INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS

2001 EDITION

FRONT END PROCESSES

Table of Contents

Scope ... **1**
Difficult Challenges ... **2**
Technology Requirements and Potential Solutions ... **5**
Starting Materials ... 5
Surface Preparation ... 15
Thermal/Thin Films And Doping ... 19
Thermal/Thin-Films ... 19
Doping Technology ... 27
Source and Drain Extensions ... 27
Contact ... 27
Polysilicon Gate Electrode Doping ... 28
Channel ... 28
Front End Etch Processing ... 30
DRAM Stacked Capacitor ... 32
DRAM Trench Capacitor ... 36
Non-Volatile Memory (FLASH) ... 37
Ferroelectric Random Access Memory (FeRAM) ... 39
Memory Capacity ... 39
Cell Size ... 40
Ferroelectric Materials Alternatives ... 40
Estimated Minimum Switching Charge ... 41
Conclusion ... **43**

List of Figures

Figure 26 Front End Processes Chapter Scope ... 1
Figure 27 Starting Materials Potential Solutions ... 14
Figure 28 Surface Preparation Potential Solutions ... 18
Figure 29 Thermal / Thin Films Potential Solutions ... 26
Figure 30 Doping Potential Solutions ... 29
Figure 31 Front End Processes Etching Potential Solutions ... 32
Figure 32 DRAM Stacked Capacitor Potential Solutions ... 35
Figure 33 FeRAM Potential Solutions ... 43

LIST OF TABLES

Table 48a Front End Processes Difficult Challenges—Near-term 3
Table 48b Front End Processes Difficult Challenges—Long-term.................................. 4
Table 49a Starting Materials Technology Requirement—Near-term 6
Table 49b Starting Materials Technology Requirements—Long-term 10
Table 50a Surface Preparation Technology Requirements*—Near-term 16
Table 50b Surface Preparation Technology Requirements*—Long-term.......................17
Table 51a Thermal & Thin Film, Doping and Etching Technology Requirements—Near-term 20
Table 51b Thermal & Thin Film, Doping and Etching Technology Requirements—Long-term 24
Table 52a DRAM Stacked Capacitor Films Technology Requirements—Near-term 34
Table 52b DRAM Stacked Capacitor Films Technology Requirements—Long-term 35
Table 53a DRAM Trench Capacitor Technology Requirements—Near-term 36
Table 53b DRAM Trench Capacitor Technology Requirements—Long-term 37
Table 54a FLASH Non-volatile Memory Technology Requirements—Near-term 38
Table 54b FLASH Non-volatile Memory Technology Requirements—Long-term 39
Table 55a FeRAM Technology Requirements—Near-term 41
Table 55b FeRAM Technology Requirements—Long-term..42

FRONT END PROCESSES

SCOPE

The Front End Processes (FEP) Roadmap focuses on future process requirements and potential solutions related to scaled field effect transistors (MOSFETs), DRAM storage capacitors, as well as Flash and ferroelectric RAM (FeRAM) devices. The purpose of this chapter is to define comprehensive future requirements and potential solutions for the key front end wafer fabrication process technologies and materials associated with these devices. Hence, this Roadmap encompasses the tool, and materials, as well as the unit and integrated processes starting with the silicon wafer substrate and extending through the contact silicidation and pre-metal dielectric layer preparation processes. The following specific technology areas are covered: starting materials, surface preparation, thermal/thin films and doping for MOSFETs, pre-metal dielectric layers, front end plasma etch, as well as processes and materials for DRAM stack and trench capacitors, Flash memory gate structures, and FeRAM storage devices.

A forecast of scaling-driven technology requirements and potential solutions is provided for each technology area. The forecasted requirements tables are model-based unless otherwise noted. The potential solutions identified serve to benchmark known examples of possible solutions, and are intended for other researchers and interested parties. They are not to be considered the only approaches. Indeed, innovative**,** *novel solutions* are sought, and are identified by red colored regions of the requirements tables.

Some FEP related topics are presented in other sections of this Roadmap. The scaled device performance and structures forecasts that drive FEP requirements are covered in the *Process Integration, Devices, and Structures* chapter. The tool-related issues for plasma etch and chemical mechanical polish (CMP) for trench isolation are found in the *Interconnect* chapter because of overlap with interconnect tool issues. The crosscut needs of FEP are covered in the following chapters: *Yield Enhancement*, *Metrology*, *Environment, Safety, & Health*, and *Modeling & Simulation*. FEP factory requirements are covered in the *Factory Integration* chapter.

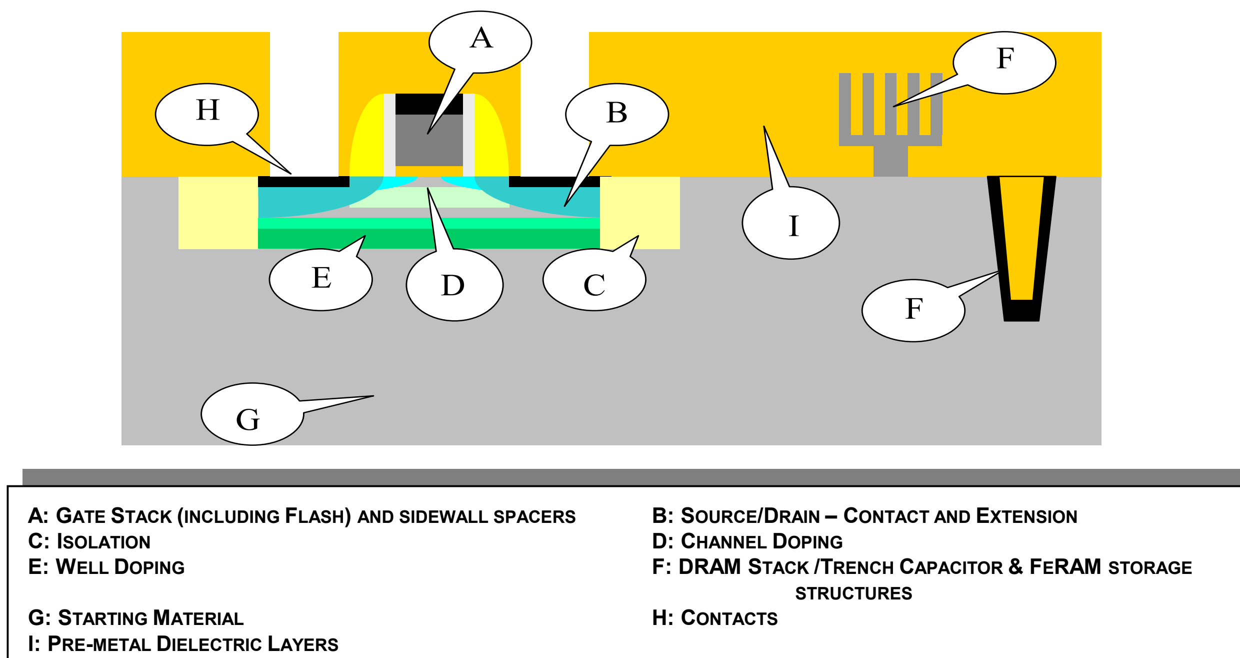


Figure 26 Front End Processes Chapter Scope

DIFFICULT CHALLENGES

THE FRONT END PROCESSES GRAND CHALLENGE—
THE FEP RESPONSE TO THE ERA OF MATERIAL LIMITED DEVICE SCALING

Device scaling has been the primary means by which the semiconductor industry has achieved the historically unprecedented gains in productivity and performance quantified by Moore's Law. These gains have traditionally been paced by the development of new lithography tools, masks, photoresist materials and critical dimension etch processes. In the past several years it has become clear that despite advances in these crucial process technologies and the resultant ability to produce ever-smaller feature sizes, front end technologies have not kept pace, and scaled device performance has been compromised. The crux of this problem comes from the fact that the traditional transistor and capacitor formation materials, silicon, silicon dioxide, and polysilicon have been pushed to fundamental material limits and continued scaling has required the introduction of new materials. *We have entered the era of material limited device scaling!*

The emergence of this era has placed new demands on virtually every front end material and unit process, starting with the silicon wafer substrate and encompassing the fundamental planar CMOS building blocks and memory storage structures. In addition, the end of planar CMOS is becoming visible within the time horizon of this Roadmap. As a consequence we must be prepared for the emergence of CMOS technology that uses non-conventional dual gate MOSFETs or alternatives. Some believe these may be needed as early as the year 2007. The challenges associated with these diverse new materials and the control of the physical interfaces associated with these materials constitutes the central theme of the FEP difficult challenges summarized in Table 48.

In no area is this issue more clear or urgent than in the MOSFET gate stack. Here, a new gate dielectric material having a higher dielectric constant is needed. This need was identified in the 1999 ITRS, linked to MOSFETs having gate lengths smaller than 65nm, which were expected to emerge in the year 2005. In the interim, the patterning technology for producing 65 nm gates has accelerated and these are now expected in the year 2001. Although promising high κ candidate materials have been identified, fundamental performance and reliability issues, as well as issues with CMOS integration are still under investigation. It is doubtful that these materials will enter production before the year 2005. In the interim, evolution of the oxynitride gate dielectric materials is expected to continue, but with compromised device performance, notably as it relates to higher off-state leakage and power consumption. It is not surprising therefore that the most urgent need for the high κ dielectric is associated with devices designed for low standby power. Here, it is expected that the high κ material will be needed by the year 2005 if off-state power consumption expectations are to be met. Looking beyond the gate dielectric, the depletion layers that exist in the doped polysilicon gate material become increasingly onerous as planar devices are scaled into the deep submicron region with the result that dual metal gates are expected to replace the dual doped polysilicon gates currently the mainstay of CMOS technology. Ultimately, scaling is expected to require the replacement of planar CMOS devices with non-standard devices.

Table 48a Front End Processes Difficult Challenges—Near-term

THROUGH 2007, MPU/ASIC PHYSICAL GATE LENGTH ≥ 25 nm	*SUMMARY OF ISSUES*
New gate stack processes and materials	Extension of Oxynitride gate dielectric materials to < 1.0 nm E.O.T for high performance MOSFETs, consistent with device reliability requirements Introduction and process integration of high κ gate stack materials and processes for low operating and low standby power MOSFETs Control of boron penetration from doped polysilicon gate electrode Minimized depletion of dual doped polysilicon electrodes Possible introduction of dual metal gate electrodes with appropriate work function (toward end of period) Metrology issues associated with gate dielectric film thickness and stack electrical and materials characterization
Critical dimension and effective channel length (L_{eff}) control	Control of gate etch process that yield a physical gate length that is smaller than the feature size printed in the resist, while maintaining <10% overall 3-sigma control of the combined lithography and etch processes Control of profile shape, edge roughness, line and space width for isolated as well as closely-spaced fine line patterns Control of self-aligned doping processes and thermal activation budgets to achieve ~15% 3σ L_{eff} control Maintenance of CD and profile control throughout the transition to new gate stack materials and processes CD and etch metrology
CMOS integration of new memory materials and processes	Development and introduction of very high κ DRAM capacitor dielectric layers Migration of DRAM capacitor structures from silicon-insulator-metal to metal-insulator-metal Integration and scaling of FeRAM ferroelectric materials Scaling of Flash inter-poly and tunnel dielectric layers may require high κ Limited temperature stability of high κ and ferroelectric materials challenges CMOS Integration
Surfaces and interfaces: structure, composition and contamination control	Contamination, composition, and structure control of channel/gate dielectric interface as well as gate dielectric/gate electrode interface Interface control for DRAM capacitor structures Maintenance of surface and interface integrity through full-flow CMOS processing Statistically significant characterization of surfaces having extremely low defect concentrations for starting materials and pre-gate clean surfaces
Scaled MOSFET dopant introduction and control	Doping and activation processes to achieve shallow source/drain regions having parasitic resistance that is less than ~16–20% of ideal channel resistance (=V_{dd}/I_{on}) Control of parasitic capacitance to achieve less than ~19–27% of gate capacitance, consistent with acceptable I_{on} and minimum short channel effect Achievement of activated dopant concentration greater than solid solubility in dual-doped polysilicon gate electrodes Formation of continuous self-aligned silicide contacts over shallow source and drain regions Metrology issues associated with 2D dopant profiling

Table 48b Front End Processes Difficult Challenges—Long-term

BEYOND 2007, MPU/ASIC PHYSICAL GATE LENGTH <25 nm	ISSUES
Continued scaling of planar CMOS devices	Higher κ gate dielectric materials including temperature constraints Metal gate electrodes with appropriate work function Sheet resistance of clad junctions Enhanced channel mobility e.g., strained layers CD and L_{eff} control Chemical, electrical and structural characterization
Introduction and CMOS integration of non-standard, double gate MOSFET devices	Devices may be needed as early as 2007 Selection and characterization of optimum device types CMOS integration with other devices, including planar MOSFETs Introduction, characterization and production hardening of new FEP unit processes Device and FEP process metrology Increased funding of long term research
Starting silicon material alternatives beyond 300 mm	Need for future productivity enhancement dictates the requirement for a next generation, large silicon substrate material Historical trends suggest that the new starting material have nominally twice the area of present generation substrates, e.g., 450 mm Economies of the incumbent Czochralski crystal pulling, wafer slicing, and polishing processes are questionable beyond 300 mm; research is required for a cost-effective substrate alternative to bulk silicon
New memory storage cells, storage devices, and memory architectures	Scaling of DRAM storage capacitor beyond $6F^2$ and ultimately beyond $4F^2$ Further scaling of Flash memory interpoly and tunnel oxide thickness FeRAM storage cell scaling Introduction of new memory types and storage concepts
Surface and interface structural, contamination, and compositional control	Achievement and maintenance of structural, chemical, and contamination control of surfaces and interfaces that may be horizontally or vertically oriented relative to the chip surface Metrology and characterization of surfaces that may be horizontally or vertically oriented relative to the chip surface Achievement of statistically significant characterization of surfaces and interfaces that may be horizontally or vertically oriented relative to the chip surface

High κ materials and new metal-insulator-metal capacitor structures are also expected to evolve in the DRAM segment where the need for reduced on-chip storage node area is driving the need for very high unit capacitances. Also, it is expected that high κ materials will be required for the Flash memory inter-poly and tunnel dielectric layers. In the memory area it is also expected that FeRAM will make a significant commercial appearance where ferroelectric storage materials would be used. The introduction of these diverse materials into the manufacturing mainstream is viewed as an important difficult challenge.

In the starting wafer area, it is expected that alternatives to bulk silicon such as strained silicon on unstrained silicon-germanium and strained silicon-germanium carbon, as well as silicon-on-insulator substrates will proliferate. These all imply FEP process architecture changes. An important difficult challenge expected to emerge within this Roadmap horizon is the need for the next generation silicon substrate. Here, it is questionable whether the incumbent techniques for crystal growth and wafer preparation can be cost-effectively scaled to the next generation. The search for potential substrate alternatives presents an important research need, with significant CoO implications for the economic growth of the industry.

The etching processes used to form the critical dimension features such as MOSFET gates, and DRAM word and bit lines continue to pose difficult challenges in terms of CD and line profile shape control. These problems are expected to become more difficult as etch techniques are increasingly employed to produce feature sizes that are smaller than those printed in the photoresist. This added process complexity and the associated variance increases must all be managed to

achieve the final physical feature size tolerance. The introduction of new materials is expected to add to the difficulty of these tasks.

The introduction of new materials is expected to impose added challenges to the methods used to dope and activate silicon. In addition to the scaling imposed need for producing very shallow, highly activated junctions, the limited thermal stability of many high κ materials is expected to place additional boundaries on the thermal budgets associated with dopant activation. In a worst-case scenario, the introduction of these new materials could have a significant impact on critical transistor parameters. This might necessitate changes not only in the detailed fabrication process procedures but, perhaps, might also require implementation of unique device architectures

TECHNOLOGY REQUIREMENTS AND POTENTIAL SOLUTIONS

STARTING MATERIALS

Starting materials selection has traditionally involved the choice of either polished Czochralski (Cz) or epitaxial silicon wafers. The Czochralski wafer menu has recently been enriched with the addition of surface-enhanced, hydrogen- or argon-heat treated wafers as well as wafers prepared from specially grown, low-defect crystals. In addition, Silicon-on-insulator (SOI) wafers now appear poised to become more than a niche technology. Commodity devices such as DRAM are commonly manufactured on lower cost Cz polished wafers, while high-performance logic ICs are generally manufactured on more costly epitaxial wafers because their use has facilitated the achievement of greater manufacturing yield (e.g., gate dielectric integrity) and robustness (e.g., latch-up suppression capability). This latter capability may no longer be as critical due to the implementation of shallow trench isolation (STI) and the development of alternate doping means for achieving latch-up suppression. Nevertheless, yield issues together with the anticipated reduction in the fabrication cost of epitaxy wafers, favors continued use of epitaxial wafers for logic applications. On the other hand, process flow simplification and enhanced high-frequency logic performance is creating an opportunity for more costly SOI wafers. This wide variety of starting materials is likely to continue into the foreseeable future and is addressed in the potential solutions depicted in *Figure 27.*

Table 49a and b forecasts requirements for wafers as received from the silicon wafer manufacturer, intended for use in the manufacture of both DRAMs and high-performance MPUs. These requirements include general wafer parameters plus specific parameters appropriate to polished, epitaxial, and SOI wafers. Fundamental barriers presently limit the rate of cost effective improvement in specific wafer characteristics such as localized light scatterers (LSS) defect densities, site flatness values, and edge exclusion dimensions. These barriers include the capability and throughput limitations of metrology tools, as well as wafer manufacturing cost and yield issues fundamental to the crystal-pulling process and subsequent wafer finishing operations.

Accordingly, a methodology has been introduced to display not only the ability of the wafer supplier to meet the parameter trends in Table 49a and b, but to also display both the criticality of the parameter from the device manufacturer's perspective, with the latter labeled as "showstoppers," "process margin improvements," or "maintenance" as well as the metrology tool readiness. For this reason, the wafer parameter trend table cells have been annotated with device performance criticality color codes on the lower left-hand side of each cell and the metrology tool readiness color codes depicted on the lower right-hand side. The marking system and meanings are shown in the key below for both DRAM and high-performance MPUs.

STARTING MATERIALS REQUIREMENTS CELL KEY

Technology requirement value and supplier manufacturing capability →

Criticality of wafer attribute to IC wafer user→

←Metrology readiness capability

STARTING MATERIALS TECHNOLOGY REQUIREMENTS TABLE LEGEND

White–Manufacturable Solutions Exist and Are Being Optimized

Yellow–Manufacturable Solutions are Known

Red–Manufacturable Solutions are NOT Known

Table 49a Starting Materials Technology Requirement—Near-term

Year of Production	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*Driver*
DRAM ½ Pitch (nm)	*130*	*115*	*100*	*90*	*80*	*70*	*65*	*D ½*
MPU / ASIC ½ Pitch (nm)	*150*	*130*	*107*	*90*	*80*	*70*	*65*	*M*
MPU Printed Gate Length (nm)	*90*	*75*	*65*	*53*	*45*	*40*	*35*	*M*
MPU Physical Gate Length (nm)	*65*	*53*	*45*	*37*	*32*	*28*	*25*	*M*
DRAM Total Chip Area (mm^2)	**127**	**100**	**118**	**93**	**147**	**116**	**183**	*D ½*
DRAM Active Transistor Area (mm^2)	**55.3**	**36.9**	**59.0**	**42.3**	**73.1**	**53.2**	**89.9**	*D ½*
MPU High-Performance Total Chip Area (mm^2)	**310**	**310**	**310**	**310**	**310**	**310**	**310**	*M*
MPU High-Performance Active Transistor Area (mm2)	**28.7**	**28.2**	**27.7**	**27.2**	**26.8**	**26.8**	**26.8**	*M*
*General Characteristics * (99% Chip Yield) [A,B,C]*								
Wafer diameter (mm) **	**300*** **	**300*** **	**300*** **	**300**	**300**	**300**	**300**	*D ½, M*
Edge exclusion (mm) ****	**3**	**3**	**2**	**2**	**2**	**2**	**2**	*D ½, M*
Front surface particle size (nm), latex sphere equivalent [D]	**≥ 90**	**≥ 90**	**≥ 90**	**≥45**	**≥40**	**≥35**	**≥33**	*D ½, M*
Particles (cm^{-2}) [E]	**≤ 0.18**	**≤ 0.27**	**≤ 0.17**	**≤ 0.24**	**≤ 0.14**	**≤ 0.19**	**≤ 0.11**	*D ½*
Particles (#/wf)	**≤ 123**	**≤ 185**	**≤ 117**	**≤ 164**	**≤ 95**	**≤ 130**	**≤ 77**	*D ½*
Critical surface metals (at/cm^2) [F]	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	*D ½, M*
Site flatness (nm) [G]	**≤ 130**	**≤ 115**	**≤ 100**	**≤ 90**	**≤ 80**	**≤ 70**	**≤ 65**	*D ½, M*
*Polished Wafer * (99% Chip Yield)*								
The LLS requirement is specified for particles only; discrimination between particles and COPs is required (see General Characteristics) [D,E]								
Total bulk Fe (at/cm^3) [H]	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	**≤ 1 E10**	*D ½, M*
Oxidation stacking faults (OSF) (DRAM) (cm^{-2}) [I]	**≤ 2.8**	**≤ 2.3**	**≤ 1.9**	**≤ 1.6**	**≤ 1.4**	**≤1.1**	**≤ 1.0**	*D ½*
Oxidation stacking faults (OSF) (MPU) (cm^{-2}) [I]	**≤ 1.0**	**≤ 0.8**	**≤ 0.6**	**≤ 0.5**	**≤ 0.4**	**≤ 0.3**	**≤ 0.3**	*M*
*Epitaxial Wafer * (99% Chip Yield)*								
Total Allowable Front Surface Defect Density is The Sum of Epitaxial Large Structural Defects, Small Structural Defects and Particles (see General Characteristics) [J,K]								
Large structural epi defects (DRAM) (cm^{-2}) [L]	**≤ 0.008**	**≤ 0.010**	**≤ 0.009**	**≤ 0.011**	**≤ 0.007**	**≤ 0.009**	**≤ 0.005**	*D ½*
Large structural epi defects (MPU) (cm^{-2}) [L]	**≤ 0.003**	**≤ 0.003**	**≤ 0.003**	**≤ 0.003**	**≤ 0.003**	**≤ 0.003**	**≤ 0.003**	*M*
Small structural epi defects (DRAM) (cm^{-2}) [M]	**≤ 0.016**	**≤ 0.020**	**≤ 0.017**	**≤ 0.022**	**≤ 0.014**	**≤ 0.017**	**≤ 0.011**	*D ½*
Small structural epi defects (MPU) (cm^{-2}) [M]	**≤ 0.006**	**≤ 0.006**	**≤ 0.006**	**≤ 0.006**	**≤ 0.006**	**≤ 0.006**	**≤ 0.006**	*M*

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow–Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known

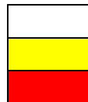
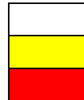


Table 49a Starting Materials Technology Requirement—-Near-term (continued)

YEAR OF PRODUCTION	2001	2002	2003	2004	2005	2006	2007	DRIVER
DRAM ½ PITCH (nm)	130	115	100	90	80	70	65	D ½
MPU / ASIC ½ PITCH (nm)	150	130	107	90	80	70	65	M
MPU PRINTED GATE LENGTH (nm)	90	75	65	53	45	40	35	M
MPU PHYSICAL GATE LENGTH (nm)	65	53	45	37	32	28	25	M
DRAM Total Chip Area (mm^2)	**127**	**100**	**118**	**93**	**147**	**116**	**183**	***D ½***
DRAM Active Transistor Area (mm^2)	**55.3**	**36.9**	**59.0**	**42.3**	**73.1**	**53.2**	**89.9**	***D ½***
MPU High-Performance Total Chip Area (mm^2)	**310**	**310**	**310**	**310**	**310**	**310**	**310**	***M***
MPU High-Performance Active Transistor Area (mm^2)	**28.7**	**28.2**	**27.7**	**27.2**	**26.8**	**26.8**	**26.8**	***M***
Silicon-On-Insulator Wafer* (95% Chip Yield)								
Wafer diameter (mm) **	**200**	**300*****	**300*****	**300**	**300**	**300**	**300**	D ½, M
Silicon final device layer thickness (Partially Depleted) (tolerance ± 5%) (nm) [N]	**98–163**	**80–133**	**68–113**	**56–93**	**48–80**	**42–70**	**38–63**	M
Silicon final device layer thickness (Fully Depleted) (tolerance ± 5%) (nm) [O]	**20–33**	**16–27**	**14–23**	**11–19**	**10–16**	**8–14**	**8–13**	M
Buried oxide (BOX) thickness (Fully Depleted) (tolerance ± 5%) (nm) [P]	**49–81**	**40–66**	**34–56**	**28–46**	**24–40**	**21–35**	**19–31**	M
D_{LASOI}, Large area SOI wafer defects (DRAM) (cm^{-2}) [Q]	**≤ 0.040**	**≤ 0.051**	**≤ 0.043**	**≤ 0.055**	**≤ 0.035**	**≤ 0.044**	**≤ 0.028**	D ½
D_{LASOI}, Large area SOI wafer defects (MPU) (cm^{-2}) [Q]	**≤ 0.017**	**≤ 0.017**	**≤ 0.017**	**≤ 0.017**	**≤ 0.017**	**≤ 0.017**	**≤ 0.017**	M
D_{SASOI}, Small area SOI wafer defects (DRAM) (cm^{-2}) [R]	**≤ 0.464**	**≤ 0.695**	**≤ 0.434**	**≤ 0.607**	**≤ 0.351**	**≤ 0.482**	**≤ 0.285**	D ½
D_{SASOI}, Small area SOI wafer defects (MPU) (cm^{-2}) [R]	**≤ 0.894**	**≤ 0.910**	**≤ 0.926**	**≤ 0.942**	**≤ 0.956**	**≤ 0.956**	**≤ 0.956**	M
D_{EC}, Extended Crystal Defects (MPU) (cm^{-2}) [S]	**1.8 E5**	**1.8 E5**	**1.9 E5**	**1.9 E5**	**1.9 E5**	**1.9 E5**	**1.9 E5**	M

White–Manufacturable Solutions Exist, and Are Being Optimized
Yellow–Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known



** Parameters define limit values; independent predictors of yield, mathematically, or empirically modeled at 99% (or 95% for SOI). Limit values rarely coincide for more than one parameter. A given wafer will generally not exhibit more than one limit value "at a time;" other parameter values most likely near median value, thereby ensuring total yield for all parameters is at least 99% (or 95% for SOI).*
*** Significant gaps in metrology and wafer manufacturing equipment need to be closed on 200 mm, especially for the 130 nm node, inasmuch as 300 mm is only now being phased in and 200 mm will still be prevalent through the 130 nm node.*
**** Numerical values are for 300 mm, although 200 mm will be the dominant polished, epitaxial, and SOI wafer diameter.*
***** Edge exclusion = 3 mm is consistent with equipment and wafer fab capabilities through 2002. Reduction in the edge exclusion beyond 2002 is desirable to increase chips per wafer.*

Table 49a and b notes:

***A.** Targets are given for the surface concentration of carbon atoms after cleaning. The 2001 target value is based on the assumption that a 10% (7.3 x 10^{13} atoms/cm^2) carbon atom coverage on a bare silicon (100) surface after cleaning can be tolerated during device fabrication. Organics/polymers are therefore modeled approximately 0.1 of a monolayer, ≤ 1 x10^{14} C atom/cm^2. Surface organic levels are highly dependent on wafer packaging, on hydrophobic or hydrophilic wafer surface conditions, and on wafer storage conditions such as temperature, time and ambient..*

***B**. Front-surface microroughness ≤ 0.10 nm (RMS) for all CD generations; instrumentation choice, target values, and spatial frequency range (scan size) are selected based on application. Power spectral density analysis is recommended to utilize full accessible range of instruments.*

***C**. The oxygen concentration may be specified depending on the particulars of the IC user, generally specified in the range of 18-31 ± 2 ppma. The range of the center-point value is based on IC process requirements. The ± tolerance is the min-max range about the center-point value. The IOC '88 oxygen concentration value is obtained by multiplying the ASTM F121-79 value by 0.652 (See ASTM Test Method F1188-00 for conversion and calibration factors)[1]. METROLOGY*

---

[1] n.a. F1188-00 Standard Test Method for Interstitial Atomic Oxygen Content of Silicon by Infrared Absorption. American Society for Testing and Materials:West Conshohocken, PA, 2001.

NOTE: P/T ratios of current measurement techniques (GFA, SIMS, FTIR) for oxygen in heavily doped silicon are inadequate to assure a tolerance of ± 2 ppma. General agreement on a single calibration factor for use in the determination of oxygen by IR absorption spectroscopy is needed.

With advanced crystal growth technologies, bulk micro defects (BMDs) can be achieved independent of the interstitial oxygen concentration. The importance of BMDs has recently again been emphasized.[2] BMD density is measured using ASTM Test Method F 1239. The BMD density may be especially required in those IC fabrication cases with low thermal budget. Nitrogen doping enhances oxygen precipitation and, thus, improves gettering efficiency in low thermal budget device processes. Not all device processes, however, require the presence of BMDs. BMDs for internally gettered or not internally gettered polished wafer may be taken as > $1 \times 10^8/cm^3$ or < $1 \times 10^7/cm^3$, respectively, after IC processing, although these values are only generic as they depend on the BMD size.

***D.** Front surface particle or crystal originated pit (COP) size = $K_1(CD)$; $K_1 = 0.5$ for design rules smaller than 100 nm. For design rules greater or equal to 100 nm, the particle or COP size = K_{11}, where $K_{11} = 90$ nm. Optical particle counters are believed to report localized scattering event (LSE) signals low by about 10% due to PSL calibration factors which also reduce particle/COP discrimination accuracy. The relationship between actual defect size and associated LSE signal depends on defect type and scanner geometry. COPs are reported larger, metal and semiconductor particles, smaller, and dielectric particles, about the same as LSE signals. One solution is defect sizing based on defect identification. Relating defect size to yield is difficult since defects are not correctly sized. State-of-the-art LSE capability is judged to be 90 nm through 2003. The current particle size capability for SOI wafers, however, is 150-200 nm, due to the altered response in the optical metrology tools, compared to polished or epitaxial wafers.*

***E.** The LLS requirement is specified for particles only; discrimination between particles and COPs is required. Front-surface particles modeled for 99% yield by $Y = \exp[-D_P R_P A_{transistor}]^3$, where $A_{transistor}$, effective chip area, $A_{eff}=2.5*CD^2T+(1-aCD^2T/A_{chip})A_{chip}*0.60$ and "a" is the DRAM cell fill factor (see Table 52a). The kill factor $R_P = 0.2$, although the kill factor may be larger for deep trench devices. T = number of transistors or bits/chip per technology generation. Assuming a pre-gate cleaning efficiency of 50% for particles in Surface Preparation, the particle values are accordingly increased by a factor 2. The analogous calculation for MPUs can also be made using the high-performance MPU A_{eff} as listed in the Table 49 (Short-Term and Long-Term) headers. We have listed the DRAM value as it is smaller than the MPU value. Detailed back-surface particle information is not included in Table 49, since, in practice, lithography concerns are being met by identifying back-surface particles visually, suggesting that only large defects are of impact. If desired, the calculations may be made using the following model for back-surface particle size and density. The front-surface height elevation, H, due to a back-surface particle of size, D, under a back-surface film of thickness, T, and a wafer thickness, W, may be expressed as $[(xD + xT + W) - (T + W)]$, which may be reduced to $[(xD) - (1-x)T]$, where $x = 0.6$ is the compression of the particle and back-surface film due to the pressure of the chuck on the wafer. Assuming a front-surface elevation of 2(CD) results in a 100% lithographic printing failure; the back-surface particle size is expressed as: $D = [(2/0.6)(CD) + (0.4/0.6)(T)]$, where CD and T are expressed in nm. In this model, T may be set equal to 100 nm, for example. Back-surface particles modeled for 99% yield: $Y = \exp[-D_P R_P A_{eff}]$ (1). $R_P = 1.0$, $A_{eff} = A_{CHIP} \times 0.03 \times 0.8$, where 0.03 corresponds to 3% of the chip area touching the chuck and 0.8 corresponds to 80% of the effective chip area that is degraded by effects of the back-surface particle on the front-surface de-focus effect. D_p, then, represents the density of defects allowable in visible inspection for back-side particles. The equation for the "killer" backside particle diameter strongly depends on two assumptions which are process dependent. The first is that a focal plane excursion of 2 CD is required for a 100% assured printing failure. Although a process window this wide may exist in many cases, some tightly specified exposures may be less tolerant to focal plane deviations. This would lead to a smaller particle becoming a backside killer. The second assumption is that the particles and film are both compressed to 60% of their original dimension. This assumption might not be true if the particle were made of a material much harder than the film or the particle was similar in hardness to silicon and there was no backside film (T=0). Either of these circumstances allows a smaller particle to become a possible backside killer. The back-side yield equation assumes that the entire chip is killed by a back-surface particle generating a front-surface focal plane deviation during lithography (the critical particle diameter is that value accordingly used in the equation, or larger). This occurs because a particle with diameter much smaller than the thickness of the wafer may create a bulge on the front surface up to 10 mm in diameter, so a significant portion of the field is out of focus, and the chip does not yield. A mitigating circumstance occurs if the particle is near the die edge, however, since the bulge at the die edge will tend to create only an apparent local tilt in the field which can be accommodated by a scanning stepper leveling system. This gives rise to the 80% effective degraded area.*

***F**. The metals are empirically grouped into three classes[4, 5]: (a) Mobile metals which may be easily cleaned such as Na and K and may be modeled by taking the flat-band shift of a capacitance-voltage (CV) test approximately 0.5 mV for a representative 1 nm EOT; (b) metals which dissolve in silicon or form silicides such as Fe, Ni, Cu, Cr, Co, and Pt; and (c) major gate-oxide-integrity (GOI) killers such as Ca. Each of these metals is taken at a maximum value of $1 \times 10^{10}/cm^2$. Ir in class (b) and Ba and Sr in class (c) are also important but not explicitly included in the $1 \times 10^{10}/cm^2$ number at this time due to the lack of sufficient industry awareness.*

***G.** The metric for site flatness should be matched to the type of exposure equipment used in leading edge applications. For the 130 nm technology node to the end of optical lithography, scanners will be utilized with rectangular fields (nominally 25 mm x 32 mm for 4X scanners) with slit width of 5 to 8 mm (varies tool to tool, still not standardized) and SFSR is the appropriate metric. Full-field steppers with square fields (nominally 22 x 22 mm) may still be utilized for non-critical levels although they are being phased out; in this case, SFQR is the appropriate metric. In either case, the metric value is approximately equal to the CD for dense lines (DRAM half pitch), although some IC houses deviate upwards by as much as 40% from this algorithm. Partial sites should be included. Note that flatness metrology requires sufficient spatial resolution to capture topographical features relevant for each technology node. This can be expressed as a bandwidth, defined as the upper spatial frequency corresponding to a specified attenuation level. Work is being done in a task force of the SEMI Global Silicon Wafer Committee (PRECOM-1) to validate the methodology and to determine appropriate values. The anticipated shift from capacitive to optical measurement of site flatness beyond the 100 nm node may cause a degradation in SOI flatness measurement capability, compared to polished or epitaxial wafers, analogous to the LLS case.*

---

[2] *K. Sueoka et al., Oxygen Precipitation Behavior and Its Optimum Condition For Internal Gettering and Mechanical Strength in Epitaxial And Polished Silicon Wafers, ECS PV **2000-17**, 164-179 (2000)*

[3] *W. Maly, H.T. Heineken, and F. Agricola. "A Simple New Yield Model." Semiconductor International, number 7, 1994, pages 148–154.*

[4] *P.W. Mertens, T. Bearda, M. Houssa, L.M. Loewenstein, I. Cornelissen, S. de Gendt, K. Kenis, I. Teerlinck, R. Vos, M. Meuris and M.M. Heynes, "Advanced Cleaning for the Growth of Ultrathin Gate Oxide," Microelectronic Engineering **48**, 199-206 (1999)*

[5] *T. Bearda, S. de Gendt, L. Loewenstein, M. Knotter, P. Mertens and M. Heyns, "Behaviour of Metallic Contaminants During MOS Processing," Solid State Phenomena**, 65-66**, 11-14 (1999)*

H. *Fe consistent with recombination lifetime, τ_r, as measured by the SPV technique (for lightly doped p-type material) at low injection level*[6]. *Note that the bulk Fe concentration (at/cm^3) cannot be converted to surface concentration (at/cm^2) via wafer thickness. Recombination lifetime $\tau_r \geq (L^2)/D_n$, where L = minority-carrier diffusion length and D_n = minority-carrier diffusion coefficient at 27°C*[7]. *The diffusion length is taken equal to the wafer thickness and the allowable lifetime is doubled to ensure a sufficient safety factor. Appropriate technique(s) to control, stabilize and passivate surface effects is required, depending on the technique (SPV, PCD, etc.), especially for a bulk lifetime greater than 20 µsec. For any technique other than SPV, the injection level must be noted. No oxygen precipitation in sample, no back-side mechanical damage, and resistivity of 5–20 ohm-cm recommended.*

I. *OSF density empirically modeled by K_3 (CD)$^{1.42}$; CD in nm; $K_3 = 2.75 \times 10^{-3}$*.[8] *The utilization of the OSF density relation by extension into CD regimes not envisioned in the original experimental analysis will require re-assessment. Test at 1100°C, 1 hour steam, strip oxide/etch; n-type material more difficult to control OSF.*

J. *Other epitaxial defects such as hillocks and mounds should also be accounted for, but an appropriate yield model is not available.*

K. *Desired epitaxial layer thickness tolerance is ± 4% for a 2 to 10 µm center-point epitaxial layer thickness target value. In the case of p^-/p^- epi, the minimum epi layer thickness is designed to avoid the possible influence of bulk grown-in defects such as COPs; this consideration is less critical for p/p^+ where the COPs are significantly reduced in the p^+ substrate compared to p^-.*

L. *Large structural epi defects (large area defects >1 µm LSE signal) modeled at 99% yield where $Y = \exp[-D_{LAD} R_{LAD} A_{CHIP}]$ (2), where $R_{LAD} = 1$ and A_{CHIP} applies to DRAM and high-performance MPU as appropriate. METROLOGY NOTE: Many current generation scanning surface inspection systems (SSIS) cannot reliably size surface features with LSE signals greater than about 0.5 µm due to the light scattering characteristics of these large structural epi defects and the optical design of the tool. Further, a metrology gap clearly exists since production worthy tools are not available that can separate large structural epi defects from other features like large particles as well as identify and count epitaxial stacking faults.*

M. *Small structural epi defects (≤ 1 µm LSE signal) modeled at 99% yield where $Y = \exp[-D_{SF} R_{SF} A_{CHIP}]$ (1), where $R_{SF} = 0.5$ and A_{CHIP} applies to DRAM and high-performance MPU as appropriate. Starting Materials uses the DRAM at production and the MPU high-performance MPU areas. METROLOGY NOTE: A metrology gap clearly exists since production worthy tools are not available that can identify and count small structural epi defects.*

N. *The silicon final device layer thickness (partially depleted) is obtained by 2 x MPU physical gate length (with a range in nominal values of ± 25%). Range of target value refers to the center point measurement with uniformity to indicate within-wafer maximum positive or negative % deviation from the center value. The final device silicon is less than incoming material due to consumption during device fabrication. It is expected that it will be difficult to maintain the partially depleted operating mode for a planar SOI-CMOS device once the silicon approaches ~ 50 nm thickness.*

O. *The silicon final device thickness (fully depleted) is obtained by 0.4 x MPU physical gate length (with a range in nominal values of ± 25%). Range of target value refers to the center point measurement with uniformity to indicate within-wafer maximum positive or negative % deviation from the center value. The final device silicon is less than incoming material due to consumption during device fabrication.*

P. *The BOX thickness for fully depleted devices is taken as the MPU physical gate length. BOX scales with gate length to help to control short channel leakage.*[9] *Range in nominal target value of ± 25% allows for tradeoff between the BOX and silicon thickness to control short channel effects in the fully depleted SOI devices. NOTE: For partially depleted SOI devices, the BOX thickness has less of a direct impact on device parameters. Considerations of BOX capacitance, circuit heat dissipation, gettering, BOX electrical integrity, SOI wafer manufacturing capabilities, wafer quality and wafer cost have driven the choice of the BOX thickness values. The BOX thickness is expected to remain between 100-200 nm for the time-frame of partially depleted SOI devices. It is expected that it will be difficult to maintain the partially depleted operating mode for a planar SOI-CMOS device once the silicon final device layer thickness approaches ~ 50 nm.*

Q. *Large area SOI (LASOI) wafer defects with yield of 95%; $Y = \exp[-D_{LASOI} R_{LASOI} A_{chip}]$ (2), D_{LASOI} = LASOI defect density, R_{LASOI} =1.0 (best present estimate).*

R. *Small area SOI (SASOI) wafer defects with yield of 95%; $Y = \exp[-D_{SASOI} R_{SASOI} A_{eff}]$ (2), D_{SASOI} =SASOI defect density, R_{SASOI} = 0.2 (best present estimate). Sources of SASOI can include COPs, metal silicides, or local SiO_2 islands in the top silicon layer. These SASOI defects may also be detected by localized light scattering (LLS) measurements.*[10, 11, 12]

S. *Yields comparable to bulk devices have been achieved with extended crystal defects, D_{EC}. $D_{EC} = -\ln(0.95) / R_{EC}A_{eff}$, R_{EC}=0.000001.*

---

[6] *G. Zoth and W. Bergholtz, "A Fast, Preparation-Free Method to Detect Iron in Silicon," J. Appl. Phys., **67**, 6764-6771, (1990).*

[7] *W. Shockley, Electrons and Holes in Semiconductors, p.69, Princeton: D. Van Nostrand Co., Inc. 1950.*

[8] *M. Kamoshida. "Trends of Silicon Wafer Specifications vs. Design Rules in ULSI Device Fabrication. Particles, Flatness and Impurity Distribution Deviations." DENKA KAGAKU, number 3, pages 194–204, 1995.*

[9] *D.K. Sadana, J. Lasky, H.J. Hovel, K. Petrillo and P. Roitman. "Nano-Defects in Commercial Bonded SOI and SIMOX." 1994 IEEE International SOI Conference Proceedings, pp. 1111–1112, Nantucket Island, MA (1994).*

[10] *Y. Omura, S. Nakashima, K. Izumi, and T. Ishii, IEDM Tech. Digest, 0.1-µm-Gate, Ultrathin-Film CMOS Devices Using SIMOX Substrate with 80-nm-Thick Buried Oxide Layer, p. 675-678 (1991)*

[11] *W. P. Maszara, R. Dockerty, C.F.H. Gondran and P.K. Vasudev. "SOI Materials for Mainstream CMOS Technology." in Silicon-On-Insulator Technology and Devices VIII, S. Cristoloveanu, P.L.F. Hemment, K. Izumi and S. Wilson editors, PV **97-23**, pp. 15–26, The Electrochemical Society Proceeding Series, Pennington, NJ (1997).*

[12] *H. Aga, M. Nakano and K. Mitani. "Study of HF Defects in Thin Bonded SOI Dependent on Original Wafers." Extended Abstracts of the 1998 International Conference on Solid State Devices and Materials, pp. 304–305, Hiroshima (1998).*

Table 49b Starting Materials Technology Requirements—Long-term

YEAR OF PRODUCTION	2010	2013	2016	DRIVER
DRAM ½ PITCH (nm)	45	32	22	D ½
MPU / ASIC ½ PITCH (nm)	50	35	25	M
MPU PRINTED GATE LENGTH (nm)	25	18	13	M
MPU PHYSICAL GATE LENGTH (nm)	18	13	9	M
DRAM Total Chip Area (mm^2)	181	239	238	D ½
DRAM Active Transistor Area (mm^2)	89.5	146.9	166.1	D ½
MPU High-Performance Total Chip Area (mm^2)	310	310	310	M
MPU High-Performance Active Transistor Area (mm^2)	26.8	26.8	26.8	M
*General Characteristics * (99% Chip Yield) [A,B,C]*				
Wafer diameter (mm) **	300	300*****	450	D ½,M
Edge exclusion (mm) ****	2	2	2	D ½,M
Front surface particle size (nm), latex sphere equivalent [D]	≥ 23	≥ 16	≥ 11	D ½
Particles (cm^{-2})[E]	≤ 0.11	≤ 0.07	≤ 0.06	D ½
Particles (#/wf)	≤ 77	≤ 47	≤ 95	D ½
Critical surface metals (at/cm^2) [F]	≤ 1 E10	≤ 1 E10	≤ 1 E10	D ½,M
Site flatness (nm) [G]	≤ 45	≤ 32	≤ 22	D ½,M
*Polished Wafer * (99% Chip Yield)*				
The LLS requirement is specified for particles only; discrimination between particles and COPs is required (see General Characteristics) [D,E]				
Total bulk Fe (at/cm^3) [H]	≤ 1 E10	≤ 1 E10	≤ 1 E10	D ½,M
Oxidation stacking faults (OSF) (DRAM) (cm^{-2}) [I]	≤0.6	≤ 0.4	≤ 0.2	D ½
Oxidation stacking faults (OSF) (MPU) (cm^{-2}) [I]	≤ 0.2	≤ 0.1	≤ 0.1	M
*Epitaxial Wafer * (99% Chip Yield)*				
Total Allowable Front Surface Defect Density is The Sum of Epitaxial Large Structural Defects, Small Structural Defects and Particles (see General Characteristics) [J,K]				
Large structural epi defects (DRAM) (cm^{-2}) [L]	≤ 0.006	≤ 0.004	≤ 0.004	D ½
Large structural epi defects (MPU) (cm^{-2}) [L]	≤ 0.003	≤ 0.003	≤ 0.003	M
Small structural epi defects (DRAM) (cm^{-2}) [M]	≤ 0.011	≤ 0.008	≤ 0.008	D ½
Small structural epi defects (MPU) (cm^{-2}) [M]	≤ 0.006	≤ 0.006	≤ 0.006	M

White–Manufacturable Solutions Exist, and Are Being Optimized
Yellow–Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known

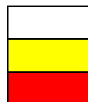


Table 49b Starting Materials Technology Requirements—Long-term (continued)

YEAR OF PRODUCTION	2010	2013	2016	DRIVER
DRAM ½ PITCH (nm)	45	32	22	D ½
MPU / ASIC ½ PITCH (nm)	50	35	25	M
MPU PRINTED GATE LENGTH (nm)	25	18	13	M
MPU PHYSICAL GATE LENGTH (nm)	18	13	9	M
DRAM Total Chip Area (mm^2)	**181**	**239**	**238**	D ½
DRAM Active Transistor Area (mm^2)	**89.5**	**146.9**	**166.1**	D ½
MPU High-Performance Total Chip Area (mm^2)	**310**	**310**	**310**	M
MPU High-Performance Active Transistor Area (mm^2)	**26.8**	**26.8**	**26.8**	M
Silicon-On-Insulator Wafer (95% Chip Yield)*				
Wafer diameter (mm)	**300**	**300*******	**450**	D ½, M
Silicon final device layer thickness (Partially Depleted) (tolerance ± 5%) (nm) [N]	**27–45**	**20–33**	**14–23**	M
Silicon final device layer thickness (Fully Depleted) (tolerance ± 5%) (nm) [O]	**5–9**	**4–7**	**3–5**	M
Buried oxide (BOX) thickness (Fully Depleted) (tolerance ± 5%) (nm) [P]	**14–23**	**10–16**	**7–11**	M
D_{LASOI}, Large area SOI wafer defects (DRAM) (cm^{-2}) [Q]	**≤ 0.028**	**≤ 0.021**	**≤ 0.022**	D ½
D_{LASOI}, Large area SOI wafer defects (MPU) (cm^{-2}) [Q]	**≤ 0.017**	**≤ 0.017**	**≤ 0.017**	M
D_{SASOI}, Small area SOI wafer defects (DRAM) (cm^{-2}) [R]	**≤ 0.287**	**≤ 0.175**	**≤ 0.154**	D ½
D_{SASOI}, Small area SOI wafer defects (MPU) (cm^{-2}) [R]	**≤ 0.956**	**≤ 0.956**	**≤ 0.956**	M
D_{EC}, Extended Crystal Defects (MPU) (cm^{-2}) [S]	**1.9 E5**	**1.9 E5**	**1.9 E5**	M

White–Manufacturable Solutions Exist, and Are Being Optimized
Yellow--Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known

* *Parameters define limit values; independent predictors of yield, mathematically or empirically modeled at 99% (or 95% for SOI). Limit values rarely coincide for more than one parameter. A given wafer will generally not exhibit more than one limit value "at a time;" other parameter values most likely near median value, thereby ensuring total yield for all parameters is at least 99% (or 95% for SOI).*

** *Significant gaps in metrology and wafer manufacturing equipment need to be closed on 200 mm, especially for the 130 nm node, inasmuch as 300 mm is only now being phased in and 200 mm will still be prevalent through the 130 nm node.*

*** *Numerical values are for 300 mm, although 200 mm will be the dominant polished, epitaxial and SOI wafer diameter.*

**** *Edge exclusion = 3 mm is consistent with equipment and wafer fab capabilities through 2002. Reduction in the edge exclusion beyond 2002 is desirable to increase chips per wafer.*

***** *International Roadmap Committee has suggested that 450 mm may be appropriate for 2013.*

The wafer requirements have been selected to ensure that in any given year each parameter value contribute no more than 1% to leading-edge chip yield loss, except for SOI—a rather new mainstream material system—for which we allow up to 5% and accordingly, results in several yellow parameter values for 2001/2002 and red in 2004 for SOI. The values in the tables are generally, but not exclusively, derived from statistical yield-defect models, taking into account leading edge technology parameters such as critical dimension (CD), bit density, transistor density, and chip size. The validity of these derived values is limited by the sometimes questionable accuracy and predictability of the underlying models based on the assumptions utilized. With the onset of the mesoscopic era characterized by nanometer device dimensions, compliance with these model-based values can be very costly and, in some cases, unjustified. For this reason, detailed re-examination of the costs incurred versus the value derived from achieving compliance often suggests limiting the scope of these models via appropriate truncation.

Model-based parameter requirements do not comprehend the distribution of parameter values intrinsic to the wafer manufacturing process where either of two statistical distributions commonly apply. Parameter values distributed symmetrically around a central or mean value, such as thickness, can often be described by the familiar normal distribution. The values of zero-bounded parameters (such as site flatness, particle density, and surface metal concentration) can usually be approximated by a log-normal distribution, in which the logarithms of the parameter values are normally distributed. The latter distribution is highly asymmetric with long tails at the upper end of the distribution.

The ideal methodology for management of material-contributed yield loss would be to allocate loss by defect type such that these defects do not contribute more than 1% to the overall IC fabrication yield loss. Yield loss for a particular defect is equal to the integral of the product of (1) the probability of failure due to a given value of the parameter (as established by the appropriate yield model) and (2) the fraction of wafers having that value (as established by the normal or log-normal distribution function). By applying this methodology, one could determine acceptable product distributions, which could then be utilized as materials acceptance criteria. Rather than a table of allowable parameter limits, this methodology would set specifications around the parameter distributions demonstrated by individual suppliers. By far, the greatest barrier to the acceptance of this approach is the validation of the yield models, which remains elusive, despite the experience of forty years of IC manufacture. As a result, wafer standards are usually expressed as "goal-post" specifications (maximum/minimum values allowed) which invariably increases the silicon supplier manufacturing and inspection costs and ultimately, the IC wafer user CoO.

Defining wafer specifications using the methodology of statistical distributions is being actively investigated as an alternative to goal-post specs. Successful implementation of a distributional specification requires that the silicon supplier's process is sufficiently well understood, under control and capable of meeting the IC user requirements.

These issues are addressed as potential solutions in Figure 27. Until these ideals can be achieved, yield models based on the best available information are used and parameter limits assigned based on a 99% yield requirement for that parameter (except for SOI which is based on a 95% yield requirement). It is further assumed that the yield loss from any individual wafer parameter does not significantly contribute to the yield loss from any other parameters, i.e., that the defect yield impacts are statistically independent. Where validation data is available, this empirical approximation has been shown to result in requirement values nearly equal to the limit values obtained from the aforementioned methodology using parameter distributions.

As the acceptance values for many parameters approach metrology limits, enhanced cooperation between wafer suppliers and IC manufacturers is essential for establishing and maintaining acceptable product distributions and costs. Further development and validation of IC yield/defect models is required. However, it is essential to balance the "best wafer possible" against the CoO opportunity of not driving wafer requirements to the detection limit, but instead to some less stringent value consistent with achieving high IC yield. For example, the surface metal and particle contamination requirements for starting wafers are less stringent than the pre-gate values given in Surface Preparation (see *Tables 50a and b*) because it is assumed that a minimum cleaning efficiency of 50% (actually 95% has been reported for surface iron removal) results during IC processing steps such as the pre-gate clean. It is also noted that the chemical nature of the surface produced by the wafer supplier (hydrophobic versus hydrophilic) and the wafer-carrier interaction during shipment as well as the humidity in the storage room are important in affecting the subsequent adsorption of impurities and particles on the wafer surface.

In recent years the physical structure as well as the chemical nature of the wafer front surface has emerged as a critical concern. Both polished and epitaxial wafers exhibit specific defects that must be controlled. Polished wafer defects include metal and organic surface chemical residues, particles, and grown-in microdefects, such as "crystal-originated pits" (COPs). Epitaxial wafer defects include large structural defects (> 1 µm) and small structural defects (< 1µm). For this reason, starting material requirements are expressed in terms of specific types of surface defects for different wafer types. The removal and prevention of surface defects is a current state-of-the-art challenge for silicon wafer technology. The development of laser scanning and other instrumentation to count, size, and determine the composition and morphology of these defects is a critical metrology challenge, and is addressed in the *Metrology* chapter. It is important to note that the total defect count is the sum of the various constituents.

Additionally, the back surface of the wafer is increasingly being polished to reduce particle contamination, improve wafer flatness, and increase wafer strength. The polished back surface more readily exhibits microscopic contamination and wafer handling damage. As a result, back-surface cleanliness requirements may emerge and drive the need for more stringent robotic handler standards. However, based on a Starting Materials IC Users Survey, site flatness degradation due

to the presence of back-side particle, does not appear to be of significance and has been dropped from this edition of the ITRS. Many external gettering techniques also degrade the quality of the polished back surface and the front-surface site flatness due to non-uniformities in the deposited back-surface gettering film and consequently may no longer be allowable.

Polished wafer—As device geometries continue to shrink, ultra-low material defect concentrations and surface micro-property variations are becoming more important. With advanced crystal growth techniques, however, bulk microdefects (BMDs) can be controlled independently of the interstitial oxygen concentration. Nevertheless, the formation of uncontrolled SiO_x precipitates is likely to result in excessive device leakage current, necessitating greater attention to internal gettering in polished wafers. The magnitude and uniformity requirements will impose stringent bulk defect control to achieve homogeneous internal gettering capabilities.

The dependence of gate dielectric integrity on the crystal growth parameters and the related role of point defects and agglomerates have been extensively documented. The resulting defect density (D_o) parameter has served effectively as a measure of material quality for several device generations. However, for devices with < 2 nm equivalent oxide thicknesses this parameter is no longer an indicator of device yield and performance and has accordingly been deleted from Tables 49a and b as a requirement. It should be noted, however, that with the potential introduction of high κ gate dielectric materials, different pre- and post-gate surface preparation methods and cleanliness requirements might emerge.

Epitaxial wafer—The improved gate dielectric integrity in epitaxial material appears due to the absence of residual polishing micro-damage and grown-in micro-defects, such as COPs and near-surface oxygen precipitates, that are present in polished wafers. However, the presence of extended epitaxial defects and other large area defects as well as small structural epitaxial defects must be controlled in order to realize these benefits in gate dielectric integrity. For present applications, lightly doped p-type epitaxial layers grown on a heavily p-doped Cz substrate (p/p^+ and p/p^{++}) continue to be the mainstream epitaxial wafer type. Here, the heavily doped substrate serves as both an impurity getter (solubility-enhanced gettering) as well as a ground plane for latch-up suppression. For epitaxial layer resistivities greater than a few Ω-cm, a back-surface seal may be required for the p/p^{++} structures. The potential replacement of the epitaxial p/p^+ wafer with a polished wafer having an implanted high dose ground plane (high-dose buried layer) to prevent latch-up and getter impurities is also receiving attention although here the presence and control of COPs remains an issue. The utilization of p/p^- is also receiving attention for advanced IC applications, because of the reduced system capacitance that may be achieved as compared to heavily doped substrates. However, in this case, with the absence of solubility-enhanced gettering, the role of oxygen and BMDs may have to be reassessed. It is well known that oxygen precipitates more slowly in p^- as compared with p^+ and p^{++} material. Because the optimum oxygen concentration significantly depends on the employed IC thermal process sequence, many factors must be considered when selecting the optimal oxygen concentration, especially in the case where the shallow trench isolation depth (STI) is greater than the epitaxial layer thickness. On the other hand, BMD control independent of the oxygen concentration may be expected to become even more significant. These various types of epitaxial structural configurations may be expected to lead to some product consolidation as illustrated in the Potential Solutions in Figure 27.

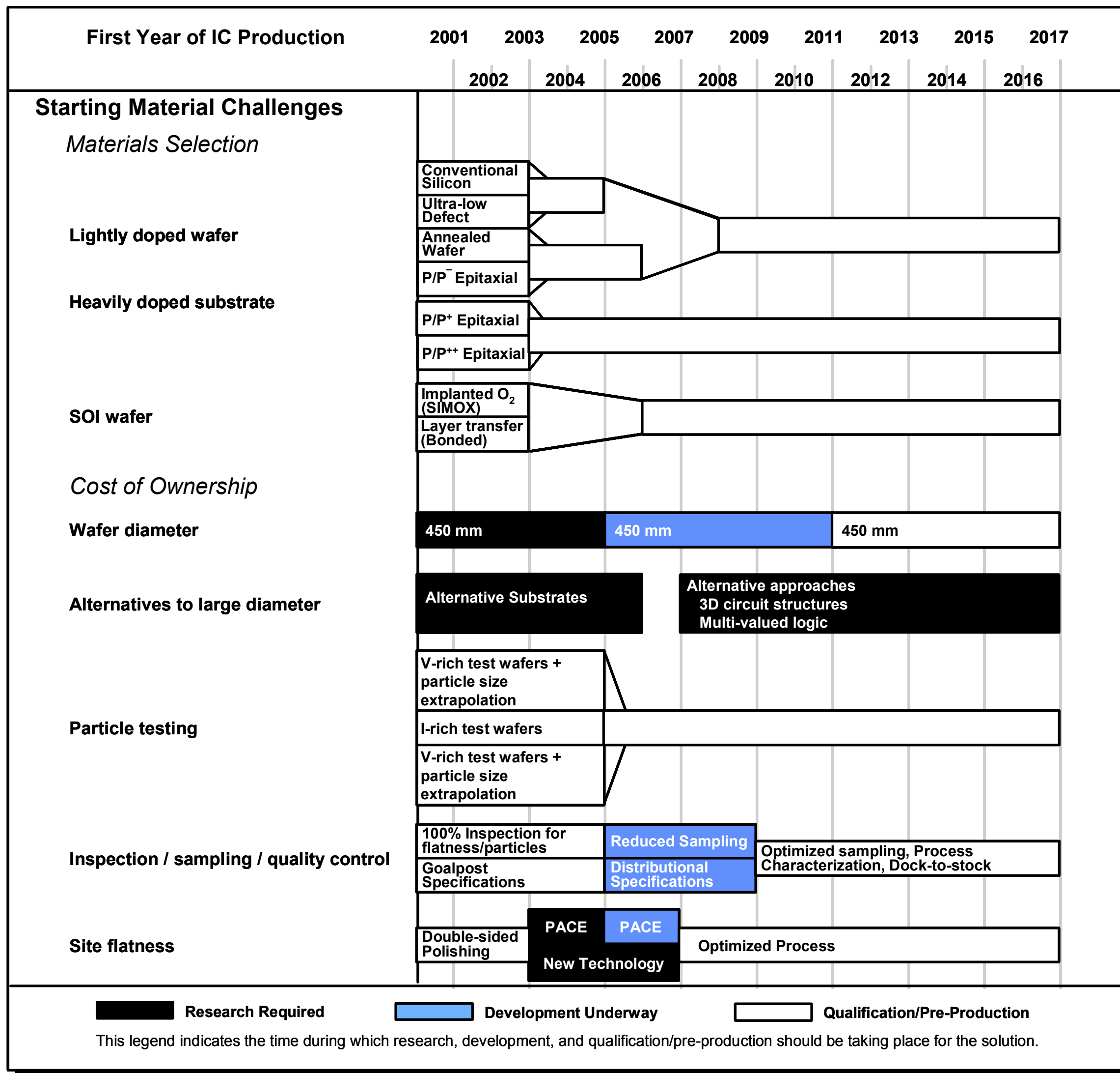


Figure 27 Starting Materials Potential Solutions

Silicon-on-insulator wafer—SOI offers the potential for high speed, low-power consumption, soft-error and latch-up immunity, manufacturing process simplification, and smaller chip size. Current SOI manufacturing techniques result in wafers that are more costly than polished or epitaxial wafers. It has been estimated that SOI permits a one node extension of a wafer factory's process capability and the device performance benefits will motivate the use of SOI for those applications where the benefits outweigh the higher wafer cost. Some bulk IC designs may be transferred directly to SOI substrates. However, optimization frequently requires mask set and process sequence modifications. The broad variety of today's IC applications such as micro-processors, servers, smart power, and RF signal processors require a considerable range of Si-device layer and buried oxide (BOX) layer thicknesses. A number of SOI wafer fabrication approaches are more fully entering into production to serve this range of SOI applications.

Structural and defect characterization of these various SOI wafer fabrication techniques and correlation among the SOI properties, device characteristics, and yield continues to be essential. The relative immaturity of SOI materials compared to bulk and epitaxial leads to an additional challenge for understanding of SOI-specific defects and their impact on device performance and yield in a production environment. The learning achieved since the 1999 ITRS update has resulted in an improved categorization of SOI defect types. Instead of categorizing all point defects in the silicon layer together, and distinguishing micro-voids and pinholes as BOX defects dependent on the SOI wafer manufacturing method, an approach that categorizes large-area defects and small area defect in SOI wafers is now being pursued. Large-area defects,

extending over the scale of several to many microns, include Si-layer voids and large bonding voids. These large defects are judged to have a serious effect on chip yield and are assigned a kill rate of 100%. Smaller defects, such as COPs, metal silicides or local SiO_2 islands in the top silicon layer (measured in tens of nanometers to tenths of microns) are believed to have a less severe impact on device performance and thus the allowable density is calculated based on a lower kill rate. The treatment of extended crystal defects (i.e., threading dislocations) has been retained as a separate parameter due to their substantially lower impact and kill rate.

Metrology for SOI wafers is a significant challenge. The particle and site flatness metrology readiness grades listed for general characteristics are not applicable for SOI wafers. Interference effects arising from multiple reflections from the Si and BOX layers fundamentally alter the response of optical metrology tools compared to polished and epitaxial wafers, generally degrading the measurement capability. For instance, the current particle size capability for SOI wafers is 150-200 nm as compared to 90 nm for polished or epitaxial wafers. Accordingly, the metrology readiness grade for SOI wafer particles is yellow through the 100 nm node turning to red beyond this node. The anticipated shift from capacitive to optical measurement of wafer site flatness beyond the 100 nm node may cause a similar degradation of site flatness measurement capability, requiring a change in the metrology readiness grade from white to red for SOI applications after the 100 nm node. Metrology methods for many of the SOI defect categories call for destructive chemical etching that decorate but do not uniquely distinguish various types of crystal defects. These various defects may not all have the same origin, size, or impact on the device yield and, therefore, may exhibit different kill rates. Non-destructive and fast-turn around methods are also needed for measurement of electrical properties and structural defects in SOI materials.

Large diameter wafer—The conversion to 300 mm diameter wafers, that began slowly in 1999, is expected to peak in 2002–2003, driven by the need to continuously enhance IC manufacturing productivity. International cooperation and standardization of wafers, carriers, and factory protocols continue to be critical for achieving this conversion in a cost-effective and timely manner, as is the availability of cost-effective 300 mm wafers. From the perspective of cost-effective wafer availability, business issues are the primary migration concern as it appears that the engineering issues associated with cost-effective 300 mm crystal growth, finishing and wafer handling have largely been addressed. The relative contribution of the 300 mm wafer to the overall IC manufacturing cost structure, however, has been noted to be at least three times as high as the corresponding 200 mm wafers, suggesting the need for significant CoO improvements.

Projections of wafer diameter beyond 300 mm suggest that 450 mm may be the next appropriate size, perhaps as early as 2013. In reality, however, the term “450 mm wafer” is a metaphor for a major productivity enhancement required by the industry to stay on our present business and economic growth trend. An actual wafer diameter change to 450 mm, however, will be needed if no other new productivity improvements are available, although these wafer size changes are always fraught with both economic and technical issues. Continuing to maintain a two-year technology cycle, as tactically envisioned in the ITRS, provides partial mitigation of the need for the next wafer size change, potentially delaying it on nearly a year-by-year basis. Assessment of these high-level ITRS strategies should be possible using the industry economic and productivity models presently under development at International SEMATECH.

Nevertheless, extension of historical wafer-diameter trends suggests the implementation of research in 2011 to facilitate the introduction of 675 mm wafers in ~ 2020. However, it is far from clear that 450 mm or 675 mm diameter wafers will be economically viable. The engineering issues associated with crystal growth, finishing and handling of such large wafers appear to be very substantial. A paradigm shift in the preparation of cost-effective silicon substrates may be required to mitigate the escalating costs associated with conventional silicon substrate materials. One possible approach is the introduction of cost-effective SOI wafers. Another approach is the fabrication of an electrically active silicon material layer on an appropriate supporting substrate that has been optimized for compatibility with factory substrate-handling and chip- packaging requirements.

SURFACE PREPARATION

Wafer cleaning and surface preparation continue to evolve in concert with implementation of new materials and processes, while retaining certain long-held characteristics. In the front-end, research and development have historically focused on maximizing the quality of the gate dielectric, and these efforts continue as the industry moves toward high-k logic gates. New DRAM capacitor materials will also drive a need for compatible surface preparation processes.

Technology requirements for surface preparation are shown in Tables 50a and b; more details for the data are available in the *supplemental material*. Front-end predictions are problematic due to the present lack of definition of future high κ materials and properties. However, it is clear that particulate contamination will continue to be a concern at increasingly

demanding levels. Furthermore, interface control including organic contamination and interfacial oxide is expected to be important both for high κ gate dielectrics and epitaxial Si and SiGe. For the gate, high κ dielectrics may loosen the requirements for metal contaminant control, due to their increased physical thickness versus current thin oxides. After gate stack formation, which is expected in the future to include metal gate electrodes, residues from etching the stack must be removed using chemistries compatible with the gate stack materials. New MPU and DRAM materials coupled with tightening material budgets will increase the need for highly selective etching chemistries and processes, and these must be introduced without deleterious ESH effects. Finally, increased attention is being given to backside particles due to their effect on lithography alignment and cross-contamination to adjacent wafer front sides in cleaning tools.

Table 50a Surface Preparation Technology Requirements—Near-term*

YEAR OF PRODUCTION	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*DRIVER*
DRAM ½ PITCH (nm)	*130*	*115*	*100*	*90*	*80*	*70*	*65*	*D ½*
MPU / ASIC ½ PITCH (nm)	*150*	*130*	*107*	*90*	*80*	*70*	*65*	*M*
MPU PRINTED GATE LENGTH (nm)	*90*	*75*	*65*	*53*	*45*	*40*	*35*	*M*
MPU PHYSICAL GATE LENGTH (nm)	*65*	*53*	*45*	*37*	*32*	*28*	*25*	*M*
Wafer diameter (mm)	**300**	**300**	**300**	**300**	**300**	**300**	**300**	***D ½, M***
Wafer edge exclusion (mm)	**3**	**3**	**1**	**1**	**1**	**1**	**1**	***D ½, M***
Front surface particles								
Critical Particle Size (nm) [A]	**65**	**58**	**50**	**45**	**40**	**35**	**33**	*D ½*
Particles (cm^{-2}) [B]	**0.091**	**0.136**	**0.085**	**0.119**	**0.069**	**0.094**	**0.056**	*D ½*
Particles (#/wafer) [C]	**63**	**94**	**60**	**83**	**48**	**66**	**39**	*D ½*
Back surface particle size (nm), latex sphere equivalent [D]	**250**	**200**	**200**	**200**	**200**	**200**	**200**	*D ½, M*
Particles (cm^{-2}) [E]	**0.68**	**0.68**	**0.68**	**0.68**	**0.68**	**0.68**	**0.68**	*MPU*
Particles (#/wafer) [C]	**468**	**468**	**468**	**468**	**468**	**468**	**468**	*MPU*
Critical GOI surface metals (E+9 ions/cm^2) [F]	**5.0**	**5.0**	**5.0**	**5.0**	**5.0**	**5.0**	**5.0**	*MPU*
Critical Other surface metals (E+10 ions/cm^2) [F]	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	*MPU*
Mobile ions (E+10 ions/cm^2) [G]	**1.5**	**2.1**	**2.7**	**2.5**	**3.1**	**2.6**	**2.8**	*D ½*
Residual interface carbon contamination (E+13 C at/cm^2) [H]	**2.6**	**2.1**	**1.8**	**1.5**	**1.3**	**1.1**	**1.0**	*LOP*
AFM Surface Roughness nm [I]	**0.20**	**0.20**	**0.20**	**0.20**	**0.20**	**0.20**	**0.20**	*D ½, M*
Water Marks (#/wafer) [J]	**<1**	**<1**	**<1**	**<1**	**<1**	**<1**	**<1**	*D ½*
Residual interfacial oxygen (O at/cm^2) [K]	**<1 E14**	**<1 E14**	**<1 E14**	**<1 E14**	**<1 E14**	**<1 E14**	**<1 E14**	*D ½, M*

** Requirements are listed as maximum allowable levels such that they contribute to no more than 1% yield loss.*

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow--Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known

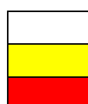


Table 50a and b Notes:

(A) The critical particle size St = 1/2 DRAM 1/2 pitch.

(B) $Y=0.99=\exp[-D_pR_pTA(CD)^2]$. *For DRAM, this version replaces* $TA(CD)^2$ *with the effective chip area,* $A_{eff}=2.5F^2T+(1-F^2T/A_{chip})*0.6A_{chip}$. *The kill factor,* R_p, *is taken as 0.2.* $D_p=-\ln(0.99)/[R_pTA(CD)^2]$ *for DRAMs and* $D_p= -\ln(0.99)/[R_pTA(GL)^2]$ *for MPUs. If a different critical Particle size (Dx) is used for measurement purposes then* D_p *should be adjusted by* $(Dx/.5CD)^2$.

(C) To obtain these values, the allowable particle density is multiplied by the fixed quality area $[DW=DP_p(R_{wafer}-ee)2]$, *and rounded to the nearest whole number.*

(D) The critical particle size is taken as the allowable step height which raises the surface out of the lithographic plane of focus. The critical size is based the budgeted portion of the total allowable DOP variation.

(E) Backside particles are modeled for 99% Yield by $Y=\exp[-DpRpA_{eff}]$ *(1), Rp = 1.0 and* $A_{eff} = A_{chip} \times 0.03 \times 0.8$, *where 0.03 corresponds to only 3% of the back side of the chip area touches the chuck and 0.8 corresponds to only 80% of the effective chip area is degraded by effects of the back-surface particle on the front-surface defocus effect.* $Dp = -\ln(.99)/.03*.8*A_{chip}$.

(F) In past roadmaps, metal contamination targets have been based on an empirically derived model predicting failure due to metal contamination as a function of gate oxide thickness. However, the oxides used in the experiments from which this model was derived were far thicker than gate oxide thicknesses used today. More recent data suggest an updated approach is appropriate. The metals are empirically grouped into three classes [13], [14]: (a) Mobile metals which may be easily cleaned such as Na and K and may be modeled by taking the flat-band shift of a capacitance-voltage (CV) test less than or equal to 50 mV; (b) metals which dissolve in silicon or form silicides such as Fe, Ni, Cu, Cr, Co, and Pt; and (c) major gate-oxide-integrity (GOI) killers such as Ca. Targets for mobile ions are based on allowable threshold voltage shift from a CV test. Current targets for GOI "killers" and other metals are based on empirical data. For extrapolation to future years, there may be reason to predict less stringent targets because effects should scale with respect to physical dielectric thickness (not EOT) which will increase upon introduction of high-k gate dielectrics. However, in the absence of data to corroborate such a prediction, as well as predictions of physical dielectric thickness, the targets are left constant for future years.

(G) Based on $Di=1/q*C_{gate}*DV*.10$, where C_{gate} is computed for electrically equivalent to SiO_2 gate dielectric thickness and DV is the allowable threshold voltage variability for this node. It is assumed that 10% of allowable shift is attributable to Di. $Di = ((3.9*8.85)/1.6)*(DV/EOT)*E+9$

(H) Residual carbon in a gate stack resulting from organic contamination after surface preparation. Dc at the 180nm corresponded to 10% carbon atom coverage of a bare silicon wafer (7.3E+13atoms/cm^2). Dc for subsequent nodes were scaled linearly with the ratio of CD to 180nm. Dc = (CD/180)(7.3E+13)

(I) Device channel mobility cannot be degraded >10% due to surface preparation induced surface roughness. Current technologies are successfully manufactured with AFM based determination of 0.2 nm RMS of surface micro-roughness. This RMS value is an average over the range of spatial frequencies sampled by the AFM. The surface micro-roughness that affects carrier mobility occurs at smaller spatial frequencies than those that are sampled in the typical AFM micro-roughness measurement.

(J) Water marks are generally large enough to touch more than one die and they result in failure of each of the die they touch. Therefore a single water mark will exceed the allowable die loss of 1%. Hence the specification is <1 water mark per wafer.

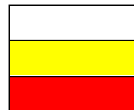
(K) Residual interfacial oxygen resulting from inadequate passivation after surface preparation. Oxygen concentrations up to <1E+14 atoms/cm^2 are acceptable for all processes until selective epitaxial for raised source/drains and high κ gates are implemented.

Table 50b Surface Preparation Technology Requirements—Long-term*

YEAR OF PRODUCTION	2010	2013	2016	DRIVER
DRAM ½ PITCH (nm)	45	32	22	D ½
MPU / ASIC ½ PITCH (nm)	50	35	25	M
MPU PRINTED GATE LENGTH (nm)	25	18	13	M
MPU PHYSICAL GATE LENGTH (nm)	18	13	9	M
Wafer diameter (mm)	300	300	450	D ½, M
Wafer edge exclusion (mm)	1	1	1	D ½, M
Front surface particle				
Critical Particle Size (nm) [A])	23	16	11	D ½
Particles (cm^{-2}) [B]	0.056	0.034	0.030	D ½
Particles (#/wafer) [C]	39	24	48	D ½
Back surface particle size (nm), latex sphere equivalent [D]	200	200	200	D ½, M
Particles (cm^{-2}) [E]	0.68	0.68	0.68	MPU
Particles (#/wafer) [F]	468	468	468	D ½
Critical GOI surface metals (E+9 ions/cm^2) [G]	5.0	5.0	5.0	MPU
Critical Other surface metals (E+10 ions/cm^2) [G])	5.0	5.0	5.0	MPU
Mobile ions (E+10 ions/cm^2) [H]	3.12	3.36	3.48	D ½
Organic contamination (E+13 C at/cm^2)[I]	0.73	0.53	0.37	LOP
Surface Roughness [J]	0.20	0.20	0.20	D ½, M
Water Marks (#/wafer) [K]	<1	<1	<1	D ½
Surface Oxygen (O at/cm^2) [L]	<1 E12	<1 E12	<1 E12	D ½, M

** Requirements are listed as maximum allowable levels such that they contribute to no more than 1% yield loss.*

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow–Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known



[13] Mertens, P. W., "Advanced Cleaning Technology," UCPSS 2000, Ostende, Belgium, invited tutorial, pp. 31-48 (2000).
[14] Mertens, P. W., et. al., "Recent Advances in Wafer Cleaning Technology," Semicon Europa Front End Technology Conference, Munich, 24 April (2001).

Surface preparation challenges can be found in the *supplemental material* and potential solutions are shown in Figure 28. Wet chemical cleaning is still favored because many inherent properties of liquid solutions facilitate the removal of metals (high solubility in liquid chemistries) and particles (zeta-potential control, shear stress and efficient energy transfer by megasonics). Need for other techniques may arise, however, to provide interfacial control for advanced gates. However, broadly effective and non-damaging particle and residue removal using non-liquid techniques has yet to be demonstrated. Single wafer cleaning, either wet or dry, is expected to see increased implementation due to process integration and cycle time concerns, but it remains unclear if or when its use will become widespread. Increased attention will need to be given to mask and reticle cleaning, in which defect density requirements are even more stringent than at the wafer level.

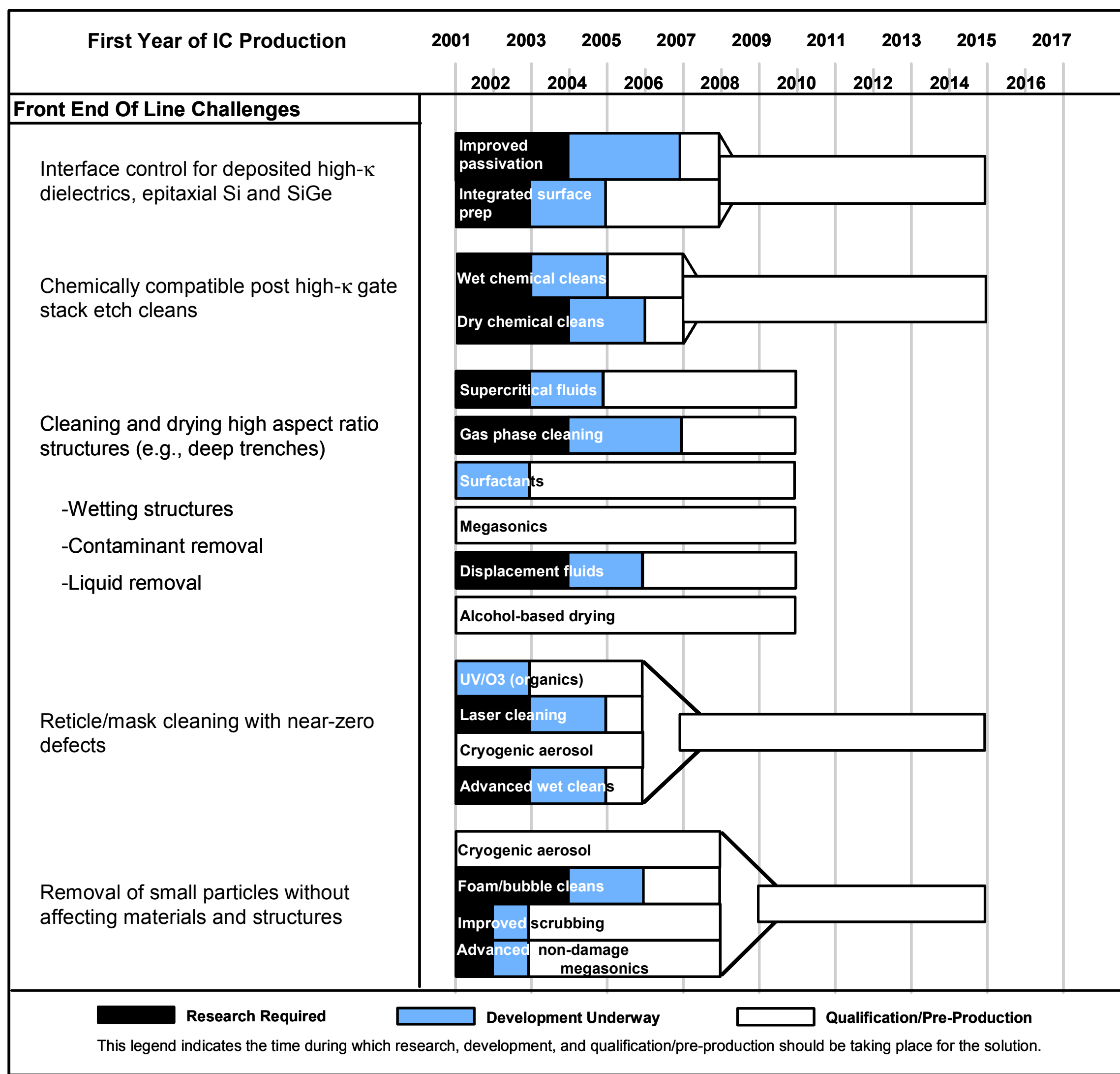


Figure 28 Surface Preparation Potential Solutions

Other technologies, such as ESH and Yield Enhancement, overlap surface preparation. Reduced chemical use, chemical and water recycling, and alternative processes using less harmful chemistries can offer ESH and CoO benefits. Dilute RCA cleans remain common. Ozonated water processes are being implemented as replacements for sulfuric acid based resist strips and post-cleans. Efforts in chemical and water usage reduction should continue. Automated process monitoring and control can also reduce CoO, and their increased use is expected particularly for 300mm and larger wafer sizes where the cost of monitor wafers and process excursions become excessive. Surface preparation overlaps with defect reduction technology in the need for defining appropriate purity levels in chemicals and DI water. To minimize CoO,

aggressive purity targets should be adopted only where a technological justification exists. In all areas of surface preparation, a balance must be achieved between process and defect performance, cost, and environmental, health, and safety issues.

THERMAL/THIN FILMS AND DOPING

Front end processing requires the growth and deposition of high quality, uniform, defect-free films. These films may be insulators, silicon, or conductors. The precision doping of the underlying substrate, consistent with CoO considerations, and deposited layers is normally required. In addition to reduced thermal budgets required by device scaling, several other difficult challenges have been identified for this area including the growth or deposition of reliable very thin (≤ 1.0 nm) gate dielectric layers; the identification and development of alternate high dielectric constant layers including suitable interface layers for both gate and DRAM capacitor applications; the formation of and making contact to ultra-shallow junction devices; and the development of alternate, depletion-free, low-resistivity gate electrode materials. Other important challenges include the achievement of abrupt channel doping profiles, defect management for minimum post-implant leakage under small thermal budget environments, and sidewall spacer formation. Due to high channel doping and the collapse of the overdrive voltage, device drive currents will tend to degrade. To preserve high-drive current, mobility-enhanced channels, such as strained Si (with unstrained Si-Ge layers) as well as strained Si-Ge layers may be initially needed, but ultimately new device structures will need to be developed. CMOS-compatible innovative solutions are needed in all of these areas. Underlying all these concerns is the inevitable increase of leakage currents associated with threshold voltage scaling, gate dielectric tunneling, and junction leakage.

THERMAL/THIN-FILMS

The gate dielectric has emerged as one of the most *difficult challenges* for future device scaling. Requirements summarized in Tables 51a and b indicate an oxide equivalent thickness progressing to substantially less than 1 nm. Direct tunneling currents and boron penetration preclude the use of SiO_2 dielectric layers below ≈ 1.5 nm thickness. In high-performance applications that have high allowable leakage, oxynitrides may be scaled to much thinner films; however their thickness control and reliability limits have not been established. For low-power applications where the allowable gate leakage is very low, higher dielectric constant materials will be needed *as early as the 80 nm node***.** To date, no suitable alternative high κ material and interface layer has been identified with the stability, reliability, and interface characteristics to serve as a gate dielectric for these applications. A significant, global research and development effort has been implemented to identify and qualify a suitable alternative gate dielectric material. The near-term gate dielectric solution requires the fabrication and use of ultra-thin silicon oxide, oxynitride films, or silicon nitride films. The last film shows attractive boron diffusion penetration resistance and a moderately higher dielectric constant value approaching 7. Near-term solutions will impose severe restraints on surface preparation, pre-and post-process ambient control, silicon compatible materials development (e.g., gate electrodes and contacts), and post-processing thermal budgets. Similar problems are anticipated with the DRAM storage capacitor dielectric, anticipated to occur at an earlier technology node.

Table 51a Thermal & Thin Film, Doping and Etching Technology Requirements—Near-term

YEAR OF PRODUCTION	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*DRIVER*
DRAM ½ PITCH (nm)	*130*	*115*	*100*	*90*	*80*	*70*	*65*	*DRAM*
MPU / ASIC ½ PITCH (nm)	*150*	*130*	*107*	*90*	*80*	*70*	*65*	*MPU*
MPU PRINTED GATE LENGTH (nm)	*90*	*75*	*65*	*53*	*45*	*40*	*35*	*MPU*
MPU PHYSICAL GATE LENGTH (nm)	*65*	*53*	*45*	*37*	*32*	*28*	*25*	*MPU*
Equivalent physical oxide thickness for MPU/ASIC T_{ox} (nm) [A]	**1.3–16**	**1.2–1.5**	**1.1–1.6**	**0.9–1.4**	**0.8–1.3**	**0.7–1.2**	**0.6–1.1**	*MPU*
Gate dielectric leakage at 100°C (nA/µm) High-performance [B]	**10**	**30**	**70**	**100**	**300**	**700**	**1000**	*MPU*
Physical gate length low power (nm)	**90**	**80**	**65**	**53**	**45**	**37**	**32**	*LOW POWER*
Equivalent physical oxide thickness for low operating power T_{ox} (nm) [A]	**2.0-24**	**1.8-2.2**	**1.6-2.0**	**1.4-1.8**	**1.2-1.6**	**1.1-1.5**	**1.0-1.4**	*LOP*
Gate dielectric leakage (pA/µm) LOP [B]	**100**	**100**	**100**	**300**	**300**	**300**	**700**	*LOP*
Equivalent physical oxide thickness for low standby power T_{ox} (nm) [A]	**2.4-2.8**	**2.2-2.6**	**2.0-2.4**	**1.8-2.2**	**1.6-2.0**	**1.4-1.8**	**1.2-1.6**	*LSTP*
Gate dielectric leakage (pA/µm) LSTP [B]	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	*LSTP*
Equivalent physical oxide thickness for DRAM transfer device T_{ox} (nm) [A]	**5**	**4.5**	**4.1**	**3.6**	**3.3**	**3.0**	**2.7**	*DRAM*
Gate dielectric leakage at 100°C (fA/µm) DRAM transfer device [B]	**4.1**	**4.6**	**2.6**	**2.4**	**1.4**	**1.6**	**1.4**	*DRAM*
Thickness control EOT (% 3σ) [C]	**<± 4**	**<± 4**	**<± 4**	**<± 4**	**<± 4**	**<± 4**	**<± 4**	*MPU/ASIC*
Gate Etch Bias (nm) [D-1]	**25.2**	**21.2**	**17.8**	**15.0**	**12.6**	**11.2**	**10.0**	*MPU/ASIC*
L_{gate} 3σ variation (nm) [D]	**6.31**	**5.30**	**4.46**	**3.75**	**3.15**	**2.81**	**2.5**	*MPU/ASIC*
Total maximum allowable lithography 3σ (nm) [D-2]	**5.15**	**4.33**	**3.64**	**3.06**	**2.57**	**2.29**	**2.04**	*MPU/ASIC*
Total maximum allowable etch 3σ (nm), including photoresist trim and gate etch [D-2]	**3.64**	**3.06**	**2.57**	**2.17**	**1.82**	**1.62**	**1.44**	*MPU/ASIC*
Resist trim maximum allowable 3σ (nm) [D-3]	**2.10**	**1.77**	**1.49**	**1.25**	**1.05**	**0.94**	**0.83**	*MPU/ASIC*
Gate etch maximum allowable 3σ (nm) [D-3]	**2.50**	**2.10**	**1.77**	**1.48**	**1.32**	**1.18**	**1.05**	*MPU/ASIC*
CD bias between dense and isolated lines [E]	**≤ 15%**	**≤ 15%**	**≤ 15%**	**≤ 15%**	**≤ 15%**	**≤ 15%**	**≤ 15%**	*MPU/ASIC*
Minimum measurable gate dielectric remaining (post gate etch clean) [F]	**>0**	**>0**	**>0**	**>0**	**>0**	**>0**	**>0**	*MPU/ASIC*
Profile control (side wall angle) [FF]	**>89**	**>89**	**>89**	**90**	**90**	**90**	**90**	*MPU/ASIC*
Drain extension Xj (nm) [G]	**27-45**	**22-36**	**19-31**	**15-25**	**13-22**	**12-19**	**10-17**	*MPU/ASIC*
Maximum drain extension sheet resistance (PMOS) (Ω/sq) [H]	**400.0**	**460.0**	**550.0**	**660.0**	**770.0**	**830.0**	**760.0**	*MPU/ASIC*
Extension lateral abruptness (nm/decade) [I]	**7.2**	**5.8**	**5.0**	**4.1**	**3.5**	**3.1**	**2.8**	*MPU/ASIC*
Contact Xj (nm) [J]	**48–95**	**39–78**	**33–66**	**27–45**	**24–47**	**21–42**	**18–37**	*MPU/ASIC*
Sidewall spacer thickness (nm) extension structure [K]	**48–95**	**39–78**	**33–66**	**27–45**	**24–47**	**21–42**	**18–37**	*MPU/ASIC*
Spacer thickness, elevated contact [L]	**—**	**—**	**—**	**15–25**	**13–22**	**12–19**	**10–17**	*MPU/ASIC*
Spacer thickness, single contact [M]	**—**	**—**	**—**	**—**	**—**	**—**	**5-8.5**	*MPU/ASIC*
Maximum silicon consumption (nm) [N]	**23–46**	**19–38**	**16–32**	**13–26**	**11–23**	**10–20**	**9–18**	*MPU/ASIC*
Silicide thickness (nm) [O]	**35.8**	**29.2**	**24.8**	**20.4**	**17.6**	**15.4**	**13.8**	*MPU/ASIC*
Contact silicide sheet Rs (Ω/sq) [P]	**4.2**	**5.1**	**6.1**	**7.4**	**8.5**	**9.7**	**10.9**	*MPU/ASIC*
Contact maximum resistivity (Ω-cm^2) [Q]	**4.10E-07**	**3.20E-07**	**2.70E-07**	**2.10E-07**	**1.80E-07**	**1.60E-07**	**1.10E-07**	*MPU/ASIC*
Gate electrode thickness [R]	**65–130**	**53–106**	**45–90**	**37–74**	**32–64**	**30–60**	**25–50**	*MPU/ASIC*
Active poly doping for 25% depletion allowance (cm^{-3}) [S]	**9.2 E19**	**9.2 E19**	**1.14 E20**	**1.50 E20**	**1.66 E20**	**1.66 E20**	**1.87 E20**	*MPU/ASIC*
Average gate electrode sheet Rs (Ω/sq) from table [T]	**5**	**5**	**5**	**5**	**5**	**5**	**5**	*MPU/ASIC*
Channel concentration for Wdepletion <1/4L_{eff} (cm^{-3}) [U]	**4.0 E18**	**6.0 E18**	**8.0 E18**	**1.1E 19**	**1.4E 19**	**1.6 E19**	**2.3 E19**	*MPU/ASIC*
Uniform channel concentration (cm^{-3}), for V_t=0.4 [V]	**0.8–1.5 E18**	**0.8–1.5 E18**	**1.5–2.5 E18**	**1.5–2.5 E18**	**1.5–2.5 E18**	**2.0–4.0 E18**	**2.5–5.0 E18**	*MPU/ASIC*
Retrograde channel depth (nm) [W]	**21-30**	**19-27**	**16-23**	**15-21**	**13-19**	**11-16**	**10-15**	*MPU/ASIC*
Undoped dielectric layer thickness (nm) [X]	**500**	**465**	**435**	**400**	**365**	**335**	**300**	*MPU/ASIC*
Alkali diffusion barrier layer thickness (nm)	**50**	**46**	**43**	**40**	**36**	**33**	**30**	*MPU/ASIC*
Undoped bit line dielectric layer thickness (nm) [Y]	**200**	**193**	**187**	**180**	**174**	**168**	**162**	*DRAM*
Cell dielectric layer thickness (nm) [Z]	**200**	**193**	**187**	**180**	**174**	**168**	**162**	*DRAM*
M-1 dielectric layer thickness (nm) [AA]	**250**	**242**	**233**	**225**	**217**	**210**	**203**	*DRAM*

Notes for Table 51a and b Thermal / Thin Films, Gate Etch, and Doping Requirements

[A]. This number represents the effective thickness of the dielectric alone, at the maximum operating frequency of the technology, without substrate or electrode effects, and is determined through an electrical measurement of capacitance corrected for substrate (quantum) and electrode (depletion) effects. A more detailed discussion of the measurement of EOT is on a separate workbook page. The color coding of each node considers the ability of known dielectrics to meet gate leakage, uniformity, and reliability requirements. If there is no known solution to even one of these requirements for the mid-value EOT, the node is coded red. Likewise, the node can be coded yellow if a solution is being pursued for any one of these criteria. For Low Power technologies, it is expected that EOT values used by different companies could range +/- 0.2 nm , i.e., about +/- 10% for short term nodes to +/- 25% for long term nodes.

[A1] Model 1 for EOT is 0.024 times the Physical Gate length

[A2] Model 2 for EOT is the thickness that gives 8 MV/cm field if all of the power supply voltage drops across the dielectric, i.e., Vdd/8MV/cm.

[B] The gate leakage, specified at 100°C, is taken to be the same as the transistor subthreshold leakage at room temperature. This leakage is specified in the PIDS chapter section on Logic—High Performance and Low Power Technology Requirements as the off-state leakage (excluding the junction and the gate leakage components) at room temperature. Since the device subthreshold leakage is expected to increase by a factor of roughly 100x between room temperature and 100°C, the gate leakage is expected to be only about 1% of the total leakage under worst case conditions. Equating the gate leakage to the device sub-theshold leakage was assumed to be satisfactory from a circuits operation standpoint, but it should be noted that not all design approaches (companies) will allow such a high gate leakage. The gate leakage is measured on the minimum nominal device, and the specification is taken to apply to all transistor bias configurations, that is, both when Vg = Vs = 0 and Vd = Vdd as well as when Vs =Vd » 0 and Vg = Vdd. Numbers for low standby power devices come from the Japan PIDS TWG. Numbers for DRAM assume all of the allowed cell leakage (in FEP Table 52) comes from the transfer device.

[B1] The areal gate leakage is modeled as the allowable gate leakage divided by the physical gate length. However, it should be noted that the total gate leakage is the sum of three leakage components: 1) leakage between the source and the gate in the gate-source overlap area, 2) leakage between the channel and the gate over the channel region, and 3) leakage between the gate and the drain in the gate-drain overlap area. The magnitude of each of these three components will depend on the gate, source, and drain biasing conditions. The color coding of leakage nodes is based on UTQUANT simulations of tunneling current from an inversion channel to the gate for the mid-point EOT. These simulation results are given in a separate worksheet. It should be emphasized that the tunneling current density will generally be much higher between the junction and gate than between an inversion channel and gate. Thus these simulations represent a best case (lowest leakage) condition, where the gate-to-junction overlap area is minimal. When oxide will meet the leakage specification, the node is coded white. Based on the literature, optimized oxynitride dielectrics are expected to have a leakage current about 100 times lower than oxide; nodes are coded yellow when optimized oxynitride is needed to meet the leakage specs. Nodes requiring alternate, high k dielectrics are coded red.

[B2] The unmanaged gate leakage power is the total static chip power that would occur if all the devices on a chip had gate leakage equal to the maximum allowable value. Power management will require the extensive use of power reduction techniques, such as power-down or multiple Vt devices to achieve an acceptable static power level.

[C] From Modeling of Manufacturing Sensitivity and of Statistically Based Process Control [15]

Requirements for 0.18 micron NMOS device

[D1] Bias is defined as the difference between the printed gate length and the final post-etch gate length

[D] The total gate length 3σ variation encompasses all random process variation including point to point on a wafer, wafer to wafer, and lot to lot variations. It excludes systematic variations such as lithography proximity effects, and etch variations such as CD bias between densely spaced and isolated lines. This total variability is taken to be less than or equal to 10% of the final feature size. A conventional MOS structure is the basis for these calculations. MOS transistor structures which vary in any way from the conventional structure (e.g. Vertical MOS transistors) will have different technical challenges and will not fall within these calculations. The data is computed taking into account lithographic errors during resist patterning and combined etch errors due to both resist trim and gate etch

[D2] The allowable lithography variance s2L is limited to 2/3 of the total variance, s2T of the combined lithography and etch processes. It is further assumed that the lithographic and etch processes are statistically independent and therefore that the total variance is the sum of the etch and lithography variances. This implies among other things that the printed features in the resist have vertical wall profiles and be sufficiently thick to with-stand the etch process with loss of dimensional fidelity.

[D3] It is assumed that the resist trim and gate etch processes are statistically independent and therefore that the respective variances, s2, of the two processes are additive. 1/3 of the combined trim-etch variance is allocated to the trim process, with the remaining 2/3 allocated to the etch process.

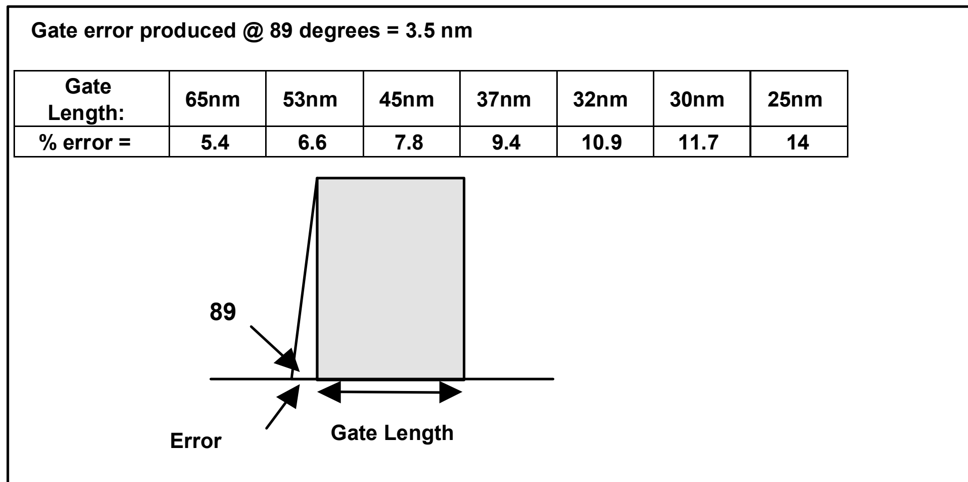
[E] 15% CD budget is a combination of measurements from Etch, Lithography and Metrology. By 2004 the 2nm dense-iso values achieved today will no longer be within specification. Work is currently underway to improve upon this.

[F] It is important that some dielectric remains after the gate etch clean step. Between technology nodes the dielectric thickness decreases and there is an onset of using high k materials (2004) to replace the gate dielectric. Both advances represent challenges to ensure there is an amount of remaining dielectric and the ability to measure the remaining material.

---

[15] *P. Zeitzoff and A. Tasch, "Modeling of Manufacturing Sensitivity and of Statistically Based Process Control Requirements for 0.18 micron NMOS device," Characterization and Metrology for ULSI Technology: 1998 International Conference, D.G. Seiler, et al. eds., page 73*

[FF] Profile can be a major contributor to etch errors (see inset). Accurate measurement of vertical profiles remains difficult. Long term, the effect of edge roughness on device performance needs to be addressed and methodology of the measurement determined.



Gate error produced @ 89 degrees = 3.5 nm

Gate Length:	65nm	53nm	45nm	37nm	32nm	30nm	25nm
% error =	5.4	6.6	7.8	9.4	10.9	11.7	14

*[G] Xj at Channel (Extension Junction) = 0.55*Physical Gate Length (with a range of about ±25%).*

[H] The maximum drain extension sheet resistance is modeled by allocating 7% of the allowable source and drain parasitic resistances to the drain extensions. See worksheet of RsXj. The drain extension sheet resistance value must be optimized together with the contact resistance and junction lateral abruptness (which effects spreading resistance), in order to meet the overall parasitic resistance requirements.

*[I] Channel abruptness is in nm per decade drop-off in doping concentration = 0.11 * Isolated Line (nm) - based on Short Channel effect.[16]*

*[J] Contact Junction Depth = 1.1*Physical Gate Length (with a range of about ±33%).*

[K] Spacer thickness (width) is taken as the same as the Contact Junction Depth. Validity established using response surface methodology in "Response Surface Based Optimization of 0.1 µm PMOSFETs with Ultra-Thin Oxide Dielectrics"[17]

[L] Spacer thickness for an elevated junction (where there is no deeper contacting junction) is taken as the extension junction depth (with a range of about ±25%)

[M] Spacer thickness for an single junction device (where the extension junction is formed after the spacer is in place) is taken to be one half of the extension junction depth (with a range of about ±25%). Provided the lateral junction motion is 0.6 x Xj.

[N] Silicon consumption is based on having the silicide thickness equal to half the contact junction depth

[O] Silicide thickness is taken to be 1/2 of the center Contact Xj to avoid consumption-induced increase in contact resistivity. Less than half of the junction can be consumed.[18]

[P] Contact silicide sheet resistance: assumes 15 µW-cm silicide resistivity , i.e., $TiSi_2$ or $CoSi_2$

[Q] The Si/Silicide maximum interfacial contact resistivity values were calculated assuming that 100% of the PIDS total allowed MOSFET Source/Drain resistance is allocated to the contact resistivity. It further that the transistor contact length is taken to be twice the MPU half pitch, where length is in the direction of current flow. Since the PIDS allocation is in terms of Rs xW, The equation for the contact resistivity rhoc is: rhoc = Rs x W x M. These values should be appropriately modified if different transistor contact lengths are assumed. (See worksheet on Contact Rs). Note that this contact resistivity is the maximum allowable and cannot be used for real devices. The values of contact resistivity, drain extension sheet resistance and drain extension lateral abruptness must be co-optimized in order to meet the overall parasitic resistance requirements.

[R] Gate thickness is takes between one and two times the physical gate length

[S] See Poly Doping worksheet

[T] Gate sheet resistance taken as a constant of 5 ohm/sq for near-term nodes for circuit performance.

---

[16] *Y. Taur, :25 nm CMOS Design Considerations," IEDM 1998, Technical Digest, IEEE, Dec. 1998, pages 789-792*

[17] *A. Srivastava and C.M.Osburn, "Response Surface Based Optimization of 0.1 µm PMOSFETs with Ultra-Thin Oxide Dielectrics," SPIE Proc., vol. 3506, (1998), page 253*

[18] *C.M. Osburn, J.Y. Tsai and J. Sun, "Metal Silicides: Active Elements of ULSI Contacts," J. Electronic Mater., vol. 25(11), (1996), page 1725*

[T1] From the gate sheet resistance and film thickness requirements (average). Assumes metal gate at and beyond 100 nm node

[U] Drain Extension Concentration for W depletion < 1/4 Logic Half Pitch. Values were interpolated from calculations done for the 1999 ITRS.[19]

[V] Uniform channel concentration for Vt = 0.4.[20] *Values were interpolated from calculations done for the 1999 ITRS. Neither quantum mechanical nor potential increase in short channel effects were used in this calculations. These effects do, however, tend to offset each other. NOTE: The assumption of a constant threshold voltage of 0.4 V may not be consistent with the leakage current criteria. To reach the leakage current criteria may result in unacceptably large threshold voltages for the scaled power supplies resulting in severe performance degradation. In addition, high concentration channels could severely impact drain currents due to impurity scattering.*

[W]. The retrograde well profile must be less than 0.5 times the drain extension depth to improve short channel effects.[21, 22] *Values were modeled as 33% to 45% of the gate length.*

[X] Assumes 10% uniformity and stress <2 E9 dyne/cm^2.

[Y] Assumes dielectric constant of 4, 10% uniformity, and stress <2 E9 dyne/cm^2. Post CMP flatness taken as <10% of initial thickness.

[Z] Assumes dielectric constant of 4, 10% uniformity and stress <2 E9 dyne/cm^2. Post CMP flatness taken as <15% of initial thickness.

[AA] Assumes dielectric constant of 4, 10% uniformity and stress <2 E9 dyne/cm^2. Post CMP flatness taken as <15% of initial thickness.

---

[19] *B.G. Streetman, "Solid State Electronic Devices," 4th ed., Englewood Cliffs, NJ: Prentice Hall, page 174*

[20] *R. Muller and T. Kamins. Device Electronics for Integrated Circuits, New York, NY: John Wiley and Sons, Inc., 1977, page 324*

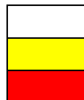
[21] *S. Thompson, P. Packan, and M. Bohr, "Linear versus Saturated Drive Current: Tradeoffs in Super Steep Retrograde Engineering," VLSI Technology Digest, (1996), page 154*

[22] *I. De and C.M. Osburn, "Impact of Super-steep-retrograde Channel Doping Profiles on the Performance of Scaled Devices," IEEE Trans. Elec. Dev., vol. 46, no.8, (1999), page 1711*

Table 51b Thermal & Thin Film, Doping and Etching Technology Requirements—Long-term

Year of Production	2010	2013	2016	Driver
DRAM ½ Pitch (nm)	45	32	22	DRAM
MPU / ASIC ½ Pitch (nm)	50	35	25	MPU
MPU Printed Gate Length (nm)	25	18	13	MPU
MPU Physical Gate Length (nm)	18	13	9	MPU
Equivalent physical oxide thickness for MPU/ASIC Tox (nm) [A]	0.5–0.8	0.4–0.6	0.4–0.5	MPU/ASIC
Gate dielectric leakage at 100°C (µA/µm) High-performance [B]	3.0	7.0	10	MPU/ASIC
Physical gate length low power (nm)	22	16	11	Low Power
Equivalent physical oxide thickness for low operating power T_{ox} (nm) [A]	0.8-1.2	0.7-1.1	0.6-1.0	LOP
Gate dielectric leakage (pA/µm) LOP [B]	1000	3000	10000	LOP
Equivalent physical oxide thickness for low standby power T_{ox} (nm) [A]	0.9-1.3	0.8-1.2	07-1.1	LSP
Gate dielectric leakage (pA/µm) LSTP [B]	3.0	7.0	10.0	LSTP
Equivalent physical oxide thickness for DRAM transfer device T_{ox} (nm) [A]	1.55	1.05	0.55	DRAM
Gate dielectric leakage at 100°C (nA/µm) DRAM transfer device [B]	0.7	0.3	0.2	DRAM
Thickness control EOT (% 3σ) [C]	<± 4	<± 4	<± 4	MPU/ASIC
Leffective control	≤ 20%	≤ 20%	≤ 20%	MPU/ASIC
Gate etch bias (nm) [D-1]	7.1	5.0	3.6	MPU/ASIC
L_{gate} 3σ variation (nm) [D]	1.77	1.25	0.88	MPU/ASIC
Total allowable lithography 3σ (nm) [D-2]	1.45	1.02	0.72	MPU/ASIC
Total allowable etch 3σ (nm), including photoresist trim and gate etch [D-2]	1.02	0.72	0.51	MPU/ASIC
Resist trim allowable 3σ (nm) [D-3]	0.59	0.42	0.29	MPU/ASIC
Gate etch allowable 3σ (nm) [D-3]	0.83	0.59	0.41	MPU/ASIC
CD bias between dense and isolated lines [E]	≤ 15%	≤ 15%	≤ 15%	MPU/ASIC
Minimum measurable gate dielectric remaining (post gate etch clean) [F]	>0	>0	>0	MPU/ASIC
Profile control (side wall angle) [FF]	90	90	90	MPU/ASIC
Drain extension Xj (nm) [G]	7–12	5–9	4–6	MPU/ASIC
Maximum drain extension sheet resistance (PMOS) (Ω/sq) [H]	830.0	940.0	1210.0	MPU/ASIC
Extension lateral abruptness (nm/decade) [I]	2.0	1.4	1.0	MPU/ASIC
Contact Xj (nm) [J]	13–26	10–19	7–13	MPU/ASIC
Sidewall spacer thickness (nm) extension structure [K]	13–26	10–19	7–13	MPU/ASIC
Spacer thickness, elevated contact [L]	7–12	5–9	4–6	MPU/ASIC
Spacer thickness, single contact [M]	3.5–6	2.5–4.5	2–3	MPU/ASIC
Maximum silicon consumption (nm) [N]	6–13	5–9	3–6	MPU/ASIC
Silicide thickness (nm) [O]	9.9	7.2	5.0	MPU/ASIC
Contact silicide sheet Rs (Ω/sq) [P]	15.2	21.0	30.3	MPU/ASIC
Contact maximum resistivity (Ω-cm^2) [Q]	6.40E-08	3.80E-08	2.40E-08	MPU/ASIC
Gate electrode thickness [R]	18–36	13–26	9–18	MPU/ASIC
Active poly doping for 25% depletion allowance (cm^{-3}) [S]	1.8E20	2.5E20	2.99E20	MPU/ASIC
Average gate electrode sheet Rs (Ω/sq) from Table [T]	5	6	7	MPU/ASIC
Channel concentration for Wdepletion <1/4L_{eff} (cm^{-3}) [U]	5.0E19	1.3E20	5.0E20	MPU/ASIC
Uniform channel concentration (cm^{-3}), for V_t=0.4 [V]	5.0–9.0 E18	0.9–1.8 E19	1.5–3.0 E19	MPU/ASIC
Retrograde channel depth (nm) [W]	7–10	5–7	3–5	MPU/ASIC
Undoped dielectric layer thickness (nm) [X]	300	250	200	MPU/ASIC
Alkali diffusion barrier layer thickness (nm)	30	25	20	MPU/ASIC
Undoped bit line dielectric layer thickness (nm) [Y]	146	131	118	DRAM
Cell dielectric layer thickness (nm) [Z]	146	131	118	DRAM
M-1 dielectric layer thickness (nm) [AA]	182	164	148	DRAM

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow--Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known



Intermediate and long-term solutions require the identification of materials with a higher dielectric constant (>10 suggested for intermediate term and >20 for long term) with other electrical characteristics (such as stability and interface-state densities) and reliability approaching that of high quality gate SiO_2. A major problem with a material other than SiO_2 is the anticipation that a very thin SiO_2 layer may still be required at the channel interface to preserve interface-state characteristics and channel mobility. This would severely degrade any benefits that accrue from the use of the high κ dielectric, thereby increasing the equivalent oxide thickness. The presence of a single molecular layer of Si-O bonding to

bridge between the silicon substrate and a high κ material is expected to present a physical limit to scaling of equivalent oxide thickness to no less than 0.3 nm. It is also anticipated that an appropriate material may be required between the high κ material and the gate electrode. Improved thickness control and uniformity will also be essential to achieve V_t control for 300 mm wafers. Sensitivity to post-gate, process-induced damage associated with ion implant and plasma etching is expected to increase, especially as it relates to leakage current associated with the gate dielectric perimeter.

Another challenge is the achievement of acceptably low electrical leakage in order to meet both the gate leakage specification and the reliability requirements. To achieve these needs, the high κ dielectric must have a band gap of 4–5 eV with a barrier height of >1 eV to limit thermionic emission and direct tunneling. In addition, the candidate dielectric material must have negligible trap densities to be stable and to suppress Frenkle-Poole tunneling. Finally, the material must have excellent diffusion barrier properties to prevent gate electrode material or gate electrode dopant contamination of the transistor channel

The gate electrode also represents a major challenge for future scaling. Channel autodoping associated with boron out-diffusion and polysilicon depletion will eventually require the phase-out of dual-doped polysilicon gate material. A long-term solution such as low resistivity gate material (s) has not been sufficiently addressed and years of research will be required to identify and qualify an alternative gate electrode material.

Work function, resistivity, and compatibility with CMOS technology are key parameters for the new candidate gate electrode materials. The last issue requires that different materials be used for the PMOS and NMOS transistor gate electrodes to achieve acceptable threshold voltages; the former having a Fermi level near the silicon valence band, and the latter having a Fermi level near the conduction band. Sheet resistance considerations may ultimately require the use of cladded gate electrodes, where an interface layer is used to achieve the desired gate work function and the second layer is used to lower the overall gate sheet resistance. Alternatives are the use of SOI or dual-junction isolated transistors where a mid band-gap Fermi level may be used in conjunction with substrate or well biasing. Near-term potential solutions to the gate electrode problem include improvements in the doped polycide gate stack or the use of boron-doped Si-Ge gate electrodes. The development of enhanced means of activating the doped polysilicon for achieving tighter control of the work function as well as boron penetration resistance of the gate dielectric (e.g., the use of silicon nitride) is of great importance. (Link to the *PIDS* chapter)

In order to maintain high device drive currents over the long term, technology improvements are required to ultimately increase the mobility observed in traditional bulk CMOS devices The use of strained channel layers, such as strained Si on relaxed Si-Ge for NMOS and strained Si-Ge for PMOS will help in achieving this objective but will require considerable process optimization. Alternate devices such as non-standard, double gate devices may also provide a longer-term solution.

The incorporation of mobility-enhanced channels, alternate interfacial layers, high κ dielectrics, and new gate electrode materials into CMOS configurations poses a significant integration challenge. The thermal stability of many of the candidate material systems is inadequate to allow typical junction annealing cycles to be used after gate formation. The use of these new materials may require either junction annealing temperatures to be dramatically lowered or a reversal in the sequence of gate stack and junction formation, such as the "replacement gate scheme" or gate-last processes. These schemes increase manufacturing complexity, CoO, and may impact device performance and reliability by altering the gate to junction overlap and strenuous efforts are in place to retain the conventional CMOS process procedure.

Sidewall spacers are currently used to achieve isolation between the gate and source/drain regions, as well as to facilitate the fabrication of self-aligned, drain-engineered dopant structures. The robustness of the sidewall spacer limits the gate and source/drain contacting structure and processes that can be used. Sidewall spacers have traditionally been formed from deposited oxides, thermal oxidation of polysilicon, deposited nitrides, and various combinations of the above. These processes will continue to be used at least until the time when elevated or raised source/drain structures are required, at which time the compatibility with the side-wall spacer will become essential. Single drain elevated structures will require sidewall thicknesses in the range of 3–10 nm and gate dielectric-like reliability and stability to manage parasitic series resistance. Below a physical gate length of about 20 nm, even the best thermal oxides are susceptible to defect formation when subjected to selective epitaxial silicon or silicide processes. Nitrides or oxynitrides may offer a better alternative than oxide; however, additional research is needed to find and qualify an acceptable sidewall spacer.

Thermal and thin films are also very important for filling trench isolation as well as for pre-metal dielectrics. Trends for increasing vertical trench sidewalls, high aspect ratio gaps, top and bottom corner profile control, uniform filling of dense/isolated structures, and control are the key requirements for this application.

The key thermal/doping integration issues are maintaining shallow junction profiles, junction abruptness, obtaining high dopant activation, ensuring thermal compatibility of materials, and controlling how these issues impact device electrical performance. A potential solutions roadmap for Thermal/Thin Films is given in Figure 29.

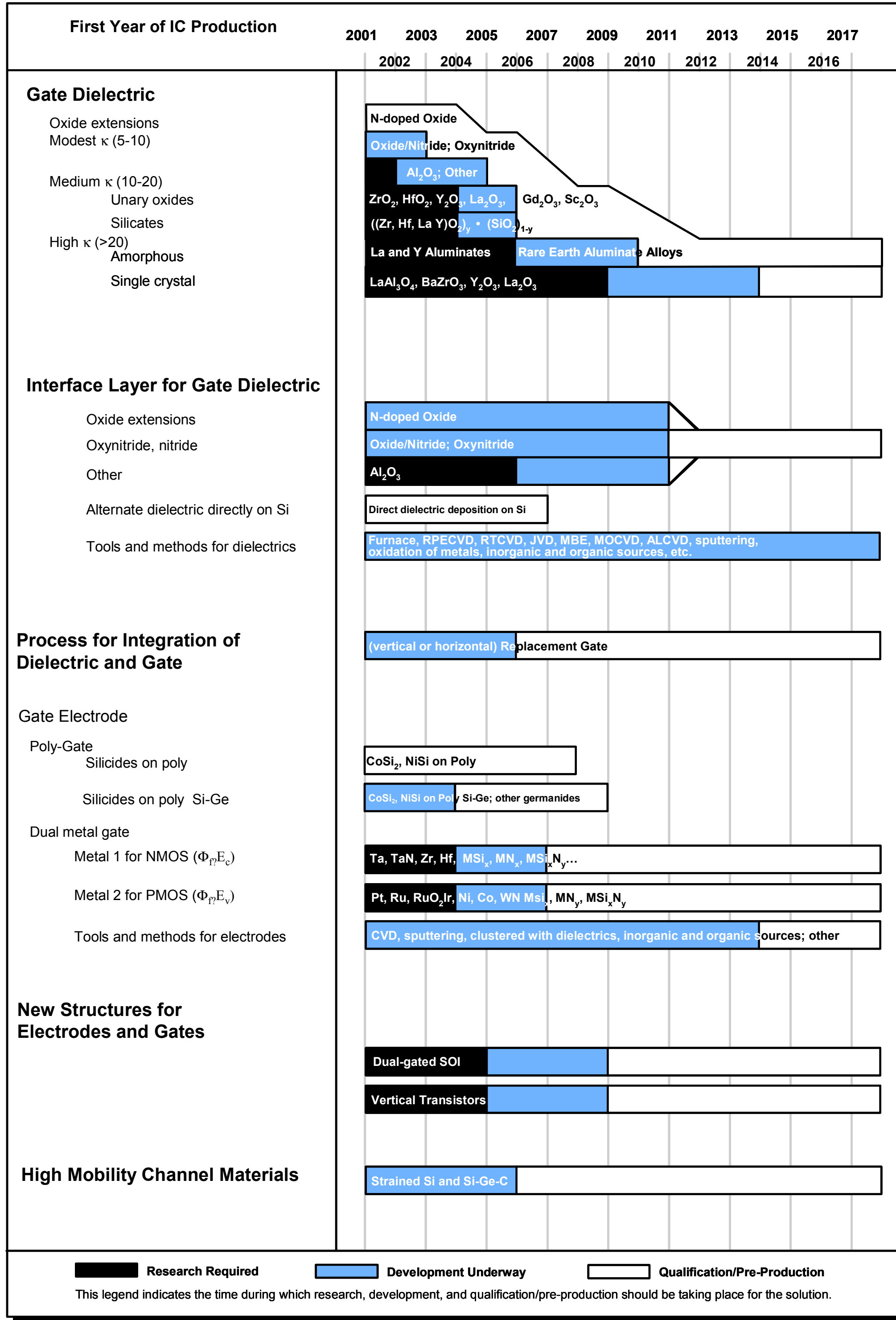


Figure 29 Thermal / Thin Films Potential Solutions

DOPING TECHNOLOGY

The near-term *difficult challenge* facing doping technology is the extension of conventional methods used to fabricate ultra-shallow junctions into the deep submicron regime. Intrinsic to this challenge is concommitantly achieving both the shallowness of the scaled drain extension junction and the necessary high lateral abruptness. These requirements all derive from the need to achieve a low parasitic source and drain series resistance, which may not exceed 16–35% of the channel resistance. (See also the *PIDS* Chapter Logic Technology Requirements, High Performance Near-term Table.)

The conflicting demands of producing highly doped, shallow source/drain and drain extension junctions and contacting these with low sheet resistance contact material will also challenge conventional salicide processes. Several other difficult challenges have been identified involving the formation of, and making contact to, ultra-shallow device junctions. These include activation technology to achieve junction depths as shallow as 10 nm with a sheet resistance of about 600 Ω/sq (refer to Thermal and Thin Film Doping and Etching Requireements, *FEP Table 51a and b*) concommitant with high depth- and lateral- abruptness; fabrication of silicide contacts to n^+ and p^+ regions with specific contact resistivities of $\leq 1\times10^{-7}\Omega\text{-cm}^2$; 2D dopant profiling of ultra-shallow junctions; and the placement of halo dopant implants with the necessary precision. Another challenge is the achievement of p+ polysilicon doping with an active concentration of ~ $2\times10^{20}/\text{cm}^3$.

Beyond the 65 nm node, the grand challenge is stated more simply as "Transistor Structure." Planar CMOS might be extended beyond the 65nm node using high κ gate dielectric layers, dual metal gate electrodes and elevated contact structures. As an alternate, new transistor structures such as non-standard double gate MOSFETs together with CMOS compatible processing may be needed to achieve transistor performance goals, and these will need to be developed and qualified in time for production. These structures and processes will need to be temperature and process compatible with the high κ gate stack. Here, there is a similar set of underlying challenges—thermal process compatibility with the dielectric (activation versus high κ); formation of low contact resistance structures compatible with a metal gate electrode; high conductance source/drain extensions, with very high levels of dopant activation and lateral abruptness; and the understanding and application of processes for metastable dopant activation. Underlying all these concerns is the inevitable rise of leakage currents associated with threshold scaling, gate dielectric tunneling, and junction tunneling.

SOURCE AND DRAIN EXTENSIONS

The management of short channel effects is expected to have a significant impact on processes used for doping drain extensions, channels, and channel edges. Drain extension doping levels are expected to increase, driven by the need to reduce junction depth while minimizing parasitic resistance. Similarly, drain extension doping profiles, which with earlier technology nodes required lateral grading for minimum hot carrier damage, will need to become more laterally abrupt to support low voltage operation. The complexion of the abruptness challenge depends on the device type. For P-channel devices, sensitivity simulations for abruptness indicate that above a critical abruptness value, narrow gate length devices are harder to build and there is only a marginal reduction in parasitic resistance. Improving the abruptness beyond some critical value therefore gives only minor improvements. On the other hand, for N-channel devices, a more abrupt source extension junction leads to a higher source injection velocity and higher resulting drive current. Therefore, for NMOS higher abruptness values are desirable. It is also noted that the present models from which the drain extension sheet resistivity and lateral abruptness values are derived remain rather primitive, and that further future changes in these requirements are likely to evolve as these models become more sophisticated.

Sub-nanometer spatially resolved 2D metrology is needed to monitor the location and shape of both vertical and lateral dopant profiles in the extension region. The extension depth, abruptness and resistivity values in Table 51 are given for the PMOS boron junctions. These are the critical junctions challenged by continued device shrinks. *NMOS junction requirements and challenges are detailed in supplemental material, in the R_SX_J worksheet.*

Potential doping solutions for these problems include ultra low energy ion implantation or plasma immersion ion implantation (PIII), combined with Spike annealing, and in time, alternative anneals such as LTA [Laser Thermal Annealing] may also be required. Another potential long-term solution is the use of epitaxially deposited doped layers (using techniques such as selective epitaxial deposition, or atomic layer epitaxy, ALE). (Refer to *Figure 30, Doping Potential Solutions*)

CONTACT

The conventional means of shunting the contact regions with a self-aligned silicide will be challenged by the need to achieve low gate sheet resistivity while at the same time continuously minimizing silicon consumption in the source drain

contact region. In addition, scaling of the contact area will result in unacceptable parasitic resistance if the specific contact resistivity of the silicon/silicide interface cannot be correspondingly reduced. The achievement of an acceptable contact resistivity will require that the maximum concentration of activated dopant be achieved in the silicon at the silicide interface, and may also require that the metal/semiconductor barrier height be reduced. In the near term, the doping and contact shunting processes are becoming increasingly interdependent, resulting in a complex materials science problem for which innovative integrated CMOS-compatible solutions are required. Thus, the interim extension of conventional salicided source/drain structures will challenge process capability and manufacturability of incumbent methods for producing doped contact regions.

Contact shunting potential solutions given in Figure 30 include progressively scaling of the incumbent salicide process; selectively depositing silicide; the siliciding of deposited sacrificial silicon; and the selective deposition of an alternate low resistivity contact metal. Ultimately, a robust contact process will require development of innovative new device structures in which the source and drain contact regions are elevated with respect to the channel. However, current methods, under investigation for producing elevated contact structures that may be needed for the sub-100 nm nodes, are still in the research stage and are not yet production qualified.

POLYSILICON GATE ELECTRODE DOPING

The polysilicon gate electrode also represents a major challenge for near-term scaling. Both the effective increase in the gate dielectric thickness associated with depletion and the channel autodoping associated with boron out-diffusion from the p^+ polysilicon gate will eventually require the phase-out beyond the 65 nm technology node of the currently used dual-doped polysilicon gate material. The model for the poly doping requirement is based on an allotment of the allowable poly depletion layer in relation to the EOT requirement (PIDS chapter). This is expressed in Table 51a and b as a 25% EOT allotment. This allotment results in less than optimal transistor performance and suggests that metal gates may be needed in the near future, but are unlikely to be implementable before the year 2007. Potential solutions are also shown in Figure 30. Near term potential solutions to the gate electrode problem include improvements in the doped polycide gate stack where the boron penetration resistance of the gate dielectric candidates becomes of great importance. Similarly, the development of enhanced means of activating the doped polysilicon for minimizing depletion and achieving tighter control of the work function is important and has possible solutions through materials engineering.

CHANNEL

The maintenance of acceptable off-state leakage with continually decreasing channel lengths will require that channel doping levels increase to offset the degradation in short channel effects for extremely small devices. Although both vertical and lateral channel engineering can be used to minimize degradation due to these effects, it will not be possible to scale threshold voltage without a large increase in off-state leakage. In the near term these scaling problems are exacerbated by the less aggressive scaling of the gate dielectric, the direct result of the unavailability of a higher κ gate dielectric material. This places even greater demands on channel doping engineering to optimize device performance. An added consequence is that for high performance devices, threshold voltage is indeed forecasted to scale, leading to very high sub-threshold leakage currents. In addition, since the thickness of the gate oxide and gate oxynitride materials are expected to scale to thickness values where direct tunneling currents are very large, off-state power consumption becomes a major cause for concern. Solutions to these problems, also discussed in the PIDS chapter, include the use of less aggressive channel length scaling for low power devices, as well as the use of multiple gate lengths and gate dielectric thicknesses on the same chip to optimize speed/power performance. Doping processes must evolve to accommodate this new complexity. Refer to the *PIDS* chapter for further discussion.

Lateral channel engineering remains an effective vehicle to improve device performance. Typically, lateral engineering is accomplished by angled implants into the channel region following gate patterning. The lateral implants cause an increase in doping concentrations near the SDE regions. The resulting channel profile is device size dependent. Smaller devices have higher average channel doping levels. This can be used to tailor the relationship between threshold voltage and device size to improve device performance. For lateral channel engineering to be effective, the lateral profile must vary over the length of the channel. This will become extremely difficult as channel lengths continue to scale to the sub-50 nm regime.

The scaling of both the vertical and lateral channel profiles is extremely difficult due to diffusion effects that degrade the abruptness of the profiles. These profiles must be extremely localized to maximize the beneficial effect on device performance. Transient enhanced diffusion, (TED), high concentration diffusion enhancement effects (such as BED) and normal thermal diffusion must all be managed to produce the required channel profiles. The need for extremely abrupt,

well-defined channel profiles may require the use of epitaxial deposition in conjunction with extremely low temperature processing. Even with this added capability, circuit designers may need to trade-off high off-state leakage versus markedly lower drive currents.

The first three sections of potential solutions in Figure 30 show a set of options for methods to produce abrupt, ultrashallow, junctions. The combination of ion implantation and anneal is expected to be the dominant solution into the second decade of this roadmap. There are, however, several options for both implant and anneal which are shown as being competitively evaluated and proposed for manufacturing. Correlated with this is the potential solution of “Defect Engineering.” This approach may form new ways to form abrupt, ultrashallow junctions, by mitigating or using the transient enhanced diffusion effects that make boron diffusion so difficult. Also noted though, is that “low resistance deposited and thermal doping” as a solution category includes all of the competitive technologies for ultrashallow junction formation. The solution involving deposited layers is expected to be the ultrashallow junction replacement technology near the end of the Roadmap.

The contact category of Figure 30 notes that self aligned metal silicides remain the incumbent technology, but that there are a number of solutions in the research phase. This is appropriate as the contact resistance challenge is on a par with the ultrashallow junction/abruptness challenge.

The device structure and channel engineering category of Figure 30 serves to remind us of truly innovative alternative solutions currently being explored. It is easy to propose that raised source/drain types of structure may be a solution to provide some relief from the ultrashallow junction/abruptness challenge and to the contact resistivity challenge. However, as discussed in the PIDS chapter, there are multiple potential alternative devices that in the longer term may change both the requirements and the prioritization of the potential solutions.

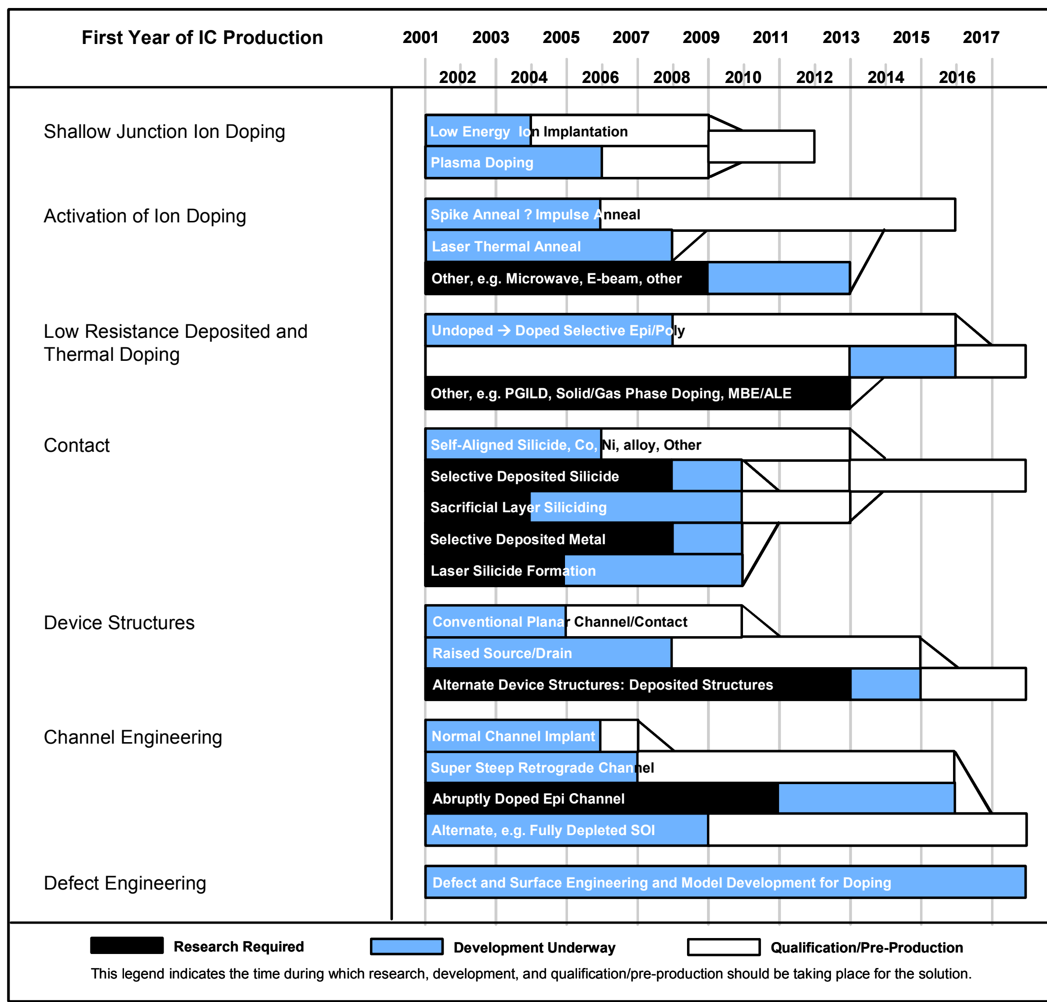


Figure 30 Doping Potential Solutions

FRONT END ETCH PROCESSING

Reduction in critical dimension (CD) and process control remain key challenges for FEP etch technology. These, coupled with new materials such as high κ gate dielectrics and metal gates; larger diameter (300mm) wafers and new generations of photoresist, make the challenges truly formidable. In addition, other CD reduction technologies, such as Resist Trim are also now being used in production as an alternative to, or in combination with enhanced lithography approaches such as Optical Proximity Correction (OPC) and Phase Shift Masks (PSM). Refer to the *Etch worksheet in supplemental material* for more details.

To achieve the level of control required, etch modules must have a number of fundamental design attributes (refer to *FEP Table 51a*), CD etch uniformity is a strong function of chamber design which fundamentally must address both uniform gas distribution and particularly uniform plasma distribution consistent with relatively low bias voltage. Although compensation effects can be used to improve uniformity, these introduce unacceptable repeatability risks because they create an inherently narrow process window. Wafer edge profile consistently remains an issue. Edge profile control is independent of wafer size and can possibly be considered one of the major challenges to equipment design generally. Even though very uniform plasmas may be formed across the chamber diameter, the uniformity at the wafer edge must be modified by edge compensation to account for edge anomalies. Theoretically, CD control at the sub 3nm level can be achieved in a variety of ways but the end result must be a vertical, smooth edge profile, achieved with a damage-free process that exhibits good selectivity control and minimum micro-loading. Of particular importance is the need for damage free processing especially as it relates to controlled completion of the gate etch process without damage to the underlying silicon. Meeting this requirement becomes increasingly difficult as critical dimensions shrink and new gate dielectric layers are introduced (See Figure 31). *In situ* etch monitoring to achieve etch stop on very thin gate dielectric layers and feed-forward/feed-back integrated metrology for profile control may well become standard techniques used to achieve sub-1nm CD control.

The required CD control and etch characteristics mentioned above must also be achieved with new materials, such as high k dielectric films and metal gate structures (See Figure 31). Here, the etch equipment and process must not only deal with these new materials but may also need to be integrated into potentially new process architectures dictated by their limited thermal stability. Many plasma sources have been developed with the intent to offer improved etch technology. Generally speaking however, the plasma density for optimum results is in the $\sim 10^{11}/cm^3$ regime. Equipment and process development may well take multiple paths. Evolution of ECR and ICP processing is expected to continue and may well develop the necessary features to deal with new gate materials. It is also possible that new etch approaches may be required to deal with the particular non-volatile by-products produced by etching metal gate electrodes. Such developments will also have collateral impacts on overall equipment robustness and particularly on MTBC and MTTC. For high-k gate dielectric layers these developments must be complete by 2005 when such dielectric layers will be needed for low stand-by power devices. It is expected that metal gate electrode etch processes and equipment will be required shortly thereafter. To deal with sensitivity to damage, Chemical Downstream Etching , Neutral Stream or other innovative etching techniques should be investigated as a possible finish- or over-etch step. Work using Pulsed Plasmas is already underway with this in mind. Ideally, for cost reasons, if not from a technical viewpoint, development should lead to an etch tool which can deal with both new gate materials and possible stringent damage requirements which may be brought about by advanced chip architecture (See Figure 31).

As linewidths shrink, the presence of line edge roughness (LER) is becoming increasingly important to CD control along with the angle the slope of the etched gate. There is also some evidence that LER contributes to gate leakage. Both lithography and etch can contribute to LER. The choice of gate material, photoresist type and etch chemistry all contribute to the degree of LER. To set a control target for this quantity, the impact of LER to device performance has to be better understood and the associated measurement methodology and equipment will also have to be developed.

Resist Trim of printed features is being used in production environments to reduce gate physical dimensions as an alternative to, or in addition to lithographic techniques such as OPC and PSM (*link to illustration*). Trimming also allows compensation for within wafer and dense/isolated line with variation in subsequent steps to enable the meeting of overall profile and CD requirements. A flexible FEP etch reactor and process is essential here. It is important to note that in addition to reducing the width of the resist uniformly across the wafer, the height of the resist must not be excessively reduced or selectivity issues will arise when transferring the pattern into the underlying hard mask. Another consideration is corner faceting. While the overall resist height may well remain intact, faceting reduces the overall effective height and the selectivity requirement becomes more difficult to achieve. Much of the current learning has been obtained on 248nm resist types. For 193nm and 157nm generation resist materials, the task will be much more difficult since both the height

of the resist material decreases significantly (~700nm[248], ~400nm[193], ~200nm[157]), and because the inferior etch resistance of these materials creates added problems with etch rate and selectivity. In addition to resist trim issues, many of the new resist materials at 193nm and 157nm can still fail during etch due to reticulation at much lower temperatures, structural inconsistency, surface irregularities and general lack of robustness. Multi-level resist techniques are also being developed to enable smaller feature transfer into underlying materials.

Changes in gate stack materials will probably occur in two phases. Firstly, the introduction of high κ gate dielectric materials, which with the use of nitrided oxide gate dielectric layers, is already starting to occur and secondly, the introduction of metal or metal nitride gates electrodes. As scaling reduces thermal oxides to the 13nm regime, direct charge carrier tunneling leads to unacceptably high gate leakage. High κ gate materials under development, such as Zr and Hf oxides, can reduce such leakage. It is well known that the interaction between gate materials and lithography is a vital aspect to achieving good CDs. These new high κ gate dielectric materials will be more difficult to etch because of their different chemical and physical properties. Wet etching will also be difficult since increased thickness (compared to SiO or SiN materials) will probably result in unacceptable undercut profiles. Clearly, as gate dielectrics with increased κ value are introduced, new etch challenges will be uncovered. Metal gate electrode materials will also pose a CMOS integration challenge. Because of work function requirements, the candidate metal gate materials for P+ Polysilicon replacement (Pt, Ir, Ni, Mn, Co), will be different from those that replace the N+ Polysilicon (Ta, Zr, Hf, Ti). These materials generally have less volatile by-products than their doped polysilicon counterparts and in addition each is expected to have its own distinct etch process requirements. Therefore CMOS integration of the etching processes becomes more difficult. This brings into question whether a simultaneous etch of both gates is possible. One solution might be the use of protective resist overcoat/masks similar to the ion implant masks used in selective CMOS doping . The combined damage-free etching of dual metal gates together with damage-free stopping on a high κ oxide remains the ultimate goal.

The introduction of new gate materials will also impact defectivity. Strict FEP etch requirements related to defect density and plasma damage must also be met. With existing device designs, plasma damage introduced by various tunneling phenomena, hot carriers and charging is fairly well understood and characterized with existing device designs and materials. With the onset of new materials, new issues will arise relative to new damage mechanisms (See Figure 31). To meet future defect density requirements the plasma processes and etch tools must generate considerably fewer and smaller particles. Improvements will be required in the etch chemistries, the control of deposition in the etch chamber, and the cleaning procedures used for the etch chamber maintenance. These requirements will have to be met, consistent with acceptable wafer processing cost and tool uptime. Plasma etch tool design and plasma processing conditions must be developed that do not cause charging damage. Alternative high-κ and/or stacked gate dielectric materials will require development of multi-step etch processes. This requirement may translate into the need to change gas chemistries in the same etch module in order to etch a variety of stacked materials, or the need to etch the bulk material in a main etch step, followed by completion in a finishing etch, followed by an over-etch step. It is highly desirable to ascertain the amount of material remaining prior to completion of the main etch through the use of interferometry or a similar sensing technique so that a pre-emptive endpoint can be determined. Again, a highly selective non-damaging process is required.

Shallow Trench Isolation (STI) also has some challenging integration issues as we move towards the 100nm regime and beyond. Here, many device manufacturers are beginning to use etch processes, rather than classic thermal processing, to round the top corner of the STI trench as a means for alleviating the classic transistor double-hump effect. Etching also has the advantage of not encroaching into the active area. For this application, integration challenges are top- and bottom-corner rounding radius control, STI wall slope control and the void–free filling of the trench with high quality oxide. These characteristics must be controlled for both variable STI gap widths and variable density of STI features while maintaining CD control.

The scaling of sidewall spacer width and its dimensional control presents another challenge to plasma etch. The spacer width and its sensitivity to over-etch is governed by the gate electrode profile, the thickness control and conformality of the spacer dielectric deposition process, as well as the anisotropy of spacer etch process (refer to *FEP Table 51a and b*). An accurate assessment of the scalability of the sidewall spacer from an etch perspective is hindered by the limited availability or process control data. Here the use of feed-forward/feed-back integrated metrology may well provide a breakthrough.

For future DRAM device generations using stacked capacitor structures, the development and introduction of High Aspect Ratio Contact via etches (HARC) of ~15:1 together with efficient post-etch etch residue removal processes remains crucial. The ability to maintain CD and selectivity consistent with appropriate etch stopping and low damage to

the shallow contact junctions will be key technical challenges. For ultra shallow junctions a small and controlled degree of silicon contact etching is desirable for optimum device contact resistivity and leakage.

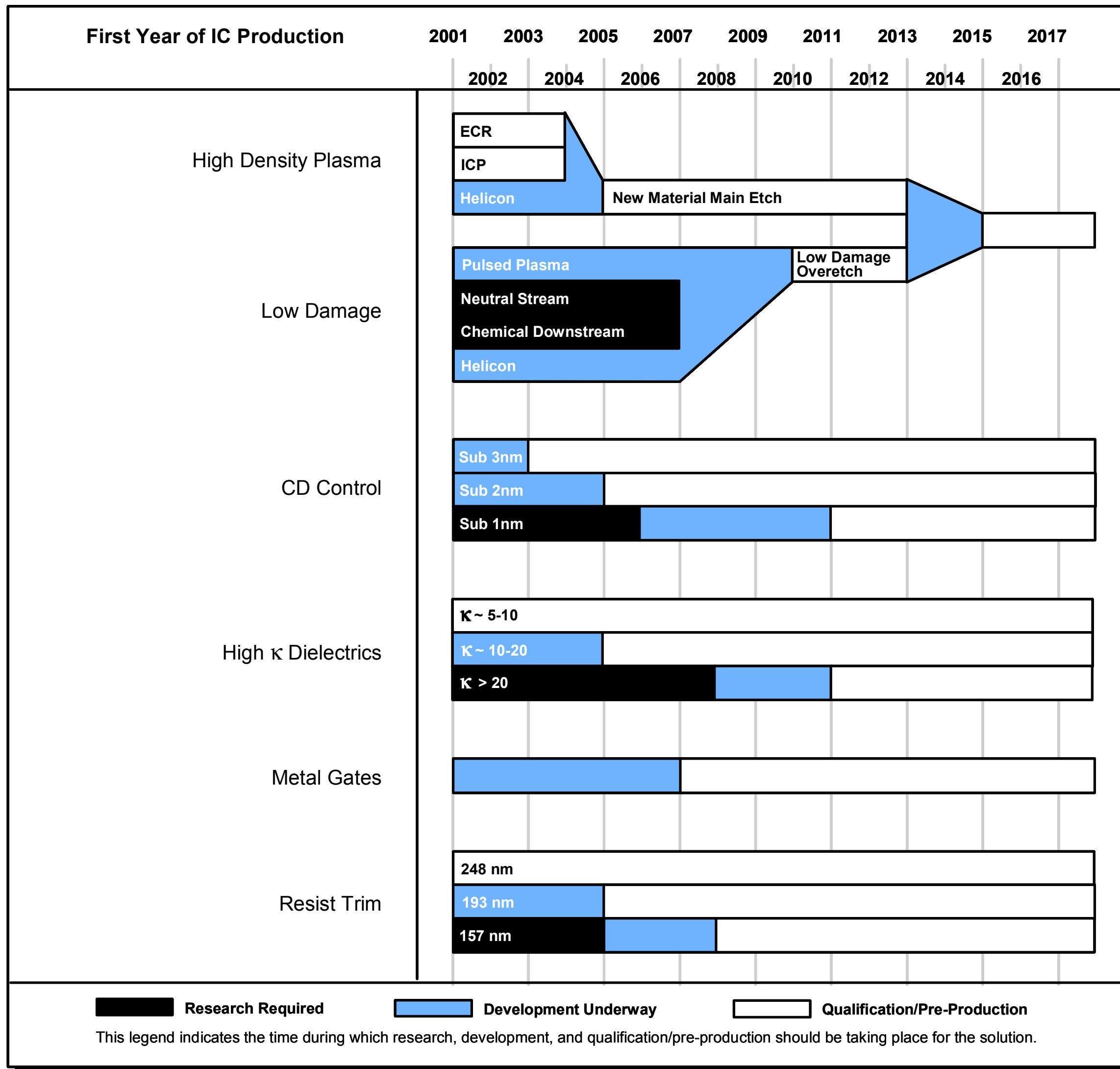


Figure 31 Front End Processes Etching Potential Solutions

DRAM Stacked Capacitor

Historically, the quadrupling capacity of DRAM products every three years has been based on the following:

1. Reduction of the minimum feature size (2×)
2. Expansion of the chip size (about 1.4×)
3. Improvement of cell area factor and cell efficiency (about 1.4×)

However, continued chip size growth is inhibited due to economic reasons and cell factor improvement is impaired due to the physical limit of the cell layout. Instead of quadrupling, DRAM products of intermediate capacity such as 128Mb and 512Mb, have appeared respectively after the 64Mb and 256Mb DRAM generations. DRAM capacitor technology faces new challenges of introducing new storage capacitor dielectric and electrode materials. Table 52 summarizes the technology requirements for the DRAM stacked capacitor. The DRAM cell size is being scaled down faster than the

general design rule. An area of at least $8F^2$ (F is feature size) is realized at the 180nm node, which is the smallest cell size with a folded bit line architecture. Each of the target values in Table 52 is based on the assumption that a cell capacitance retains at least 25fF/cell to assure stable circuit function and sufficient soft-error immunity.

With the onset of the mega-bit era, nitride/oxide dielectric films, together with 3-dimensional polysilicon capacitor structures, have been used to keep the cell capacitance sufficiently high for sensing and noise immunity. However, it is difficult to keep a sufficiently high cell capacitance value using these materials and structure at the 130nm node and beyond. Thus, alternative high κ dielectrics and electrodes are most likely to be introduced beyond this technology node. Ta_2O_5, with a range of dielectric constants, is one of the most promising dielectrics. For the 130nm node, a poly-silicon bottom electrode and a 3D capacitor cell with high κ dielectric and metal counter-electrode are needed (this is an example of a Metal Insulator Silicon (MIM) structure). A Ta_2O_5 stack with an effective dielectric constant of 22 is a promising choice for this node. However, due to the growth of an oxide layer at the interface during thermal annealing of the capacitor dielectric, this structure is not viable beyond the 130nm node. On the other hand, if metals such as Ru and Pt are used as the bottom electrode of the storage node (MIM), a Ta_2O_5 stack exhibits an effective dielectric constant of more than 50 because the metal electrode is free from oxidation and can provide a highly oriented crystal microstructure.[23] Therefore, a MIM structure is required beyond the 130nm node.

At the 90nm node, metals or conductive metallic oxides such as Pt, Ru, RuO_2 and IrO_2 have to be used as the storage node bottom electrode primarily to improve the immunity to oxidation and provide a template for preferred microstructure. From the thermal budget viewpoint, these electrode materials should be deposited at low temperature by using CVD based methods. However, relatively higher temperature annealing in oxygen ambient will most likely be required. Lowering the process temperature is needed if metals are used for bit lines to minimize device performance degradation.

Reducing leakage current at lower processing temperatures is another difficult challenge for DRAM capacitor technology for the 90nm node and beyond. Careful process integration is required to prevent capacitor film degradation caused by plasma damage and the oxide reducing processes used in the BEOL (Back End Of Line).

Beyond the 90nm node, a higher κ material such as BST with a dielectric constant of about 200–250 will be needed in order to reduce the aspect ratio of the storage node and the High Aspect Contact (HAC) hole.

Near the 45nm node, new ultra high κ materials with a dielectric constant in the range of 700, such as epi-BST, will be required. However, even if such new materials are successfully developed, an upper electrode deposition process for a very high aspect ratio storage node may limit capacitor integration. Therefore, in addition to material and process development, new memory cell concepts such as a gain cell architecture will be required at the 45nm node and beyond.

At and beyond the 32nm node, cell size has to be reduced below $6F^2$ to keep a reasonable chip size. New cell architecture such as a cross-point cell will be required. In addition, a dielectric constant of 1500 is called for. However, there is no known material with such a high κ value. Thus innovation of “super high κ” value is called for or a new device structure that uses ferroelectrics in the DRAM context may be attempted.

It is noted that by amending retention time, the color of this row is no longer red beyond the 65nm node since cell leakage currents are adequately met. However, silicon diffusion limited reverse bias leakage now becomes a fundamental limitation for retention time, indicating the potential need for SOI substrates for 65nm DRAM and beyond.

The process technology requirements for system-on-a-chip (SOC) with embedded DRAM exhibit many variations depending on the ratio of logic area and memory area. Cell capacitance requirements for embedded DRAM may be smaller than those for stand-alone DRAM. One of the serious problems for SOC is contact via formation. In general, the stacked capacitor DRAM processes require relatively deep contact vias as compared with those in logic processes. Therefore, the contact via size of DRAM has to be enlarged to minimize aspect ratio. For this reason it will be difficult to achieve the same metal line pitch for the logic section using the same DRAM design rule. In the logic-based SOC, cell size expansion is needed to reduce the capacitor height and to decrease the contact via aspect ratio. On the other hand, in the memory-based SOC, the metal line pitch has to be adjusted so that the DRAM contact via size may be kept large enough. Therefore, some additional break-through in SOC is required to solve this contact via density issue.

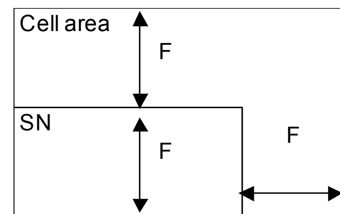
[23] *K.Kishiro, et al. Jpn. J. Appl. Phys. Vol.37 (1998) pp.1336-1339*

Table 52a DRAM Stacked Capacitor Films Technology Requirements—Near-term

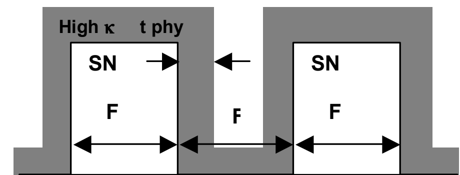
YEAR OF PRODUCTION	2001	2002	2003	2004	2005	2006	2007
DRAM ½ PITCH (nm)	130	115	100	90	80	70	65
MPU / ASIC ½ PITCH (nm)	150	130	107	90	80	70	65
MPU PRINTED GATE LENGTH (nm)	90	75	65	53	45	40	35
MPU PHYSICAL GATE LENGTH (nm)	65	53	45	37	32	28	25
DRAM Product [A]	512M		1G		2G		4G
Cell size factor a [B]	8.0	8.0	6.0	6.0	6.0	6.0	6.0
Cell size [μm^2] [C]	0.14 =0.26*0.52	0.11 =0.23*0.46	0.06 =0.2*0.3	0.049 =0.18*0.27	0.038 =0.16*0.24	0.029 =0.14*0.21	0.03 =0.13*0.2
Storage node size (μm^2) [D]	0.051 =0.13*0.39	0.040 =0.115*0.35	0.020 =0.1*0.2	0.016 =0.09*0.18	0.013 =0.08*0.16	0.010 =0.07*0.14	0.008 =0.065*0.13
Capacitor Structure	Cylinder MIS Ta_2O_5	Cylinder MIS Ta_2O_5	Pedestal MIM Ta_2O_5 (Ref. U)	Pedestal MIM Ta_2O_5 (Ref. U)	Pedestal MIM BST	Pedestal MIM BST	Pedestal MIM BST
Dielectric constant	22	22	50 Ref. U	50 Ref. U	250	300	450
SN height H (μm)	0.9	0.9	0.9	0.9	0.65	0.53	0.38
Cylinder factor [E]	1.5	1.5	1.5	1.5	1.0	1.0	1.0
Roughness factor	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Total capacitor area (μm^2)	1.48	1.30	0.87	0.72	0.32	0.23	0.16
Structural Coefficient [F]	10.9	12.3	14.5	14.8	8.5	7.9	6.2
teq@25fF (nm) [G]	2.0	1.80	1.20	1.00	0.45	0.32	0.22
t phy.@25fF (nm) [H]	5.9	4.5	15.3	12.8	28.7	24.7	25.0
A/R of SN (OUT) for cell plate depo. [I]	7.6	8.5	13.4	13.3	28.9	25.7	25.2
HAC diameter (μm) [J]	0.16	0.14	0.12	0.11	0.10	0.08	0.08
Total interlevel insulator and metal thickness except SN (μm) [K]	1.00	0.96	0.93	0.90	0.86	0.84	0.81
HAC depth (μm) [L]	1.90	1.86	1.86	1.76	1.51	1.37	1.19
HAC A/R	12.2	13.5	15.5	16.3	15.8	16.3	15.3
V capacitor(V)	1.8	1.8	1.8	1.5	1.5	1.5	1.2
Retention time (ms) [M]	64	64	64	64	64	64	64
Leak current (fA/cell) [N]	1.05	1.05	1.05	0.88	0.88	0.88	0.70
Leak current density (nA/cm^2)	71.3	81.0	121.6	121.9	270.6	378.2	448.9
Deposition temp.(degree C)	~ 500	~ 500	~ 500	~ 500	~ 500	~ 500	< 500
Film anneal temp. (degree C)	~ 800	~ 800	~ 750	~ 750	~ 750	~ 750	< 750
Word line Rs (ohm/sq.)	5.0	3.0	2.0	2.0	2.0	2.0	2.0

Notes for 52a and b:

[A] 2001 Overall Roadmap Technology Characteristics , Table 1a and b
[B] a = (Cell size)/F^2 (F : minimum feature size)
[C] Cell size = aF^2 (Cell shorter side =2F)*
[D] SN size =(a/2 -1)F^2 (SN shorter side = F)*
[E] Cylinder structure increase the capacitor area by a factor of 1.5.
[F] SC = (total Capacitor area) / (Cell size)
*[G] teq=3.9*E0*(total Capacitor area)/25fF*
*[H] t phy. = teq*Er/3.9 If polysilicon is used as a bottom electrode. t phy. =(teq-1)*Er/3.9*
[I] A/R of SN (OUT) = (SN height) / (F - 2 t phy.)*
*[J] HAC diameter =1.2*F (HAC : High Aspect Contact)*
[K] The thickness is assumed to be 1.05µm@180nm. (10% reduction by each generation)
[L] HAC depth = SN height + Total Interlevel insulator and metal thickness
[M] DRAM Retention time (PIDS)
*[N] (Sense Limit*C*V_{dd}/2) / (Retention Time * MARGIN) (Sense limit=30% leak, MARGIN=100)*



Notes[C] & [D] Cell area and Projected SN



Note [I] A/R of SN (OUT)

Table 52b DRAM Stacked Capacitor Films Technology Requirements—Long-term

YEAR OF PRODUCTION	2010	2013	2016
DRAM ½ PITCH (nm)	45	32	22
MPU / ASIC ½ PITCH (nm)	50	35	25
MPU PRINTED GATE LENGTH (nm)	25	18	13
MPU PHYSICAL GATE LENGTH (nm)	18	13	9
DRAM Product [A]	8G	32G	64G
Cell size factor a [B][O]	6.0	4.0	4.0
Cell size (μm^2) [C]	0.012 =0.09*0.14	0.0041 =0.064*0.06	0.0019 =0.044*0.04
Storage node size (μm^2) [D]	0.004 =0.045*0.09	0.0010 =0.032*0.03	0.0005 =0.022*0.02
Capacitor structure	Pedestal MIM BST	Pedestal MIM ???	Pedestal MIM ???
Dielectric constant	800	1500	3000
SN height H (µm)	0.21	0.15	0.08
Cylinder factor [E]	1.0	1.0	1.0
Roughness factor	1.0	1.0	1.0
Total capacitor area (μm^2)	0.06	0.02	0.008
Structural coefficient [F]	5.0	4.9	3.9
teq@25fF (nm) [G]	0.08	0.028	0.010
t phy.@25fF (nm) [H]	17.2	10.7	8.0
A/R of SN (OUT) for cell plate depo. [I]	19.8	14.3	13.3
HAC diameter (µm) [J]	0.05	0.04	0.03
Total interlevel insulator and metal thickness except SN (um) [K]	0.73	0.66	0.59
HAC depth (µm) [L]	0.94	0.81	0.67
HAC A/R	17.4	21.0	25.4
V capacitor (V)	0.9	0.6	0.5
Retention time (ms) [M]	64	64	64
Leak current (fA/cell) [N]	0.53	0.35	0.29
Leak current density (nA/cm^2)	868.1	1738.3	3893.8
Deposition temp. (degree C)	< 500	< 500	< 500
Film anneal temp. (degree C)	~ 650	< 650	< 650
Word line Rs (ohm/sq.)	2.0	2.0	2.0

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow--Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known

Year of First Product Shipment Technology Node	2001 130 nm	2002	2003	2004 90 nm	2005	2006	2007 65 nm	2010 45 nm	2013 32 nm	2016 22 nm
Upper Electrode	metal									
High κ dielectric	ON	Ta$_2$O$_5$, Al$_2$O$_3$			BST, STO		Epi-BST			
Bottom Electrode	poly-Si		metal				perovskite			

(A) Metal : Ti, TiN, W, Pt, Ru, RuO_2, IrO_2 (B) Perovskite : $SrRuO_3$; N.[24].

Figure 32 DRAM Stacked Capacitor Potential Solutions

[24] Fukushima et al., IEDM Technical Digest, pp. 257-260, 1997.

DRAM TRENCH CAPACITOR

Tables 53a and b summarize the technology requirements for DRAM trench capacitor technology. The target values are based on the assumption of a 35fF capacitance per DRAM cell. For storage nodes down to the 100 nm design rule, trench capacitor technology is characterized by conventional thickness scaling of the nitride/oxide capacitor dielectric, in combination with surface enhancement techniques such as bottle-shaped trenches. As a result of ground-rule shrinking, the aspect ratio (trench depth to trench width) will increase up to values of ~60 for the 100 nm design rule. It is expected that a new high κ dielectric material will be required by 90nm technology node.

For embedded applications, the trench technology with its capacitor buried in the substrate enables a planar transition between the DRAM cell array and the logic circuit. Thus, a critical thinning of individual metal lines of the multilevel metallization, or reduced lithographic resolution at the transition area is avoided. Also avoided is the need for deep, high aspect ratio contact holes. In addition, since the capacitor is processed prior to the transfer device, degradation of device performance from the capacitor-forming thermal cycle is not encountered.

The scaling of the cell area factors as a consequence of the new chip size model will require a further optimization of the layout and the areas consumed by each element of the DRAM storage cell. This implies measures such as the replacement of the conventional planar transfer device by a vertical transistor, or the reduction in storage node area by further increases in the aspect ratio of the capacitor structure. Finally, all current DRAM technologies with their lithographic defined word-, and bit-line pitch, will reach a theoretical scaling limit at a cell size of $4F^2$ given by the 2F pitch for each metal line. For smaller values, new concepts such as multi-bit circuits or multi-layer DRAMs are required. Currently, there are no known solutions such fundamental changes.

Table 53a DRAM Trench Capacitor Technology Requirements—Near-term

YEAR OF PRODUCTION	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*
DRAM ½ PITCH (nm)	*130*	*115*	*100*	*90*	*80*	*70*	*65*
MPU / ASIC ½ PITCH (nm)	*150*	*130*	*107*	*90*	*80*	*70*	*65*
MPU PRINTED GATE LENGTH (nm)	*90*	*75*	*65*	*53*	*45*	*40*	*35*
MPU PHYSICAL GATE LENGTH (nm)	*65*	*53*	*45*	*37*	*32*	*28*	*25*
Storage node area, top view [(DRAM 1/2 pitch)^2]	**2**	**2**	**2**	**<2**	**<2**	**<2**	**<2**
Trench structure	**bottled**	**bottled**	**bottled**	**bottled**	**bottled**	**bottled**	**bottled**
Trench area enhancement factor	**1.3**	**1.5**	**1.5**	**1.7**	**1.7**	**1.7**	**1.7**
Capacitor dielectric equivalent oxide thickness (nm)	**4.5**	**4.5**	**4.3**	**3.6**	**3.2**	**2.8**	**2.6**
Trench depth [µm], (at 35fF)	**7.6**	**7.5**	**8.2**	**6.9**	**6.9**	**6.8**	**6.9**
Aspect ratio (trench depth / trench width)	**43**	**48**	**61**	**64**	**72**	**81**	**88**
Upper electrode	**Poly-Silicon**	**Poly-Silicon**	**Poly-Silicon**	**Poly-Silicon**	**Poly-Silicon**	**Metal/Poly-Si**	**Metal/Poly-Si**
Dielectric material	**NO**	**NO**	**NO**	**High-κ**	**High-κ**	**High-κ**	**High-κ**
Bottom electrode	**Silicon**	**Silicon**	**Silicon**	**Silicon**	**Silicon**	**Silicon**	**Silicon**
Capacitor Structure / dielectric	**Silicon-Insulator-Silicon / NO**			**Silicon-Insulator-Silicon / High-κ**		**Metal Silicon-Insulator-Silicon / High-κ**	

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow–-Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known

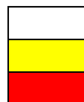


Table 53b DRAM Trench Capacitor Technology Requirements—Long-term

Year of Production	*2010*	*2013*	*2016*
DRAM ½ Pitch (nm)	*45*	*32*	*22*
MPU / ASIC ½ Pitch (nm)	*50*	*35*	*25*
MPU Printed Gate Length (nm)	*25*	*18*	*13*
MPU Physical Gate Length (nm)	*18*	*13*	*9*
Storage node area, top view [(DRAM 1/2 pitch)^2]	**<2**	**<2**	**<2**
Trench structure	**bottled**	**bottled**	**bottled**
Trench area enhancement factor	**1.7**	**1.7**	**1.7**
Capacitor dielectric equivalent oxide thickness (nm)	**1.2**	**0.6**	**0.3**
Trench depth (µm), (at 35fF)	**4.9**	**4**	**3**
Aspect ratio (trench depth / trench width)	**91**	**104**	**114**
Upper electrode	**Metal**	**Metal**	**Metal**
Dielectric material	**1: Epi-high-κ / 2: High-κ**	**1: Epi-high-κ / 2: High-κ**	**1: Epi-high-κ / 2: High-κ**
Bottom electrode	**1: Silicon**	**1: Silicon**	**1: Silicon**
	2: Metal	**2: Metal**	**2: Metal**
Capacitor structure / dielectric	**1: MIS / (Epi)-high-κ** **2: MIM / High-κ**		

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow–-Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known

Non-Volatile Memory (FLASH)

FLASH Memory is a new addition to the 2001 ITRS, and is the result of collaboration between the FEP and PIDS technology working groups. Non-volatile semiconductor memories, such as flash electrically erasable programmable read-only memories (flash EEPROMs), typically comprise a floating gate memory cell, which includes a source region, a drain region and a channel region formed in the silicon substrate, and a floating gate, formed above the substrate between the channel region and a control gate. The floating gate is separated from the substrate by the tunnel (oxide) dielectric and from the control gate by the interpoly dielectric. A voltage differential is created in the cell when a high voltage is applied to the control gate while the channel region is kept at a low voltage. This voltage difference causes electrons to move from the channel region to the floating gate through a phenomenon known as tunneling, thus charging the floating gate. This movement of electrons is referred to as programming. An opposite voltage difference is created between the control gate and the drain or the channel region, causing electrons to move from the floating gate to the drain or channel region through a tunneling mechanism, thus discharging the floating gate. This movement of electrons is referred to as erasing.

Scaling of semiconductor devices like non volatile flash memories requires continuous thickness reduction of the tunnel dielectric layer in order to improve the program/erase performance and to reduce the applied voltage. Scaling also requires a thickness reduction of the interpoly dielectric to avoid the degradation of the coupling ratio factor of the memory cell. The most important issues related to these requirements are the capacity of the tunnel oxide dielectric to sustain the required number of program/erase cycles and, for both dielectrics, to assure the required charge retention properties of the device in all operating conditions.

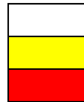
Present technologies, mainly based on thermal and CVD oxynitrides for both dielectrics, are not expected to satisfy the aforementioned dielectric scaling needs, thus requiring the implementation of new technologies and new materials. The critical technology requirements presented in Table 54a and b are the consensus of ITRS Flash experts. It is important to note that due to challenges in the tunnel oxide and interpoly dielectric layers, red shows up as early as 2005 when high κ dielectrics may be required.

At the present time, international consensus has not been achieved relative to the potential solutions needed to address these scaling issues. It is expected that this issue will be addressed in the 2002 ITRS update.

Table 54a FLASH Non-volatile Memory Technology Requirements—Near-term

YEAR OF PRODUCTION	2001	2002	2003	2004	2005	2006	2007
DRAM ½ PITCH (nm)	130	115	100	90	80	70	65
MPU / ASIC ½ PITCH (nm)	150	130	107	90	80	70	65
MPU PRINTED GATE LENGTH (nm)	90	75	65	53	45	40	35
MPU PHYSICAL GATE LENGTH (nm)	65	53	45	37	32	28	25
Flash technology node (nm) [A]	150	130	115	100	90	80	70
Flash NOR tunnel oxide thickness (EOT - nm) [B]	9.5–10.5	9.5–10	9–10	9–10	8.5–9.5	8.5–9.5	8.5–9.5
Flash NAND tunnel oxide thickness (EOT - nm) [B]	8.5–9.5	8.5–9	8–9	8–9	8–9	7.5–8	7.5–8
Flash tunnel oxide thickness control EOT (% 3σ) [C]	<± 4	<± 4	<± 3.5	<± 3.5	<± 3	<± 3	<± 3
Flash tunnel oxide minimum Q_{BD} @ 1×10^{-2} A/cm^2 (C/cm^2) [D]	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Flash tunnel oxide defectivity @ minimum Q_{BD} (def/cm^2) [E]	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Flash tunnel low field leakage (nA/5V) [F]	100	100	100	100	100	100	100
Flash program/erase window ΔV_T (V) [G]	>3	>3	>3	>3	>3	>3	>3
Flash erase time degradation t_{max}/t_0 [H]	<2	<2	<2	<2	<2	<2	<2
Flash program time degradation t_{max}/t_0 [I]	<2	<2	<2	<2	<2	<2	<2
Flash NOR interpoly dielectric thickness (EOT-nm) [J]	13-15	12-14	11-13	11-13	10-12	9-11	9-11
Flash NAND interpoly dielectric thickness (EOT-nm) [J]	14-16	13-15	12-14	12-14	12-14	11-13	10-12
Flash interpoly dielectric thickness control EOT (% 3σ) [K]	<± 7	<± 7	<± 6	<± 6	<± 6	<± 6	<± 6
Flash interpoly dielectric T_{max} of formation $t > 5' / < 5'$ (°C) [L]	800 / 900	800 / 900	750 / 900	750 / 900	700 / 850	700 / 850	650 / 800
Flash interpoly dielectric conformality on floating gate EOT_{min}/EOT_{max} [M]	>0.92	>0.95	>0.95	>0.95	>0.95	>0.98	>0.98
Flash maximum charge loss 10 years @Room Temp (V) – single/dual bit (%) [N]	20 / 10	20 / 10	20 / 10	20 / 10	20 / 10	20 / 10	20 / 10

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow—-Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known



Notes for Table 54a and b:

[A] Flash devices tend to lag the current CMOS technology node. This entry provides the F value for designs in the indicated time period.

[B] Tunnel oxides must be thick enough to assure retention but thin enough to allow ease of erase/write. This difficult problem hinders scaling.

[C] Tunnel oxide thickness control must guarantee correct program/erase windows.

[D] Minimum QBD value (Constant Current Stress) to guarantee device write/erase cycling

[E] Tunnel oxide defectivity to guarantee device write/erase cycling.

[F] Leakage value to guarantee device charge retention.

[G] Between minimum and maximum values of the program/erase distributions.

[H] Time degradation after maximum specification number of write/erase cycles.

[I] Time degradation after maximum specification number of write/erase cycles.

[J] Interpoly dielectric must be thick enough to assure retention but thin enough to assure an almost constant coupling ratio. Charge retention with scaling down is the major issue.

[K] Thickness control to assure correct coupling ratio and minimum thickness for charge retention.

[L] For long (>5 min) and short (<5 min) thermal processes to avoid tunnel oxide and device degradation

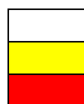
[M] Uniform step coverage is important for charge retention.

[N] To assure device functionality.

Table 54b FLASH Non-volatile Memory Technology Requirements—Long-term

YEAR OF PRODUCTION	2010	2013	2016
DRAM ½ PITCH (nm)	45	32	22
MPU / ASIC ½ PITCH (nm)	50	35	25
MPU PRINTED GATE LENGTH (nm)	25	18	13
MPU PHYSICAL GATE LENGTH (nm)	18	13	9
Flash technology node (nm) [A]	**50**	**35**	**25**
Flash NOR tunnel oxide thickness (EOT - nm) [B]	**8-9**	**8**	**8**
Flash NAND tunnel oxide thickness (EOT - nm) [B]	**6–7**	**6–7**	**6–7**
Flash tunnel oxide thickness control EOT (% 3σ) [C]	**<± 2.5**	**<± 2.5**	**<± 2**
Flash tunnel oxide minimum Q_{BD} @ 1×10^{-2} A/cm^2 (C/cm^2) [D]	**0.3**	**0.3**	**0.4**
Flash tunnel oxide defectivity @ minimum Q_{BD} (def/cm^2) [E]	**<0.01**	**<0.01**	**<0.01**
Flash tunnel low field leakage (nA/5V) [F]	**100**	**100**	**100**
Flash program/erase window ΔV_T (V) [G]	**>3**	**>3**	**>3**
Flash erase time degradation t_{max}/t_0 [H]	**<2**	**<2**	**<2**
Flash program time degradation t_{max}/t_0 [I]	**<2**	**<2**	**<2**
Flash NOR interpoly dielectric thickness (EOT-nm) [J]	**8-10**	**6-8**	**4-6**
Flash NAND interpoly dielectric thickness (EOT-nm) [J]	**10-12**	**9-11**	**9-11**
Flash interpoly dielectric thickness control EOT (% 3σ) [K]	**<± 5**	**<± 5**	**<± 5**
Flash interpoly dielectric T_{max} of formation $t > 5' / < 5'$ (^{o}C) [L]	**600 / 750**	**600 / 700**	**600 / 700**
Flash interpoly dielectric conformality on floating gate EOT_{min}/EOT_{max} [M]	**>0.98**	**>0.98**	**>0.98**
Flash maximum charge loss 10 years @Room Temp (V) – single/dual bit (%) [N]	**20 / 10**	**20 / 10**	**20 / 10**

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow–Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known



FERROELECTRIC RANDOM ACCESS MEMORY (FERAM)

FeRAM is a new addition to the 2001 ITRS, and is the result of collaboration between the FEP and PIDS technology working groups. This roadmap was prepared using information gathered from conference papers and other sources together with current information provided by FeRAM researchers. Review and commentary by ITRS FeRAM experts resulted in the final document.

Historically speaking, FeRAM devices had been proposed much earlier than semiconductor memory devices.[25] At present however, memory capacity is limited to ~1/500 that of commodity DRAM, due to limited ferroelectric film reliability and to difficulties associated with capacitor fabrication. FeRAMs depend substantially on the continued development of materials such as ferroelectric films making the forecasts presented here somewhat speculative. Nevertheless, the roadmap covers the years 2001 to 2016 in order to provide a strategic overview of the technology directions and the challenges that must be overcome. This section deals with 1) *Memory capacity*, 2) *Cell size*, 3) *Ferroelectric materials*, and 4) *Minimum switching charge estimation*.

MEMORY CAPACITY

In the longer term, Memory capacity growth is forecasted to quadruple every three years concurrent with a 0.7× technology node reduction. Currently, FeRAM process technology considerably lags leading edge memory. In the near term therefore (2001–2004), it is forecasted that FeRAM scaling will occur at an accelerated rate, with a yearly quadrupling and concurrent 0.7× reduction. This forecasted scaling rate yields a 64 Mb standard memory capacity in 2004 together with an embedded 16 Mb capacity, manufactured using 180nm technology. Embedded capacity, which is

[25] *J. L. Moll and Y. Tarui, IEEE Trans. Electron Devices, ED10, 338, 1963.*

strongly applications dependent is assumed to be nominally ¼ that of standard memory throughout the entire forecast period. *See FEP Table 55a and b.*

The aforementioned yearly near term scaling pace is interrupted in year 2005, holding the 180nm node, with capacity also holding versus 2004. This anticipated slowing reflects the introduction of a new ferroelectric material and the anticipated greater fabrication complexity. Thereafter, scaling accelerates from the 2004–2006 pause yielding a 256Mb device made with 130nm technology, after which the long term scaling rate (4× capacity, 0.7× scaling every three years) maintains.

CELL SIZE

Two factors should be considered regarding cell size. First is the cell structure. The One Transistor-One Capacitor (1T-1C) cell will appear in 2002, replacing the current 2T-2C cell which is presently needed to ensure stable data read out. It is anticipated that this will reduce cell size by 60%. The 1T-1C configuration is mandatory for the realization of large capacity FeRAM. The second factor is the capacitor structure. The change from the current planar capacitor type to a stack configuration in 2003 is anticipated to result in a further 60% cell size reduction. A *3D type cell* is assumed to appear in 2007 when the normal Stack structure cannot provide the minimum needed switching charge. The different capacitor configurations are shown in the drawing accompanying *Table 55*. The cell size factor "a" of 60 is still very large in 2001 (cell size = aF^2 where F=feature size) . The aforementioned cell structure and capacitor configuration changes are forecasted to reduce the cell "a" factor to ten in years 2005 and 2007 after which the cell size factor is fixed at the value of eight for the period 2010 to 2016. Smaller cell size factors such as 6 may appear following the results of learning experiences with leading DRAM technology.

FERROELECTRIC MATERIALS ALTERNATIVES

There are several ferroelectric materials under evaluation at the present time, and represent the most important attribute of this device. At present there is no clear decisive material choice. Two current materials contenders are PZT, or $Pb(Zr,Ti)O_3$ and SBT, or $SrBi_2Ta_2O_9$. SBT has superior endurance characteristics with a Pt bottom electrode and is more suitable for low voltage operation because of its smaller coercive field (Ec). SBT is therefore favored to replace PZT, which was first used in production. However, compared to PZT, SBT has a smaller switching charge per unit area, Q_{SW}, which is important since it is more difficult to maintain minimum switching charge when scaling. Also, degradation of film characteristics due to processes after the film fabrication may hamper such replacement. It is also reported that PZT has superior imprint characteristics. (Imprint is defined as a resistance to polarization reversal which develops after repeated cycling of the memory capacitor.)

The most important issues with PZT and SBT films are suppression of film deterioration that is attributed to oxygen loss, the achievement of stable data read/write characteristics, and data retention. Process improvements are also required for embedding FeRAM. It is important to avoid high temperature annealing or hydrogen incorporation into ferroelectric films after the oxygen anneal used to crystallization the films. For example, low temperature MOCVD ferroelectric film depositions after metal wiring processes, which avoids high temperature anneals, or hydrogen barrier layers may be used. Also, conductive oxides such as IrO_2 or $SrRuO_3$ (SRO) are often used as capacitor electrodes since their use improves ferroelectric film quality.

Physical Vapor Deposition (PVD) and Chemical Solution Deposition (CSD) including Sol-Gel methods are currently the most commonly used methods for ferroelectric film deposition. However, continued scaling dictates the need to shift to methods with better step coverage such as MOCVD. Etching of capacitor electrodes is very difficult to be done with RIE since the most suitable capacitor electrodes do not have volatile etch products. Therefore sputter etching is widely used. This limits CD control and makes scaling more difficult. A RIE process using a hard mask is thus being developed to overcome this difficulty.

In the year 2005 and beyond, new materials may be introduced to maintain the minimum switching charge by using a material with a larger Q_{SW}. One of the promising materials is BLT or $(Bi,La)_4Ti_3O_{12}$ (2).[26] The use of 3D capacitors having a larger capacitor area in conjunction with conventional ferroelectric materials with smaller Q_{SW}, is another scaling option for years 2005 and beyond.

PZT and SBT are often doped. For instance PZT may be doped with lanthanum, and SBT with niobium. Doping is used to achieve the following film enhancements: leakage current suppression, improved endurance or imprint characteristics, suppression of post process film degradation, and others

[26] *B. H. Park, B. S. Kang, S.D. Bu, T. W. Noh, J. Lee, and W. Jo, 682, Nature, 1999.*

ESTIMATED MINIMUM SWITCHING CHARGE

The estimated minimum switching charge has been derived as follows. The sense amplifier for FeRAM is assumed to be basically the same as that of DRAM. Therefore, the bitline signal voltage was calculated using DRAM data from the 1999 ITRS. These data provide that the capacitance Cs remain constant at 25fF/cell independent of technology node, and the bitline capacitance is 320fF at the 1Gb or 0.18µm node. Based on this data with the further assumption that bitline capacitance is proportional to $F^{2/3}$, where F is the feature size[27] (3) allows for the calculation of $\Delta V_{bitline}$. The $\Delta V_{bitline}$ is about 140mV, and we assume that this is needed for the sense amplifier circuit independent on technology nodes. Multiplying $\Delta V_{bitline}$ (140mV) with $C_{bitline}$ then gives the minimum switching charge.

Dividing the minimum switching charge value derived above by the ferroelectric film switching charge per unit area, Q_{SW}, (assumed to be 20-40 µC/cm^2) then yields the desired capacitor area. If this area is larger than the projected capacitor size, then a 3D capacitor should be adopted. Based on this, a 3D capacitor will be needed by year 2007. (See Table 55). A conventional structure can be used in 2005 on the condition that Q_{SW} is not less than 34.5µC/cm^2. This would probably require the introduction of a new ferroelectric material such as the aforementioned BLT.

The FeRAM forecast of Table 55 is based on these assumptions and calculations. "Red brick walls" begin to appear in 2005 at the earliest, and become more widespread in 2007. The first priority to break through these walls, is the development of highly reliable ferroelectric materials that exhibit negligible post-process degradation.

Table 55a FeRAM Technology Requirements—Near-term

YEAR OF PRODUCTION	2001	2002	2003	2004	2005	2006	2007
DRAM ½ PITCH (nm)	130	115	100	90	80	70	65
MPU / ASIC ½ PITCH (nm)	150	130	107	90	80	70	65
MPU PRINTED GATE LENGTH (nm)	90	75	65	53	45	40	35
MPU PHYSICAL GATE LENGTH (nm)	65	53	45	37	32	28	25
Feature Size (µm): F [A]	0.5	0.35	0.25	0.18	0.18	0.18	0.13
FeRAM Generation (Mass production)[B]							
Standard memory (bit)	1Mb	4Mb	16Mb	64Mb	64Mb	128Mb	256Mb
Embedded memory (Byte)	32KB (256Kb)	128KB (1Mb)	512KB (4Mb)	2MB (16Mb)	2MB (16Mb)	4MB (32Mb)	8MB (64Mb)
Access time (ns) [C]	80	65	55	40	30	30	20
Cycle time (ns) [D]	130	100	80	70	50	50	32
Cell area factor a [E]	60	40	24	16	10	10	10
Cell size (µm^2) [F]	15.000	4.900	1.500	0.518	0.324	0.324	0.169
Total cell area (mm^2) for standard memory [G]	15.73	20.55	25.17	34.79	21.74	43.49	45.37
Total cell area (mm^2) for embedded memory [H]	3.93	5.14	6.29	8.70	5.44	10.87	11.34
Projected capacitor size (µm^2)[I]	2.00	0.98	0.50	0.26	0.13	0.13	0.07
Capacitor area (µm^2) [J]	2.00	0.98	0.50	0.26	0.13	0.13	0.09
Cap area/proj cap size[K]	1.00	1.00	1.00	1.00	1.00	1.00	1.34
Height of bottom electrode/F (for 3D capacitor)[L]	n/a	n/a	n/a	n/a	n/a	n/a	0.17
Capacitor structure [M]	planar	planar	stack	stack	stack	stack	3D
T2C or 1T1C [N]	2T2C	1T1C	1T1C	1T1C	1T1C	1T1C	1T1C
Vop (Volt) [O]	3.0	3.0	2.5	1.8	1.5	1.5	1.2
Minimum switching charge density (µC/cm^2) @Vop[P]	4.4	7.1	11.2	17.2	34.5	34.5	40.0
Minimum switching charge per cell (fC/cell) @Vop[Q]	88.5	69.8	55.8	44.8	44.8	44.8	36.1
Retention @85℃ (Years)[R]	10 Years	10 Years	10 Years	10 Years	10 Years	10 Years	10 Years
Fatigue with assuring retention[S]	1.0 E12	1.0 E13	1.0 E14	1.0 E15	>1.0 E16	>1.0 E16	>1.0 E16

White–Manufacturable Solutions Exist and Are Being Optimized
Yellow–Manufacturable Solutions are Known
Red–Manufacturable Solutions are NOT Known

[27] *A. Nitayama, Y. Kohyama, ad K. Hieda, 355, IEDM 1998.*

Table 55a and b Notes:

[A] Feature size "F" is defined as the critical dimension in the cell.

[B] Embedded Memory (Byte) strongly depends on applications, and assumed to be 1/4 of Standard Memory (bit) here.

[C] ,[D] Values for1Mb are estimated based on Ramtron FM1808 (256kb): 70ns/130ns and Fujitsu's ISSCC 2001 paper (1Mb,accesstime=80ns).

[E] a = Cell size/F^2. Assumptions: planar --> stack (×60%), 2T2C --> 1T1C (×60%).

[F] Cell size = a F^2*

*[G] Cell area*Memory size (bit).*

*[H] Cell area*Memory size (bit) . cf. JIS (Japanese Industrial Standard) demands chip area < 20mm^2 for IC card applications.*

[I] Should be doubled for 2T2C. 2001-2004: 8F^2, 2005-2008: 4F^2, 2011-2017: 3F^2 are assumed.

[J] Should be doubled for 2T2C. 3D will be a pedestal structure.

[K] More than 1 for 3D capacitors, otherwise: 1.

*[L] For instance, 0.17 means that the height is 0.17*F.*

[M] See figures (right).

[N] Besides cell structures, configurations are being investigated; ex. Chain-FeRAM.

[O] Vop=operational voltage. Low voltage operation is a key issue.

[P] This value can be calculated by 17) divided by 10).

*[Q] Calculated by ΔVbitline*Cbitline with the assumptions that ΔVbitline= 140mV is needed and Cbitline is the same as DRAM.*

[R] Depends on applications. 85C comes from the specifications for IC cards.

*[S] 100MHz*10 years=3E+16. Some 1E+15 is required to compete with SRAM and DRAM.*

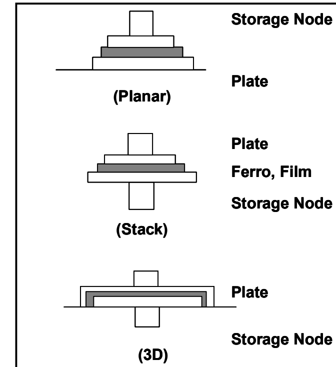


Table 55b FeRAM Technology Requirements—Long-term

YEAR OF PRODUCTION	2010	2013	2016
DRAM ½ PITCH (nm)	45	32	22
MPU / ASIC ½ PITCH (nm)	50	35	25
MPU PRINTED GATE LENGTH (nm)	25	18	13
MPU PHYSICAL GATE LENGTH (nm)	18	13	9
Feature Size (µm): F [A]	0.1	0.07	0.05
FeRAM Generation (Mass production)[B]			
Standard memory (bit)	1Gb	4Gb	16Gb
Embedded memory (Byte)	32MB (256Mb)	128MB (1Gb)	512MB (4Gb)
Access time (ns) [C]	10	8	6
Cycle time (ns) [D]	16	12	10
Cell area factor a [E]	8	8	8
Cell size ($µm^2$) [F]	0.080	0.039	0.020
Total cell area (mm^2) for standard memory [G]	85.90	168.36	343.60
Total cell area (mm^2) for embedded memory [H]	21.47	42.09	85.90
Projected capacitor size ($µm^2$)[I]	0.03	0.015	0.0075
Capacitor area ($µm^2$) [J]	0.08	0.06	0.05
Cap area/proj cap size[K]	2.53	4.06	6.37
Height of bottom electrode/F (for 3D capacitor)[L]	0.57	1.15	2.01
Capacitor structure [M]	3D	3D	3D
T2C or 1T1C [N]	1T1C	1T1C	1T1C
Vop (Volt) [O]	1.0	0.7	0.7
Minimum switching charge density (µC/cm^2) @Vop[P]	40.0	40.0	40.0
Minimum switching charge per cell (fC/cell) @Vop[Q]	30.3	23.9	19.1
Retention @85℃ (Years)[R]	10 Years	10 Years	10 Years
Fatigue with assuring retention[S]	>1.0 E16	>1.0 E16	>1.0 E16

Year of First Product Shipment	2001	2002	2003	2004	2005	2007	2010	2013	2016
Technology Node	130nm	115nm	100nm	90nm	80nm	65nm	45nm	32nm	22nm
Ferroelectric Materials		PZT*, SBT				PZT, SBT, New Materials (BLT, etc.)			
Deposition Methods		PVD, CSD#		PVD, CSD, MOCVD		MOCVD, New Methods			

**) Since PZT contains lead, it may pose a problem from the viewpoint of ESH.*

#) Chemical Solution Deposition

Figure 33 FeRAM Potential Solutions

An endurance of 10^{15} read/write cycles is required to replace other RAMs such as SRAM and DRAM. In order to confirm such endurance values, testing within a practical time period is very critical, since the FeRAM temperature acceleration factor is rather small.

Presently, reliability and cost are the greatest barriers to wide-spread use of FeRAM devices. For these reasons FeRAM applications are limited to portable low-power uses such as IC cards, etc. However, it is important to point out that the FeRAM market can expand if these problems can be overcome. FeRAM could then begin to replace Flash and SRAM , which devices have found use in similar applications segments. It is important to note again the following outstanding features of FeRAM as an optimum memory for multimedia applications.

- Non-volatility
- Low voltage (power) operation
- High speed
- High endurance
- Capacity for high levels of integration. (Cell structure is similar to DRAM.) An encouraging fact is that the storage capacity of commodity Flash memory has dramatically increased and is currently almost equal to that of DRAM. This has occurred because of market demand for large capacity, nonvolatile memory. FeRAM could also satisfy this market demand and therefore could be another Flash. Global efforts by researchers for FeRAM development are highly encouraged.

CONCLUSION

This chapter has reviewed the ITRS Front End Process challenges, requirements, and potential solutions associated with the era of materials limited device scaling. This coming decade is expected to require the development and introduction of a host of new materials and unit processes ranging from the starting substrate materials, and encompassing virtually all facets of front end processing. In contrast with interconnect, where low-κ materials present a great challenge, FEP requires the introduction of a variety of high κ materials for applications as diverse as MOSFET gate dielectric layers, DRAM storage capacitors, and Flash memory tunnel, and interpoly dielectric layers. In addition, the growth of the FeRAM market requires the development and optimization of a broad class of ferroelectric thin film materials.

In addition to the above, the advances in lithography and critical dimension etch technology have resulted in the ability to produce short channel MOSFET devices before the availability of critical gate stack materials needed to fully exploit the benefits of these scaled devices. Accordingly, in addition to the urgent need for the aforementioned high-κ gate dielectric materials, dual metal gate devices are required, and the traditional silicon oxy-nitride gate dielectric layers will be scaled to dimensions well beyond previous expectations, creating concerns about overall device reliability, as well as off-state power consumption. These issues are expected to result in more complex FEP process flows where multiple gate lengths and gate dielectric layer thicknesses will be integrated on the same chip. These issues are also expected to create pressure for technology advances in doping to deal with the greater short channel sensitivity that arises with the extended scaling of the silicon nitride gate dielectric.

Also, as previously discussed, it is expected that non-standard dual gate devices will enter production within the Roadmap time horizon. The manufacture of these devices is expected to have a fundamental impact on the FEP materials, unit processes, and process architectures.

Layered on top of these crucial issues is the need for the next generation silicon starting material, the introduction of which is expected to pose significant cost and technology challenges.

The road to the future for FEP is marked by many barriers and challenges that are significant but not insurmountable. Success will require the appropriate generation of fundamental knowledge about materials and process technologies and the timely conversion of this fundamental knowledge into production materials and process suitable for the cost-effective manufacture of these next generation devices. FEP is indeed a rich area for research and innovation.