltage, and all transistor and isolation parameters. Fig. 9 shows distributions of threshold voltages of an N-channel transistor which are essentially identical for eight epi and PE lots. Reliability data such as latch-up, burn-in, and ESD data on various products on a 0.18-μm chipset technology were equivalent. Additionally, initial evaluations of PE on 0.13 μm logic technology have been promising in terms of die yield and E-test parameters. Reliability tests are ongoing for the 0.13-μm logic technology.

VI. Production Plans/Cost Savings

Product pilot and conversion plans are in place and the full qualification of products on the 0.18-μm chipset process is

Authorized licensed use limited to: Teknillinen Korkeakoulu. Downloaded on October 21, 2009 at 06:00 from IEEE Xplore. Restrictions apply.

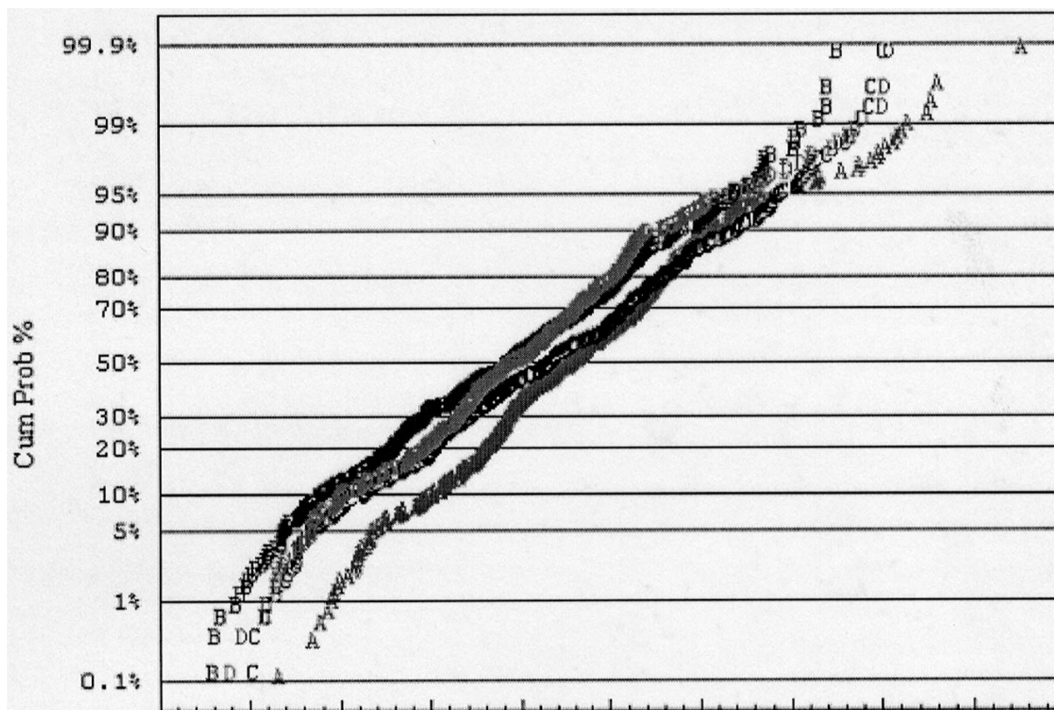


Fig. 9. N-channel relative threshold voltage (Vt) distribution. Symbol A is epi and the rest are PE wafers. Epi and PE Vt distributions are the same for multiple lots.

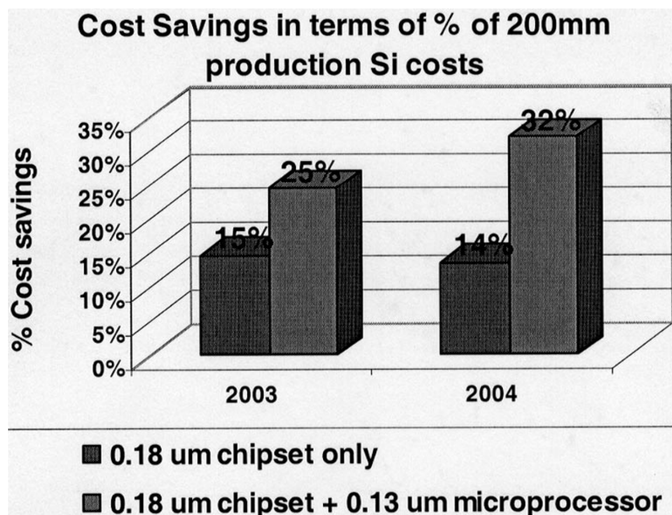


Fig. 10. Cost savings using PE in terms of percent of 200-mm production Si costs.

ongoing. PE will save 20%–25% in starting silicon material cost. We expect a materials cost savings equivalent to 15% of our 2003 silicon production costs just for the 0.18 μm chipset process (Fig. 10). Initial feasibility evaluations of our 0.13-μm CPU technology with PE have shown favorable yield and E-test results. The qualification and conversion of 0.18 μm chipset products would build additional confidence to qualify and convert the 0.13-μm CPU technology. Once implemented, this will save 25% of total silicon production costs when PE is extended to 0.13-μm microprocessor technology in 2003. This savings is expected to increase to 32% in 2004 for these two technologies.

VII. PE Challenges for 300-mm Wafers

Because the wafer thickness changes little from 200 to 300 mm, the gravitational stress on a 300-mm wafer is significantly higher than 200-mm wafers in a high-temperature annealing furnace. In order to have manufacturable 300-mm hydrogen anneal processes, the conventional three or four point support boats will no longer be effective. Ring support or full wafer support is necessary to avoid slip [7]. This adds significantly to the manufacturing cost of PE relative to 200 mm. The challenge for high-temperature furnace suppliers is to make the hydrogen anneal process have about 50% of the added cost of the 300-mm epi process.

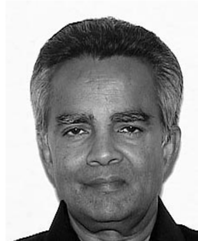
References

[1] "300mm factory costs," Intel Corp. internal memo, July 2002.
[2] S. Nadahara, H. Kubota, and S. Samata, "Hydrogen annealed silicon wafer," in *Diffusion and Defect Data: Solid State Phenomena* Zurich, Switzerland, 1997, vol. 57–58.
[3] J. Fulford, D. Wristers, and M. Gardner, "Effect of hydrogen denudation on thin oxides, device performance, and epitaxial elimination," in *Proc. 11th Biennial Univ./Government/Industry Microelectronics Symp.*, Austin, TX, May 1995.
[4] K. V. Ravi, L. Ling, and S. Hu, "Silicon wafers for CMOS and other integrated circuits," US Patent 6 423 615 B1, July 23, 2002.
[5] H. Abe, I. Suzuki, and H. Koya, "Effect of hydrogen annealing on oxygen precipitation behavior and gate oxide integrity in Czochralski Si wafers," *J. Electrochem. Soc.*, vol. 144, Jan. 1997.
[6] S. Kawamoto, K. Izunome, K. Kohtari, M. Kurokawa, T. Nakatani, T. Sakamoto, and K. Kashimo, "Nanotopography characterization in Czochralski-grown silicon wafer in hydrogen annealing," in *Extended Abstract 3rd Int. Conf. Materials Microelectronics*, Ireland: Inst. Physics, 2000, no. P16.
[7] R. Takeda, P. Xin, J. Yoshikawa, Y. Kirino, Y. Matsushita, Y. Hosoki, N. Tsuchiya, and O. Fujii, "300 mm diameter hydrogen annealed silicon wafers," *J. Electrochem. Soc.*, vol. 144, Oct. 1997.
[8] Semiconductor Industry Assoc., "World semiconductor forecasts,", June 2002.
[9] K. V. Ravi, "Low cost and high quality—The challenge for silicon wafers," in *Proc. Annu. Electronic Materials Symp.*, Apr. 2000, (Fig. 7).

Authorized licensed use limited to: Teknillinen Korkeakoulu. Downloaded on October 21, 2009 at 06:00 from IEEE Xplore. Restrictions apply.

Mehran Aminzadeh received the B.S.E.E. degree from Ohio State University, in 1983, and the M.S.E.E. and Ph.D. degrees from Oregon State University, in 1985 and 1988, with emphasis on electrical characterization of materials.

He joined Intel Corporation in 1988 and worked on electrical characterization of silicon wafers. From 1996 to 1999, he pursued work in the field of flat panel display in candescent. In 1999, he rejoined Intel in the materials organization and has been working as a Silicon Technologist on materials cost savings projects.



K. V. Ravi received the the B.S. degree in metallurgy from the Indian Institute of Science and B.Sc. degree in chemistry from the University of Madras, India, the M.S. degree in materials science from the University of California, Berkeley, and the Ph.D. degree in materials science from Case Western University.

He has been with Intel Corporation for approximately five years focusing on new materials development as they relate to silicon wafers. His prior affiliations include Applied Materials, Lockheed-Martin Corporation, Crystallume, Mobil Solar Energy Corporation, Motorola, Inc., and Texas Instruments Inc. He is the author of over 100 technical publications in the fields of metals, semiconductors, photovoltaics and CVD diamond technology and the author of the book *Imperfections and Impurities in Semiconductor Silicon*.

George E. Sery (S'76–M'78) received the B.S. and M.S. degrees in electrical engineering from the University of Minnesota, Mineapolis, in 1976 and 1978, respectively.

He is an Intel Fellow and Director of Device Technology Optimization in Intel's California Technology and Manufacturing Group. He is currently responsible for directing process characterization, performance improvement, and capability enhancement for Intel's 0.13-μm CMOS logic technology. He joined Intel in 1978 as part of the SRAM Technology Development Group. He has been involved with development of NMOS and CMOS technologies for logic, SRAM, and Flash memory applications. For each technology, he has led the device physics team responsible for device development and process characterization. He has co-authored numerous publications related to logic and Flash memory development and holds patents in Flash technology.

Dr. Sery has received three Intel Achievement Awards for Flash memory and logic development improvements.

Amiad Conley received the B.S. degree in material engineering and the B.A. degree in physics from the Technion - Israel Institute of Technology, in 1994. He received the M.Sc. degree in applied physics from the Hebrew University of Jerusalem, in 1996.

He joined Intel Corporation, in 1994, and worked in yield analysis and process integration areas until 2000. Since then he manages the product engineering and device group in one of Intel's fabs.

Matt Casperson received the B.S.E.E. degree from Sacramento State University, in 1995.

He joined Intel Corporation in 1995 and has worked in product development engineering since that time. He is currently working on several cost reduction projects.

Authorized licensed use limited to: Teknillinen Korkeakoulu. Downloaded on October 21, 2009 at 06:00 from IEEE Xplore. Restrictions apply.