International Technology Roadmap for Semiconductors

2003 Edition

Front End Processes

The ITRS is devised and intended for technology assessment only and is without regard to any commercial considerations pertaining to individual products or equipment.

TABLE OF CONTENTS

Scope 1

Difficult Challenges 2

The Front End Processes Grand Challenge—
The FEP Response to the Era of Material-Limited Device Scaling 2

Technology Requirements and Potential Solutions 5

Starting Materials 5

Surface Preparation 17

Thermal/Thin Films and Doping 23

Thermal/Thin Films 23

Doping Technology 34

Difficult Challenges 34

Source and Drain Extensions 34

Contact 35

Gate Electrode 36

Channel 36

Front End Etch Processing 39

DRAM Stacked Capacitor 41

DRAM Trench Capacitor 45

Non-Volatile Memory (FLASH) 47

Ferroelectric Random Access Memory (FeRAM) 50

Memory Capacity 51

Cell Size 51

Ferroelectric Materials Alternatives 51

Estimated Minimum Switching Charge 52

Cross-cut Issues 56

FEP Metrology Cross-cut Issues 56

FEP Modeling and Simulation Cross-cut Issues 56

FEP Environment, Safety, and Health Cross-cut Issues 57

Inter-focus ITWG Discussion 57

Impact of Future Emerging Research Devices 57

Conclusion 58

LIST OF FIGURES

Figure 43 Front End Processes Chapter Scope 1
Figure 44 Starting Materials Potential Solutions 15
Figure 45 Surface Preparation Potential Solutions 22
Figure 46 Thermal/Thin Films Potential Solutions 33
Figure 47 Doping Potential Solutions 38
Figure 48 Front End Processes Etching Potential Solutions 41
Figure 49 DRAM Stacked Capacitor Potential Solutions 45
Figure 50 Minimum Feature Size of NAND Flash Memory 48
Figure 51 Minimum Feature Size of NOR Flash Memory 48
Figure 52 FeRAM Potential Solutions 55

LIST OF TABLES

Table 68a Front End Processes Difficult Challenges—Near-term 3
Table 68b Front End Processes Difficult Challenges—Long-term 4
Table 69a Starting Materials Technology Requirement—Near-term 6
Table 69b Starting Materials Technology Requirements—Long-term 11
Table 70a Surface Preparation Technology Requirements—Near-term 18
Table 70b Surface Preparation Technology Requirements—Long-term 19
Table 71a Thermal and Thin Film, Doping and Etching Technology Requirements—Near-term 24
Table 71b Thermal and Thin Film, Doping and Etching Technology Requirements—Long-term 29
Table 72a DRAM Stacked Capacitor Films Technology Requirements—Near-term 43
Table 72b DRAM Stacked Capacitor Films Technology Requirements—Long-term 44
Table 73a DRAM Trench Capacitor Technology Requirements—Near-term 46
Table 73b DRAM Trench Capacitor Technology Requirements—Long-term 47
Table 74a FLASH Non-volatile Memory Technology Requirements—Near-term 49
Table 74b FLASH Non-volatile Memory Technology Requirements—Long-term 50
Table 75a FeRAM Technology Requirements—Near-term 53
Table 75b FeRAM Technology Requirements—Long-term 55

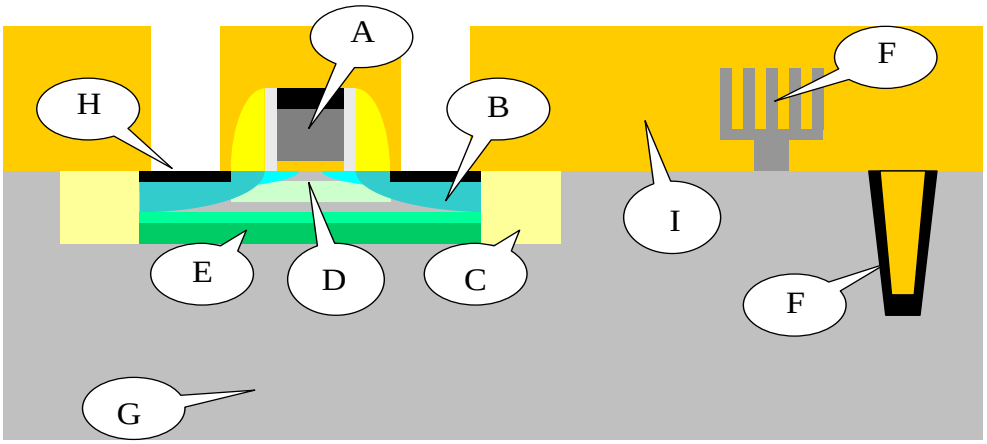
FRONT END PROCESSES

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 tools, and materials, as well as the unit and integrated processes starting with the silicon wafer substrate and extending through the contact silicidation processes. The following specific technology areas are covered: *starting materials*, *surface preparation*, *thermal/thin films, shallow trench isolation (STI), and doping* for MOSFETs, *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. Issues for copper/low-κ dielectrics cleaning and surface preparation, 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.



A: GATE STACK (INCLUDING FLASH) AND SIDEWALL SPACERS
B: SOURCE/DRAIN – CONTACT AND EXTENSION
C: ISOLATION
D: CHANNEL DOPING
E: WELL DOPING
F: DRAM STACK /TRENCH CAPACITOR & FERAM STORAGE STRUCTURES
G: STARTING MATERIAL
H: CONTACTS
I: PRE-METAL DIELECTRIC LAYERS

Figure 43 Front End Processes Chapter Scope

DIFFICULT CHALLENGES

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

MOSFET 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 process 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 bulk 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 such as planar fully depleted SOI devices. An overview of the device alternatives is presented in the *Emerging Research Devices section of the PIDS chapter*. Some believe these may be needed as early as the year 2008. 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 68.

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 65 nm, which were at that time expected to emerge in the year 2005. In the interim, the patterning technology for producing 65 nm gates has accelerated and these have been achieved in 2001. Advances have also been made that permit the extension of silicon oxynitride gate dielectric materials, with the result that the need for high-κ has been delayed. 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 2006. 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 operating and low standby power. Here, it is expected that the high-κ material will be needed by the year 2006 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. Also, near-term measures such as the use of strained silicon channels are expected to provide needed boosts to device speed, but ultimately, scaling is expected to require the replacement of planar CMOS devices with non-standard dual gate devices and/or fully depleted planar devices. The introduction of these devices will require the replacement of bulk silicon substrates with silicon-on-insulator (SOI) substrates.

Table 68a Front End Processes Difficult Challenges—Near-term

MPU/ASIC Physical Gate Length ≥20 nm/ Through 2009	*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 high performance, low operating and low standby power MOSFETs CMOS integration of enhanced channel mobility, e.g., strained layers Control of boron penetration from doped polysilicon gate electrode Minimized depletion of dual-doped polysilicon electrodes Introduction of dual metal gate electrodes with appropriate work function (toward end of period) Metrology issues associated with gate dielectric film thickness and gate stack electrical and materials characterization
Critical dimension and effective channel length (L_{eff}) control	Control of gate etch processes 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 L_{eff} control Maintenance of CD and profile control throughout the transition to new gate stack materials and processes CD and etch metrology Site flatness to ensure effective lithographic printing
Introduction and 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 ~17–33% of ideal channel resistance (=V_{dd}/I_{on}) Control of parasitic capacitance to achieve less than ~23–29% of gate capacitance, consistent with acceptable Ion 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 68b Front End Processes Difficult Challenges—Long-term

MPU/ASIC Physical Gate Length <20 nm/ Beyond 2009	*Summary of 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 CD and L_{eff} control Chemical, electrical, and structural characterization
Introduction and CMOS integration of non-standard, double gate MOSFET devices	Devices are needed starting from 2011 and may be needed as early as 2007 (this is a backup for high-κ materials and metal gates on standard CMOS) 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 Role of SOI utilization (and structural configuration) for advanced non-classical CMOS
Starting silicon material alternatives greater than 300 mm diameter require the start of wafer manufacturing development in year 2003	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 If 450 mm wafers are to become available for production in 2011 as currently forecasted, wafer-manufacturing development should be implemented eight years earlier, e.g., 2003.
New memory storage cells, storage devices, and memory architectures	Scaling of DRAM storage capacitor beyond $6F^2$ Further scaling of Flash memory interpoly and tunnel oxide thickness FeRAM storage cell scaling Introduction of new memory types and storage concepts (Candidates—MRAM, phase-change memory for 2010, and single electron, molecular, nano-floating products beyond 2010)
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 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 and MRAM will make a significant commercial appearance where ferroelectric and ferromagnetic storage materials would be used. The introduction of these diverse materials into the manufacturing mainstream are viewed as important difficult challenges.

In the starting wafer area, it is expected that alternatives to bulk silicon such as various strained silicon alternatives on bulk, as well as silicon-on-insulator substrates will proliferate. These all imply FEP process architecture changes. An important difficult challenge expected to emerge within the Roadmap horizon is the need for the next generation 450 mm silicon substrate. Here, it is questionable whether the incumbent techniques for wafer preparation can be cost-effectively scaled to the next generation. It is also questionable whether this substrate will be bulk silicon or SOI and whether strained silicon will be the required active layer material. The search for potential substrate alternatives presents an important research need that must commence in 2003, if this new substrate material is to be ready for device manufacture in the year 2011.

Front end cleaning processes will be impacted by the introduction of new front end materials. In addition, scaled devices are expected to become increasingly shallow, requiring that cleaning processes become completely benign in terms of substrate material removal and surface roughening. Also, the scaled and new device structures that will be introduced will become increasingly fragile, limiting the aggressiveness of the cleaning processes that may be employed. DRAM stacked and trench storage capacitor structures will show increasing aspect ratios making sidewall contamination removal increasingly difficult.

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 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 most high-κ materials is expected to place new boundaries on thermal budgets associated with dopant activation. In a worst-case scenario, the introduction of these materials could have a significant impact on the overall CMOS process architecture.

TECHNOLOGY REQUIREMENTS AND POTENTIAL SOLUTIONS

STARTING MATERIALS

Technology Requirements—Tables 69a and 69b forecast trends 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 wafer characteristics such as localized light scatterers (LLS) 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 Tables 69a and 69b, but to also display the metrology tool readiness. For this reason, the wafer parameter trend table cells have been annotated with the silicon supplier criticality color code on the left-hand side of each cell and the metrology tool readiness color code depicted on the right-hand side. The marking system and meanings are shown in the tables for both DRAM and high-performance MPUs. This legend applies only to Tables 69a and 69b.

Technology requirements value, supplier manufacturing capability cells and metrology readiness capability cells.	
Technology Requirements Value and Supplier Manufacturing Capability	**Metrology Readiness Capability**
Manufacturable solutions exist, and are being optimized	
Manufacturable solutions are known	
Manufacturable solutions are NOT known	

Wafer Types—Starting materials selection has traditionally involved the choice of either polished Czochralski (CZ) or epitaxial silicon wafers. In addition, silicon-on-insulator (SOI) wafers are evolving to become more than a niche technology. The opportunity for these more costly SOI wafers to be used in mainstream high-volume applications is being driven by process flow simplification in some cases, and by improved high-frequency logic performance, including enhanced device performance via unique device configurations such as the FinFET. Commodity devices such as DRAM are commonly manufactured on lower cost CZ polished wafers, although “crystal originated pits” (COP)-free wafers are increasingly required to avoid interference with inline inspections used for defect reduction. High-performance logic ICs are generally manufactured on more costly epitaxial wafers because their use has facilitated the achievement of greater robustness (e.g., soft error immunity and 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 epitaxial wafers, favors continued use of epitaxial wafers for logic applications. This wide variety of starting materials will likely continue into the foreseeable future and is the reason for inclusion of polished, epitaxial, and SOI wafers in Tables 69a and 69b.

Parameter Values—The wafer requirements have been selected to ensure that in any given year each parameter value contributes no more than 1% to leading-edge chip yield loss, including 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)— taken as the DRAM half-pitch (i.e., the technology generation)— 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 for both the gate dielectric equivalent oxide thickness (EOT) and the device physical channel length, compliance with these model-based values can be very costly and, in some cases, requires re-examination. For this reason, detailed re-assessment of the costs incurred versus the value derived from achieving compliance often suggests limiting the scope of these models via appropriate truncation.

Table 69a Starting Materials Technology Requirement—Near-term

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*Driver*
Technology Node		*hp90*			*hp65*			
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*	*D ½*
MPU/ASIC Physical Gate Length (nm)	**45**	**38**	**32**	**28**	**25**	**23**	**20**	*M*
DRAM Total Chip Area (mm^2)	**139**	**110**	**82**	**122**	**97**	**131**	**104**	*D ½*
DRAM Active Transistor Area (mm^2)	**36.7**	**29.1**	**22.7**	**35.5**	**28.2**	**43.3**	**34.3**	*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)	**32.6**	**32.1**	**31.7**	**31.7**	**31.7**	**31.7**	**31.7**	*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)	**2**	**2**	**2**	**2**	**2**	**2**	**2**	*D ½, M*
Front surface particle size (nm), latex sphere equivalent [D]	**≥90**	**≥90**	**≥90**	**≥90**	**≥90**	**≥90**	**≥90**	*D ½, M*
Particles (cm^{-2}) [E]	**≤0.35**	**≤0.35**	**≤0.35**	**≤0.18**	**≤0.18**	**≤0.09**	**≤0.09**	*D ½*
Particles (#/wf)	**≤238**	**≤238**	**≤241**	**≤123**	**≤123**	**≤63**	**≤63**	*D ½*
Site flatness (nm), SFQR 26 mm × 8 mm site size [F, R]	**≤101**	**≤90**	**≤80**	**≤71**	**≤64**	**≤57**	**≤51**	*D ½, M*
Nanotopography, p-v, 2 mm diameter analysis area [Q]	**≤25**	**≤23**	**≤20**	**≤18**	**≤16**	**≤14**	**≤13**	*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]								
Oxidation stacking faults (OSF) (DRAM) (cm^{-2}) [G]	**≤1.9**	**≤1.6**	**≤1.4**	**≤1.2**	**≤1.0**	**≤0.8**	**≤0.7**	*D ½*
Oxidation stacking faults (OSF) (MPU) (cm^{-2}) [G]	**≤0.6**	**≤0.5**	**≤0.4**	**≤0.3**	**≤0.3**	**≤0.2**	**≤0.2**	*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) [H, I]								
Large structural epi defects (DRAM) (cm^{-2}) [J]	**≤0.007**	**≤0.009**	**≤0.012**	**≤0.008**	**≤0.010**	**≤0.008**	**≤0.010**	*D ½*
Large structural epi defects (MPU) (cm^{-2}) [J]	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	*M*
Small structural epi defects (DRAM) (cm^{-2}) [K]	**≤0.014**	**≤0.018**	**≤0.024**	**≤0.017**	**≤0.021**	**≤0.015**	**≤0.019**	*D ½*
Small structural epi defects (MPU) (cm^{-2}) [K]	**≤0.006**	**≤0.006**	**≤0.006**	**≤0.006**	**≤0.006**	**≤0.006**	**≤0.006**	*M*

Technology requirements value, supplier manufacturing capability cells and metrology readiness capability cells.	
Technology Requirements Value and Supplier Manufacturing Capability	**Metrology Readiness Capability**
Manufacturable solutions exist, and are being optimized	
Manufacturable solutions are known	
Manufacturable solutions are NOT known	

Table 69a Starting Materials Technology Requirement—Near-term (continued)

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*Driver*
Technology Node		*hp90*			*hp65*			
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*	*D ½*
MPU/ASIC Physical Gate Length (nm)	**45**	**38**	**32**	**28**	**25**	**23**	**20**	*M*
Silicon-On-Insulator Wafer (99% Chip Yield)[R]*								
Starting silicon layer thickness (Partially Depleted) (tolerance ± 5%, 3σ) (nm) [L]	**78–133**	**67–115**	**58–100**	**53–91**	**48–83**	**44–76**	**40–70**	*M*
Starting silicon layer thickness (Fully Depleted) (tolerance ± 5%, 3σ) (nm) [M]	**24–43**	**21–39**	**20–36**	**19–34**	**18–33**	**17–31**	**16–30**	*M*
Buried oxide (BOX) thickness (Fully Depleted) (tolerance ± 5%, 3σ) (nm) [N]	**68–112**	**56–94**	**48–80**	**42–70**	**38–64**	**34–56**	**30–50**	*M*
D_{LASOI}, Large area SOI wafer defects (DRAM) (cm^{-2}) [O]	**≤0.007**	**≤0.009**	**≤0.012**	**≤0.008**	**≤0.010**	**≤0.008**	**≤0.010**	*D ½*
D_{LASOI}, Large area SOI wafer defects (MPU) (cm^{-2}) [O]	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	*M*
D_{SASOI}, Small area SOI wafer defects (DRAM) (cm^{-2}) [P]	**≤0.137**	**≤0.173**	**≤0.221**	**≤0.142**	**≤0.178**	**≤0.116**	**≤0.146**	*D ½*
D_{SASOI}, Small area SOI wafer defects (MPU) (cm^{-2}) [P]	**≤0.154**	**≤0.156**	**≤0.159**	**≤0.159**	**≤0.159**	**≤0.159**	**≤0.159**	*M*

** Parameters define limit values; independent predictors of yield, mathematically or empirically modeled at 99%. 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%.*

*** Significant gaps in metrology and wafer manufacturing equipment need to be closed on 200 mm, especially for the 90 nm node, inasmuch as 300 mm is only now being phased in and 200 mm will still be prevalent through the 90 nm node.*

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

Technology requirements value, supplier manufacturing capability cells and metrology readiness capability cells.	
Technology Requirements Value and Supplier Manufacturing Capability	**Metrology Readiness Capability**
Manufacturable solutions exist, and are being optimized	
Manufacturable solutions are known	
Manufacturable solutions are NOT known	

Tables 69a and 69b notes:

[A] Surface metals are empirically grouped into three classes [1 2]:(a) Mobile metals that 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 that dissolve in silicon or form silicides such as Fe, Ni, Cu, Cr, Co; Al, Zn; and (c) major gate-oxide-integrity (GOI) killers such as Ca. Each of these metals is taken at a maximum value of $1 \times 10^{10}/cm^2$ for all subsequent technology generations. The surface concentration of carbon atoms after cleaning is based on the assumption that a 10% (7.3×10^{13} atoms/cm^2) carbon atom coverage on a bare silicon (100) surface can be tolerated during device fabrication. Organics/polymers are therefore modeled approximately 0.1 of a monolayer, $\leq 1 \times 10^{14}$C atoms/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.

Total bulk Fe consistent with recombination lifetime, τ_r, as measured by the SPV technique (for lightly doped p-type material) at low injection level is held at $1 \times 10^{10}/cm^3$ for all subsequent technology generations.[3] 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[4]. 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.

---

[1] *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).*
[2] *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).*
[3] *G. Zoth and W. Bergholtz, "A Fast, Preparation-Free Method to Detect Iron in Silicon," J. Appl. Phys., 67, 6764–6771, (1990).*
[4] *W. Shockley, Electrons and Holes in Semiconductors, p.69, Princeton: D. Van Nostrand Co., Inc. 1950.*

[B] Instrumentation choice, target values, and spatial frequency range (scan size) for front-surface microroughness are selected based on application. Power spectral density analysis is recommended to utilize full accessible range of instruments. A typical value for polished wafers is ≤ 0.10 nm (RMS) for all CD generations. Epitaxial, annealed and SOI wafers have values that are typically higher than polished wafers while still meeting the user's requirement.

[C] The oxygen concentration may be specified depending on the particulars of the IC user based on IC process requirements and is generally in the range of 18–31 ppma (SEMI M44-0702, refer to ASTM F121-79).[5] With advanced crystal growth technologies, bulk micro defects (BMDs) can be achieved independent of the interstitial oxygen concentration. The importance of BMDs for gettering has recently again been emphasized and may be especially important in those IC fabrication cases with low thermal budgets.[6] Co-doping techniques (such as nitrogen and carbon) can be used to enhance oxygen precipitation so may be particularly well suited for low thermal budget device processes. Additionally, certain growth methods coupled with heat treatments can also enhance the precipitation of oxygen. Not all device processes, however, require the presence of BMDs. BMDs for internally gettered polished wafers may be generically taken as greater than approximately $1 \times 10^8/cm^3$ after IC processing. BMD density is measured using ASTM F-1239.

[D] Critical front surface particle size = K_1F, [K_1=1] where F is the DRAM half-pitch and is used to calculate required particle densities at the given technology generation. Particle sizes reported in Tables 69a and 69b, however, are held constant at 90 nm for technology generations 90 nm through 45 nm, due to metrology capability. Particle densities are extracted from the conventional Maly Yield Equation [7] {$Y = \exp[-(D_P R_P) A_{eff}]$, where A_{eff}, is the effective chip area $A_{eff}=2.5*F^2T+(1-aF^2T/A_{chip})A_{chip}*0.18$, "a" is the DRAM cell fill factor (see Table 72a) and T = number of transistors or bits/chip per technology generation}. The kill factor R_P is assumed to be 0.2, although the kill factor may be very dependent upon the specifics of the DRAM fab. The relationship between actual defect size and associated LSE (latex sphere equivalent) size depends on defect type and scanner geometry. The current particle size threshold capability for SOI wafers is 120–150 nm, due to the altered response in the optical metrology tools, compared to polished or epitaxial wafers.

[E] Detailed back-surface particle information is not included in Table 69, since, in practice, lithography concerns are being met by identifying these defects visually. This perhaps suggests 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) (F) + (0.4/0.6) (T)], where F 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: $Y = \exp 9(-D_P R_P A_{eff})$ [5]. $R_P = 1.0$, $A_{eff} = A_{chip} \times 0.03 \times 0.8$, 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 backside particles. The equation for the "killer" backside particle diameter strongly depends on two assumptions that 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 maybe 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 backside 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 that can be accommodated by a scanning stepper leveling system. This gives rise to the 80% effective degraded area.

[F] The metric for site flatness should be matched to the type of exposure equipment used in leading edge applications, which implies scanning steppers for critical levels. While SFSR may be the most appropriate metric, it has failed to gain appreciable support in the industry. Historical reference to SFQR remains strong and it appears inevitable that this metric will continue to be used in the future. To more closely emulate practical experience of the scanning stepper, the effective site size for local site flatness is being modified to 26 mm × 8 mm accordingly. Full-field steppers with square fields (nominally 22 × 22 mm) may still be utilized for non-critical levels although these are increasingly being phased out. In either case, the metric value is approximately equal to F for dense lines (DRAM half pitch). Partial sites should be included. Also note that flatness metrology requires sufficient spatial resolution to capture topographical features relevant for each technology node.

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

[H] Other epitaxial defects such as hillocks and mounds should also be accounted for, but an appropriate yield model is not available. Accurate segregation based upon defect morphology is also not generally available with today's metrology.

---

[5] *n.a., SEMI M44-0702, "Guide to Conversion Factors for Interstitial Oxygen in Silicon", SEMI (Semiconductor Equipment and Materials International), San Jose, CA, July 2002.*

[6] *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).*

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

[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.*

[I] Desired epitaxial layer thickness tolerance is ±4% for a 2 to 10 mm center-point epitaxial layer thickness target value but may be affected on p/p^+ structures due to lack of autodoping suppression via backside film deposition, resulting from incompatibility with 300 mm wafers. 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^-.

[J] Large structural epi defects (large area defects >1 μm LSE signal) modeled at 99% yield where $Y = \exp(-D_{LAD} R_{LAD} A_{chip})$,[6] 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.

[K] Small structural epi defects (≤1μm LSE signal) modeled at 99% yield where $Y = \exp(-D_{SF} R_{SF} A_{chip})$[6] 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.

[L] The silicon final device layer thickness (partially depleted) is obtained by 2 × 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 Si consumption during device fabrication. In the table, the starting material layer thickness is shown. For years 2003 to 2009 it is obtained by adding 10 nm to the lower value in the range and 20 nm to the higher value in the range. After 2009, 10 nm is added to both values in the range in order to translate the device thickness into the starting material thickness. Si loss depends on processing conditions used—it is assumed here that processing parameters are controlled more tightly after 2009.

[M] The silicon final device thickness (fully depleted) is obtained by 0.4 × 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. In the table the starting material layer thickness is shown. For years 2003 to 2009 it is obtained by adding 10 nm to the lower value in the range and 20 nm to the higher value in the range.

[N] The BOX thickness for fully depleted devices is taken as the 2 × MPU physical gate length. BOX scales with gate length to help to control short channel effects and heat dissipation. Range in nominal target value of ±25% allows for trade-off 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 timeframe of partially depleted SOI devices.

[O] Large area SOI (LASOI) wafer defects with yield of 99%; $Y = \exp(-D_{LASOI} R_{LASOI} A_{chip})$,[6] D_{LASOI} = LASOI defect density, $R_{LASOI} = 1.0$ (best present estimate).

[P] Small area SOI (SASOI) wafer defects with yield of 99%; $Y = \exp(-D_{SASOI} R_{SASOI} A_{eff})$,[6] 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.[9 10 11]

[Q] Peak-to-valley threshold, 2 mm diameter analysis area. Maximum p-v reading taken as CD/4, based on extrapolation of wafer supplier process capability for 180–90 nm technology generations, plus published data on linewidth distortion for sub-100 nm critical dimensions.

---

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

[10] *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).*

[11] *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).*

[R] The magnitude of within-wafer variation of various wafer parameters changes over different length scales, depending on the nature of the mechanisms that produce them. The impact on subsequent device manufacturing caused by these variations, which occur at different spatial wavelengths, also depends on the nature of the fab processes and resulting devices. For instance, parameters governed by gas flow and temperature gradient variations, such as CVD layer thickness, typically vary appreciably only over fairly long distances, of the order of millimeters to centimeters. It is often adequate to measure such slowly varying parameters at only a modest number of locations on the wafer, using a metrology tool with a fairly low spatial resolution, in order to control such processes. Other parameters, such as wafer surface topography, vary on multiple length scales with different impacts in the fab. At very large length scale (tens of centimeters), wafer surface height variations are many microns in magnitude (e.g. bow and warp), and can affect various mechanical properties of the wafer. At length scales on the order of one centimeter, the surface variations are fractions of a micron in height. These variations (i.e. site flatness) generally are not critical to mechanical shape of the wafer, but are vital to depth of focus in lithography. At still smaller length scales of a few millimeters or less, the surface height variations are on the order of tens of nanometers high. They do not cause focus failures in lithography, but can produce line width variations in gate lengths and polishing removal uniformity problems in CMP. On the length scale of microns, surface roughness variations are of the order of Angstroms, but can cause gate oxide integrity problems. As another example, in fully depleted SOI wafers, thickness variations of the top silicon layer can cause transistor threshold voltage variation die-to-die (at centimeter length scale), within-die (at millimeter length scale), and conceivably, even transistor-to-transistor (on a sub-micron scale). To control parameter variations across a large range of spatial wavelengths requires a tool capable of measuring the whole wafer to capture long wavelength components, but with a very high density of data points (with correspondingly small sampled area) to capture small wavelength components. The spatial wavelength requirements thus have a profound effect on metrology capability. Methods that work well at long spatial wavelengths may become unsuitable at small spatial wavelengths due to measurement throughput limitations and/or inadequate spatial resolution. Metrology grades in this table reflect current spatial wavelength requirements. Future process and device developments that demand measurement at shorter spatial wavelengths may alter these capability grades in unforeseen ways.

Table 69b Starting Materials Technology Requirements—Long-term

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*	*Driver*
Technology Node	*hp45*		*hp32*		*hp22*		
DRAM ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*	*D ½*
MPU/ASIC Physical Gate Length (nm)	**18**	**14**	**13**	**10**	**9**	**7**	*M*
DRAM Total Chip Area (mm^2)	**83**	**104**	**83**	**104**	**138**	**87**	*D ½*
DRAM Active Transistor Area (mm^2)	**27.3**	**34.3**	**27.3**	**34.3**	**52.7**	**27.7**	*D ½*
MPU High-Performance Total Chip Area (mm^2)	**310**	**310**	**310**	**310**	**310**	**310**	*M*
MPU High-Performance Active Transistor Area (mm^2)	**31.7**	**31.7**	**31.7**	**31.7**	**31.7**	**31.7**	*M*
*General Characteristics * (99% Chip Yield) [A, B, C]*							
Wafer diameter (mm)	**300**	**300**	**300**	**450**	**450**	**450**	***D ½,M***
Edge exclusion (mm)	**2**	**2**	**2**	**2**	**2**	**2**	***D ½, M***
Front surface particle size (nm), latex sphere equivalent [D]	**≥90**	**≥65**	**≥65**	**≥65**	**≥45**	**≥45**	***D ½,M***
Particles (cm^{-2}) [E]	**≤0.09**	**≤0.09**	**≤0.09**	**≤0.04**	**≤0.05**	**≤0.06**	***D ½,M***
Particles (#/wf)	**≤63**	**≤61**	**≤61**	**≤69**	**≤75**	**≤89**	***D ½,M***
Site flatness (nm), SFQR 26 mm × 8 mm site size [F, R]	**≤45**	**≤36**	**≤32**	**≤25**	**≤23**	**≤18**	***D ½,M***
Nanotopography, p-v, 2 mm diameter analysis area [Q]	**≤11**	**≤9**	**≤8**	**≤6**	**≤6**	**≤4**	***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]							
Oxidation stacking faults (OSF) (DRAM) (cm^{-2}) [G]	**≤0.6**	**≤0.4**	**≤0.4**	**≤0.3**	**≤0.2**	**≤0.2**	*D ½*
Oxidation stacking faults (OSF) (MPU) (cm^{-2}) [G]	**≤0.2**	**≤0.1**	**≤0.1**	**≤0.07**	**≤0.06**	**0.04**	*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) [H, I]							
Large structural epi defects (DRAM) (cm^{-2}) [J]	**≤0.012**	**≤0.010**	**≤0.012**	**≤0.010**	**≤0.007**	**≤0.012**	*D ½*
Large structural epi defects (MPU) (cm^{-2}) [J]	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	*M*
Small structural epi defects (DRAM) (cm^{-2}) [K]	**≤0.024**	**≤0.019**	**≤0.024**	**≤0.019**	**≤0.015**	**≤0.023**	*D ½*
Small structural epi defects (MPU) (cm^{-2}) [K]	**≤0.006**	**≤0.006**	**≤0.006**	**≤0.006**	**≤0.006**	**≤0.006**	***M***
Silicon-on-Insulator Wafer (99% Chip Yield) [R]*							
Wafer diameter (mm)	**300**	**300**	**300**	**450**	**450**	**450**	*D ½, M*
Starting silicon layer thickness (Partially Depleted) (tolerance ± 5%, 3σ) (nm) [L]	**37–55**	**31–45**	**29–42**	**25–35**	**23–32**	**21–28**	*M*
Starting silicon layer thickness (Fully Depleted) (tolerance ± 5%, 3σ) (nm) [M]	**15–19**	**14–17**	**14–16**	**13–15**	**13–14**	**12–14**	*M*
Buried oxide (BOX) thickness (Fully Depleted) (tolerance ± 5%, 3σ) (nm) [N]	**26–44**	**22–36**	**18–32**	**16–26**	**14–22**	**10–18**	*M*
D_{LASOI}, Large area SOI wafer defects (DRAM) (cm^{-2}) [O]	**≤0.012**	**≤0.010**	**≤0.012**	**≤0.010**	**≤0.007**	**≤ 0.012**	*D ½*
D_{LASOI}, Large area SOI wafer defects (MPU) (cm^{-2}) [O]	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤0.003**	**≤ 0.003**	*M*
D_{SASOI}, Small area SOI wafer defects (DRAM) (cm^{-2}) [P]	**≤0.184**	**≤0.146**	**≤0.184**	**≤0.146**	**≤0.095**	**≤0.181**	*D ½*
D_{SASOI}, Small area SOI wafer defects (MPU) (cm^{-2}) [P]	**≤0.159**	**≤0.159**	**≤0.159**	**≤0.159**	**≤0.159**	**≤0.159**	*M*

**Parameters define limit values; independent predictors of yield, mathematically or empirically modeled at 99%. 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%.*

Technology requirements value, supplier manufacturing capability cells and metrology readiness capability cells.	
Technology Requirements Value and Supplier Manufacturing Capability	**Metrology Readiness Capability**
Manufacturable solutions exist, and are being optimized	
Manufacturable solutions are known	
Manufacturable solutions are NOT known	

Model Limitations—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 lognormal distribution, in which the logarithms of the parameter values are normally distributed. The latter distribution is skewed with a long tail 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 lognormal 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 more than forty years of IC manufacturing. As a result, wafer standards are usually expressed as "goal-post" specifications (maximum/minimum values allowed) that invariably increase the silicon supplier manufacturing and inspection costs and ultimately, the IC wafer user cost of ownership (CoO). 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. Until these ideals can be achieved, however, Poisson distribution yield models based on the best available information are used and parameter limits assigned based on a 99% yield requirement for that parameter. 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.

Cost of Ownership—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 defined by acceptable metrology practice, 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 70a and 70b) 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 requested by the IC manufacturer (hydrophilic versus hydrophobic) 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. Further emphasis on the CoO has been ascertained by developing a model examining the viability of a 100% wafer inspection to a particular parameter (i.e., site flatness). This model considers the additional wafer supplier's cost of ensuring 100% compliance to the IC manufacturer's specification relative to the potential loss associated with processing a die with a high probability of failing if a 100% inspection is not done. This methodology has been submitted to an appropriate trade journal, with a listing of the internet addresses whereby the relevant worksheets can be accessed so that each IC manufacturer can analyze the trade-off appropriate for their wafer specifications and product family of interest. *Link to the models paper and the excel spread sheet.*

Wafer Parameter Selection—Both the chemical nature and the physical structure of the wafer front surface are of critical concern, and wafer parameters related to both are included in Tables 69a and 69b. Chemical defects include metal and organic particles and surface chemical residues. These defects are equally significant for polished, epitaxial, and SOI

wafers, although there is some concern that the detrimental effects of surface metals may be magnified in ultra-thin SOI films when the metals diffuse into a small silicon volume. Organic contamination strongly depends on environmental conditions during wafer storage and transportation, and accordingly is not included in Tables 69a and 69b, although a footnote lists a suggested value.

Attention is also being given to particles on the back surface of the wafer because wafers are increasingly being double-side polished to improve both the chemical and physical characteristics of the wafer. 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 backside particles does not currently appear to be of significance and, again, has not been included in this edition of the ITRS. In addition, any back-surface treatments may degrade the quality of both the polished back- and front-surfaces (e.g., extrinsic gettering and oxide back seal.)

Important physical characteristics of the wafer front surface include wafer topography and structural defects. Wafer topography encompasses various wafer shape categories that are classified according to their spatial frequency, including site flatness, surface waviness, nanotopography, and surface micro-roughness. Front surface site flatness and nanotopography are believed to be the most critical of the topographic parameters and are addressed in this ITRS revision. Back surface topography also has begun to receive attention recently, particularly in view of possible wafer interactions with stepper chucks, but the technology for quantifying this interaction is still in its infancy and this parameter is not included in the 2003 Tables.

Polished, epitaxial, and SOI wafers all exhibit specific structural defects that must be controlled. Polished wafer defects include grown-in microdefects, such as (COPs) and bulk microdefects (BMDs). With advanced crystal growth techniques, BMDs can be controlled independently of the interstitial oxygen concentration. Nevertheless, the formation of uncontrolled SiOx precipitates in polished or epitaxial wafers is likely to result in excessive device leakage current, necessitating greater attention to internal gettering in polished and, to some extent, in epitaxial wafers. The magnitude and uniformity requirements will impose stringent bulk defect control to achieve homogeneous internal gettering capabilities.

The dependence of gate dielectric integrity and other yield detractors on crystal growth parameters as well as the related role of point defects and agglomerates have been extensively documented. The resulting defect density (D_0) parameter has served effectively as a measure of material quality for several device generations. However, for devices with EOT <2 nm, this parameter is no longer an indicator of device yield and performance and has accordingly been deleted from Tables 69a and 69b as a requirement. It should be noted, however, that starting material cleanliness requirements might change if pre- and post-gate surface preparation methods are modified when high-κ gate dielectric materials are introduced (see Surface Preparation section.).

Epitaxial and SOI wafer defects include large structural defects (>1 µm) and small structural defects (<1 µm). For this reason, starting material requirements are expressed in terms of specific types of surface defects for different wafer types. Epitaxial and SOI materials appear to exhibit few surface defects 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 benefits in device yield. The removal and prevention of surface defects continues to be a state-of-the-art challenge for silicon wafer technology. The relative immaturity of SOI materials, compared to bulk and epitaxial, leads to an additional challenge for the understanding of SOI-specific defects and their impact on device performance and yield in a production environment. The learning achieved since the 1999 ITRS edition 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 defects in SOI wafers continues to be 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 deleted as a separate SOI parameter in this edition of the ITRS due to their substantially lower impact and kill rate. 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.

Potential Solutions—Figure 44 lists the most significant starting materials challenges and shows potential solutions that have been identified, along with the necessary timing for development of these solutions and their transfer into high-volume production. *Link to the Starting Materials Roll Up of Difficult Challenges (near and long term)*. In alignment with Tables 69a and 69b, Figure 44 reflects the requirements of leading edge DRAMs and high-performance MPUs, built on 300 mm or larger diameter wafers. It should be noted, however, that the adoption of 300 mm wafers has not been as rapid as expected, causing 200 mm wafers to be utilized for technology generations beyond what was originally planned. This has resulted in an expectation that 200 mm wafers should meet the parameter requirements originally forecast for 300 mm wafers. Since 300 mm wafers have the advantage of being produced with double-sided polish and other processing improvements, while 200 mm wafers generally do not, improving the quality of 200 mm wafers to an equivalent level presents a significant challenge to the wafer suppliers.

Metrology for SOI wafers is a significant challenge. The particle 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 120-150 nm as compared to 90 nm for the 90 nm technology generation for polished or epitaxial wafers in a production fab. Accordingly, the metrology readiness grade for SOI wafer particles is yellow through the 90 nm technology generation turning to red beyond this node. The anticipated shift from capacitive to optical measurement of wafer site flatness at and beyond the 90 nm technology generation appears to have been resolved for site flatness measurements. 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 the measurement of electrical properties and structural defects in SOI materials. Finally, the metrology issues for the various strained silicon configurations (spatially varying Si:Ge composition content, threading dislocations and associated defects as well as unique surface roughness issues) will require significant attentiveness (see Emerging Materials section below).

Layer thickness and uniformity are included in Tables 69a and 69b for both epitaxial and SOI wafers. For SOI wafers, the broad variety of today's IC applications such as microprocessors, 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 now more fully entering into production to serve this range of SOI applications. The BOX thickness has been doubled for this edition of the ITRS based on leading edge research, although the anticipated decreased thermal dissipation should be taken into account. The 2001 ITRS revision listed the final SOI device thickness (after fab processing), but the 2003 revision gives the incoming silicon thickness for both partially depleted (PD) and fully depleted (FD) devices, with the relevant device thickness relegated to the appropriate footnote.

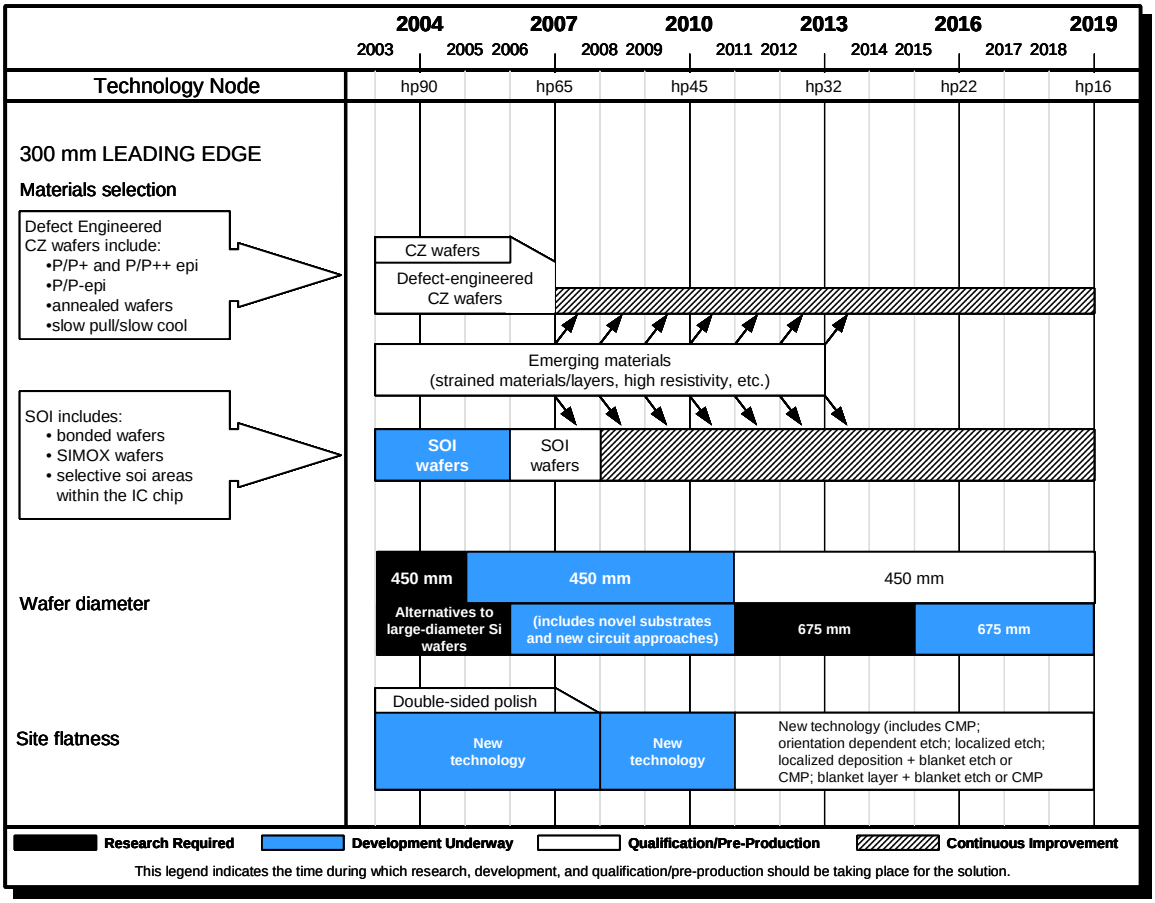


Figure 44 Starting Materials Potential Solutions

Material Selection—Material Selection is the first category of starting materials challenges listed in Figure 44, and the first type of wafer listed is polished CZ substrates, which are produced from crystals grown by the CZ method without any attempt to reduce or eliminate COP formation during crystal growth, and without any further attempt to remove COPs before the wafers are shipped to the customer. These wafers are used today predominantly by DRAM manufacturers, but such usage is increasingly threatened by the difficulty of performing meaningful inspections for yield enhancement, due to the presence of a large number of COPs. As a result, these wafers are expected to largely cease being used for DRAMs and high-performance MPUs at about the 70 nm technology generation, and defect-engineered wafers will take their place. Defect-engineered wafers include several varieties of both polished and annealed wafers. For polished wafers, the modifications used to reduce or eliminate structural defects can be accomplished either during the crystal-pulling process, or during the subsequent wafer finishing operations. An example of the former is wafers prepared from specially grown, low-defect crystals; and examples of the latter include surface-enhanced, hydrogen or argon-heat treated wafers. In addition, high-resistivity wafers (~100 Ohm-cm) are gaining acceptance for analog applications and strained silicon and related materials are under consideration for enhanced device performance *(see Emerging Research Devices section of the PIDS chapter).*

For logic 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 has been considered useful for the p/p^+ and 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 in selected regions) 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 increased 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, the role of oxygen and BMDs must be reassessed due to the absence of solubility-

enhanced gettering. 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 for the case where the STI depth is greater than the epitaxial layer thickness. On the other hand, BMD control independent of the interstitial oxygen concentration may be expected to become even more significant.

Defect-engineered wafers may be supplanted in high-performance applications by SOI wafers if SOI is found to be the best solution for scaling issues with classical planar CMOS on epitaxial or polished wafers. Although current SOI manufacturing techniques result in wafers that are more costly than polished or epitaxial wafers, SOI may become the dominant wafer type because it offers a solution to CMOS scaling issues (short-channel effects) by serving as an acceptable substrate for non-classical CMOS configurations such as FinFET and other multi-gate structures. At present it appears likely that SOI will be required at or before the 45 nm technology generation, for high-performance MPUs and devices with similar performance requirements to remain on Moore's Law. However, the requirement for SOI wafers is not as clear as for the other types of IC devices such as various types of semiconductor memory and random logic circuits, so defect-engineered wafers are still shown to persist until the end of the Roadmap.

Emerging Materials—The utilization of Si-based materials solutions that extend the Starting Materials section of the ITRS (i.e., Moore's Law) have become critically important to the future of the silicon industry. Several examples include strained device channels, high-resistivity silicon, isotopically engineered silicon, and monolithically integrated optoelectronics with silicon. The utilization of a strained device channel for enhanced device performance, for example, has been discussed in the PIDS chapter of the ITRS. Some of these technologies have the flexibility to be contained in several embodiments. For example, at least three major strained silicon structures are considered at the present time. These include bulk strained silicon (epitaxial strained silicon on a uniform content Si:Ge layer, whose lattice constant is engineered by a compositionally graded Si:Ge layer) and strained Si on Si:Ge on insulator (SGOI), where instead of a compositionally graded Si:Ge layer, an insulating layer resides (both partially depleted (PD) and fully depleted (FD) structures can be fabricated). Additional configurations include strained silicon on insulator (SSOI), excluding the Si:Ge layer for strain generation, again with PD and FD configurations. These emerging technologies, although apparently providing technical solutions to critical device issues to ensure the continuance of Moore's law, lack the maturity to include detailed specifications in Tables 69a and 69b for this year's ITRS revision. Further, since some of these technologies are applicable only to SOI structures, and other of these technologies are suitable for use on bulk polished or epitaxial wafers, Figure 44 shows these Emerging Materials as a potential solution for problems confronting both SOI and defect-engineered CZ wafers. *Link to "Emerging Materials" details.*

Large Diameter Wafer—The conversion to 300 mm diameter wafers that began slowly in 1999 is expected to accelerate in the next several years, driven by the need to continuously enhance IC manufacturing productivity. International cooperation and standardization of wafers, carriers, and factory protocols were critical for enhancing 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 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 2011, suggesting 450 mm research and development to be implemented by 2005. Previous research by the silicon supplier community in 400 mm wafers in the mid-late 1990s is expected to be quite useful in enhancing this initiative. 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. Indeed, new measures of fab productivity appear to be warranted. 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 are being aggressively pursued throughout the industry, (including the upgrading of 200 mm process technologies, single-wafer processing as *the* fundamental process change (rather than increased wafer diameter) as well as unique, non-classical device configurations (*see the PIDS chapter*).) Utilization of the industry economic and productivity models developed at International SEMATECH are also being extensively scrutinized for their applicability to a number of fabs.

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 continues to be the escalation 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.

Site Flatness—Unlike materials selection issues and wafer diameter, a lack of significant progress in making site flatness improvements is not likely to cause problems in maintaining Moore's Law, but site flatness is a major concern for semiconductor IC manufacturers as well as for silicon wafer manufacturers, so the Starting Materials Sub-TWG has included it in Figure 44.

The industry made a substantial gain in site flatness process capability by going to double-sided polish for 300 mm wafers. Incremental improvements on this basic gain are expected to satisfy IC manufacturers' requirements to approximately the 65 nm technology generation. Continued improvement beyond that point may require the implementation of new flatness-improvement technologies, including those discussed in Figure 44 and its accompanying text.

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 front-end surface preparation, research and development have historically focused on maximizing the quality of the gate dielectric. This focus continues as the industry moves toward high-κ gate dielectrics. However, another focus for surface preparation is emerging, which is centered on cleaning after the increasing number of implant masks required for all of the different transistor types that are now being integrated onto a single chip. Metal gate electrode materials as well as new transistor structures will eventually drive new requirements for front-end surface preparation.

Technology requirements for surface preparation are shown in Tables 70a and 70b; more details for the data are available in the *supplemental material*. Front-end predictions continue to be problematic due to the lack of data associated with future dielectric and gate electrode materials and their properties. Despite these uncertainties, it is clear that particulate contamination will continue to be a concern at increasingly demanding levels. Requirements for silicon and oxide loss are new additions in 2003. Control of particulates will become more challenging as the need to minimize oxide and silicon loss becomes more critical. Control of particle levels without damage to structures or etching of material is seen as a formidable challenge. However, devices in 2008 will probably be fully depleted, fabricated on SOI substrates with high-κ gate dielectrics and metal gate electrodes, therefore the number of implant mask steps may decrease and this may change the amount of allowable oxide and Si loss per post-implant mask cleaning step. In addition, the introduction of SOI may also affect the allowable levels of metal contamination as there is evidence metals may build up at the buried oxide layer interface. It is not yet clear how this will affect allowable metals level and has not been accounted for in these tables.

Interface control is expected to become increasingly critical as devices begin to utilize deposited gate dielectric materials and epitaxial Si and SiGe for strained channel formation. Deposited high-κ gate dielectrics may require an oxidized or nitrided surface prior to deposition, whereas epitaxial Si will require an oxide-free surface. High-κ gate dielectrics may, however, lead to a loosening of requirements for metal contaminant control as gates become physically thicker. After gate formation, post-etch cleans must be introduced which are compatible with high-κ dielectrics and metal gate electrodes. 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. Refer to the *Environment, Safety, and Health* chapter for comprehensive information and link to a new chemical screening tool (Chemical Restrictions Table).

There is universal understanding that watermarks cannot be tolerated on a cleaned surface and so this requirement is no longer itemized in the surface preparation tables. Also, the requirements for back surface particles have been deleted from the surface preparation tables in 2003. Although it is understood that a low level of back surface particles is desirable, there are no data or models available that can link the size or density of back surface particles to yield on the front surface of the wafer. Please see the table footnotes for further explanation.

Table 70a Surface Preparation Technology Requirements—Near-term

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*Driver*
Technology Node		*hp90*			*hp65*			
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*	*D ½*
MPU/ASIC ½ Pitch (nm)	*107*	*90*	*80*	*70*	*65*	*57*	*50*	*M*
MPU Printed Gate Length (nm)	*65*	*53*	*45*	*40*	*35*	*32*	*28*	*M*
MPU Physical Gate Length (nm)	*45*	*37*	*32*	*28*	*25*	*22*	*20*	*M*
Wafer diameter (mm)	**300**	**300**	**300**	**300**	**300**	**300**	**300**	*D ½, M*
Wafer edge exclusion (mm)	**2**	**2**	**2**	**2**	**2**	**2**	**2**	*D ½, M*
Front surface particles								
Killer defect density, D_pR_p (#/cm^2) [A]	**0.0172**	**0.0217**	**0.0283**	**0.0185**	**0.0233**	**0.0158**	**0.0199**	*D ½*
Critical particle diameter, d_c (nm) [B]	**50**	**45**	**40**	**35**	**32.5**	**28.5**	**25**	*D ½*
Critical particle count, D_{pw} (#/wafer) [C]	**59**	**75**	**97**	**64**	**80**	**54**	**68**	*D ½*
Back surface particles [D][E]*	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**
Critical GOI surface metals (10^{10} atoms/cm^2) [F]	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	*M*
Critical other surface metals (10^{10} atoms/cm^2) [F]	**1**	**1**	**1**	**1**	**1**	**1**	**1**	*M*
Mobile ions (10^{10} atoms/cm^2) [G]	**1.8**	**1.9**	**1.9**	**2**	**2.2**	**2.4**	**2.5**	*M*
Surface carbon (10^{13} atoms/cm^2) [H]	**1.8**	**1.6**	**1.4**	**1.3**	**1.2**	**1**	**0.9**	*D ½, M*
Surface oxygen (10^{13} atoms/cm^2) [I]	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	*M*
Surface roughness LVGX, RMS (Å) [J]	**4**	**4**	**4**	**4**	**4**	**4**	**4**	*M*
Silicon loss (Å) per cleaning step [K]	**1.2**	**1.0**	**0.8**	**0.7**	**0.5**	**0.4**	**0.4**	*M*
Oxide loss (Å) per cleaning step [L]	**1.2**	**1.0**	**0.8**	**0.7**	**0.5**	**0.4**	**0.4**	*M*

**Values removed from table for 2003. See table notes [D] and [E].*

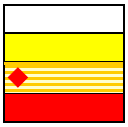
Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known

Table 70b Surface Preparation Technology Requirements—Long-term

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*	*Driver*
Technology Node	*hp45*		*hp32*		*hp22*		
DRAM ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*	*D ½*
MPU/ASIC ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*	*M*
MPU Printed Gate Length (nm)	*25*	*20*	*18*	*14*	*13*	*10*	*M*
MPU Physical Gate Length (nm)	*18*	*14*	*13*	*10*	*9*	*7*	*M*
Wafer diameter (mm)	**300**	**450**	**450**	**450**	**450**	**450**	*D ½, M*
Wafer edge exclusion (mm)	**2**	**2**	**2**	**2**	**2**	**2**	*D ½, M*
Front surface particle							
Killer defect density, D_pR_p (#/cm^2) [A]	**0.0250**	**0.0199**	**0.0250**	**0.0199**	**0.0136**	**0.0215**	*D ½*
Critical particle diameter, d_c (nm) [B]	**22.5**	**17.5**	**16**	**12.5**	**11**	**9**	*D ½*
Critical particle count, D_{pw} (#/wafer) [C]	**86**	**155**	**195**	**155**	**106**	**168**	*D ½*
Back surface particles [D][E]*	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	
Critical GOI surface metals (10^{10} atoms/cm^2) [F]	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	*MPU*
Critical other surface metals (10^{10} atoms/cm^2) [F]	**1**	**1**	**1**	**1**	**1**	**1**	*MPU*
Mobile ions (10^{10} atoms/cm^2) [G]	**2.5**	**2.5**	**2.4**	**2.4**	**2.3**	**2.3**	*MPU*
Surface carbon (10^{13} atoms/cm^2) [H]	**0.9**	**0.9**	**0.9**	**0.9**	**0.9**	**0.9**	
Surface oxygen (10^{13} atoms/cm^2) [I]	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	*D ½, M*
Surface roughness LVGX, RMS (Å) [J]	**2**	**2**	**2**	**2**	**2**	**2**	
Silicon loss (Å) per cleaning step [K]	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	*M*
Oxide loss (Å) per cleaning step [L]	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	*M*

**Values removed from table for 2003. See table notes [D] and [E]*

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known



Notes for Tables 70a and 70b:

[A] Killer defect density is calculated from the formula for 99% yield, $Y=0.99=exp[-D_pR_pA_{eff}]$. A_{eff} is the effective chip area, D_p is the defect density, and R_p is a defect kill factor indicating the probability that a given defect will kill the device. The product D_pR_p is the density of device-killing defects on the wafer. R_p is dependent on numerous things including the size and shape of the particle, the composition of the particle, and specifics of the device layout. In previous years, R_p was assumed to be 0.2 for any particle > the critical particle size, d_c. For DRAM, $A_{eff}= 2.5F^2T+(1-aF^2T/A_{chip})*0.6A_{chip}$. where F is the minimum feature size, a is the cell fill factor, T is the number of DRAM bits (transistors) per chip, and A_{chip} is the DRAM chip size. For MPUs, $A_{eff}=aT(GL)^2$, where GL is the gate length. Because A_{eff} can increase or decrease with each successive technology node, D_pR_p does not always decrease over time.

[B] Critical particle diameter, d_c, is defined by Yield Enhancement as ½ of the metal ½-pitch dimension. This should be considered an "effective" particle diameter as most particulate contamination is irregular in shape.

[C] An example is provided which assumes that the kill factor, R_p, is 0.2 for all particles larger than the critical particle size. This is the assumption made in previous versions of the roadmap, but is not universally valid and is included only for purposes of an example calculation. Particles/wafer is calculated using $[R_p*3.14159*(\text{wafer radius-edge exclusion})^2]$. To convert from particles/wafer at the critical particle size to particles/wafer at an alterative size, a suggested conversion formula is: $D_{alternate}=D_{critical}*(d_{critical}/d_{alternate})^2$.

** [D] and [E] Metrics for critical back surface particle size and back surface particle count are not being listed in 2003. While it is recognized that back surface particles are important to control and are assessed during equipment qualification, there is no clear empirical or theoretical model which links back surface particles to device yield. In the past, arguments have been made that back surface particles affect device yield mainly at the lithographic steps by causing the front surface of the wafer to move out of the focal plane leading to critical dimension variations. However, it is not clear how the limited back surface contact achievable with pin chucks interacts with back surface particle density to cause front surface flatness variations. In addition, it is also not clear how lithographic depth-of-focus (DOF) will change from year to year as this is not specified in the lithography roadmap. In general, it is felt that a good rule is to control back surface particles at a critical diameter equal to one-half of the DOF for critical lithographic steps. In 2003, DOF is about 0.4 micron, so the critical back surface particle diameter is generally considered to be 0.2 micron. It is not possible to measure absolute levels of back surface particles on in-process wafers due to large variations in back surface finish and films. A generally accepted practice is to process wafers with the polished front surface down in order to assess back surface particle adders for a particular process or operation. Current best practice indicates that back surface particle adders for any particular process step in 2003 should be less than 400 at 0.2 micron.*

[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 [12], [13] : (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 Ni, Cu, Cr, Co, Hf and Pt; and (c) major gate-oxide-integrity (GOI) killers such as Ca, Ba, Sr and Fe. 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.[14] 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) that 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. In addition, the introduction of SOI may also affect the allowable levels of metal contamination as there is evidence metals may build up at the buried oxide layer interface. It is not yet clear how this will affect allowable metals level and has not been accounted for in these tables.

[G] The model for mobile ions, D_i, calculates the number of ions that will create a threshold voltage shift that is within a portion of the Allowable Threshold Voltage Variability (ATVV). ATVV was specified in Row 15 of Table 28a in the 1999 ITRS, but is no longer specified. For the purposes of the mobile ion model in 2003, it is assumed that the ATVV is 3% of the Nominal Power Supply Voltage for Low Operating Power or Low Standby Power Technology (see PIDS chapter). The portion of ATVV allocated to mobile ions is assumed to be 5%. Therefore, $D_i=1/q(C_{gate}*ATVV*0.05)$, where C_{gate} is computed for an electrically equivalent SiO_2 gate dielectric thickness and q is the charge of an electron. This model reduces to $D_i = ((3.9*8.85)/1.6)*(0.05*ATVV/EOT)*10^9$, where ATVV is in mV and EOT is in nm (also from Low Operating Power or Low Standby Power Technology Requirements Table in the PIDS chapter) , and the oxide dielectric constant is 3.9. Note that the year-to-year value for D_i does not always decrease because D_i is not only proportional to ATVV, but is also inversely proportional to EOT.

[H] Residual carbon resulting from organic contamination after surface preparation. Surface Carbon at the 180 nm corresponded to 10% carbon atom coverage of a bare silicon wafer (7.3E+13 atoms/cm^2). Surface carbon for subsequent nodes was scaled linearly with the ratio of CD (½ DRAM ½ pitch) to 180 nm. $D_c=(CD/180)(7.3E+13)$

[I] Surface oxygen requirements are driven by the needs of pre-epitaxial cleaning. Epitaxial deposition of Si and SiGe is used for some devices, now, but will find more widespread use with the implementation of strained silicon channel technology. While some level of oxide can be removed in-situ, prior to epitaxial deposition, the trend towards lower deposition temperature will preclude the use of a higher temperature hydrogen pre-bake. Surface oxygen concentrations up to <1E+13 atoms/cm^2 are acceptable for processes such as pre-silicide cleaning. Current pre-gate cleaning does not require an oxide-free surface, but the pre-gate surface should be either fully passivated by a continuous oxide layer or have <1E+13/cm^2, as an intermediate level of oxygen will be unstable. It is unclear whether high-κ gate dielectrics will require oxide-free or oxide-passivated surfaces prior to deposition.

[J] In the 2001 ITRS, it was assumed channel mobility cannot be degraded by >10% due to surface preparation induced surface roughness. It was further claimed that current technologies were successfully manufactured with AFM based determination of 2 Å RMS of surface micro-roughness. Where this is still approximately true for surface preparation induced, i.e. additive roughness, it is more direct to simply measure roughness on product immediately after the low voltage gate oxidation (LVGX) pre-clean. In this case, the total surface micro-roughness takes into account starting substrate roughness, plus the additional micro roughness induced by pre-cleans and strips of initial oxidation, any implant screen oxidations, dummy or sacrificial oxidations the first portions of the high voltage gate oxidation (dual gate flows) and any additional roughness brought about by plasma nitridations. With this taken into account, product has recently been successfully built with 4 Å RMS surface micro-roughness. This may in part be explained by TCAD predictions that show carrier mobility being mainly affected by spatial frequencies smaller than those that are typically sampled by AFM micro-roughness metrology tools. The reason the required number is not scaling over time near term is based off comparisons to the PIDS roadmap where carrier mobility is held constant over this same time period.

[K] The values for silicon loss are driven by requirements of high performance Logic in the portion of the flow where source/drain extensions are fabricated. Specific values are relative to silicon loss measured optically on blanket polysilicon test wafers. Actual consumption on product will vary driven by damage from plasma etch/ash, ion implantation, and dopant concentrations. Decreasing values are in response to the requirement to control negative impact on drive currents (Ids). If the silicon under the source/drain extensions is recessed, this changes the junction profile increasing the source/drain extension resistance and decreasing drive currents. It is not yet possible to express a rigorous model connecting this metric with table parameters such as implant mask levels, junction depth, and critical dimension. IC manufacturers are currently targeting silicon loss to be 1.0 Å per cleaning step for the 90 nm node and 0.5Å per cleaning step for the 65 nm. Other node values are extrapolated or interpolated from those two values. It is not clear what will be required or possible in the longer term years, so the value is set at 0.4Å in 2008 and held constant after that.

---

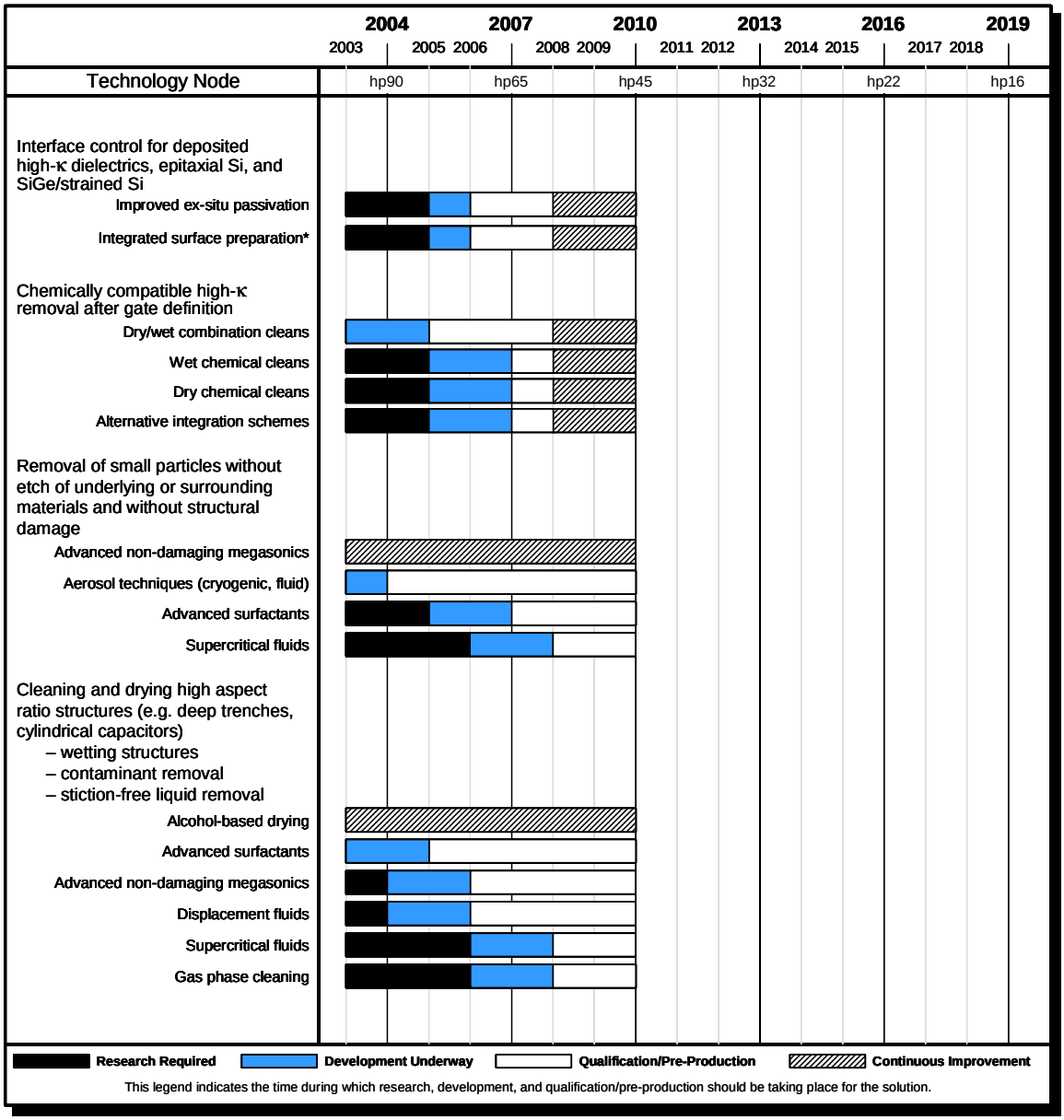
[12] Mertens, P. W., "Advanced Cleaning Technology," UCPSS 2000, Ostende, Belgium, invited tutorial, pp. 31–48 (2000).

[13] Mertens, P. W., et. al., "Recent Advances in Wafer Cleaning Technology," Semicon Europa Front End Technology Conference, Munich, 24 April (2001).

[14] P. W. Mertens, T. Bearda, M. Houssa, L. M. Loewenstein, I. Teerlinck, R. Vos, I. Cornelissen, S. De Gendt, K. Kenis, M. Meuris and M. M. Heyns; " Advanced cleaning for the growth of ultrathin gate oxide", presented at the 11th International Conference on Insulating Films on Semiconductors, Erlangen, Germany, June 16–19, 1999, in Microelectronic Engineering, 48, p. 199 (1999).

[L] The values for oxide consumption are driven by requirements of high performance Logic in the portion of the flow where source/drain extensions are fabricated. Specific values are relative to thermal oxide consumption on blanket test wafers. Actual consumption on product will vary driven by damage from plasma etch/ash, ion implantation, and dopant concentrations. Decreasing values are in response to the requirement to control negative impact on drive currents (Ids). If the silicon under the source/drain extensions is recessed, this changes the junction profile increasing the source/drain extension resistance and decreasing drive currents. By not consuming the oxide, assuming similar processing, this reduces the ability of subsequent processes to further oxidize and consume silicon. Less oxidized silicon equates to less silicon recess under the source/drain extensions. Also, consumption of deposited oxide in the isolation areas is a concern. It is not yet possible to express a rigorous model connecting this metric with table parameters such as implant mask levels, junction depth, and critical dimension. IC manufacturers are currently targeting oxide consumption to be 1.0 Å per cleaning step for the 90 nm node and 0.5 Å per cleaning step for the 65 nm. Other node values are extrapolated or interpolated from those two values. It is not clear what will be required or possible in the longer term years, so the value is set at 0.4 Å in 2008 and held constant after that.

Surface preparation challenges along with potential solutions are shown in Figure 45. 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). The need for other techniques will arise, however, to provide interfacial control for advanced gates as well as non-etching, damage-free particle removal. At this time, broadly effective and non-damaging particle and residue removal using non-liquid techniques are under development. Single wafer cleaning, both wet and dry, is expected to see increased implementation due to process integration and cycle time concerns, but it remains unclear when its use will become widespread. Supercritical CO_2 processes are experiencing a high level of research and development and will most likely be implemented on a single wafer system. Potential solutions are only indicated for the near-term years (through 2009) as it is unclear what challenges will exist for surface preparation at the 45 nm node. As in the past, it is expected that current and future surface preparation processes will be the subject of continuous improvement efforts.



** Integrated Surface Preparation techniques include various techniques that can be coupled to the deposition chamber and allow processing to continue with minimal exposure to the atmosphere. This includes, but is not limited to, UV-based cleaning, gas phase techniques, and single wafer wet techniques.*

Figure 45 Surface Preparation Potential Solutions

Other thrusts, 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 some sulfuric acid based resist strips and post-cleans. Implementation of ozonated water processes is dependent on individual fab situations and requirements for return on capital investments. In addition, new resist formulations for 193 nm and 157 nm lithography may pose challenges for ozonated water resist stripping as well as for sulfuric acid based resist stripping. 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 300 mm and larger wafer sizes where the cost of monitor wafers and process excursions become excessive. New cleaning requirements will arise related to immersion lithography, but will be tied to the implementation of that lithographic method and should be itemized by the lithography technology working group in the future. 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. Refer to the *Environment, Safety, and Health* Chapter for comprehensive information and link to a new chemical screening tool (Chemical Restrictions Table).

THERMAL/THIN FILMS AND DOPING

Front end processing requires the growth, deposition, etching and doping of high quality, uniform, defect-free films. These films may be insulators, conductors or semiconductors (e.g., silicon). The difficult challenges for front end processing include the following: 1) the growth or deposition of reliable very thin (with electrical equivalent thickness 1.0 nm) gate dielectric layers; 2) the development of alternate high dielectric constant layers, including suitable interface layers, for both logic and DRAM capacitor applications; 3) the development of depletion-free, low-resistivity gate electrode materials, and 4) the formation of low resistivity contacts to ultra-shallow junction devices. Other important challenges include the achievement of abrupt channel doping profiles, defect management for minimum post-implant leakage under reduced thermal budget environments, and formation of precise sidewall spacer structures.

An array of *technology innovations* are expected to be required to sustain the trend for increased transistor performance for deeply scaled devices, as detailed in the *PIDS Chapter*. The technology innovations are expected to be introduced within the next five years; and include new materials such as strained-Si channels (to boost carrier mobility and drive currents), high-κ gate dielectrics (to reduce gate leakage and control short channel effects) and metallic gates (to eliminate doped polysilicon depletion effects which limit practical scaling of gate stack layers). Even with the successful introduction of these new materials and structures, the limitations of planar bulk CMOS transistors, particularly the increased sub-threshold leakage currents exhibited at reduced threshold and drive voltages, will drive the introduction of new device structures such as fully depleted, single-, and multi-gate transistors. This rapid introduction of new materials and device structures in the next five to seven years constitutes an unprecedented multiplicity of challenges to develop, and also to integrate these developments into effective, cost-efficient production technologies.

THERMAL/THIN FILMS

The gate dielectric has emerged as one of the most *difficult challenges* for future device scaling. Requirements summarized in Tables 71a and b indicate an equivalent oxide thickness (EOT) progressing to substantially less than 1 nm. Direct tunneling currents and boron penetration (from the polysilicon layer) 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 too much thinner films (about 1 nm); however, their thickness control and reliability limits have not been meaningfully established. For low- power applications where the allowable gate leakage is very low, higher dielectric constant materials may be needed as early as the 70 nm technology generation (2006). These materials will be needed the following year for high performance applications, which require very low equivalent oxide thickness. 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 oxynitride films (perhaps doped with an appropriate rare-earth metal). The Hf-based family of high-κ gate dielectrics has been significantly studied during the past several years. Nevertheless, 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 71a Thermal and Thin Film, Doping and Etching Technology Requirements—Near-term

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*Driver*
Technology Node		*hp90*			*hp65*			
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*	*DRAM*
MPU/ASIC ½ Pitch (nm)	*107*	*90*	*80*	*70*	*65*	*57*	*50*	*MPU*
MPU Printed Gate Length (nm)	*65*	*53*	*45*	*40*	*35*	*32*	*28*	*MPU*
MPU Physical Gate Length (nm)	*45*	*37*	*32*	*28*	*25*	*22*	*20*	*MPU*
Equivalent physical oxide thickness for MPU/ASIC T_{ox} (nm) [A, A1]	**1.3**	**1.2**	**1.1**	**1.0**	**0.9**	**0.8**	**0.8**	*MPU*
Gate dielectric leakage at 100°C (nA/µm) High-performance [B, B1, B2]	**100**	**170**	**170**	**170**	**230**	**230**	**230**	*MPU*
Physical gate length low operating power (LOP) (nm)	**65**	**53**	**45**	**37**	**32**	**28**	**25**	*Low Power*
Physical gate length low standby power (LSTP) (nm)	**75**	**65**	**53**	**45**	**37**	**32**	**28**	*LSTP*
Equivalent physical oxide thickness for low operating power T_{ox} (nm) [A, A1]	**1.6**	**1.5**	**1.4**	**1.3**	**1.2**	**1.1**	**1.0**	*LOP*
Gate dielectric leakage (nA/µm) LOP [B,B1, B2]	**0.33**	**1.0**	**1.0**	**1.0**	**1.67**	**1.67**	**1.67**	*LOP*
Equivalent physical oxide thickness for low standby power T_{ox} (nm) [A, A1]	**2.2**	**2.1**	**2.1**	**1.9**	**1.6**	**1.5**	**1.4**	*LSTP*
Gate dielectric leakage (pA/µm) LSTP [B, B1, B2]	**3**	**3**	**5**	**7**	**8**	**10**	**13**	*LSTP*
Thickness control EOT (% 3σ) [C]	**<±4**	**<±4**	**<±4**	**<±4**	**<±4**	**<±4**	**<±4**	*MPU/ASIC*
Gate etch bias (nm) [D]	**20**	**16**	**14**	**12**	**10**	**10**	**8**	*MPU/ASIC*
L_{gate} 3σ variation (nm) [E]	**4.46**	**3.75**	**3.15**	**2.81**	**2.5**	**2.2**	**2**	*MPU/ASIC*
Total maximum allowable lithography 3σ(nm) [F]	**3.99**	**3.35**	**2.82**	**2.51**	**2.24**	**1.97**	**1.79**	*MPU/ASIC*
Total maximum allowable etch 3σ(nm), including photoresist trim and gate etch [F]	◆ **1.99**	◆ **1.68**	**1.41**	**1.26**	**1.12**	**0.98**	**0.89**	*MPU/ASIC*
Resist trim maximum allowable 3σ(nm) [G]	◆ **1.16**	◆ **0.97**	**0.82**	**0.73**	**0.65**	**0.57**	**0.52**	*MPU/ASIC*
Gate etch maximum allowable 3σ(nm) [G]	**1.62**	◆ **1.37**	**1.15**	**1.02**	**0.91**	**0.80**	**0.73**	*MPU/ASIC*
CD bias between dense and isolated lines [H]	**≤15%**	◆ **≤15**	**≤15%**	**≤15%**	**≤15%**	**≤15%**	**≤15%**	*MPU/ASIC*
Minimum measurable gate dielectric remaining (post gate etch clean) [I]	**>0**	◆ **>0**	**>0**	**>0**	**>0**	**>0**	**>0**	*MPU/ASIC*
Profile control (side wall angle) [J]	**>89**	◆ **90**	**90**	**90**	**90**	**90**	**90**	*MPU/ASIC*
*PIDS Assumed Device Structure**	**Enhanced Planar Bulk CMOS**					**FDSOI, Elev. Contact ***		
Drain extension Xj (nm) [K]	**24.8**	**20.4**	**17.6**	**15.4**	**13.8**	**8.8**	**8.0**	*MPU/ASIC*
Maximum drain extension sheet resistance (PMOS) (Ω/sq) [L]	**545**	**663**	**767**	**833**	**884**	**1739**	**1800**	*MPU/ASIC*
Maximum drain extension sheet resistance (NMOS) (Ω/sq) [L]	**255**	**310**	**358**	**389**	**412**	**811**	**840**	*MPU/ASIC*

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known ◆
Manufacturable solutions are NOT known

Table 71a Thermal and Thin Film, Doping and Etching Technology Requirements—Near-term (continued)

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*Driver*
Technology Node		*hp90*			*hp65*			
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*	*DRAM*
MPU/ASIC ½ Pitch (nm)	*107*	*90*	*80*	*70*	*65*	*57*	*50*	*MPU*
MPU Printed Gate Length (nm)	*65*	*53*	*45*	*40*	*35*	*32*	*28*	*MPU*
MPU Physical Gate Length (nm)	*45*	*37*	*32*	*28*	*25*	*22*	*20*	*MPU*
Extension lateral abruptness (nm/decade) [M]	**5**	**4.1**	**3.5**	**3.1**	**2.8**	**TBD**	**TBD**	*MPU/ASIC*
Contact Xj (nm) [N]	**49.5**	**40.7**	**35.2**	**30.8**	**27.5**	**NA**	**NA**	*MPU/ASIC*
Sidewall spacer thickness (nm) extension structure [O]	**49.5**	**40.7**	**35.2**	**30.8**	**27.7**	**NA**	**NA**	*MPU/ASIC*
Spacer thickness, elevated contact [P]	**—**	**20.4**	**17.6**	**15.4**	**13.8**	**13.2**	**12.0**	*MPU/ASIC*
Maximum silicon consumption (nm) [Q]	**24.8**	**20.4**	**17.6**	**15.4**	**13.8**	**13.2**	**12.0**	*MPU/ASIC*
Silicide thickness (nm) [R]	**25**	**20**	**21**	**19**	**17**	**16**	**14**	*MPU/ASIC*
Contact silicide sheet Rs (Ω/sq) [S]	**6.5**	**7.9**	**7.5**	**8.6**	**9.6**	**10.0**	**11.1**	*MPU/ASIC*
Contact maximum resistivity (Ω-cm^2) [T]	**◆1.93E-07**	**1.62E-07**	**1.44E-07**	**1.20E-07**	**1.05E-07**	**0.87E-07**	**0.72E-07**	*MPU/ASIC*
Gate electrode thickness (nm) [U]	**45–90**	**37–34**	**32–64**	**30–60**	**25–50**	**22–44**	**20–40**	*MPU/ASIC*
Gate depletion: required EOT based on gate choice [V]								
For the case of 1E20/cm^3 poly doping [V]	**1.42**	**1.32**	**1.09**	**0.97**	**0.88**	**0.41**	**0.41**	*MPU/ASIC*
For the case of 2E20/cm^3 poly doping [V]	**1.69**	**1.59**	**1.38**	**1.25**	**1.15**	**0.74**	**0.74**	*MPU/ASIC*
For the case of metal gate [V]	**2.04**	**1.94**	**1.74**	**1.64**	**1.54**	**1.15**	**1.15**	*MPU/ASIC*
Uniform channel concentration (cm^{-3}), for V_t=0.4 [W]	**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**	**NA**	**NA**	*MPU/ASIC*
STI depth (nm) [X]	**400**	**384**	**367**	**359**	**353**	**339**	**335**	*Bulk*
Trench width at top (nm) [Y]	**100**	**90**	**80**	**70**	**65**	**57**	**50**	*Bulk*
Trench sidewall angle (degrees) [Z}	**>86.4**	**>86.6**	**>86.9**	**>87.2**	**>87.4**	**>87.6**	**>87.9**	*Bulk*
Trench fill aspect ratio	**4.5**	**4.8**	**5.1**	**5.6**	**5.9**	**6.5**	**7.2**	*Bulk*
STI depth (nm) [AA]	**36**	**30**	**26**	**22**	**20**	**18**	**16**	*SOI*
Trench aspect ratio	**0.9**	**0.8**	**0.8**	**0.8**	**0.8**	**0.8**	**0.8**	*SOI*

** Refer to supplemental material worksheets, Contact Rs and RsXj for a more complete description of the modeled devices.*

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known

Notes for Tables 71a and 71b:

[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 sho rt term nodes to ±25% for long term nodes.

[A1] Model for EOT is (the 2003 node value–0.4 nm) times the physical gate length plus 0.4 nm.

[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. Models are provided in supplemental material. Since the device subthreshold leakage is expected to increase by a factor of roughly 100× 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-threshold leakage was assumed to be satisfactory from a circuit's 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 $V_g = V_s = 0$ and $V_d = V_{dd}$ as well as when $V_s = V_d \gg 0$ and $V_g = V_{dd}$. 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 72) 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.

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

[E] 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 that 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.

[F] The allowable lithography variance σ^2_L is limited to 4/5 of the total variance, σ^2_T 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 withstand the etch process with loss of dimensional fidelity. Refer to Etch Supplemental Materials.

[G] It is assumed that the resist trim and gate etch processes are statistically independent and therefore that the respective variances, σ^2, 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.

[H] 15% dense-iso CD budget is a combination of measurements from Etch, Lithography and Metrology.

[I] 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-κ materials (2006) 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.

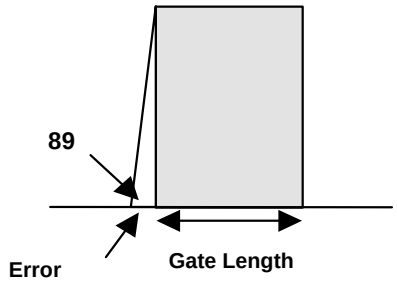
[J] Profile can be a major contributor to etch errors (see inset). Accurate measurement of vertical profiles remains difficult. Long term, the effect of edge roughness on device performance needs to be addressed and methodology of the measurement determined.

---

[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.*

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



*[K] Xj at Channel (Extension Junction) = 0.55*Physical Gate Length (with a range of ±25%) through 2007. At that time, the target device changes to a fully depleted SOI, single gate device (2008–2010), then to a multi-gate fully depleted SOI device thereafter [link to PIDS Chapter]. For these advanced devices(beyond 2007) the X_j is set equal to the channel thickness. See worksheet labeled RsXj.*

[L] 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 labeled 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. This is a relatively crude model and the resultant sheet resistance values should only be used as a guide.

*[M] Channel abruptness in nm per decade drop-off in doping concentration) = 0.11 * Physical Gate Length (nm – based on Short Channel effect.[16] Note discussion of the integration choices in the supplemental material and that this modeling has not been carried out for Advanced MOSFET configurations so years past 2007 are noted as Not Available.*

*[N] Contact Junction Depth = 1.1*Physical Gate Length (with a range of ±33%) through year 2007. Beyond 2007 this has no significance since in our taxonomy contact is assumed to be with the drain extension.*

[O] Spacer thickness (width) is taken as the same as the Contact Junction Depth for bulk devices, through 2007 . Validity established using response surface methodology in "Response Surface Based Optimization of 0.1 µm PMOSFETs with Ultra-Thin Oxide Dielectrics"[17]. Beyond 2007, the spacer width is equal to the thickness of the silicon added to the contact region, reference worksheet RsXj.

[P] Spacer thickness for an elevated junction (where there is no deeper contacting junction) is taken as the extension junction depth for bulk devices through 2007. For advanced fully depleted devices it is taken as 1.5× the channel thickness with a range of ±25%. Reference worksheet RsXj.

[Q] Silicon consumption is based on having the silicide thickness equal to half the contact junction depth, for bulk devices through 2007. For advanced fully depleted devices, having elevated contacts, the silicide thickness is such that the silicide/silicon interface is coplanar with the channel/gate dielectric interface. The silicon consumption is equal to the added silicon thickness above the plane of the channel/gate dielectric interface. In the model assumed here this added silicon is equal to 1.5× of the channel thickness. Reference worksheet RsXj.

[R] Silicide thickness is taken to be 1/2 of the Contact Xj midpoint to avoid consumption-induced increase in contact leakage for bulk devices through 2007. Less than half of the junction can be consumed.[18] For advanced devices, beyond 2007, having elevated contact structures, the silicide thickness is that thickness yielded by consumption of the contact silicon added above the plane of the gate dielectric/channel interface, i.e., 1.5× of the channel thickness. For cobalt and titanium di-silicide layers this silicide thickness is nominally equal to the silicon consumed. For nickel mono-silicide the silicide thickness is equal to 2.22/1.84× of the silicon consumed. In the table we have assumed NiSi implementation in 2005. See worksheet RsXj.

[S] Contact silicide sheet resistance: assumes 15 µΩ-cm silicide resistivity for $TiSi_2$ or $CoSi_2$ and 16 µΩ-cm silicide resistivity for NiSi.

[T] 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 assumes 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 ×W, the equation for the contact resistivity rhoc is: rhoc = Rs × W × 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.

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

---

[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.*

[V] These lines replace the Poly Doping concentration allowance of previous ITRS versions. They show that for a given Poly Doping and Channel Doping, there is a maximum EOT necessary to achieve the needed CET as defined by PIDS. This is a different requirement than that expressed in the EOT lines near the top of Table 71. These differences arise from different model assumptions. Please refer to the supplemental material for a more complete discussion.

[W] Uniform channel concentration for Vt = 0.4.[19] *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. This model applies only for bulk devices.*

[X] Assumes that the trench depth for bulk is proportional to the contact junction depth plus depletion width into the well. The constant of proportionality was determined by setting the 2003 node value equal to 400 nm.

[Y] Assumes a minimum trench width equal to the MPU half-pitch.

[Z] Assumes that the trench width is reduced by no more than half of the top dimension.

[AA] Assumes that the trench depth for SOI is equal to the Si film thickness for partially depleted SOI, i.e., 0.8 × physical gate length.

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

Table 71b Thermal and Thin Film, Doping and Etching Technology Requirements—Long-term

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*	*Driver*
Technology Node	*hp45*		*hp32*		*hp22*		
DRAM ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*	*DRAM*
MPU/ASIC ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*	*MPU*
MPU Printed Gate Length (nm)	*25*	*20*	*18*	*14*	*13*	*10*	*MPU*
MPU Physical Gate Length (nm)	*18*	*14*	*13*	*10*	*9*	*7*	*MPU*
Equivalent physical oxide thickness for MPU/ASIC T_{ox} (nm) [A, A1]	**0.7**	**0.7**	**0.6**	**0.6**	**0.5**	**0.5**	*MPU/ASIC*
Gate dielectric leakage at 100°C (µA/µm) high-performance [B, B1, B2]	**0.33**	**0.33**	**1**	**1.00**	**1.67**	**1.67**	*MPU/ASIC*
Physical gate length operating low operating power (LOP) (nm)	**22**	**18**	**16**	**13**	**11**	**9**	*LOP*
Physical gate length low standby power (LSTP) (nm)	**25**	**20**	**18**	**14**	**13**	**10**	*LSTP*
Equivalent physical oxide thickness for low operating power T_{ox} (nm) [A, A1]	**0.9**	**0.9**	**0.8**	**0.8**	**0.7**	**0.7**	*LOP*
Gate dielectric leakage (nA/µm) LOP [B, B1, B2]	**2.33**	**2.33**	**3.33**	**3.33**	**10**	**10**	*LOP*
Equivalent physical oxide thickness for low standby power T_{ox} (nm) [A, A1]	**1.3**	**1.2**	**1.1**	**1.1**	**1.0**	**0.9**	
Gate dielectric leakage (pA/µm) LSTP [B, B1, B2]	**20**	**20**	**27**	**27**	**33**	**33**	
Thickness control EOT (% 3σ) [C]	**<±4**	**<±4**	**<±4**	**<±4**	**<±4**	**<±4**	*MPU/ASIC*
Gate etch bias (nm) [D]	**7.1**	**6**	**5**	**4**	**3.6**	**3**	*MPU/ASIC*
L_{gate} 3σ variation (nm) [E]	**1.8**	**1.40**	**1.30**	**1.00**	**0.90**	**0.70**	*MPU/ASIC*
Total allowable lithography 3σ (nm) [F]	**1.61**	**1.25**	**1.16**	**0.89**	**0.80**	**0.63**	*MPU/ASIC*
Total allowable etch 3σ (nm), including photoresist trim and gate etch [F]	**0.8**	**0.63**	**0.58**	**0.45**	**0.40**	**0.31**	*MPU/ASIC*
Resist trim allowable 3σ (nm) [G]	**0.46**	**0.36**	**0.34**	**0.26**	**0.23**	**0.18**	*MPU/ASIC*
Gate etch allowable 3σ (nm) [G]	**0.66**	**0.51**	**0.47**	**0.37**	**0.33**	**0.26**	*MPU/ASIC*
CD bias between dense and isolated lines [H]	**≤15%**	**≤15%**	**≤15%**	**≤15%**	**≤15%**	**≤15%**	*MPU/ASIC*
Minimum measurable gate dielectric remaining (post gate etch clean) [I]	**>0**	**>0**	**>0**	**>0**	**>0**	**>0**	*MPU/ASIC*
Profile control (side wall angle) [J]	**90**	**90**	**90**	**90**	**90**	**90**	*MPU/ASIC*
*PIDS Assumed Device Structure **	**FDSOI ***	**FDSOI, Multi-Gate ***					
Drain extension Xj (nm) [K]	**7.2**	**11.2**	**10.4**	**8.0**	**7.2**	**5.1**	*MPU/ASIC*
Maximum drain extension sheet resistance (PMOS) (Ω/sq) [L]	**1875**	**518**	**514**	**550**	**549**	**584**	*MPU/ASIC*
Maximum drain extension sheet resistance (NMOS) (Ω/sq) [L]	**875**	**242**	**240**	**257**	**256**	**272**	*MPU/ASIC*

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known

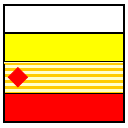
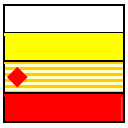


Table 71b Thermal and Thin Film, Doping and Etching Technology Requirements—Long-term (continued)

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*	*Driver*
Technology Node	*hp45*		*hp32*		*hp22*		
DRAM ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*	*DRAM*
MPU/ASIC ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*	*MPU*
MPU Printed Gate Length (nm)	*25*	*20*	*18*	*14*	*13*	*10*	*MPU*
MPU Physical Gate Length (nm)	*18*	*14*	*13*	*10*	*9*	*7*	*MPU*
Extension lateral abruptness (nm/decade) [M]	**TBD**	**TBD**	**TBD**	**TBD**	**TBD**	**TBD**	*MPU/ASIC*
Contact Xj (nm) [N]	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	*MPU/ASIC*
Sidewall spacer thickness (nm) extension structure [O]	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	*MPU/ASIC*
Spacer thickness, elevated contact [P]	**10.8**	**16.8**	**15.6**	**12.0**	**10.8**	**7.7**	*MPU/ASIC*
Maximum silicon consumption (nm) [Q]	**10.8**	**16.8**	**15.6**	**12.0**	**10.8**	**7.7**	*MPU/ASIC*
Silicide thickness (nm) [R]	**13**	**20**	**19**	**14**	**13**	**9**	*MPU/ASIC*
Contact silicide sheet Rs (Ω/sq) [S]	**12.3**	**7.9**	**8.5**	**11.1**	**12.3**	**17.2**	*MPU/ASIC*
Contact maximum resistivity (Ω-cm^2) [T]	**6.08E-08**	**2.03E-08**	**1.71E-08**	**1.10E-08**	**8.69E-09**	**5.40E-09**	*MPU/ASIC*
Gate electrode thickness [U]	**18–36**	**14-28**	**13–26**	**10–20**	**9–18**	**7–14**	*MPU/ASIC*
Gate depletion: required EOT based on gate choice [V]							*MPU/ASIC*
For the case of 1E20/cm^3 poly doping [V]	**TBD**	**0.33**	**TBD**	**0.23**	**TBD**	**0.13**	
For the case of 2E20/cm^3 poly doping [V]	**TBD**	**0.63**	**TBD**	**0.50**	**TBD**	**0.39**	
For the case of metal gate [V]	**TBD**	**1.05**	**TBD**	**0.95**	**TBD**	**0.85**	
Uniform channel concentration (cm^{-3}), for V_t=0.4 [W]	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	*MPU/ASIC*
STI depth (nm) [X]	**331**	**316**	**314**	**300**	**198**	**286**	*Bulk*
Trench width at top (nm) [Y]	**45**	**35**	**32**	**25**	**22**	**18**	
Trench sidewall angle (degrees) [Z}	**>88.1**	**>88.4**	**>88.5**	**>88.8**	**>88.9**	**>89.1**	*Bulk*
Trench fill aspect ratio	**7.9**	**9.5**	**10.3**	**12.5**	**14**	**16.4**	*Bulk*
STI depth (nm) [AA]	**14**	**11**	**10**	**8**	**7**	**6**	*SOI*
Trench aspect ratio	**0.8**	**0.8**	**0.8**	**0.8**	**0.8**	**0.8**	*SOI*

** Refer to supplemental material worksheets, 2003 Contact Rs and 2003 RsXj for a more complete description of the modeled devices*

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
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 interface layer would increase the equivalent oxide thickness and severely degrade any benefits that accrue from the use of the high-κ dielectric.

The presence of an intermediate layer of O-Si-O bonding to bridge between the silicon substrate and high-κ metal ions is expected to limit the scaling of equivalent oxide thickness to nominally 0.4 nm. It is also anticipated that an appropriate material may be required between the high-κ material and the gate electrode to minimize mutual interaction and to inhibit the growth of additional dielectric layers during subsequent processing. 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 realization of dielectric properties that 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 contamination of the transistor channel by gate electrode material or gate electrode dopant.

The gate electrode also represents a major challenge for future scaling. 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).* 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 much more research will be required to identify and qualify an alternative gate electrode material, although a significantly enhanced emphasis has been placed on this problem.

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. The introduction of single-gate, fully-depleted ultra thin body SOI devices with intrinsic channels will change the optimal values of the gate work functions to near mid-gap, suggesting that tunable workfunction systems are of high importance.

In order to maintain high device drive currents, technology improvements are required to increase channel mobility of traditional bulk CMOS devices, as well as partially depleted, and fully depleted SOI devices. The use of strained channel layers, such as strained Si on relaxed Si-Ge for NMOS and strained Si on strained Si-Ge for PMOS will help in achieving this objective but will require considerable process optimization. These enhanced mobility, e.g., strained, channel devices may be needed in conjunction with oxynitride gate dielectrics, before the introduction of high-κ materials. Alternate devices such as non-standard, double gate devices anticipated in the longer-term would also benefit from strained silicon channels.

The incorporation of mobility-enhanced channels, alternate interfacial layers, high-κ dielectrics, and new gate electrode materials into CMOS configurations pose significant integration challenges. The limited thermal stability of many of the candidate material systems is incompatible with junction annealing cycles typically used after gate formation. The use of these new materials may require that either junction annealing temperatures be dramatically lowered, or alternate processes be used that reverse the sequence of gate stack and junction formation. Examples of these include the "replacement gate" or gate-last processes. These schemes increase manufacturing complexity, CoO, and may impact device performance and reliability. Consequently strenuous efforts are in place to retain the conventional CMOS process architecture.

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, source/drain-engineered dopant structures. The robustness of the sidewall spacer limits the gate and source/drain contacting structure and processes that can be used to form these contacts. Sidewall spacers have traditionally been formed from deposited oxides, thermal oxidation of polysilicon, deposited nitrides, and various combinations thereof. Traditional sidewall processes will continue to be used at least until the time (~2008) when elevated or raised source/drain structures are required, at which time process compatibility with the side-wall spacer will become critical. Fully depleted SOI devices will require thin, robust sidewalls having gate dielectric-like reliability and stability. In addition, they must be optimized to minimize parasitic capacitance and series resistance. Below a physical gate length of about 20 nm, even the best state-of the-art thermal oxides are susceptible to defect formation when subjected to selective epitaxial silicon or silicide processes anticipated for elevated contact structures. Nitrides or oxynitrides may offer a better alternative than oxide; however, additional research is needed to find and qualify an acceptable sidewall spacer, compatible with the high-κ gate dielectric.

Thermal and deposited thin films are also very important for filling shallow isolation trenches as well as for pre-metal dielectrics. Trends for decreasing trench width, and higher aspect ratio gaps, suggest that top and bottom corner profile control and controlled uniform filling of dense/isolated structures are the key requirements for this application. In the fabrication of shallow trench isolation structures, the top corner of the active region is generally exposed by HF etching of

pad and sacrificial oxides prior to the growth or deposition of the gate dielectric. The gate conforms to this corner, forming a region of higher electric field and potentially high defectivity. This region can be thought of as a transistor in parallel with the bulk transistor, with both a lower threshold voltage and saturation current. This leads to a 'hump' in the I_d/V_g characteristics and higher subthreshold leakage. Accordingly, the top corner of the STI trench is rounded, usually by oxidation prior to the deposition of the isolation oxide. Increasing the radius of curvature of this corner increases the V_t of the parasitic transistor and decreases the magnitude of this "hump." However, unless new processes are used, device scaling will lead to a decreased radius of curvature.

The magnitude of the parasitic drain current also depends on the degree of recession of the field oxide adjacent the active edge, since that will in part determine the cross section of the edge "transistor." Therefore, as the radius of curvature is scaled down with the isolation width, hopefully so is the recession of the field oxide, resulting in at least partial mitigation of the degradation associated with the decrease of the radius of curvature. The recession of this oxide depends on the 'hardness' of the deposited isolation oxide to CMP processing and to HF dipping, as well as to the thickness of the pad and sacrificial oxides, all of which are process design choices that are optimized at each technology node.

The key thermal/doping integration issues are maintaining shallow junction profiles, junction abruptness, obtaining high dopant activation, ensuring thermal compatibility of materials, and controlling the impact of these issues on device electrical performance. A potential solutions roadmap for Thermal/Thin Films is given in Figure 46. The technology changes associated with incorporation of strained substrates, high-κ dielectrics, metal electrodes and non-bulk CMOS are sufficiently major that two years of process qualification and pre-production will likely be needed before they are ready for full production. For example, extraordinary amounts of reliability data will be needed before totally new gate stack materials would be released for sale to customers. This is in contrast to previous, less radical changes, which only required a year for qualification.

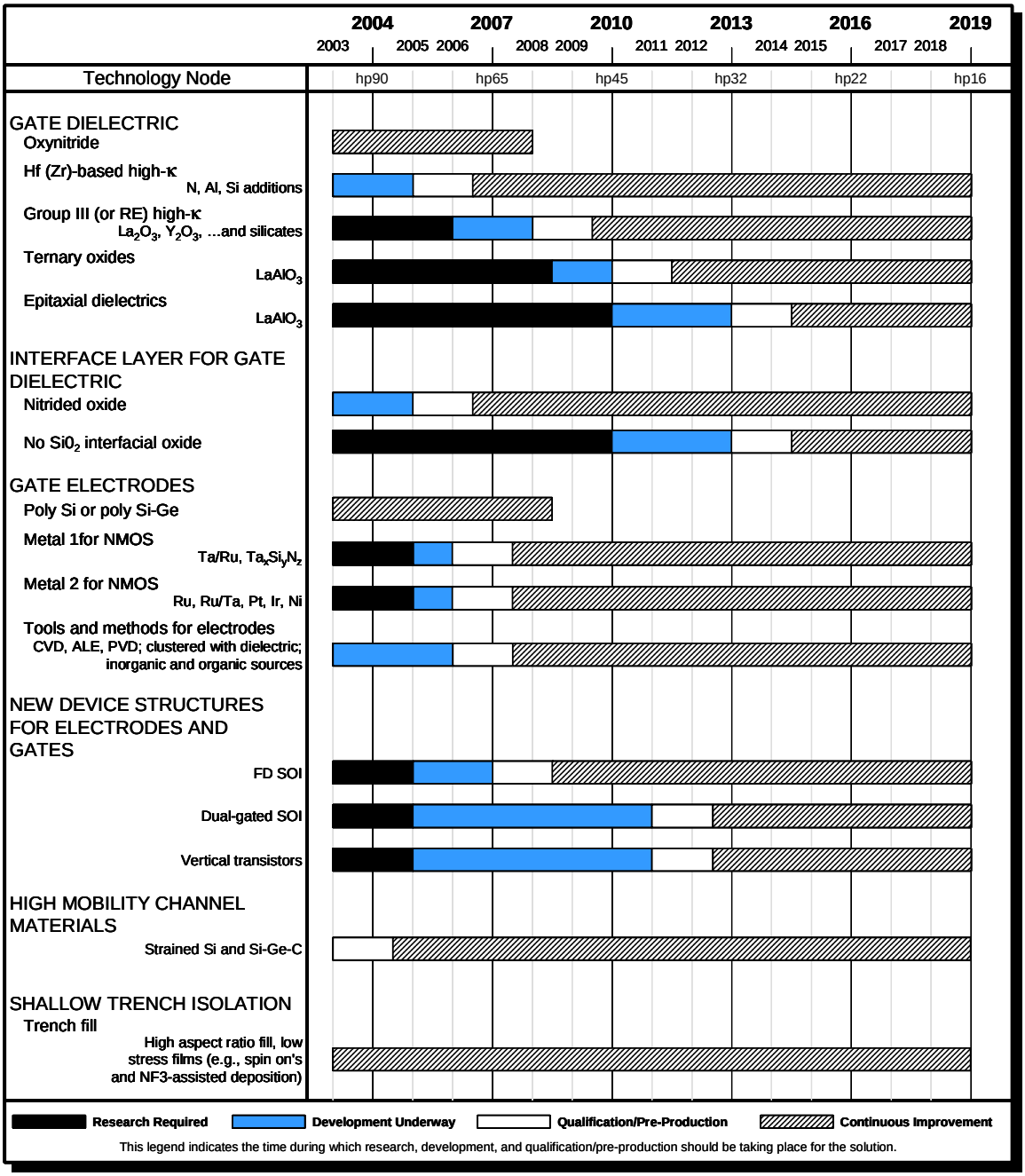


Figure 46 Thermal/Thin Films Potential Solutions

DOPING TECHNOLOGY

As previously mentioned, the traditional scaling of bulk CMOS devices that has served this industry for over two decades is becoming increasingly unworkable due to escalating off-state power consumption, with the consequence that the introduction of new materials and device structures is anticipated within the next ten years. This is discussed in the PIDS chapter where the following device scenario may be inferred:

Years 2003 through 2007—bulk silicon MOSFETS with the following enhancements:

- Optimized oxynitride gate dielectric—2003–2006
- High-κ gate dielectric layers —2006 for low-operating power devices
  —2007 for high-performance devices
- Metal gates —2007 for high performance devices and
  —2008 for low power devices

Years 2008–2010—Fully depleted SOI single gate planar devices with elevated contacts

Years 2011 and thereafter—Fully depleted double or triple gate devices e.g. FINFET, tri-gate

It is noted that this is just one scenario discussed by PIDS that will achieve the historical pace of transistor performance improvement. Other scenarios are possible, but the doping technical requirements tables are based on this scenario. It is noted that the vast majority of development activity in the industry is focused on extending the capabilities of the incumbent technology, bulk silicon CMOS. Therefore it is quite possible that other enhancements not anticipated by the FEP and PIDS technology working groups will extend the life of this technology. The reader may infer the doping requirements of extended bulk CMOS by extrapolating the requirements from the 2003–2007 period. The table notes provide all the guidance needed for such extrapolation.

DIFFICULT CHALLENGES

Through 2007 the difficult challenges for doping of CMOS transistors are 1) extending the concentration of active p-type and n-type polysilicon gate doping beyond presently known limits in order to limit depletion layer thickness in poly-Si gates; 2) achieving doping profiles in the source/drain extension regions to attain progressively shallower junction depths needed for control of short-channel effects, while concomitantly optimizing the sheet resistance-junction depth product, doping abruptness at the extension-channel junction, and extension-gate overlap; 3) developing activation technologies that permit the realization of 10 nm junction depths that exhibit sheet resistance values of nominally 500 Ohm/square; 4) achieving controlled doping profiles in the channel region to set the threshold voltage while concomitantly minimizing short channel effect and maximizing carrier mobility, and 5) the formation of, and making contact to shallow, highly doped source/drain regions.

Beyond 2007, the grand challenge is more directly stated as "Transistor Structure." Planar, bulk CMOS will be extended to the 65 nm node using the aforementioned innovations such as high-κ gate dielectric layers, high-mobility, strained channels, metal gate electrodes, and elevated contact structures. Beyond the 65 nm node, new transistor structures are anticipated to replace bulk silicon devices when continued use of bulk-Si CMOS imposes unacceptably high leakage current penalties.

SOURCE AND DRAIN EXTENSIONS

For planar bulk CMOS, 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 concomitantly minimizing parasitic resistance. These requirements derive from the need to achieve a low parasitic source and drain series resistance that may not exceed 17–33% of the ideal channel resistance. *(See also the PIDS Chapter Logic Technology Requirements, High Performance Near Term Tables)* Similarly, drain extension doping profiles, which with earlier technology nodes had required lateral grading for minimum hot carrier damage, now need to become more laterally abrupt to support these series resistance requirements. In a bulk planar MOSFET the as-implanted (vertical) junction depth with its proportional lateral straggle strongly influences subsequent lateral diffusion and encroachment of the channel. Short channel behavior is therefore strongly linked to the vertical junction depth, and the drain extension resistance strongly linked to doping concentration and lateral abruptness. The complexion of the abruptness challenge depends on the device type. In addition, increased lateral abruptness, while it reduces the accumulation resistance of the drain extension, also influences the spreading resistance with the result that increased lateral abruptness needs to be weighed against other factors that contribute to the overall series resistance. The

effort to model the requirements for sheet resistance, junction depth and junction abruptness has led to an appreciation of the complexity of the interdependence between these parameters with each other and their combined influence on the overall transistor design. Therefore, the process that collectively optimizes junction depth, doping concentration and lateral abruptness, requires the virtual design of the complete transistor characteristics for each node. This is a task beyond the scope of this roadmap. To that end therefore, all three requirements in the technical requirements tables have been indicated as "guidance" rather than well-defined requirements. *[Link to "Abruptness" report from Gossmann to be included in the supporting material].* In general terms however, for P-channel devices, sensitivity simulations indicate that above a critical abruptness value there is only a marginal reduction in parasitic resistance. Improving the abruptness beyond some critical value therefore gives only minor improvements. Some authors claim that a minimum abruptness exists beyond which further improvement is questionable. 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 devices higher abruptness values continue to be 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 simulation models become more sophisticated. This topic is addressed at length in the supporting material *[link to this material].*

The realization of ultrashallow source and drain extension junction depths, that are vertically and laterally abrupt, requires not only the development of new and enhanced methods for implanting the doping species, but requires as well the development of thermal activation processes that have an extremely small thermal budget. This is required to truncate the enhanced diffusion that accompanies the activation of the implanted dopant species. The current methods under investigation are identified in the potential solutions Figure 47. These methods may introduce significant cost adders to the CMOS process flow. Therefore, one should carefully evaluate the incremental benefits in lateral and vertical abruptness that these processes deliver versus the costs incurred. 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.

For non-bulk, MOSFETs, envisioned in year 2008 and beyond, doping processes will require modification for optimized device drive current and threshold voltage stability. Critical doping junction depth parameters will be determined by the thickness of the active silicon layer and therefore become less challenging. For example for fully depleted SOI devices, (also called ultra-thin body devices) the required silicon active layer thickness is nominally in the range of 0.3–0.4 L_g. (Lg = the physical gate length). Double gate devices such as FINFET require nominally twice this active layer thickness. These devices do not require channel doping to manage the short channel effect, and therefore may be implemented using intrinsic, undoped silicon channels. However, the precise control of doping around the gate edge to optimize gate/drain overlap (or underlap) and the management of parasitic resistance remain important technology challenges.

CONTACT

Scaling of contact area, source/drain junction depth, and contact silicide thickness will lead to increases in parasitic resistance effects unless new materials and processes are developed for producing the contact as well as the self-aligned silicide contact shunt. Several interdependent scaling issues arise that require mutual optimization between contact junction depth, silicide thickness, and silicon/silicide contact resistivity. The contact junction depth, despite the strategic placement of halo implants must still scale with gate length, as shown in Table 71. Because of this, progressively less of the contact depth remains available for silicide formation. To avoid contact leakage, no more than half the contact depth can be consumed in the formation of the silicide. Therefore with scaled contacts, the silicide must become progressively thinner to accommodate the progressively more shallow contact junction. This practice cannot be continued beyond a certain point because the silicide will tend to become discontinuous and therefore not adequately shunt the contact.

The introduction of self-aligned nickel monosilicide contacts has provided a near-term solution to this problem. This material consumes nominally half the silicon during silicide formation than the predecessor, cobalt disilicide. In the longer term, this problem can be overcome by the selective deposition of silicon epitaxial layers in the contact region thereby making more silicon available for the silicide formation process. However, as previously discussed, selective epitaxial deposition places increased demands for perfection and robustness on the sidewall spacer. Another contact scaling problem arises form the lateral scaling of the contact area. As a consequence, the contact resistivity associated with the interface between the silicide and the doped contact silicon becomes an important component of the overall source/drain parasitic resistance.

The control of this issue requires that dopant concentration at the interface be maximized, or that a lower barrier height material such as silicon/germanium be used as the contact junction. Here, the use of selectively deposited silicon:germanium materials in the contact and the controlled profiling of the contact dopant provide solutions to these

problems. However, the CMOS integration of these processes that mandate different dopant species for the p-channel and n-channel devices makes this a significant challenge. These integration challenges are made more difficult by the fact that the transistor gates are also doped and silicided together with the contact regions. Another challenge is posed by the introduction of high-κ gate dielectric materials also anticipated in the near term. The limited thermal budget of the candidate high-κ materials, will significantly impact the contact formation and shunting strategies. Similarly the introduction of metal gates (two metals of different work function will be needed for the p-channel and n-channel devices), is expected to further limit the options for contact formation and shunting.

Beyond the 65 nm node, the expected changes in CMOS transistor structures to planar single gate fully-depleted and multi-gate fully depleted devices will present a new array of challenges for formation of contacts to thin, vertical multi-channel arrays linked to heavily-doped contact bus structures. Mastering the intricacies of formation of reliable contacts to these 3D structures will require an additional set of rapid innovations in contact technology. Here, the management of the series resistance of the contact structure remains a major challenge. For the planar single gate devices, the introduction of elevated contacts cannot be avoided without incurring major resistance penalties. Similarly, the research literature contains many references to the strategic use of selective epitaxial shunting of the contact regions of double gate devices in order to realize the required reduction in parasitic resistance. The whole issue of CMOS integration, and its associated dual-doping requirements and how doping is accomplished on these epitaxially enhanced contacts remain a major development issue.

GATE ELECTRODE

The polysilicon gate electrode also represents a major challenge for near-term scaling. This is demonstrated in the lines on Table 71 labeled "Depletion: required EOT based on gate choice". This data was generated by C-V modeling with a variety of channel and polysilicon gate doping concentrations, as well as with metal gates. Curve fitting then permitted extraction of the polysilicon depletion layer EOT contribution as well as the channel quantum layer EOT contribution. Hence, the EOT for a particular node represents the equivalent oxide thickness that is required to obtain that node's capacitive equivalent thickness (CET) for a given channel and gate polysilicon doping. This approach allows the trade-offs that exist between poly-doping and gate dielectric EOT to be made explicit. The red coloration of the EOT values is associated with the need to introduce the high-k gate dielectric material. From Table 71 it becomes clear that, according to this modeling, the introduction of metal gates has the potential to delay the required introduction of the high-κ gate dielectric to year 2009, whereas with polysilicon gate electrodes at a $1\times10^{20}/cm^3$ doping level, the high-κ gate dielectric material requires introduction in 2006. The inference is that if the polysilicon doping level can be increased, then the need for the high-κ material may be delayed. Eventually the increase in the gate dielectric effective thickness associated with polysilicon depletion and the channel autodoping associated with boron out-diffusion from the p^+ polysilicon gate will require the phase-out of the currently used dual-doped polysilicon gate material beyond the 65 nm technology generation.

CHANNEL

The maintenance of acceptable off-state leakage with continually decreasing channel lengths will require channel doping levels for planar CMOS transistors to increase in order to control short channel effects for extremely small devices. Increasingly precise control of both vertical and lateral channel doping profiles is required to deal with short channel effects, introducing new challenges for doping tools, process and metrology. In the near term, these scaling problems are exacerbated by the slower than ideal scaling of the gate stack, which is the result of the unavailability of a practical material solution for higher κ gate dielectric layers and the presence of gate depletion effects due to limitations to doping levels in poly-Si gates. The circuit speed advantages of increased drive current for high-performance logic applications will drive the near-term (at the 90 nm technology generation in 2004) introduction of strained-Si channel materials to increase carrier mobility and to counter the trend towards lower carrier mobility driven by increased channel doping for control of short channel effects.

The substantial increases in off-state leakage currents that accompany scaling of threshold voltages in bulk-Si and partially-depleted CMOS transistors will ultimately drive major changes in channel materials and transistor structure, to some form of fully-depleted, multi-gate transistor. The introduction of fully depleted CMOS (see fuller discussion in the *PIDS chapter*) is anticipated to occur beyond the 65 nm technology generation in about 2008. Channel designs for fully depleted CMOS favor the use of intrinsic, undoped silicon. This may be achieved in single gate and double gate fully depleted structures, because the electric field lines from the drain terminate at the gate, thereby shielding the source from DIBL and associated short channel effect. This approach avoids the carrier mobility degradation associated with channel doping but requires that threshold voltage be exclusively controlled by the work function of the gate electrode. Dual work

function gates are required for CMOS integration but since both work functions are near mid-gap, a potential solution might consist of a single metal with work function modified by subtle changes in composition, e.g. doping.

In single-gate, fully depleted CMOS designs, the channel thickness follows similar scaling as the source/drain extension in bulk-Si and partially depleted transistors. Control of channel thickness (of the order of 10 nm for sub-25 nm gate lengths) to the 5% range (equal to the lattice constant of Si) desired for control of threshold voltage is a significant challenge for SOI wafer fabrication and process controls for manufacturing of fully-depleted CMOS devices. Shifting of single-gate, fully depleted CMOS to multi-gate structures leads to a relaxation of the channel thickness scaling. This is reflected in the doubling of the channel thickness trend, identified as the progression of the drain extension Xj in Table 71, with the anticipation of the use of a dual-gate architecture for the 40 nm technology node (2011). Extension of fully-depleted channel designs to tri-gate and surround-gate architectures promise additional improvements in drive current at the cost of increased complexity in manufacturing process.

Optimization of the doping transition from highly-doped contact regions to intrinsic channels, reduction of sub-threshold leakage at high-field channel edges in multi-gate designs and solution to a large number of process integration issues arising from a fully 3-D transistor design must be dealt with before the successful introduction of multi-gate, fully-depleted CMOS production. These challenges, added to the anticipated shifts to high-κ gate dielectrics and dual-workfunction, metal-gate materials, constitute a revolutionary change in transistor technology in the coming 5 to 7 years.

POTENTIAL SOLUTIONS

The first two sections of the Doping Potential Solutions Table in Figure 47 list approaches to realizing the "Key Innovations" discussed in the *PIDS chapter*. The "Device Structures" category highlights anticipated process solutions to the potential doping and contact integration issues arising through the introduction of new gate stack and substrate materials and non-planar device geometries. The primary impact of these changes is expected through modification of the timing or application of solutions already present on the table rather than through the introduction of revolutionary process technologies. The "Channel Engineering" category identifies specific doping and deposition solutions for simultaneously meeting threshold voltage, channel mobility, and subthreshold current requirements of the evolving device structures. We expect strained channel engineering to apply to virtually all device approaches in the future, with highly doped and super steep retrograde channels giving way to lower concentrations for FDSOI and multiple gate devices.

The next four categories of Potential Solutions focus on the production of abrupt, ultrashallow junctions. The evolution of beamline ion implantation and anneal will continue for the foreseeable future; however, the need for additional, possible complementary options is expected for alternative device structures. Evolutions of beamline ion implantation will include the ability to implant high doses at higher angles for higher aspect ratio structures, and may include implantation of larger ion clusters. The annealing roadmap will be increasingly driven by abruptness requirements in addition to sheet resistance and junction depth with shorter gate lengths. Annealing technologies have a roadmap of decreasing process times, which are the clearest potential solution to near diffusionless junctions. In the near term, spike anneals are clear leaders for junction formation while millisecond anneals are well into development. Plasma doping combined with annealing and selectively deposited junctions appear the most likely alternatives for high aspect ratio, three-dimensional structures. Correlated with this is the category of "Defect Engineering," highlighting the approach of mitigating or using the various enhanced diffusion and dopant segregation effects that to date have mainly confounded boron diffusion in particular.

Also note that "Low Resistance Deposited and Thermal Doping" as a solution category includes selective deposition not only for ultrashallow junctions, but also for raised contacts, strain engineering, etc. It is in these applications rather than ultrashallow junction engineering that selective deposition is anticipated at the 90 nm node. In addition, this section lists some of the candidate technologies for the introduction of dopants toward the far end of the current table.

The contact category of Figure 47 notes that self-aligned metal silicides remain the incumbent technology, albeit with a progression in the silicide material of choice. While NiSi is expected to be widely implemented by the 65 nm node, the material roadmap beyond NiSi remains unclear. Migration to a metal contact scheme to meet the contact resistance challenge with scaling is not expected in this decade.

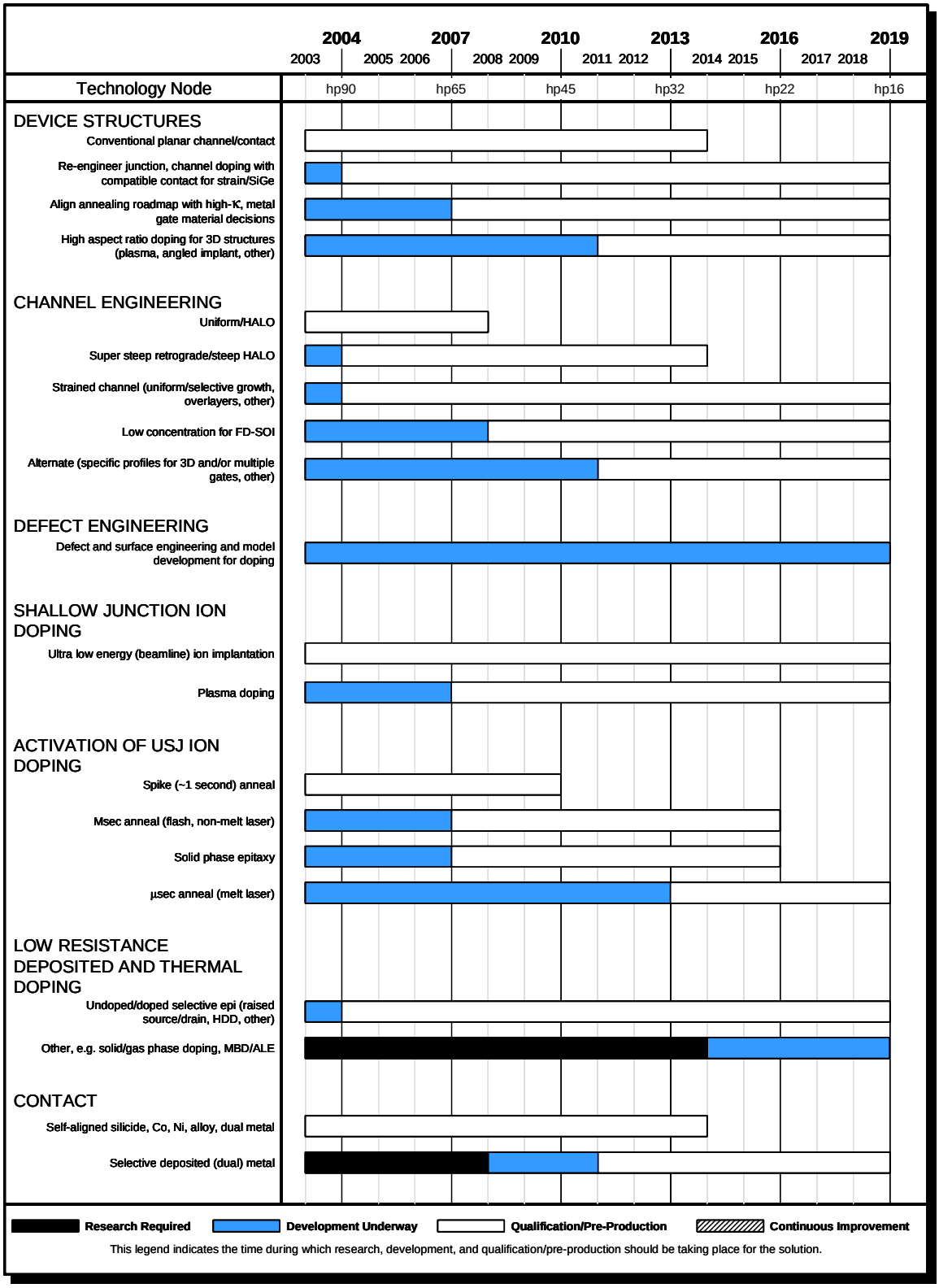


Figure 47 Doping Potential Solutions

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 (300 mm) 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 71a*). 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-2 nm 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 48). *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-1 nm CD control.

The required CD control and etch characteristics mentioned above must also be achieved with new materials, such as high-κ dielectric films and metal gate structures (See Figure 48). 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 ~$10^{11}/cm^3$ 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-κ gate dielectric layers these developments must be complete by 2006 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 that may be brought about by advanced chip architecture (See Figure 48).

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 248 nm resist types, but quite a bit of production experience has been gained on 193 nm resists as well. For 157 nm generation resist materials, the task will be much more difficult since both the height of the resist material decreases significantly

(~200 nm [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 157 nm 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 occurring and secondly, the introduction of metal or metal nitride gates electrodes. High-κ gate dielectrics other than silicon oxides and nitrides are expected in 2006 on the low standby power devices. As scaling reduces thermal oxides to the 13 angstrom regime, direct charge carrier tunneling leads to unacceptably high gate leakage. High-κ gate materials under development, such as 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. On the other hand, stopping on this more robust dielectric layer may be easier. Wet etching will be difficult since increased thickness (compared to SiO or SiN materials) may 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. 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 90 nm 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 Tables 71a 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.

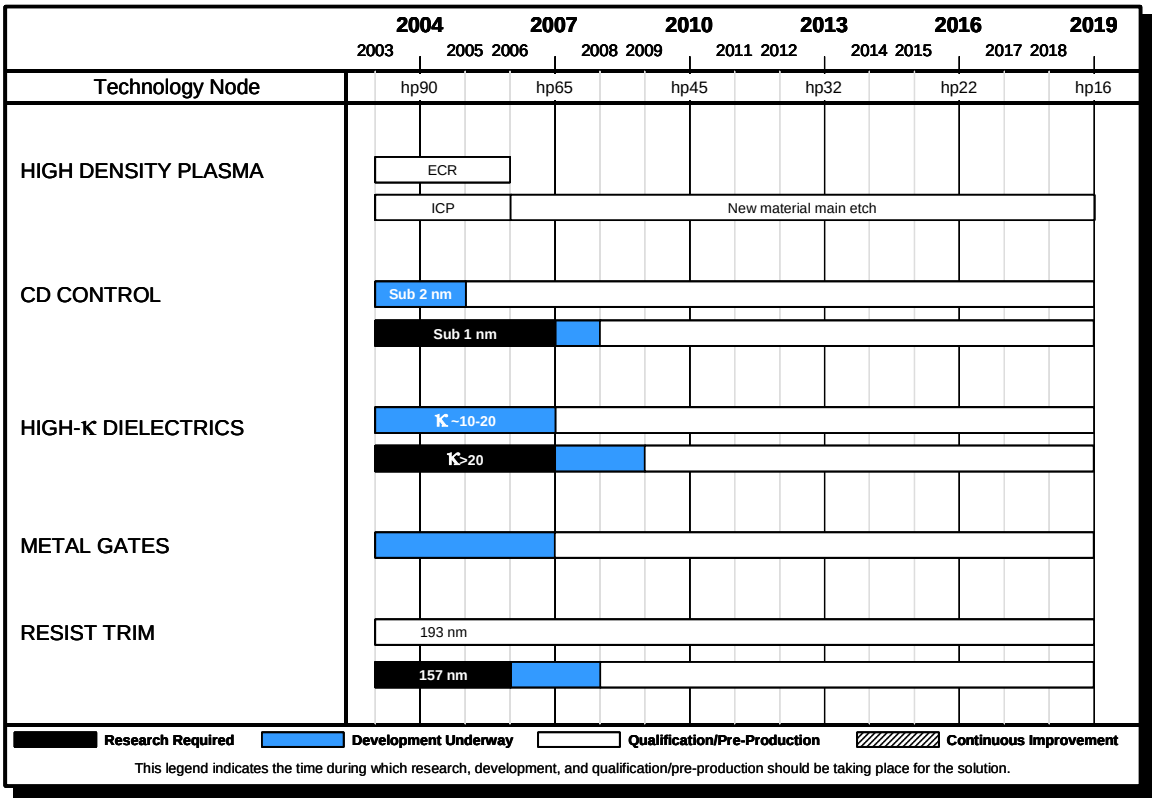


Figure 48 Front End Processes Etching Potential Solutions

DRAM Stacked Capacitor

Historically, a quadrupling of DRAM product capacity 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 128 Mb and 512 Mb, have appeared respectively after the 64Mb and 256 Mb DRAM generations. DRAM capacitor technology faces new challenges of introducing new storage capacitor dielectric and electrode materials. Table 72 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) has been realized at the 180 nm node, which is the smallest cell size with a folded bit line architecture. Each of the target values in Table 72 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 was difficult to keep a sufficiently high cell capacitance value using these materials and structure at the 130 nm node and beyond. Thus, alternative high-κ dielectrics such as Ta_2O_5 and Al_2O_3 were introduced at this technology node. Ta_2O_5, with a range of dielectric constants, is one of the most promising dielectrics mentioned above. For the 130 nm node, a poly-silicon bottom electrode and a 3D capacitor cell with high-κ dielectric and metal counter-electrode have been used (this is an example of a metal-insulator-silicon (MIS) structure). This MIS stack, using Ta_2O_5 as a dielectric, has an effective dielectric constant of 22. 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 130 nm 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.[20] Therefore, a MIM structure is required beyond the 130 nm node.

At and beyond the 90 nm node, metals or conductive metallic nitrides/oxides such as Pt, Ru, TiN, 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 90 nm 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 45 nm node, new ultra high-κ materials with a dielectric constant over 60 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 45 nm node and beyond.

Beyond the 22 nm node, cell size has to be reduced below $6F^2$ to keep a reasonable chip size. New cell architecture such a cross-point cell will be required. In addition, a dielectric constant of 100 is needed.

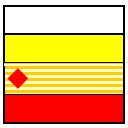
The process technology requirements for system-on-a-chip (SOC) with embedded DRAM exhibit many variations depending on the ratio of logic area and memory area. Cell capacitance requirements for embedded DRAM may be smaller than those for stand-alone DRAM. One of the serious problems for SOC is contact 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.

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

Table 72a DRAM Stacked Capacitor Films Technology Requirements—Near-term

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*
Technology Node		*hp90*			*hp65*		
DRAM ½ Pitch (nm) [A]	*100*	*90*	*80*	*70*	*65*	*57*	*50*
Cell size factor a [B]	**8.0**	**8.0**	**7.5**	**7.0**	**7.0**	**6.0**	**6.0**
Cell size (μm^2) [C]	**0.08 =0.2×0.4**	**0.065 =0.18×0.36**	**0.048 =0.16×0.3**	**0.034 =0.14×0.25**	**0.030 =0.13×0.23**	**0.019 =0.114×0.17**	**0.015 =0.1×0.15**
Storage node size (μm^2) [D]	**0.030 =0.1×0.3**	**0.024 =0.09×0.27**	**0.018 =0.08×0.22**	**0.012 =0.07×0.18**	**0.011 =0.065×0.16**	**0.006 =0.057×0.11**	**0.005 =0.05×0.1**
Capacitor structure	**Cylinder MIS**	**Cylinder MIS/MIM**	**Cylinder /Pedestal MIM**	**Cylinder /Pedestal MIM**	**Pedestal MIM**	**Pedestal MIM**	**Pedestal MIM**
Capacitor dielectric material	**ALO/TAO**	**ALO/TAO**	**ALO/TAO**	**ALO/TAO /others**	**ALO/TAO /others**	**ALO/TAO /others**	**ALO/TAO /others**
teq@ 25fF (nm) [G]	**3.5**	**2.3**	**1.8**	**1.4**	**0.8**	**0.8**	**0.8**
Dielectric constant	**22**	**22**	**40**	**50**	**50**	**50**	**50**
SN height (μm)	**2.1**	**1.5**	**1.4**	**1.4**	**1.2**	**1.6**	**1.8**
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)	**2.57**	**1.66**	**1.29**	**1.01**	**0.56**	**0.55**	**0.55**
Structural coefficient [F]	**32.1**	**25.6**	**26.8**	**29.5**	**18.8**	**28.4**	**36.3**
t phy.@25fF (nm) [H]	**14.3**	**7.3**	**18.2**	**17.9**	**9.9**	**9.8**	**9.6**
A/R of SN (OUT) for cell plate depo. [I]	**29.4**	**19.9**	**32.1**	**39.4**	**26.5**	**42.8**	**58.6**
HAC diameter (μm) [J]	**0.12**	**0.11**	**0.10**	**0.08**	**0.08**	**0.07**	**0.06**
Total interlevel insulator and metal thickness except SN (μm) [K]	**0.90**	**0.86**	**0.84**	**0.81**	**0.78**	**0.75**	**0.73**
HAC depth (μm) [L]	**3.00**	**2.36**	**2.24**	**2.16**	**1.98**	**2.35**	**2.53**
HAC A/R	**25.0**	**21.9**	**23.3**	**25.7**	**25.4**	**34.4**	**42.2**
Vcapacitor (Volts)	**1.8**	**1.7**	**1.6**	**1.5**	**1.4**	**1.3**	**1.2**
Retention time (ms) [M]	**64**	**64**	**64**	**64**	**64**	**64**	**64**
Leak current (fA/cell) [N]	**1.05**	**1.00**	**0.94**	**0.88**	**0.82**	**0.76**	**0.70**
Leak current density (nA/cm^2)	**41.1**	**60.1**	**72.9**	**87.0**	**147.4**	**137.6**	**129.0**
Deposition temperature (degree C)	**~500**	**~500**	**~500**	**~500**	**~500**	**~500**	**~500**
Film anneal temperature (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**

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known



Notes for Tables 72a and 72 b:

[A] 2003 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.)*

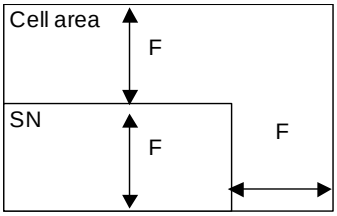
*[J] HAC diameter = 1.2*F (HAC : High Aspect Contact)*

[K] The thickness is assumed to be 1.05 µm@180 nm. (10% reduction by each generation)

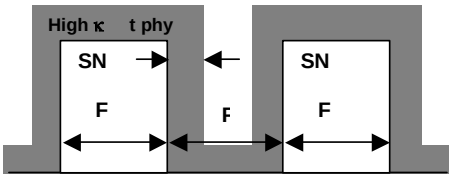
[L] HAC depth = SN height + total interlevel insulator and metal thickness

[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 72b DRAM Stacked Capacitor Films Technology Requirements—Long-term

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*
Technology Node	*hp45*		*hp32*		*hp22*	
DRAM ½ Pitch (nm) [A]	*45*	*35*	*32*	*25*	*22*	*18*
Cell size factor a [B]	**6.0**	**6.0**	**6.0**	**6.0**	**6.0**	**5.0**
Cell size (µm^2) [C]	**0.012 =0.09×0.14**	**0.007 =0.07×0.11**	**0.006 =0.064×0.1**	**0.004 =0.05×0.08**	**0.003 =0.044×0.07**	**0.0016 =0.036×0.05**
Storage node size (µm^2) [D]	**0.004 =0.045×0.09**	**0.002 =0.035×0.07**	**0.002 =0.032×0.06**	**0.001 =0.025×0.05**	**0.001 =0.022×0.04**	**0.0005 =0.018×0.03**
Capacitor structure	**Pedestal MIM**	**Pedestal MIM**	**Pedestal MIM**	**Pedestal MIM**	**Pedestal MIM**	**Pedestal MIM**
Capacitor dielectric material	**ALO/TAO /others**	**ALO/TAO /others**	**ALO/TAO /others**	**new material**	**new material**	**new material**
teq@25fF (nm) [G]	**0.7**	**0.6**	**0.5**	**0.4**	**0.4**	**0.25**
Dielectric constant	**50**	**60**	**60**	**80**	**80**	**100**
SN height (µm)	**1.9**	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**
Cylinder factor [E]	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
Roughness factor	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
Total capacitor area (µm^2)	**0.52**	**0.42**	**0.39**	**0.30**	**0.26**	**0.180**
Structural coefficient (F)	**42.6**	**57.5**	**62.8**	**80.3**	**91.2**	**111.4**
t phy.@25fF (nm) [H]	**9.2**	**9.0**	**8.2**	**8.5**	**7.5**	**6.4**
A/R of SN (OUT) for cell plate depo. [I]	**71.2**	**117.3**	**128.2**	**252.0**	**286.0**	**383.0**
HAC diameter (µm) [J]	**0.05**	**0.04**	**0.04**	**0.03**	**0.03**	**0.02**

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known

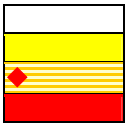
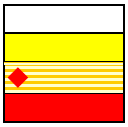


Table 72b DRAM Stacked Capacitor Films Technology Requirements—Long-term (continued)

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*
Technology Node	*hp45*		*hp32*		*hp22*	
DRAM ½ Pitch (nm) [A]	*45*	*35*	*32*	*25*	*22*	*18*
Total interlevel insulator and metal thickness except SN (µm) [K]	**0.70**	**0.66**	**0.63**	**0.59**	**0.57**	**0.53**
HAC depth (µm) [L]	**2.60**	**2.66**	**2.63**	**2.59**	**2.57**	**2.53**
HAC A/R	**48.1**	**63.2**	**68.5**	**86.3**	**97.2**	**117.2**
Vcapacitor (Volts)	**1.1**	**1**	**0.9**	**0.8**	**0.7**	**0.6**
Retention time (ms) [M]	**64**	**64**	**64**	**64**	**64**	**64**
Leak current (fA/cell) [N]	**0.64**	**0.59**	**0.53**	**0.47**	**0.41**	**0.35**
Leak current density (nA/cm^2)	**124.7**	**138.7**	**136.6**	**155.6**	**154.8**	**194.8**
Deposition temperature. (degree C)	**~500**	**~500**	**~500**	**~500**	**~500**	**~500**
Film anneal temperature (degree C)	**<750**	**~650**	**~650**	**<650**	**<650**	**<650**
Word line Rs (Ohm/sq.)	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known



Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*
DRAM ½ PITCH (nm)	**100**	**90**	**80**	**70**	**65**	**57**	**50**	**45**	**35**	**32**	**25**	**22**	**18**
Electrode [A]	**metal**												
high-κ dielectric	**ON** / **Ta_2O_5, Al_2O_3**				**Ta_2O_5, Al_2O_3, Others**						**New Material**		
Bottom Electrode [A] [B]	**poly-Si** / **metal**												

[A] Metal: Ti, TiN, W, Pt, Ru, RuO2, IrO_2
[B] Perovskite: $SrRuO_3$[21]

Figure 49 DRAM Stacked Capacitor Potential Solutions

DRAM Trench Capacitor

Tables 73a and 73b summarize the technological requirements for DRAM trench capacitors. The target values are based on the assumption of a 35 fF capacitance per DRAM cell. It is further assumed that the cell size will remain 8F^2. Attaining cell size factors <8 are considered difficult challenges for both stacked and trench capacitor based DRAM cells.

For technology nodes down to and including the 110 nm design rule, trench capacitor technology has been characterized by the use of conventional nitride/oxide capacitor dielectric in combination with trench profile optimization techniques.

The introduction of bottle-shaped trenches and trench surface roughening is anticipated at the 100 nm node. These surface area enhancement techniques will allow conventional nitride/oxide dielectric to be extended down to and including the 90 nm design rule.

It is expected that high-κ dielectric materials will be implemented at the 80 nm design rule. Al_2O_3 is the most likely candidate at the introduction node. Atomic Layer Deposition (ALD) will be required to deposit these materials into high

[21] *N. Fukushima et al., IEDM Technical Digest, pp. 257–260, 1997.*

aspect ratio trenches.[22] High-κ materials with still higher dielectric constant are under evaluation for smaller ground rules. Metal top electrodes are anticipated at the 65 nm node and will allow the replacement of SIS by MIS capacitors. Ultimately, MIM capacitors are an option at the 45 nm design rule. High-κ dielectrics are already in production at current technology nodes (≥110 nm) using stacked capacitor technology.[23] Trench DRAM technology can reserve the introduction of high-κ dielectrics for future technology generations, which is a considerable scaling advantage.

Trench technology allows the fabrication of rigid and mechanically extremely stable high aspect ratio capacitor structures. As a result of ground-rule shrinking, the trench aspect ratio (trench depth divided by trench top width after etch) will increase up to values of ~60:1 for the 80 nm design rule. Such aspect ratios have already been demonstrated.[22] Even higher aspect ratios are anticipated at smaller ground rules.

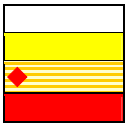
Novel cell concepts relying on the replacement of the conventional planar transfer device by a vertical array transistor are envisaged for the end of the decade. The use of a vertical transistor would greatly alleviate device-scaling issues.

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. The trench DRAM concept also avoids 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 budget is not encountered.

Table 73a DRAM Trench Capacitor Technology Requirements—Near-term

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*
Technology Node		*hp90*			*hp65*		
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*
DRAM Product	**1G**	**1G**	**2G**	**2G**	**2G**	**4G**	**4G**
Cell size factor	**8**	**8**	**8**	**8**	**8**	**8**	**8**
Cell size (μm²)	**0.08**	**0.065**	**0.051**	**0.039**	**0.034**	**0.029**	**0.02**
Trench structure	**bottled**	**bottled**	**bottled**	**bottled**	**bottled**	**bottled**	**bottled**
Trench circumference (nm)	**851**	**766**	**681**	**596**	**553**	**511**	**426**
Trench area enhancement factor (bottle) [A]	**1.65**	**1.65**	**1.65**	**1.65**	**1.65**	**1.65**	**1.65**
Trench surface roughening factor	**1.25**	**1.25**	**1.25**	**1.25**	**1.25**	**1.25**	**1**
Effective oxide thickness (CET)(nm)	**5**	**5**	**4.1**	**4.1**	**3.8**	**3.2**	**2.8**
Trench depth [μm], (at 35fF)	**5.9**	**6.4**	**6**	**6.7**	**6.6**	**6.2**	**7.7**
Aspect ratio (trench depth/trench width)	**45**	**55**	**58**	**73**	**78**	**79**	**93**
Upper electrode	**Poly-Silicon**	**Poly-Silicon**	**Poly-Silicon**	**Poly-Silicon**	**Metal**	**Metal**	**Metal**
Dielectric material	**NO**	**NO**	**High-κ**	**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-Insulator-Silicon/ High-κ**		
[A] Bottle factor = checkerboard square perimeter / conventional elliptical perimeter							

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known

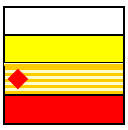


[22] *M. Gutsche et al., IEDM (2001) 411.*
[23] *Semiconductor Insights Report 0503-8222-FS41 (2003).*

Table 73b DRAM Trench Capacitor Technology Requirements—Long-term

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*
Technology Node	*hp45*		*hp32*		*hp22*	
DRAM ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*
DRAM Product	**4G**	**8G**	**16G**	**16G**	**32G**	**32G**
Cell size factor	**8**	**8**	**8**	**8**	**8**	**8**
Cell size (μm^2)	**0.016**	**0.01**	**0.008**	**0.005**	**0.004**	**0.003**
Trench structure	**bottled**	**bottled**	**bottled**	**bottled**	**bottled**	**bottled**
Trench circumference (nm)	**383**	**298**	**272**	**213**	**187**	**153**
Trench area enhancement factor	**1.65**	**1.65**	**1.65**	**1.65**	**1.65**	**1.65**
Trench surface roughening factor (bottle [A]	**1**	**1**	**1**	**1**	**1**	**1**
Effective oxide thickness (CET) (nm)	**2.2**	**1.6**	**1.4**	**0.8**	**0.6**	**0.4**
Trench depth (μm), (at 35fF)	**7**	**6.5**	**6.2**	**5**	**4.6**	**4**
Aspect ratio (trench depth/trench width)	**94**	**113**	**117**	**121**	**125**	**136**
Upper electrode	**Metal**	**Metal**	**Metal**	**Metal**	**Metal**	**Metal**
Dielectric material	**High-κ**	**1: Epi-high-κ / 2: High-κ**	**1: Epi-high-κ / 2: High-κ**	**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**	**1: Silicon**	**1: Silicon**	**1: Silicon**
	2: Metal	**2: Metal**	**2: Metal**	**2: Metal**	**2: Metal**	**2: Metal**
Capacitor structure/dielectric	**1: MIS / (Epi)-high-κ 2: MIM / High-κ**					
[A] Bottle factor = checkerboard square perimeter/conventional elliptical perimeter						

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known



NON-VOLATILE MEMORY (FLASH)

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.

The roadmap for Flash memories was first reported in the 2001 ITRS for both NAND and NOR types, but a clear definition of the technology node has yet to be given. The definition of the technology node for this device, i.e., the meaning of the minimum feature size, is not the same for NAND and NOR types, due to the different peculiarities and the different architectures.

For NAND Flash the best definition of the minimum feature size is the half-pitch of the memory cell when viewing a cross section parallel to the bit line, that is also the half pitch of the poly 2-word line. Refer to Figure 50.

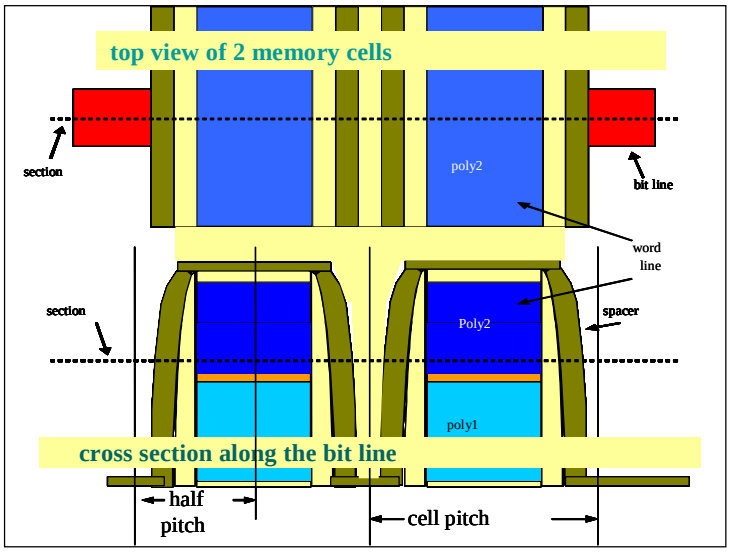


Figure 50 Minimum Feature Size of NAND Flash Memory

For NOR Flash memories the definition of the minimum feature size is not very easy and can vary among the different Flash manufacturers. Referring to Figure 51, the following are definitions of the minimum feature size specific for NOR Flash memories, as follows:

- the half pitch when viewing a cross section parallel to the poly 2 word line
- the poly 1 to poly 1 distance along the word line
- the minimum contact size

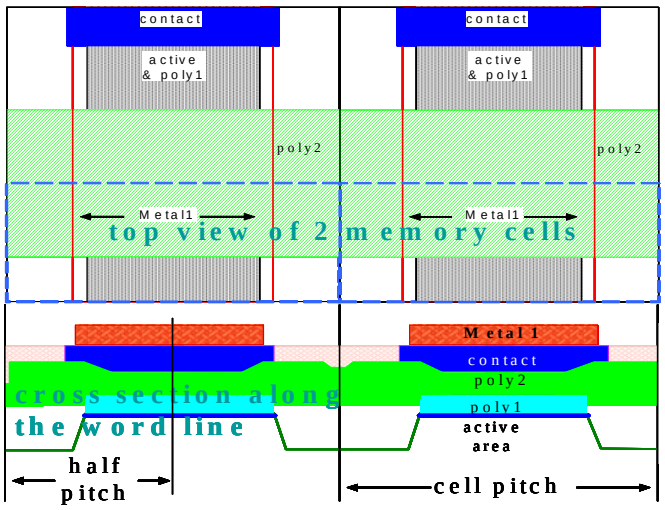


Figure 51 Minimum Feature Size of NOR Flash Memory

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.

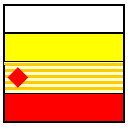
Regarding the technology roadmap, reported in Tables 74a and 74b, no major changes have been implemented relative to last year. In the 2002 ITRS update, a pull-in of the technology nodes had been made in order to reflect the acceleration occurring in Flash memory R&D and production. Now the Flash roadmap is aligned with logic in terms of technology nodes, with only a few months delay not detectable when the time unit is the year. Moreover, major changes occurred for the tunnel and the interpoly dielectric thicknesses of NAND devices. In the present ITRS revision minor adjustments have been made for the scaling down of the thickness for both the interpoly and the tunnel dielectrics.

At the present time, international consensus has not been achieved relative to the potential solutions needed to address the scaling issues for this technology. It is expected that these issues will be addressed in subsequent editions of the ITRS.

Table 74a FLASH Non-volatile Memory Technology Requirements—Near-term

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*
Technology Node		*hp90*			*hp65*		
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*
MPU/ASIC ½ Pitch (nm)	**107**	**90**	**80**	**70**	**65**	**57**	**50**
MPU Printed Gate Length (nm)	**65**	**53**	**45**	**40**	**35**	**32**	**28**
MPU Physical Gate Length (nm)	**45**	**37**	**32**	**28**	**25**	**22**	**20**
Flash technology node (nm) [A]	**107**	**90**	**80**	**70**	**65**	**57**	**50**
Flash NOR tunnel oxide thickness (EOT-nm) [B]	**9–10**	**8.5–9.5**	**8.5–9.5**	**8.0–9.0**	**8–9**	**8–9**	**8–9**
Flash NAND tunnel oxide thickness (EOT-nm) [B]	**7–8**	**7–8**	**7–8**	**6–7**	**6–7**	**6–7**	**6–7**
Flash tunnel oxide thickness control EOT (% 3σ) [C]	**<±3.5**	**<±3**	**<±3**	**<±3**	**<±3**	**<±3**	**<±2.5**
Flash tunnel oxide minimum QBD @ 1× 10^{-2} A/cm^2 (C/cm^2) [D]	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.3**	**0.3**
Flash tunnel oxide defectivity at minimum QBD (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 DVT (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**	**11–13**	**11–13**	**11–13**	**10–12**	**10–12**	**10–12**
Flash NAND interpoly dielectric thickness (EOT-nm) [J]	**13–15**	**13–15**	**13–15**	**13–15**	**10–13**	**10–13**	**10–13**
Flash interpoly dielectric thickness control EOT (% 3s) [K]	**<±6**	**<±6**	**<±6**	**<±6**	**<±6**	**<±6**	**<±6**
Flash interpoly dielectric T_{max} of formation t >5'/<5' (ºC) [L]	**750/900**	**750/900**	**750/900**	**750/900**	**700/850**	**700/850**	**700/850**
Flash interpoly dielectric conformality on floating gate EOT_{min}/EOT_{max} [M]	**>0.95**	**>0.95**	**>0.95**	**>0.98**	**>0.98**	**>0.98**	**>0.98**
Flash maximum charge loss 10 years at room temperature (V) – single/dual bit (%) [N]	**20/10**	**20/10**	**20/10**	**20/10**	**20/10**	**20/10**	**20/10**

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known



Notes for Tables 74a and 74b:

[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.

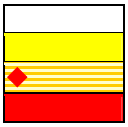
[M] Uniform step coverage is important for charge retention.

[N] To assure device functionality.

Table 74b FLASH Non-volatile Memory Technology Requirements—Long-term

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*
Technology Node	*hp45*		*hp32*		*hp22*	
DRAM ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*
MPU/ASIC ½ Pitch (nm)	**45**	**35**	**32**	**25**	**22**	**18**
MPU Printed Gate Length (nm)	**25**	**20**	**18**	**14**	**13**	**10**
MPU Physical Gate Length (nm)	**18**	**14**	**13**	**10**	**9**	**7**
Flash technology node (nm) [A]	**45**	**35**	**32**	**25**	**22**	**18**
Flash NOR tunnel oxide thickness (EOT-nm) [B]	**8–9**	**8–9**	**8–9**	**8**	**8**	**8**
Flash NAND tunnel oxide thickness (EOT-nm) [B]	**6–7**	**6–7**	**6–7**	**6–7**	**6–7**	**6–7**
Flash tunnel oxide thickness control EOT (% 3σ) [C]	**<±2.5**	**<±2.5**	**<±2.5**	**<±2.5**	**<±2.5**	**<±2.5**
Flash tunnel oxide minimum QBD at 1× 10^{-2} A/cm^2 (C/cm^2) [D]	**0.3**	**0.3**	**0.3**	**0.4**	**0.4**	**0.4**
Flash tunnel oxide defectivity at minimum QBD (def/cm^2) [E]	**<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**
Flash program/erase window DVT (V) [G]	**>3**	**>3**	**>3**	**>3**	**>3**	**>3**
Flash erase time degradation t_{max}/t_0 [H]	**<2**	**<2**	**<2**	**<2**	**<2**	**<2**
Flash program time degradation t_{max}/t_0 [I]	**<2**	**<2**	**<2**	**<2**	**<2**	**<2**
Flash NOR interpoly dielectric thickness (EOT-nm) [J]	**8–10**	**8–10**	**8–10**	**6–8**	**6–8**	**6–8**
Flash NAND interpoly dielectric thickness (EOT-nm) [J]	**10–13**	**10–13**	**9–10**	**9–10**	**9–10**	**9–10**
Flash interpoly dielectric thickness control EOT (% 3σ) [K]	**<±5**	**<±5**	**<±5**	**<±5**	**<±5**	**<±5**
Flash interpoly dielectric T_{max} of formation t >5' / < 5' (ºC) [L]	**650/800**	**650/800**	**650/800**	**600/700**	**600/700**	**600/700**
Flash interpoly dielectric conformality on floating gate EOT_{min}/EOT_{max} [M]	**>0.98**	**>0.98**	**>0.98**	**>0.98**	**>0.98**	**>0.98**
Flash maximum charge loss 10 years at room temperature (V) – single/dual bit (%) [N]	**20/10**	**20/10**	**20/10**	**20/10**	**20/10**	**20/10**

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known



FERROELECTRIC RANDOM ACCESS MEMORY (FERAM)

FeRAM was a new addition to the 2001 ITRS, and was the result of collaboration between the FEP and PIDS technology working groups. The critical requirements tables, Tables 75a and 75b and potential solutions roadmap, Figure 52, were updated using information gathered from conference papers and other sources together with current information provided by FeRAM researchers.

Historically speaking, FeRAM devices had been proposed much earlier than semiconductor memory devices.[24] At present however, memory capacity is limited to ~1/1000 that of commodity DRAM, due to limited ferroelectric film reliability and to difficulties associated with capacitor fabrication. These difficulties together with the lack of a "killer application" has constrained commercial production. The technology tables are therefore based on *demonstration samples* and not on

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

volume manufactured products, in order to show the technology level. 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 2003 to 2018 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

Table 75a shows a memory capacity of 32Mb for year 2003 demonstration samples.[25] In the longer term, beyond 2005, 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 (2003–2005), it is forecasted that FeRAM scaling will occur at an accelerated rate, with a yearly doubling and concurrent 0.7× reduction. This forecasted scaling rate yields a 128 Mb standard memory capacity in 2005, demonstrated using 130 nm technology. Since it is strongly applications dependent, embedded capacity is not shown, but it is assumed to be ¼ that of standard memory.

CELL SIZE

Currently, the mainstream cell structure is the One Transistor-One Capacitor (1T-1C) cell and is replacing the 2T-2C cell that was needed to ensure stable data read out. The 1T-1C configuration is mandatory for the realization of large capacity FeRAM. As far as the capacitor structure is concerned, the change from the planar capacitor type to a stack configuration has resulted in a 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 Tables 75a and 75b. The above-mentioned cell structure and capacitor configuration changes are forecasted to reduce the cell area factor to ten in years 2008 and 2009 after which the cell area factor is fixed at the value of eight for the period 2010 to 2018 for standard memory. Smaller cell size factors such as six 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.[26] 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 that 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 deposition 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 A reported MOCVD study has shown that a (111) oriented PZT film is very effective at yielding an improved switching charge.[27] Etching of capacitor electrodes is very difficult to do 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. High temperature etching for improving sidewall slope of the capacitor is thus being developed to overcome this difficulty.[27]

---

[25] *Y. J. Song, H. J. Joo, N. W. Jang, H. H. Kim, J. H. Park, H. Y. Kang, S. Y. Lee, and Kinam Kim, 12B-1, Symposium on VLSI Technology, 2003.*

[26] *D. J. Wouters, 28, International Conference on Solid State Devices and Materials, 2003.*

[27] *Y. Horii, Y. Hikosaka, A. Itoh, K. Matsuura, M. Kurasawa, G. Komuro, K. Maruyama, T. Eshita and S. Kashiwagi, 539, IEDM, 2002.*

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. Besides PZT and SBT, one of the promising new materials is BLT or $(Bi,La)_4Ti_3O_{12}$,[28] of which characteristics are between the foregoing two.[29] Since the characteristic of each film has been improved by efforts in recent years, it seems to be more important to master the material rather than selection.

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 1 Gb or 0.18 µm node. Based on this data with the further assumption that bitline capacitance is proportional to $F^{2/3}$, where F is the feature size[30] allows for the calculation of $\Delta V_{bitline}$. The $\Delta V_{bitline}$ is about 140 mV, and we assume that this is needed for the sense amplifier circuit independent on technology nodes. Multiplying $\Delta V_{bitline}$ (140 mV) 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.

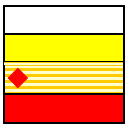
The FeRAM forecast of Tables 75a and 75b is based on these assumptions and calculations. “Red brick walls” begin to appear in 2006 at the earliest, and become more widespread in 2007. The first priority to break through these walls is the development of highly reliable ferroelectric materials that exhibit negligible post-process degradation.

[28] *B. H. Park, B. S. Kang, S.D. Bu, T. W. Noh, J. Lee, and W. Jo, 682, Nature, 1999.*
[29] *D. J. Wouters, 28, International Conference on Solid State Devices and Materials, 2003.*
[30] *A. Nitayama, Y. Kohyama, ad K. Hieda, 355, IEDM, 1998.*

Table 75a FeRAM Technology Requirements—Near-term

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*
Technology Node		*hp90*			*hp65*		
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*
MPU/ASIC ½ Pitch (nm)	**107**	**90**	**80**	**70**	**65**	**57**	**50**
MPU Printed Gate Length (nm)	**65**	**53**	**45**	**40**	**35**	**32**	**28**
MPU Physical Gate Length (nm)	**45**	**37**	**32**	**28**	**25**	**22**	**20**
Demonstration Sample Density (standard memory) [A]	**32Mb**	**64Mb**	**128Mb**	**256Mb**	**256Mb**	**512Mb**	**512Mb**
Feature Size (µm): F1 (standard memory) [B]	**0.25**	**0.18**	**0.13**	**0.13**	**0.12**	**0.11**	**0.10**
Feature Size (µm): F2 (embedded memory) [C]	**0.18**	**0.13**	**0.12**	**0.11**	**0.10**	**0.09**	**0.08**
Access time (ns) [D]	**55**	**40**	**30**	**30**	**20**	**20**	**15**
Cycle time (ns) [E]	**80**	**70**	**50**	**50**	**32**	**32**	**25**
Cell area factor a (standard memory) [F]	**15**	**15**	**12**	**12**	**12**	**10**	**10**
Cell size (μm^2) (standard memory) [G]	**0.938**	**0.486**	**0.203**	**0.203**	**0.173**	**0.121**	**0.100**
Total cell area (mm^2) for standard memory [H]	**3.93**	**8.15**	**6.80**	**13.61**	**23.19**	**32.48**	**53.69**
Cell area factor: b (embedded memory) [I]	**46**	**35**	**35**	**35**	**30**	**30**	**30**
Cell size (μm^2) for embedded memory [J]	**1.490**	**0.592**	**0.504**	**0.424**	**0.300**	**0.243**	**0.192**
Projected capacitor size (μm^2) [K]	**0.44**	**0.23**	**0.12**	**0.12**	**0.058**	**0.048**	**0.040**
Capacitor area (μm^2) [L]	**0.44**	**0.23**	**0.12**	**0.12**	**0.086**	**0.081**	**0.076**
Cap area/proj cap size [M]	**1**	**1**	**1**	**1**	**1.48**	**1.67**	**1.89**
Height of bottom electrode/F (for 3D capacitor) [N]	**NA**	**NA**	**NA**	**NA**	**0.24**	**0.33**	**0.45**
Capacitor structure [O]	**stack**	**stack**	**stack**	**stack**	**3D**	**3D**	**3D**
Cell Structure [P]	**1T1C**	**1T1C**	**1T1C**	**1T1C**	**1T1C**	**1T1C**	**1T1C**
Vop (Volt) [Q]	**2.5-3**	**1.8**	**1.5**	**1.5**	**1.2**	**1.2**	**1.2**
Minimum switching charge density ($\mu C/cm^2$) at Vop [R]	**12.8**	**19.8**	**30.5**	**30.5**	**40**	**40**	**40**
Minimum switching charge per cell (fC/cell) at Vop [S]	**55.8**	**44.8**	**36.1**	**36.1**	**34.2**	**32.3**	**30.3**
Retention at 85°C (Years) [T]	**10 Years**	**10 Years**	**10 Years**	**10 Years**	**10 Years**	**10 Years**	**10 Years**
Endurance [U]	**1.0 E13**	**1.0 E14**	**1.0 E15**	**>1.0 E16**	**>1.0 E16**	**>1.0 E16**	**>1.0 E16**

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known



Notes for Tables 75a and 75b:

[A] Embedded memory strongly depends on applications, only standard memory is thus shown here. 32 Mb at 2003: Samsung's 0.25 µm standard memory demonstrator.

[B] Feature size "F1" is defined as the critical dimension in the cell. 0.25 µm at 2003 is the value for Samsung's cell, a typical standard memory cell

[C] Feature size "F2" is defined as the critical dimension for logic area. 0.18 µm at 2003: Matsushita (sample), 0.18 µm at 2004:Fujitsu's 4Mb (mass production).

[D] 55ns at 2003: Samsung's 32Mb. Fujitsu's 0.18 µm -4Mb: 30ns, Matsushita's 0.18 µm -1Mb: 15ns.

[E] Not referenced.

[F] a = Cell size/F1^2.

[G] Cell size = a F1^2.*

*[H] Cell area*memory size (bit).*

[I] b = Cell size/F2^2.

[J] Cell size = b F2^2.*

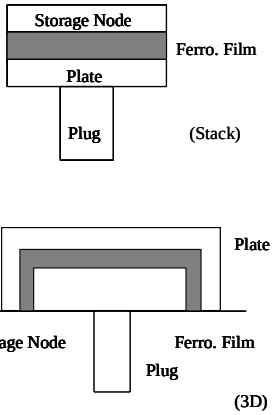
[K] 2003–2006: 7F1^2, 2007-2009: 4F1^2, 2010-2018: 3F1^2 are assumed.

[L] 3D is assumed to be a pedestal structure.

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

*[N] For instance, 0.24 means that the height is 0.24*F1.*

[O] See figures (right).



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

[Q] V_{op}=operational voltage. Low voltage operation is a key issue. Matsushita's 0.18 µm sample with SBT at 2003: 1.1V.

[R] This value can be calculated by [S] divided by [L]. This value is assumed to be 40 for 3D.

*[S] Calculated by $DV_{bitline}$*$C_{bitline}$ with the assumptions that $DV_{bitline}$=140 mV is needed and $C_{bitline}$ is as same as DRAM.*

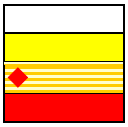
[T] Depends on applications. 85ºC comes from the specifications for IC cards.

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

Table 75b FeRAM Technology Requirements—Long-term

Year of Production	*2010*	*2012*	*2013*	*2015*	*2016*	*2018*
Technology Node	*hp45*		*hp32*		*hp22*	
DRAM ½ Pitch (nm)	*45*	*35*	*32*	*25*	*22*	*18*
MPU/ASIC ½ Pitch (nm)	**45**	**35**	**32**	**25**	**22**	**18**
MPU Printed Gate Length (nm)	**25**	**20**	**18**	**14**	**13**	**10**
MPU Physical Gate Length (nm)	**18**	**14**	**13**	**10**	**9**	**7**
FeRAM Generation (standard memory) [A]	**1Gb**	**2Gb**	**4Gb**	**8Gb**	**16Gb**	**32Gb**
Feature Size (µm): F1 (standard memory) [B]	**0.09**	**0.07**	**0.065**	**0.050**	**0.045**	**0.035**
Feature Size (µm): F2(embedded memory) [C]	**0.07**	**0.057**	**0.050**	**0.04**	**0.035**	**0.028**
Access time (ns) [D]	**10**	**10**	**8**	**8**	**6**	**6**
Cycle time (ns) [E]	**16**	**16**	**12**	**12**	**10**	**10**
Cell area factor a (standard memory) [F]	**8**	**8**	**8**	**8**	**8**	**8**
Cell size (μm^2) (standard memory) [G]	**0.065**	**0.039**	**0.034**	**0.020**	**0.016**	**0.010**
Total cell area (mm^2) for standard memory [H]	**69.58**	**84.18**	**145.17**	**171.80**	**278.31**	**336.73**
Cell area factor: b (embedded memory) [I]	**24**	**24**	**20**	**20**	**16**	**16**
Cell size (μm^2) for embedded memory [J]	**0.118**	**0.078**	**0.050**	**0.032**	**0.020**	**0.013**
Projected capacitor size (μm^2) [K]	**0.024**	**0.015**	**0.013**	**0.008**	**0.0061**	**0.0037**
Capacitor area (μm^2) [L]	**0.071**	**0.06**	**0.057**	**0.048**	**0.045**	**0.038**
Cap area/proj cap size [M]	**2.90**	**4.06**	**4.48**	**6.37**	**7.33**	**10.20**
Height of bottom electrode/F (for 3D capacitor) [N]	**0.71**	**1.15**	**1.30**	**2.01**	**2.37**	**3.45**
Capacitor structure [O]	**3D**	**3D**	**3D**	**3D**	**3D**	**3D**
Cell Structure [P]	**1T1C**	**1T1C**	**1T1C**	**1T1C**	**1T1C**	**1T1C**
Vop (Volt) [Q]	**1.0**	**1.0**	**0.7**	**0.7**	**0.7**	**0.7**
Minimum switching charge density ($\mu C/cm^2$) at Vop [R]	**40**	**40**	**40**	**40**	**40**	**40**
Minimum switching charge per cell (fC/cell) at Vop [S]	**28.2**	**23.9**	**22.7**	**19.1**	**17.8**	**15.0**
Retention at 85°C (Years)[T]	**10 Years**	**10 Years**	**10 Years**	**10 Years**	**10 Years**	**10 Years**
Endurance [U]	**>1.0 E16**	**>1.0 E16**	**>1.0 E16**	**>1.0 E16**	**>1.0 E16**	**>1.0 E16**

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known



Year of First Product Shipment Technology Node	2003	2004 hp90	2005	2006	2007 hp65	2008	2009	2010 hp45	2012	2013 hp32	2015	2016 hp22	2018
Ferroelectric Materials	PZT, SBT, BLT				PZT, SBT, BLT, New Materials								
Deposition Methods	PVD, CSD, MOCVD				MOCVD, New Methods								

CSD—Chemical Solution Deposition

*PZT—Pb(Zr, Ti)O_3** *SBT—$SrBi_2Ta_2O_9$* *BLT—$(Bi, La)_4Ti_3O_{12}$*

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

Figure 52 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. Some new ideas such as Non-destructive Read-out scheme, which is free from the limitation of read/write cycles, are being investigated to overcome endurance issues.

Presently, the lack of "killer application" as well as reliability and cost are the greatest barriers to widespread use of FeRAM devices. For these reasons FeRAM applications are mainly limited to embedded memory for 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 or even greater than that of DRAM. This increase 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.

CROSS-CUT ISSUES

FEP METROLOGY CROSS-CUT ISSUES

Advanced gate stack, wafer cleaning, and doping process technologies as well as starting materials measurement requirements continue to challenge existing metrology capability. This is highlighted by the near term and long term *Metrology Challenges Table* in the Metrology Roadmap and by the discussion of FEP Metrology in the Metrology Roadmap. The *FEP Metrology Technology Requirements Table* lists measurement precision for gate dielectric thickness and other FEP films and processes, and the FEP Metrology section indicates that meeting the stated precision is a difficult goal. It is important to note that interfacial measurements for control of gate dielectric processes will be very difficult. The Key Metrology Challenges that are based on FEP needs are:

- Impurity detection (especially particles) at levels of interest for starting materials and reduced edge exclusion for metrology tools. Control of gettering in SOI substrates.
- Measurement of complex material stacks and interfacial properties including physical and electrical properties. The specific FEP requirement is for measurement of high-κ gate stacks including metal gates and interface process control. The addition of SOI and strained silicon requires development of metrology capability.
- 3D dopant profiling

FEP MODELING AND SIMULATION CROSS-CUT ISSUES

In the 2003 ITRS the FEP challenges surround the introduction of new materials and non-standard dual gate MOSFETs. This raises various requirements on *Modeling and Simulation*. Especially, material issues need to be addressed in most modeling areas. This includes among others strained materials, so the importance of modeling of stress and strain is further growing. New device architectures request especially large progress in numerical device simulation, together with improvements of the simulation of the process steps used to fabricate these devices, e.g., the formation of shallow junctions. Process variants are getting increasingly important as device further scale—a premier example is the redistribution of variance allowance between lithography and etching in this roadmap—and simulation can and must contribute to assessing the impact of such variants on the final device and chip. High-κ dielectrics are required to be introduced before 2007, so modeling must be able to appropriately describe them as soon as possible. The formation of ultra-shallow, abrupt, highly activated drain extensions continues to be a major challenge, and support from modeling is required both to improve the physical understanding for the processes used (e.g., kinetics of dopants and point defects during annealing) and to subsequently optimize them by numerical simulation. This knowledge is also needed for defect engineering, which aims at achieving shallower junctions by the exploitation of the interaction between dopant atoms and

defects. Furthermore, the reduction of critical dimensions (CD) and their variations are generally a key issue, and it is highly desirable to use simulation to identify among the many parameters influencing CD the most important ones, in order to minimize experimental effort.

FEP ENVIRONMENT, SAFETY, AND HEALTH CROSS-CUT ISSUES

Refer to the *Environment, Safety, and Health* chapter for comprehensive information and link to a new chemical screening tool (Chemical Restrictions Table).

INTER-FOCUS ITWG DISCUSSION

It should be evident that FEP shares numerous issues and dependencies with other Focus ITWGs. Chief among those are issues surrounding gate EOT and leakage requirements with the PIDS and to some extent the Design ITWGs. Other issues with these ITWGs revolve around junction depth and sheet resistance requirements as well as requirements driven by alternate device structures. Resolution of these issues is generally attained through compromise and trade-offs. CD control remains an issue that was not resolved to everyone's satisfaction in this edition of the ITRS and must be resolved in the 2004 update by cooperation among the FEP, Lithography, PIDS and Design ITWGs. It is further expected that there will be more extensive discussions with PIDS regarding the modeling of other non-traditional MOSFETS such as FinFET and other multi-gate transistors. Other interactions include those with the Yield Enhancement ITWG to validate different statistical defect models. A very important interaction was with the Interconnect ITWG where members of the FEP surface preparation team provided technical support in the development of interconnect surface preparation and cleaning technical requirements and potential solutions.

IMPACT OF FUTURE EMERGING RESEARCH DEVICES

This 2003 FEP chapter marks the first time where device alternatives to planar bulk CMOS have been explicitly incorporated into the FEP technical requirements tables. It has been increasingly recognized that bulk CMOS faces significant challenges that threaten to prevent the historical 17% per annum increase in logic transistor speed. Various device alternatives are discussed in *the Emerging Research Devices* chapter, where the following sequence of "technology boosters" are presented as follows:

- Strained silicon channels
- Ultra-thin body (fully depleted) single gate devices (with elevated contacts) (FDSOI)
- High-κ gate dielectric layers
- Metallic gates
- Ultra-thin body (fully depleted) double gate devices
- Ballistic or quasi-ballistic carrier transport
- Reduced fringing (and/or gate-drain overlap) capacitance
- Metallic source/drain junction

Some of the requirements associated with this scenario have been discussed and modeled in the thermal films and doping sections. Specifically some requirements related to strained silicon, FD SOI, and FD Double gate devices have been modeled. The modeled double gate device is one where the channel lies in the plane of the wafer, with the gates above and below the channel (See doping supplemental material worksheet XjRs.) Another double, or multiple gate device alternative where the channel is vertically disposed with respect to the plane of the wafer, with gates on each side, e.g. FinFET, or Tri-gate has not yet been modeled, but will be addressed for the 2004 ITRS update. One important issue that requires further address is the entire area of lateral abruptness of the drain junction at the channel, as well as the degree of gate/drain overlap. This is currently an active area of research and will also be addressed in the next chapter edition.

The introduction time of these non-classical devices is pegged at 2008 with the introduction of planar single gate ultra-thin body devices, having elevated contacts, followed by double- or multiple-gate devices in 2011. However in the years prior, bulk CMOS will require significant enhancement through the introduction of devices having strained silicon channels, high-κ gate dielectric layers, and metal gates. Extensive research and development is required to execute this scenario. The traditional pace of transistor performance improvement requires this investment, which if not made will result in the inevitable slowing of device performance.

Memory devices are similarly challenged, as is evident from the "red walls" that show up in the requirements tables for DRAM, Flash and FeRAM memories. In addition to significant materials requirements, memory structures are expected to fundamentally change in order that the storage bit density and cost/bit may continue to progress at historical rates. This is another difficult challenge articulated in the Emerging Research Devices chapter, where the "realization of a manufacturable, cost-effective fabrication technology for electrically accessible high-speed, high-density, non-volatile RAM is identified, that must be integratable with CMOS logic process flows." The research and development challenges associated with continued scaling of traditional memory structures, and the introduction of new memory devices are as significant as those with the continued scaling of logic transistors, and will require equivalent investments in monetary and intellectual capital.

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 processes suitable for the cost-effective manufacture of these next generation devices. When one considers the challenges highlighted in the FEP chapter along with the potential impacts of Emerging Research Devices that will soon fall within the Roadmap time horizon, FEP is indeed a rich area for research and innovation.