

FOREWORD

The International Technology Roadmap for Semiconductors (ITRS) is the result of a worldwide consensus building process. This document predicts the main trends in the semiconductor industry spanning across 15 years into the future. The participation of experts from Europe, Japan, Korea, and Taiwan as well as the U.S.A. ensures that the *ITRS* is a valid source of guidance for the semiconductor industry as we strive to extend the historical advancement of semiconductor technology and the worldwide integrated circuit (IC) market. These five regions jointly sponsor the *ITRS*.

The Semiconductor Industry Association (SIA) coordinated the first efforts of producing what was originally *The National Technology Roadmap for Semiconductors (NTRS)*. The *NTRS* provided a 15-year outlook on the major trends of the semiconductor industry. As such, it was a good reference document for all semiconductor manufacturers. Most of all, the *NTRS* documents provided useful guidance for suppliers of equipment, materials, and software and clear targets for researchers in the outer years.

The semiconductor industry became a global industry in the 1990s, as many semiconductor chip manufacturers established manufacturing or assembly facilities in multiple regions of the world. Similarly, the suppliers to the semiconductor industry have established worldwide operations. Furthermore, alliances, joint ventures, and many forms of cooperation have been established among semiconductor manufacturers as well as among equipment, materials, and software suppliers.

The above considerations led to the realization that a Roadmap that provides guidance for the whole industry would benefit from inputs from all regions of the world that have leadership activities in the field of semiconductors. This realization has led to the creation of the *International Technology Roadmap for Semiconductors (ITRS)*. The invitation to cooperate on the *ITRS* was extended by the SIA at the World Semiconductor Council in April of 1998 to Europe, Japan, Korea, and Taiwan. The initial collaboration of these five regions produced the *ITRS 1998 Update*, which consisted of a comprehensive revision of the 1997 *NTRS* technology requirements tables. Subsequently, the five regions jointly produced *The International Technology Roadmap for Semiconductors, 1999 Edition* and jointly sponsored *The International Technology Roadmap for Semiconductor, 2000 Update*.

The most recent full revision occurred in 2001. The *2001 ITRS* presented the latest prediction of the industry needs and potential solutions through 2016. Additionally, it also began to address the limits of traditional scaling, the opportunity for “equivalent scaling” [the extension of the device scaling approach by improving electrical performance with new or improved materials], the challenges of the industry to maintain historical trends in productivity, and the introduction of the investigation and feasibility of new device candidates within the timeframe of the Roadmap (2013-2016).

In 2002, the *ITRS* international technology working groups (TWGs) reviewed the tables of the *2001 ITRS* edition. The resulting update and clarification produced the *2002 ITRS Update*. Over 100 tables are revised. Table line items are corrected to correlate with other TWG chapters. Several tables now contain new line items or indicate a change in the difficulty of achieving a particular requirement by indicating a color change in the cell.

For the first time, the *ITRS* does not predict a further acceleration in the timing of introduction of new technologies as the industry struggles through the worst recession of its history. The *2002 ITRS Update* table headers and *Overall Roadmap Technology Characteristics (ORTC)* tables 1a–1j remain unchanged with the original timing as announced in the *2001 ITRS*. However, once recovery is fully underway, the progress made by companies during these trying times will be realized. The *2003 ITRS* effort, which will be a complete review and rewrite, will indicate whether this pace is a new trend or just a moment of reflection within the industry during the phase of early recovery.

It is the purpose of the *ITRS* documents to provide a reference of requirements, potential solutions, and their timing for the semiconductor industry. This objective has been accomplished by providing a forum for international discussion, cooperation, and agreement among the leading semiconductor manufacturers and the leading suppliers of equipment, materials, and software, as well as researchers from university, consortia, and government labs.

In the last few years, the *ITRS* documents have become a truly common reference for the entire semiconductor industry. Indeed, the cooperative efforts of the *ITRS* participants have fostered cooperation among international consortia, universities, and research institutions around the world. It is hoped that the *2002 ITRS Update* will further contribute to stimulate cooperative R&D investments so that the financial burden can be more uniformly shared by the whole industry. It is also hoped that the *2002 ITRS Update* will continue to stimulate the fundamental elements that encourage innovation in individual companies.

TABLE OF CONTENTS

Foreword.....	i
Introduction [from the 2001 ITRS].....	1
Overview.....	1
Meaning of <i>ITRS</i> Technology Requirements.....	1
Position on Potential Solutions	2
Overall Roadmap Process and Structure	2
Technology Characteristics / Requirements Tables	4
Technology Nodes	6
Drivers for ITWG Technology Requirements	7
Grand Challenges.....	9
Overview.....	9
In the Near Term (through 2007)	9
Enhancing Performance.....	9
Cost-effective Manufacturing	11
In The Long Term (2008 through 2016).....	14
Enhancing Performance.....	14
Cost-effective Manufacturing	14
2002 Update Tables.....	16
Design	16
Test.....	19
Process Integration, Devices, and Structures.....	29
Front End Processes	45
Lithography	63
Interconnect.....	74
Factory Integration.....	79
Assembly and Packaging	100
Environment, Safety, and Health	112
Yield Enhancement.....	117
Metrology	134
Modeling and Simulation	143
Overall Roadmap Technology Characteristics	150
Background.....	150
Overview of 2001 Revisions	150
Definitions.....	150
Roadmap Timeline	150
Product Generations and Chip-Size Model.....	154
Chip-Size, Lithographic-Field, and Wafer-Size Trends.....	161
Performance of Packaged Chips	164
Electrical Defect Density	169
Power Supply and Power Dissipation	170
Cost	171

Glossary	174
Characteristics of Major Markets	174
Chip and Package—Physical and Electrical Attributes	176
Fabrication Attributes And Methods	177
Maximum Substrate Diameter (mm)	177
Electrical Design And Test Metrics	177
Design and Test	177
Add	177
Technology Requirements Solutions	177

LIST OF FIGURES

Figure 1 Composition of the Technology Working Group (TWG)	3
Figure 2 A Typical Production “Ramp” Curve	5
Figure 3 Technology Node Compared to Actual Production Capacity Technology Distribution	6
Figure 4 Definition of the Half Pitch and Gate Length	7
Figure 5 MOS Transistor Scaling—1974 to present	152
Figure 6 Scaling Calculator	152
Figure 7 ITRS Roadmap Acceleration Continues—Half Pitch Trends	153
Figure 8 ITRS Roadmap Acceleration Continues—Gate Length Trends	153

LIST OF INTRODUCTION TABLES

Table A Improvement Trends for ICs Enabled by Feature Scaling	1
Table B ITRS Table Structure—Key Lithography-Related Characteristics by Product Type	4
Table C Rounded versus Actual Trend Numbers	151

LIST OF OVERALL ROADMAP TECHNOLOGY CHARACTERISTICS TABLES

Table 1a	Product Generations and Chip Size Model Technology Nodes— Near-term Years	155
Table 1b	Product Generations and Chip Size Model Technology Nodes— Long-term years.....	155
Table 1c	DRAM Production Product Generations and Chip Size Model— Near-term Years	156
Table 1d	DRAM Production Product Generations and Chip Size Model— Long-term Years	156
Table 1e	DRAM Introduction Product Generations and Chip Size Model— Near-term Years	157
Table 1f	DRAM Introduction Product Generations and Chip Size Model— Long-term Years	157
Table 1g	MPU (High-volume Microprocessor) Cost-Performance Product Generations and Chip Size Model—Near-term Years	158
Table 1h	MPU (High-volume Microprocessor) Cost-Performance Product Generations and Chip Size Model—Long-term Years	159
Table 1i	High-Performance MPU and ASIC Product Generations and Chip Size Model—Near-term Years	160
Table 1j	High-Performance MPU and ASIC Product Generations and Chip Size Model—Long-term Years	160
Table 2a	Lithographic-Field and Wafer-Size Trends—Near-term Years	163
Table 2b	Lithographic-Field and Wafer Size Trends—Long-term Years	163
Table 3a	Performance of Packaged Chips: Number of Pads and Pins—Near-term Years.....	164
Table 3b	Performance of Packaged Chips: Number of Pads and Pins—Long-term Years	165
Table 4a	Performance and Package Chips: Pads, Cost— Near-term Years [Update].....	166
Table 4b	Performance and Package Chips: Pads, Cost— Long-term Years [Update].....	167
Table 4c	Performance and Package Chips: Frequency On-Chip Wiring Levels— Near-term Years [Update].....	167
Table 4d	Performance and Package Chips: Frequency, On-Chip Wiring Levels— Long-term Years [Update].....	168
Table 5a	Electrical Defects—Near-term Years.....	169
Table 5b	Electrical Defects—Long-term Years.....	169
Table 6a	Power Supply and Power Dissipation—Near-term Years.....	170
Table 6b	Power Supply and Power Dissipation—Long-term Years	170
Table 7a	Cost—Near-term Years	172
Table 7b	Cost—Long-term Years	173

LIST OF DIFFICULT CHALLENGES TABLES

Table 13 Design Difficult Challenges	16
Table 19 Test Difficult Challenges	19
Table 34 Process Integration Difficult Challenges [<i>Update</i>]	29
Table 48 Front End Processes Difficult Challenges	45
Table 56 Lithography Difficult Challenges	63
Table 61 Interconnect Difficult Challenges [<i>Update</i>]	74
Table 65 Factory Integration Difficult Challenges [<i>Update</i>]	79
Table 74 Assembly and Packaging Difficult Challenges [<i>Update</i>]	100
Table 82 Environment, Safety, and Health Difficult Challenges	112
Table 88 Yield Enhancement Difficult Challenges	117
Table 96 Metrology Difficult Challenges	134
Table 102 Modeling and Simulation Difficult Challenges [<i>Update</i>]	143

LIST OF 2002 ITRS *UPDATED* TECHNOLOGY REQUIREMENTS TABLES

SYSTEM DRIVERS

Tables 8–12 are unchanged for the 2002 Update [Refer to the [2001 ITRS System Drivers](#) chapter]

Figure 10 First Integration of Technologies on SOC with Standard CMOS Process.....	17
--	----

DESIGN

Tables 13–18 are unchanged for the 2002 Update. [Refer to the [2001 ITRS Design](#) chapter]

TEST

Tables 19–22, 24–27, 29, 31, and 33 are unchanged for the 2002 Update [Refer to the [2001 ITRS Test](#) chapter]

Table 23a High Frequency Serial Communications Test Requirements—Near-term.....	21
Table 28 DFT-BIST Device Test Requirements—Near-term**	22
Table 30a Commodity Flash Memory Test Requirements—Near-term.....	25
Table 30b Commodity Flash Memory Test Requirements—Long-term	26
Table 32a Burn-in Requirements—Near-term	27
Table 32b Burn-in Requirements—Long-term	28

PROCESS INTEGRATION

Table 34 is the update of Process Integration Difficult Challenges

Tables 39-48 are unchanged for the 2002 Update [Refer to the 2001 ITRS Process Integration chapter]

Table 35a	High-performance Logic Technology Requirements—Near-term.....	31
Table 35b	High-performance Logic Technology Requirements—Long-term	32
Table 36a	Low Operating Power (LOP) Logic Technology Requirements—Near-term	35
Table 36b	Low Operating Power (LOP) Logic Technology Requirements—Long-term.....	36
Table 36c	Low Standby Power (LSTP) Technology Requirements—Near-term	37
Table 36d	Low Standby Power (LSTP) Technology Requirements—Long-term.....	38
Table 37a	DRAM Technology Requirements—Near-term	41
Table 37b	DRAM Technology Requirements—Long-term	41
Table 38a	Non-Volatile Memory Technology Requirements—Near-term	43
Table 38b	Non-Volatile Memory Technology Requirements—Long-term	44

FRONT END PROCESSES

Tables 48, 52, 55, and 56 are unchanged for the 2002 Update

[Refer to the 2001 ITRS Front End Processes chapter]

Table 49a	Starting Materials Technology Requirements—Near-term	48
Table 49b	Starting Materials Technology Requirements—Long-term.....	50
Table 50a	Surface Preparation Technology Requirements*—Near-term.....	52
Table 50b	Surface Preparation Technology Requirements*—Long-term	53
Table 51a	Thermal & Thin Film, Doping and Etching Technology Requirements—Near-term	56
Table 51b	Thermal & Thin Film, Doping and Etching Technology Requirements—Long-term.....	58
Table 53a	DRAM Trench Capacitor Technology Requirements—Near-term	60
Table 54a	FLASH Non-volatile Memory Technology Requirements—Near-term	61
Table 54b	FLASH Non-volatile Memory Technology Requirements—Long-term.....	62

LITHOGRAPHY

Tables 58 and 60 are unchanged for the 2002 Update [Refer to the [2001 ITRS Lithography chapter](#)]

Table 57a Lithography Technology Requirements—Near-term	64
Table 57b Lithography Technology Requirements—Long-term	65
Table 59a Optical Mask Requirements.....	66
Table 59b EUVL Mask Requirements	69
Table 59c EPL Mask Requirements	71

INTERCONNECT

Table 61 is the update of Interconnect Difficult Challenges

Table 64 is unchanged for the 2002 Update [Refer to the [2001 ITRS Interconnect chapter](#)]

Table 62a MPU Interconnect Technology Requirements—Near-term	75
Table 62b MPU Interconnect Technology Requirements—Long-term	76
Table 63a DRAM Interconnect Technology Requirements—Near-term.....	77
Table 63b DRAM Interconnect Technology Requirements—Long-term	78

FACTORY INTEGRATION

Table 65 is the update of Factory Integration Difficult Challenges

Tables 72–73 are unchanged for the 2002 Update [Refer to the [2001 ITRS Factory Integration chapter](#)]

Table 66a Factory Operations Technology Requirements—Near-term	81
Table 66b Factory Operations Technology Requirements—Long-term	82
Table 67a Production Equipment Technology Requirements—Near-term.....	84
Table 67b Production Equipment Technology Requirements—Long-term	85
Table 68a Material Handling Systems Technology Requirements—Near-term (40K WSPM) [7]	88
Table 68b Material Handling Systems Technology Requirements—Long-term (40K WSPM) [7]	89
Table 69a Factory Information and Control Systems Technology Requirements—Near-term[1].....	92
Table 69b Factory Information and Control Systems Technology Requirements—Long-term[1].....	93
Table 70a Facilities Technology Requirements—Near-term	95
Table 70b Facilities Technology Requirements—Long-term.....	96
Table 71a Probe/Test Manufacturing Technology Requirements—Near-term.....	98
Table 71b Probe/Test Manufacturing Technology Requirements—Long-term	98

ASSEMBLY AND PACKAGING

Table 74 is the update of Assembly and Packaging Difficult Challenges

Table 81 is unchanged for the 2002 Update [Refer to the [2001 ITRS Assembly and Packaging](#) chapter]

Table 75a Single-chip Packaging Technology Requirements—Near-term	102
Table 75b Single-chip Packaging Technology Requirements—Long-term	105
Table 76 Comparison between Bare Chip Mounting and WLP Technology.....	108
Table 77 Chip to Next Level Potential Solutions.....	109
Table 78 Ball Grid Array Packages Potential Solutions.....	110
Table 79 Single Chip Packages Potential Solutions.....	111
Table 80 BGA and FBGA/CSP Package Substrate Potential Solutions.....	111

ENVIRONMENT, SAFETY, AND HEALTH

Tables 82 and 85–87 are unchanged for the 2002 Update [Refer to the [2001 ITRS ESH](#) chapter]

Table 83a ESH Intrinsic Requirements—Near-term.....	113
Table 84a Chemicals, Materials and Equipment Management Technology Requirements—Near-term	115

YIELD ENHANCEMENT

Tables 88 and 92 are unchanged for the 2002 Update

[Refer to the [2001 ITRS Yield Enhancement](#) chapter]

Table 89 Defect Budget Technology Requirement Assumptions	118
Table 90 Yield Model and Defect Budget MPU Technology Requirements	119
Table 91 Yield Model and Defect Budget DRAM Technology Requirements.....	121
Table 93a Defect Detection Technology Requirements—Near-term	123
Table 93b Defect Detection Technology Requirements—Long-term	124
Table 94a Yield Learning Technology Requirements—Near-term	126
Table 94b Yield Learning Technology Requirements—Long-term.....	127
Table 95a Wafer Environmental Contamination Control Technology Requirements—Near-term	129
Table 95b Wafer Environmental Contamination Control Technology Requirements—Long-term.....	131

METROLOGY

Tables 96 and 99 are unchanged for the 2002 Update [Refer to the [2001 ITRS Metrology](#) chapter]

Table 97a Metrology Technology Requirements—Near-term	135
Table 97b Metrology Technology Requirements—Long-term	136
Table 98a Lithography Wafer Metrology Technology Requirements—Near-term	137
Table 98b Lithography Metrology Technology Requirements—Long-term	138
Table 100a Front End Processes Metrology Technology Requirements—Near-term	139
Table 100b Front End Processes Metrology Technology Requirements—Long-term	140
Table 101a Interconnect Metrology Technology Requirements—Near Term	141
Table 101b Interconnect Metrology Technology Requirements—Long-term	142

MODELING AND SIMULATION

**Table 102 is the update of Modeling and Simulation Difficult Challenges
[Refer to the [2001 ITRS Modeling and Simulation](#) chapter].**

Table 103a Modeling and Simulation Technology Requirements: Capabilities—Near-term	145
Table 103b Modeling and Simulation Technology Requirements: Accuracy and Speed—Near-term	147
Table 103c Modeling and Simulation Technology Requirements: Capabilities—Long-term	149

Contact Information for the ITRS	178
---	------------

INTRODUCTION [FROM THE 2001 ITRS]

OVERVIEW

For four decades, the semiconductor industry has distinguished itself by the rapid pace of improvement in its products. The principal categories of improvement trends are shown in Table A with examples of each. Most of these trends have resulted principally from the industry's ability to exponentially decrease the minimum feature sizes used to fabricate integrated circuits. Of course, the most frequently cited trend is in integration level, which is usually expressed as Moore's Law (the number of components per chip doubles every 18 months). The most significant trend for society is the decreasing cost-per-function, which has led to significant improvements of productivity and quality of life through proliferation of computers, electronic communication, and consumer electronics.

Table A Improvement Trends for ICs Enabled by Feature Scaling

<i>TREND</i>	<i>EXAMPLE</i>
<i>Integration Level</i>	Components/chip, Moore's Law
<i>Cost</i>	Cost per function
<i>Speed</i>	Microprocessor clock rate, GHz
<i>Power</i>	Laptop or cell phone battery life
<i>Compactness</i>	Small and light-weight products
<i>Functionality</i>	Nonvolatile memory, imager

All of these improvement trends, sometimes called "scaling" trends, have been enabled by large R&D investments. In the last two decades, the growing size of the required investments has motivated industry collaboration and spawned many R&D partnerships, consortia, and other cooperative ventures. *The International Technology Roadmap for Semiconductors (ITRS)* has been an especially successful worldwide cooperation. It presents an industry-wide consensus on the "best current estimate" of the industry's research and development needs out to a 15-year horizon. As such, it provides a guide to the efforts of companies, research organizations, and governments. The *ITRS* has improved the quality of R&D investment decisions made at all levels and has helped channel research efforts to areas that truly need research breakthroughs.

The 2001 edition of the *ITRS* is the result of the continued worldwide consensus building process. The participation of semiconductor experts from Europe, Japan, Korea, Taiwan, and U.S.A. ensures that the *2001 ITRS* continues to be the definitive source of guidance for semiconductor research as we strive to extend the historical advancement of semiconductor technology and the integrated circuit market. This is the second edition of *ITRS* that has had worldwide participation throughout its two-year cycle of creation. The diverse expertise and dedicated efforts that this international effort mobilized have brought the Roadmap to a new level of worldwide consensus about future semiconductor technology requirements.

The complete 2001 ITRS and past editions of the 2000 Update and the 1999 ITRS editions are available for viewing and printing as an electronic document at the internet web site <http://public.itrs.net>.

MEANING OF *ITRS* TECHNOLOGY REQUIREMENTS

Since its inception in 1992, a basic premise of the Roadmap has been that continued scaling of microelectronics would further reduce the cost per function (averaging ~25% per year) and promote market growth for integrated circuits (averaging ~17% per year). Thus, the Roadmap has been put together in the spirit of a challenge—essentially, "*What technical capabilities need to be developed for the industry to continue to stay on Moore's Law and the other trends?*" More and more of the semiconductor industry's research effort, including consortia and collaboration with suppliers, has been shared in a precompetitive environment. The *ITRS* identifies the principal technology needs to guide the shared research. It does this in the two following ways: (1) showing the "targets" that need to be met by "technology solutions" currently under development, and (2) indicating where there are no "known manufacturable solutions" (of reasonable

2 Introduction [from the 2001 ITRS]

confidence) to continued scaling in some aspect of the semiconductor technology. This latter situation is highlighted as red cells in the Roadmap technology requirements tables, and is also referred to as the “Red Brick Wall.” The red is officially on the Roadmap to clearly warn where historical trends of progress might end if some real breakthroughs are not achieved in the future. For some Roadmap readers, the red designation may not have adequately served its purpose of highlighting serious and exciting challenges. There can be a tendency to view any number on the Roadmap as “on the road to sure implementation” regardless of its color. To do so would be a serious mistake.

An analysis of “red” usage might classify the red parameters into the following two categories:

1. where the consensus is that the particular value will ultimately be achieved (perhaps late), but for which the industry does not have much confidence in any currently proposed solution(s), or
2. where the consensus is that the value will never be achieved (for example, some “work-around” will render it irrelevant or progress will indeed end).

To achieve the red parameters of the first category, breakthroughs in research are needed. It is hoped that such breakthroughs would result in the “red” turning to “yellow” (defined as “manufacturable solutions are known”) and, ultimately, “white” (defined as manufacturable solutions are known and are being optimized”) in future editions of *ITRS*. A conservative interpretation might view the red parameters of the second category as effectively “beyond” or “off” the Roadmap.

The *ITRS* time horizon (15 years) provides a limit to what may be considered “on/off the Roadmap.” To date, each edition of the *ITRS* has been built around a view toward continued scaling of CMOS (Complementary Metal-Oxide-Silicon) technology. However, with the 2001 edition, we are reaching the point where the horizon of the Roadmap challenges the most optimistic projections for continued scaling of CMOS (for example, MOSFET channel lengths of roughly 9 nm). It is also difficult for most people in the semiconductor industry to imagine how we could continue to afford the historic trends of increase in process equipment and factory costs for another 15 years! Thus, the 2001 *ITRS* begins to address post-CMOS devices.

Another constraint by which some items may be on/off the Roadmap is the breadth of technology addressed. The scope of the 2001 *ITRS* specifically addresses detailed technology requirements for all CMOS integrated circuits, including mixed-signal products. This group constitutes over 75% of the world’s semiconductor consumption. Of course, many of the same technologies used to manufacture CMOS ICs are also used for other products such as compound-semiconductor, discrete, optical, and micro-electromechanical systems (MEMS) devices. Thus, to a large extent, the Roadmap covers many common technology requirements for most IC-technology-based micro/nanotechnologies even though that is not the explicit purpose of the Roadmap.

POSITION ON POTENTIAL SOLUTIONS

The *ITRS* strives to avoid prematurely identifying definite solutions to the future technology challenges. This is difficult, since guidance on the research needs is intended. Despite this need to provide guidance, the Roadmap participants are continually pursuing new ways to prevent the Roadmap from being interpreted as limiting the range of creative approaches to further advance microelectronics technology. One of the resulting compromises has been to only present illustrative examples of potential solutions to selected challenges in the *ITRS*. These are not to be construed even as complete lists of all solutions suggested to date, much less exhaustive lists of what should be explored. A few of the potential technical solutions are listed, where known, only to inform the readers of current thinking and efforts. Furthermore, the listing of a particular potential solution does not constitute an endorsement by the Roadmap process.

It is the intent of this document to identify the technological barriers and when the industry will likely run into them. It is *not* the intent of this document to identify the most likely solutions to be adopted, nor to focus attention on those potential solutions currently known at the expense of other new concepts. In fact, it is eagerly hoped that this Roadmap will inspire additional innovative solutions. *The semiconductor industry’s future success continues to depend on new ideas.*

OVERALL ROADMAP PROCESS AND STRUCTURE

Each technology-area chapter of the *ITRS* is written by a corresponding International Technology Working Group (ITWG). The ITWGs are of two types: *Focus* ITWGs and *Crosscut* ITWGs. The Focus ITWGs correspond to typical sub-activities that sequentially span the Design/Process/Test/Package product flow for integrated circuits. The Crosscut

ITWGs represent important supporting activities that tend to individually overlap with the “product flow” at multiple critical points.

For the 2001 ITRS, the Focus ITWGs are the following:

- Design
- Test
- Process Integration, Devices, and Structures
- Front End Processes
- Lithography
- Interconnect
- Factory Integration
- Assembly and Packaging

Crosscut ITWGs are the following:

- Environment, Safety, and Health
- Yield Enhancement
- Metrology
- Modeling and Simulation

Each ITWG receive inputs from the regional Technology Working Groups (TWGs) of the five geographical regions (Europe, Japan, Korea, Taiwan, and the U.S.A.) One-to-two representatives from each regional TWG represent the regional TWG on the corresponding ITWG. The regional TWGs are composed of experts from industry (chip-makers as well as their equipment and materials suppliers), government research organizations, and universities. In 2001, a total of 839 experts volunteered their services in the twelve TWGs in five regions. The composition of the total TWG membership is analyzed in Figure 1. For this edition, three 2-day ITRS meetings in Grenoble, France (April, 2001), San Francisco, U.S.A. (July, 2001), and Santa Clara, U.S.A. (November, 2001) provided the main forums for face-to-face discussions among the members of each ITWG and coordination among the different ITWGs. In addition, each of the TWG incorporates feedback gathered from an even larger community through “sub-TWG meetings” and public “Roadmap Workshops.” The Roadmap resulting from this broad input is, hopefully, a “best-attempt” at building the widest possible consensus on the future technology needs of the semiconductor industry.

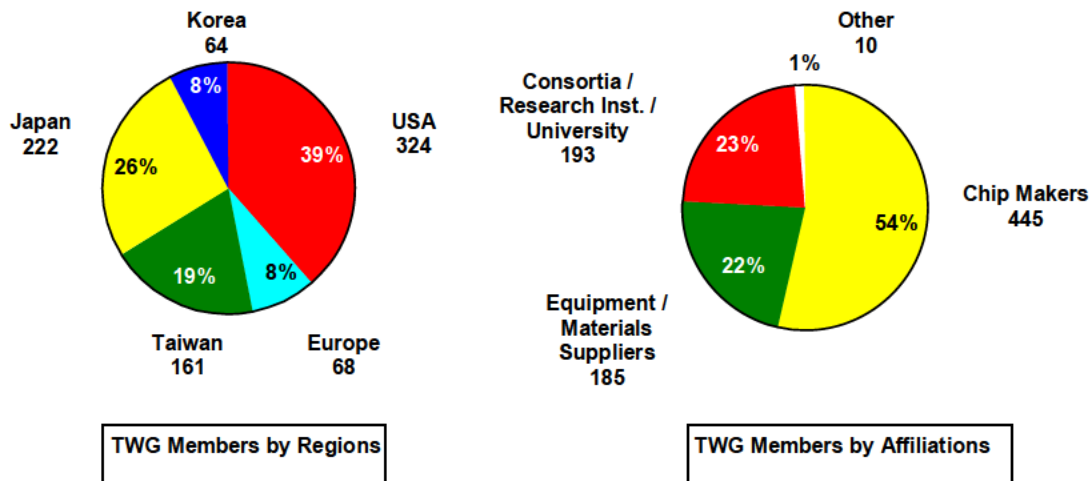


Figure 1 Composition of the Technology Working Group (TWG).

4 Introduction [from the 2001 ITRS]

Overall coordination of the *ITRS* process is the responsibility of the International Roadmap Committee (IRC), which has two-to-four members from each region (representing a regional coordinating committee such as the SIA Roadmap Coordinating Group for the U.S.A.). The principal IRC functions include

- Providing guidance/coordination for the ITWGs,
- Hosting the *ITRS* Workshops, and
- Editing the *ITRS*.

TECHNOLOGY CHARACTERISTICS / REQUIREMENTS TABLES

A central part of the IRC guidance/coordination is provided through the initial creation (as well as continued updating) of a set of Overall Roadmap Technology Characteristics (ORTC) tables. These tables summarize key high-level technology requirements, which define the future “technology nodes” and generally establish some common reference points to maintain consistency among the chapters written by individual ITWGs. The high-level targets expressed in the ORTC Tables are based, in part, on the compelling economic strategy of maintaining the historical high rate of advancement in integrated circuit technologies. Thus, the ORTC provide a “top-down business incentive” to balance the tendency for the ITWGs to become conservative in expressing their individual, detailed future requirements.

Each ITWG chapter contains several principal tables. They are individual ITWGs’ technology requirements tables patterned after the ORTC tables. For the *2001 ITRS*, the ORTC and technology requirements tables are separated into “Near-term Years” (2001, 2002... through 2007) and “Long-term Years” (2010, 2013, and 2016). This format is illustrated in Tables B, which contains a few key rows from lithography-related ORTC tables.

Table B ITRS Table Structure—Key Lithography-Related Characteristics by Product Type

Near-term Years

YEAR OF PRODUCTION	2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)	130	115	100	90	80	70	65
MPU ½ Pitch (nm)	150	130	107	90	80	70	65
MPU Printed Gate Length (nm)	90	75	65	53	45	40	35
MPU Physical Gate Length (nm)	65	53	45	37	32	28	25

Long-term Years

YEAR OF PRODUCTION	2010	2013	2016
DRAM ½ Pitch (nm)	45	32	22
MPU ½ Pitch (nm)	45	32	22
MPU Printed Gate Length (nm)	25	18	13
MPU Physical Gate Length (nm)	18	13	9

The ORTC and technology requirements tables are intended to indicate current best estimates of introduction timing for specific technology requirements. “Production” is defined as when any two companies have reached the production volume of 10k parts per month. Please refer to the Glossary for detailed definitions for Year of Introduction and Year of Production.

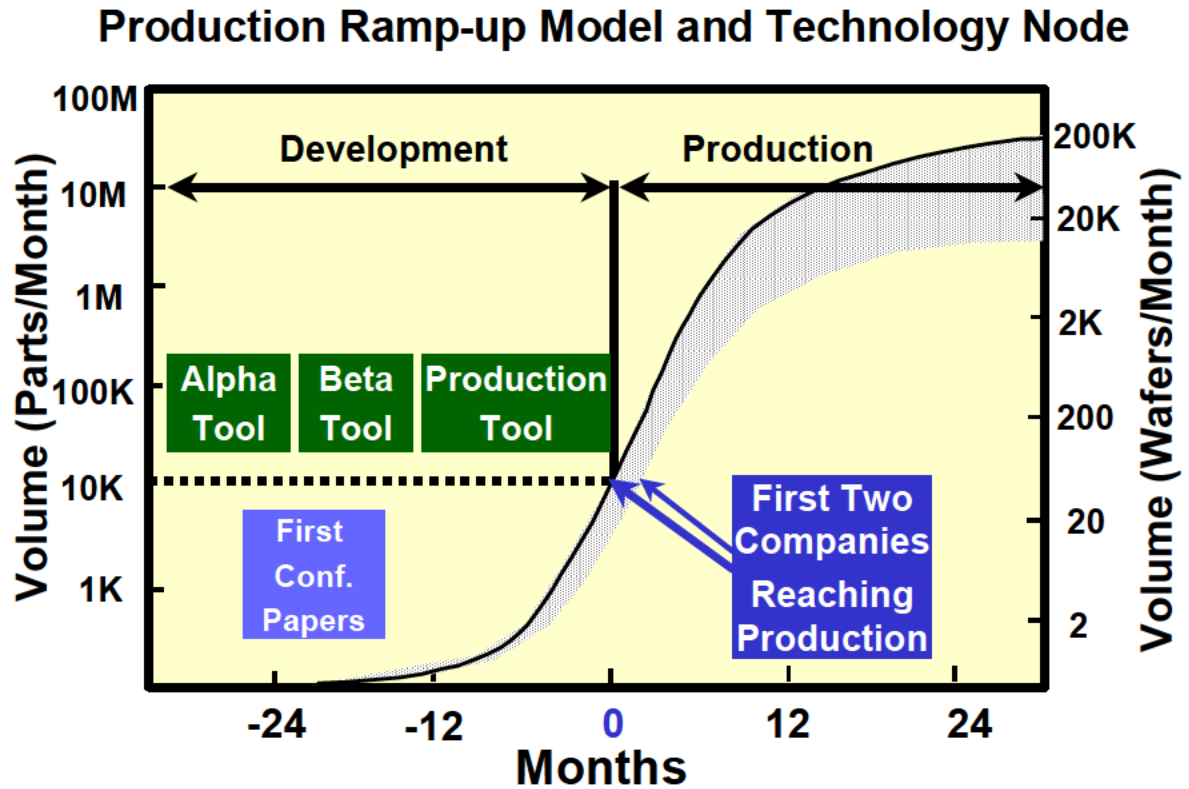


Figure 2 A Typical Production “Ramp” Curve

The ORTC and technology requirements tables are intended to indicate current best estimates of introduction time points for specific technology requirements. Ideally, the Roadmap might show multiple time points along the “research-development-prototyping-manufacturing” cycle for each requirement. However, in the interests of simplicity, usually only one point in time is estimated. The default “time of Introduction” in the ITRS is the “Year of Production,” which is defined in Figure 2.

The “Production” time in ITRS refers to the time when the first company brings a technology to production and a second company follows within three months. It is noted that the ITRS Roadmap, by its definition, focuses on forecasting the earliest introduction of the leading-edge technologies in respective fields for producing semiconductors.

Therefore it is obvious that many companies, for a variety of reasons, may choose to introduce a technology node later than the earliest introductions, hence that there is a wide variation of the technologies in actual production status from leading-edge to trailing-edge. Figure 3 shows, in horizontal bar-graph, the actual, annual worldwide wafer production capacity distributions over different process feature sizes. The distributions are quite wide-spread while the ITRS Technology Nodes, shown in small blue marks, are located exactly on the leading-edges of each of them.

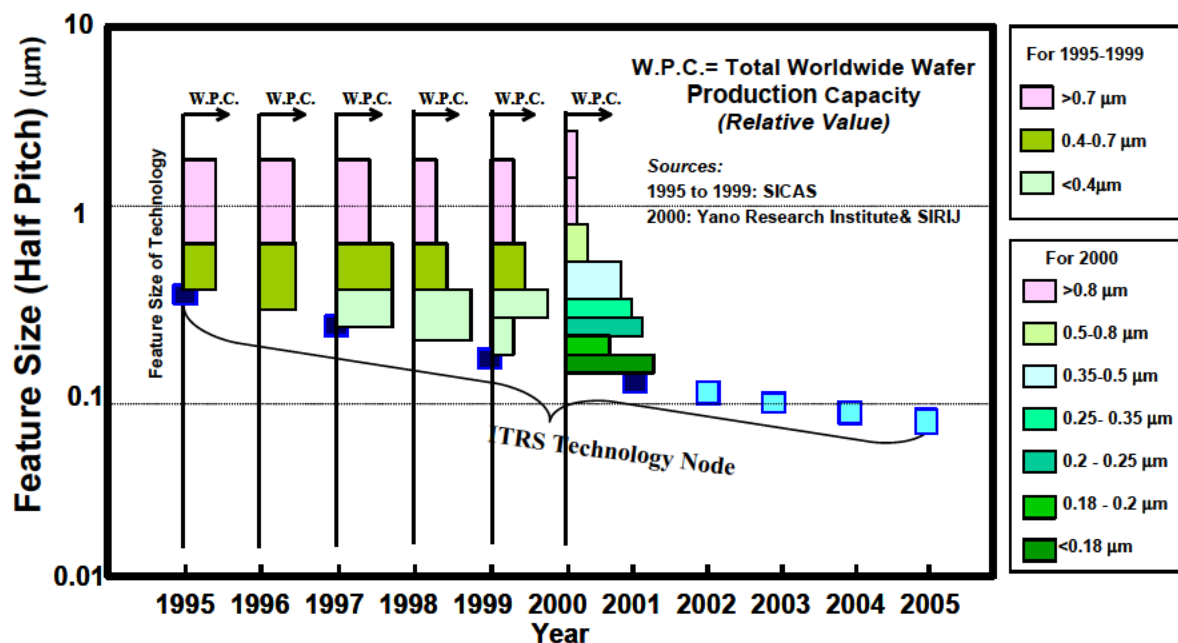


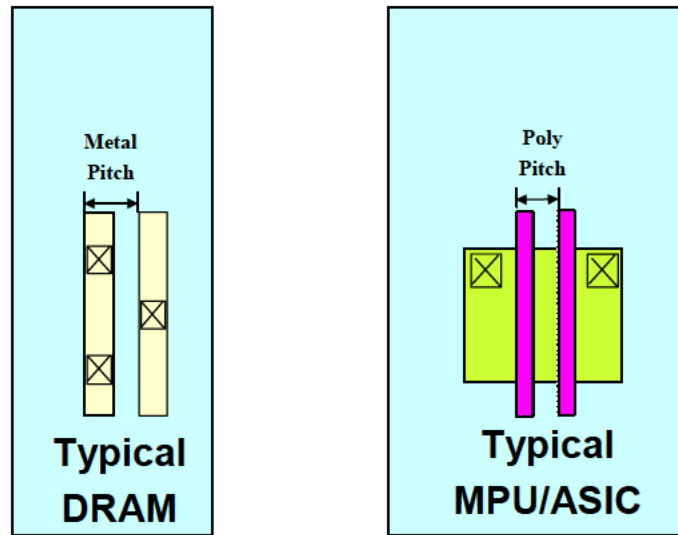
Figure 3 Technology Node Compared to Actual Production Capacity Technology Distribution

Note that some rows in the ORTC and technology requirements tables refer to other timing points, which are defined for each case (e.g., “at sample”). Of course, for the “Long-term Years,” for which the table intervals are three years, it is possible for the “best-estimate year of production” to fall in between the selected three-year intervals for some technology requirements. Also note that the “production” ramp in Figure 2 can be viewed as the time to ramp to full production wafer starts. For a fab designed for 20K wafer-starts-per-month (wspm) capacity, the time to ramp from 20 wspm to full capacity can take 9–12 months. This time would correspond to the same time for ramping device unit volume capacity from 6K units to 6M units per month if the chip size were 140 mm^2 ($430 \text{ gross die per } 300 \text{ mm wafer} \times 20\text{K wspm} \times 70\% \text{ total yield from wafer starts to finished product} = 6\text{M units/month}$).

TECHNOLOGY NODES

The concept of “technology nodes” used to be quite straightforward to understand as it has historically been linked to the introduction of new Dynamic Random Access Memory (DRAM) generations with a $4\times$ increase in bits/chip between generations. For as long as this cycle strictly followed Moore’s Law (three-year cycle for $4\times$), the technology nodes and DRAM generations were essentially synonymous. However, in recent years, a greater diversity of products serving as technology drivers, faster introduction/optimization of product-specific technology, and the general increase in business and technology complexity are all tending to de-couple the many technology parameters that have traditionally characterized “advance to the next technology node.” For example, microprocessor unit (MPU) products have recently driven the reduction of gate length at a faster pace than lithography half-pitch. While DRAM continues to drive the lithography half-pitch, MPU drives the gate length. Even the choice of basic lithography technology has tended to become more product specific (such as “pushing the wavelength as fast as possible” versus “using phase-shift masks”). Following the practice of the 1999 ITRS, the 2001 ITRS ORTC tables list the DRAM half-pitch, the MPU half-pitch, and MPU gate length, as shown in Table B. These technology parameters are defined in Figure 4. Any of the four parameters (rows) in Table B may be chosen as the driver for a given technology requirements table of a given ITWG. Nevertheless, for a point of reference, the DRAM half-pitch is still used in the 2001 ITRS as the designation of the technology node (DRAM still requires the smallest half pitch among all products.). For example, according to Table B, 2003 will be the year of production of the 100 nm node. Again, the “node designation” is defined by DRAM half pitch, not by the transistor gate length or minimum feature size characteristic of that node. Additional (and, in some cases, more precise) definitions related to the ITRS tables may be found in the *Glossary*.

Half Pitch (=Pitch/2) Definition



** DRAM pitch determines the technology “node” designation*

Figure 4 Definition of the Half Pitch and Gate Length

In recent years the “technology-development cycle” has been closer to two years than three years. On the other hand, the scaling ratio of the minimum feature size of one technology generation to the previous generation may no longer adhere to the historical value of 0.7. For example, 100 nm is not 0.7× of 130 nm. In addition, some companies may choose to introduce a “half node” (for example 150 nm may be considered a half node between the 180 nm node and the 130 nm node) with the intention of introducing the next “full” node at a later time. The 2001 ITRS acknowledges the validity of these practices and follows the 1999 ITRS in listing the near-term (2001–2007) technology requirements on yearly intervals and the long-term (2008–2016) requirements on three-year intervals as shown in Table B. Thus, we can say that 2003 is the year of production of the 100 nm node and 2004 is the year of production of the 90 nm node. One may also interpret the long-term three-year interval table as a “challenge” to extend the trend of one major new technology node every three years with scaling ratio of 0.7 between nodes through the “22 nm node” (which include 9 nm transistor gate lengths) in 2016.

DRIVERS FOR ITWG TECHNOLOGY REQUIREMENTS

The particular lithography-related rows selected for Table B from the ORTC tables are special in that any one of them may be designated by an ITWG as a “driver” for any specific row in one of their technology requirements tables. For example, the physical gate length may be the appropriate driver for the gate CD control and the source/drain junction depth. The designation of drivers for technology requirements provides some traceable and supportable assumptions for constructing the ITWG tables. It also provides useful links between the ORTC tables and the ITWG tables. Thus, as the Roadmap is updated in subsequent editions, these links will be used for creating a first-pass version of the new tables. For example, if the requirements in one of these driver rows of the ORTC tables were subsequently pulled-in by one year, it would be assumed that rows in the ITWG technology requirements tables would shift by default along with their designated ORTC driver row.

PROCESS INTEGRATION, DEVICES, AND STRUCTURES

Table 34 Process Integration Difficult Challenges [Update]

	Difficult Challenges $\geq 65\text{nm}$ / Through 2007	Summary of Issues
Was	1. High-performance applications: meeting performance and power dissipation requirements for highly scaled MOSFETs.	– Cost effectiveness, process control, and reliability of very thin oxy-nitride gate dielectrics, especially considering the high gate leakage. – Implementation of metal gate electrode by about 2007. – Need to reduce series S/D parasitic resistance. – Controlling static power dissipation in the face of rapidly increasing leakage. – Architecture and circuit design improvement and innovation will be needed
Is	1. High-performance applications: meeting performance and power dissipation requirements for highly scaled MOSFETs.	– Cost effectiveness, process control, and reliability of very thin oxy-nitride gate dielectrics, especially considering the high gate leakage. – Implementation of metal gate electrode and high κ gate dielectric by about 2007. – Need to reduce series S/D parasitic resistance. – Controlling static power dissipation in the face of rapidly increasing leakage. – Architecture and circuit design improvement and innovation will be needed
	2. Low-power applications: meeting performance and leakage requirements for highly scaled MOSFETs.	– Early availability of manufacturing-worthy high κ gate dielectrics is necessary to meet stringent gate leakage and performance requirements. – Very slow scaling of V_{dd} will make overall device scaling difficult.
	3. Implementation into manufacturing of non-classical MOSFET devices (for example, double-gate SOI).	– It is likely that these transistors will be necessary eventually to control short-channel and other effects in highly scaled devices. See Emerging Research Devices section, Non-classical CMOS, for more detail.
	4. Ensuring reliability of new materials and structures in a timely manner.	– Accelerated reliability assurance of high κ material for gate stack will be needed for early insertion into manufacturing. – Ensuring reliability of new gate electrode materials will be a challenge. – Ensuring reliability of new, non-classical CMOS structures will be a challenge. – Ensuring reliability of very thin oxy-nitrides with very high leakage current will be critical for high-performance applications. – Difficulty of screening with high leakage currents
	5. Constructing DRAM, SRAM, and high density nonvolatile memory (NVM) for scaled technologies	– DRAM main issues: adequate storage capacitance for devices with reduced feature size; access device design; holding the overall leakage to acceptably low levels; and deploying low sheet resistance materials for bit and word lines to ensure desired speed for scaled DRAMs. Also, the availability of manufacturing worthy 193 nm lithography and integrated DRAM etch capability for 100 nm half pitches in 2003. – SRAM: difficult lithography and etch as well as process integration issues. – NVM: very difficult scaling issues with tunnel and interpoly dielectrics.
	6. High-performance mixed-signal solutions for scaled technologies.	– Passive element scaling: embedded inductor densities and Q factor values. – Signal isolation. – Optimizing RF CMOS devices with scaled technologies: gate leakage is a particularly sensitive issue. – Transition to reduced analog supply voltages. – Difficulty and cost of integrating analog/RF and high-performance digital functions on a chip.

Table 34 Process Integration Difficult Challenges [[Update](#)](continued)

<i>Difficult Challenges < 65 nm, Beyond 2007</i>	<i>Summary of Issues</i>
7. Fundamental improvements in MOSFET device effective transconductance needed to maintain device performance scaling trend.	– With sharp reductions in V_{dd} and 17% annual increase in intrinsic transistor speed, basic MOSFET device performance will be inadequate to meet circuit speed requirements.
8. Dealing with atomic-level fluctuations and statistical process variations in sub-30 nm MOSFETs.	– Fundamental problems of atomic-level statistical fluctuations are not completely understood.
9. New interconnect schemes	– Eventually, copper/low κ performances will be inadequate. – Solutions (optical, microwave/RF, etc.) are currently unclear.
10 Toward the end of the Roadmap or beyond, implementation of advanced non-CMOS devices and architectures, including memory.	– Will drive major changes in process, materials, physics, design, etc. – Non-CMOS devices may coexist with CMOS: integration of the two will be difficult, especially for mixed signal. – See Emerging Research Devices sections for more discussion and detail.

Table 35a High-performance Logic Technology Requirements—Near-term

Year of Production		2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)		130	115	100	90	80	70	65
MPU / ASIC ½ Pitch (nm)		150	130	107	90	80	70	65
MPU Printed Gate Length (nm)		90	75	65	53	45	40	35
MPU Physical Gate Length (nm)		65	53	45	37	32	28	25
Physical gate length high-performance (HP) (nm) [1]		65	53	45	37	32	28	25
Was	Equivalent physical oxide thickness for high-performance T_{ox} (EOT) (nm) [2]	1.3–1.6	1.2–1.5	1.1–1.6	0.9–1.4	0.8–1.3	0.7–1.2	0.6–1.1
Is	<u>EOT: equivalent oxide thickness (physical) for high-performance (nm) [2]</u>	1.3–1.6	1.2–1.5	1.1–1.6	0.9–1.4	0.8–1.3	0.7–1.2	0.6–1.1
Gate depletion and quantum effects electrical thickness adjustment factor (nm) [3]		0.8	0.8	0.8	0.8	0.8	0.8	0.5
Was	T_{ox} electrical equivalent (nm) [4]	2.3	2.1	2	2	1.9	1.9	1.4
Is	<u>Equivalent oxide thickness (electrical) (nm) [4]</u>	2.3	2.1	2	2	1.9	1.9	1.4
Nominal power supply voltage (V_{dd}) (V) [5]		1.2	1.1	1	1	0.9	0.9	0.7
Nominal high-performance NMOS sub-threshold leakage current, $I_{sd,leak}$ (at 25° C) ($\mu A/\mu m$) [6]		0.01	0.03	0.07	0.1	0.3	0.7	1
Nominal high-performance NMOS saturation drive current, I_{dd} (at V_{dd} at 25° C) ($\mu A/\mu m$) [7]		900	900	900	900	900	900	900
Required percent current-drive "mobility/transconductance improvement" [8]		0%	0%	0%	0%	0%	0%	0%
Parasitic source/drain resistance (R_{sd}) (ohm- μm) [9]		190	180	180	180	180	170	140
Parasitic source/drain resistance (R_{sd}) percent of ideal channel resistance (V_{dd}/I_{dd}) [10]		16%	16%	17%	18%	19%	19%	20%
Parasitic capacitance percent of ideal gate capacitance [11]		19%	22%	24%	27%	29%	32%	27%
High-performance NMOS device τ ($C_{gate} * V_{dd} / I_{dd}$ -NMOS)(ps) [12]		1.6	1.3	1.1	0.99	0.83	0.76	0.68
Relative device performance [13]		1	1.2	1.5	1.6	2	2.1	2.5
Was	Energy per ($W/L_{gate}=3$) device switching transition ($C_{gate} * (3 * L_{gate}) * V^2$) (fJ/Device) [14]	0.347	0.212	0.137	0.099	0.065	0.052	0.032
Is	<u>Power-delay product for ($W/L_{gate}=3$) device [$C_{gate} * (3 * L_{gate}) * V_{dd}^2$] (fJ/Device) [14]</u>	0.347	0.212	0.137	0.099	0.065	0.052	0.032
Static power dissipation per ($W/L_{gate}=3$) device (Watts/Device) [15]		5.60E-09	6.70E-09	1.00E-08	1.10E-08	2.60E-08	5.30E-08	5.30E-08

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



Table 35b High-performance Logic Technology Requirements—Long-term

Year of Production		2010	2013	2016
DRAM ½ Pitch (nm)		45	32	22
MPU / ASIC ½ Pitch (nm)		50	35	25
MPU Printed Gate Length (nm)		25	18	13
MPU Physical Gate Length (nm)		18	13	9
Physical gate length high-performance (HP) (nm) [1]		18	13	9
Was	Equivalent physical oxide thickness for high-performance T_{ox} (EOT) (nm) [2]	0.5-0.8	0.4-0.6	0.4-0.5
Is	<u>EOT: equivalent oxide thickness (physical) for high-performance</u> (nm) [2]	0.5-0.8	0.4-0.6	0.4-0.5
Gate depletion and quantum effects electrical thickness adjustment factor (nm) [3]		0.5	0.5	0.5
Was	T_{ox} electrical equivalent (nm) [4]	1.2	1	0.9
Is	<u>Equivalent oxide thickness (electrical)</u> (nm) [4]	1.2	1	0.9
Nominal power supply voltage (V_{dd}) (V) [5]		0.6	0.5	0.4
Nominal high-performance NMOS sub threshold leakage current, $I_{sd leak}$ (at 25 °C) ($\mu A/\mu m$) [6]		3	7	10
Nominal high-performance NMOS saturation drive current, I_{dd} (at V_{dd} , at 25 °C) ($\mu A/\mu m$) [7]		1200	1500	1500
Required percent current-drive "mobility/transconductance improvement" [8]		30%	70%	100%
Parasitic source/drain resistance (R_{sd}) (ohm- μm) [9]		110	90	80
Parasitic source/drain resistance (R_{sd}) percent of ideal channel resistance (V_{dd}/I_{dd}) [10]		25%	30%	35%
Parasitic capacitance percent of ideal gate capacitance [11]		31%	36%	42%
High-performance NMOS device τ ($C_{gate} * V_{dd} / I_{dd}$ -NMOS)(ps) [12]		0.39	0.22	0.15
Relative device performance [13]		4.3	7.2	10.7
Was	Energy per ($W/L_{gate}=3$) device switching transition ($C_{gate} * (3 * L_{gate}) * V^2$) (fJ/Device) [14]	0.015	0.007	0.002
Is	<u>Power-delay product for ($W/L_{gate}=3$) device [$C_{gate} * (3 * L_{gate}) * V_{dd}^2$]</u> (fJ/Device) [14]	0.015	0.007	0.002
Static power dissipation per ($W/L_{gate}=3$) device (Watts/Device) [15]		9.70E-08	1.40E-07	1.10E-07

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



Notes for Table 35a and b:

	[1] Values set by ORTC. Gate dimensional control is set by the Lithography and FEP Etch ITWGs, and is assumed to have a three sigma value of $\pm 10\%$. Gate dimension variation is assumed to be the primary factor responsible for driving device parameter variation.
Was	[2] EOT range set by FEP TWG. Yellow/red feasibility coloring set by FEP TWG projections on thickness control and reliability capability. Calculations in rest of PIDS table (and the underlying PIDS workbook) based on approximate midpoint EOT values.
Is	[2] EOT range set by FEP TWG. Yellow/red feasibility coloring set by FEP TWG projections on thickness control and reliability capability. Calculations in rest of PIDS table (and the underlying PIDS workbook) based on approximate midpoint EOT values. For a gate dielectric of thickness T_d and relative dielectric constant k, EOT is defined by: $EOT = T_d / (k/3.9)$, where 3.9 is the relative dielectric constant of thermal silicon dioxide. The ideal gate capacitance per unit area of the gate dielectric of thickness T_d is the same as that of a gate dielectric made up of thermal silicon dioxide with a thickness of EOT.
	[3] Accounts for gate electrode depletion and inversion-layer quantum effects. Yellow feasibility coloring reflects FEP assessment of polysilicon doping capability; red feasibility coloring reflects the introduction of metal-gate electrodes by 2007 (which reduces the gate depletion value). [4] Sum of midpoint EOT and Electrical Thickness Adjustment Factor. Used in CV/I performance metric and CV^2 dynamic power metric calculations. Red/yellow feasibility coloring determined by worst-case EOT and Electrical Thickness Adjustment Factor red/yellow feasibility coloring.
Was	[5] Nominal power supply voltage has been set to maintain sufficient voltage over-drive to continue historical approximate 17% per year device performance scaling while still enabling approximate 30% per year switching energy reduction and still maintaining reasonable vertical gate dielectric electric field strengths. Actual power supply voltage values may vary $\pm 10\%$, depending on the particular circuit design application or technology optimization.
Is	[5] Nominal power supply voltage has been set to maintain sufficient voltage over-drive to continue historical approximate 17% per year device performance scaling while still enabling approximate 30% per year switching energy reduction and still maintaining reasonable vertical gate dielectric electric field strengths. Actual target power supply voltage values may vary $\pm 10\%$, depending on the particular circuit design application or technology optimization.
	[6] Nominal sub-threshold leakage current is defined as the NMOSFET source current at room temperature with the drain bias set equal to the nominal power-supply voltage and with the gate, source and substrate biases set to zero volts; all MOSFET device dimensions are assumed to be at their nominal/target values. Total NMOS off-state current is the NMOSFET drain current at room temperature, and is the sum of the NMOS sub-threshold, gate, and junction leakage current components. The sub-threshold leakage current is assumed to be larger than either the gate or junction leakage current components at either room or high-temperature conditions. The threshold voltage value (and the corresponding sub-threshold current) has been set to maintain sufficient voltage over-drive to continue historical approximate 17% per year device performance scaling. Yellow feasibility coloring by 2007 reflects the potential need for non-classical CMOS or ultra-shallow junction technology in order to control short-channel and/or high-field effects (See Emerging Research Devices section). The above sub-threshold, gate, and junction current scaling scenario also applies to PMOS devices. Note, sub-threshold current value applies to fastest MOS devices only; slower/lower-leakage MOS devices will also be available. Future systems will consist of a mix of both high and lower-leakage devices. [7] Nominal saturation current drive is defined as the MOSFET drain current at room temperature with the gate bias and the drain bias set equal to the nominal power-supply voltage; all MOSFET device dimensions are assumed to be at their nominal/target values. Nominal PMOS saturation current-drive value is assumed to be 40-50% of the nominal NMOS saturation current-drive value. Yellow/red feasibility coloring indicates the projected need for fundamental device current-drive (or transconductance/mobility) improvement by 2010 in order to continue historical approximate 17% per year device performance scaling. NMOS/PMOS current-drive targets are approximate with only 1.5 significant digits of accuracy. [8] Fundamental device mobility/transconductance improvement needed by 2010 in order to continue historical approximate 17% per year device performance scaling. Yellow/red feasibility coloring indicates the difficulty in implementing fundamental device current-drive (transconductance/mobility) improvement required in 2010 and later to continue historical approximate 17% per year device performance scaling.
Was	[9] Total parasitic device source/drain resistance R_{sd} (sum of the source and drain parasitic resistances). R_{sd} targets are consistent with FEP TWG projections. Yellow/red feasibility coloring reflects FEP TWG projections on contact resistance, salicide sheet resistance, and drain extension scaling. Similar R_{sd} values are assumed for the LOP and LSTP devices.
Is	[9] Maximum allowable parasitic device source/drain resistance R_{sd} (sum of the source and drain parasitic resistances). R_{sd} targets are consistent with FEP TWG projections. Yellow/red feasibility coloring reflects FEP TWG projections on contact resistance, salicide sheet resistance, and drain extension scaling. Similar R_{sd} values are assumed for the LOP and LSTP devices.
Was	[10] Maximum ratio of the parasitic device source/drain resistance (R_{sd}) to the ideal channel resistance (V_{dd}/I_{dd}). R_{sd} targets are consistent with FEP TWG projections. Yellow/red feasibility coloring reflects FEP TWG projections on contact resistance, salicide sheet resistance, and drain extension scaling.
Is	[10] Maximum Ratio of the parasitic device source/drain resistance (R_{sd}) to the ideal channel resistance (V_{dd}/I_{dd}). R_{sd} targets are consistent with FEP TWG projections. Yellow/red feasibility coloring reflects FEP TWG projections on contact resistance, salicide sheet resistance, and drain extension scaling.

Notes for Table 35a and b(continued)

Was	[11] Maximum ratio of the parasitic gate overlap/fringing capacitance to the ideal gate capacitance. Assume constant C-parasitic value of $2.4E-16F/\mu m$. [$3\times$ the fringing capacitance value per side, including the Miller effect]; this value is assumed to be independent of bias conditions and/or technology. Parasitic capacitance factor is included in CV/I and CV^2 performance and power metric calculations. Similar parasitic capacitance values are assumed for the LOP and LSTP devices.
Is	[11] ~~Maximum~~ ratio of the parasitic gate overlap/fringing capacitance to the ideal gate capacitance. Assume constant C-parasitic value of $2.4E-16F/\mu m$. [$3\times$ the fringing capacitance value per side, including the Miller effect]; this value is assumed to be independent of bias conditions and/or technology. Parasitic capacitance factor is included in CV/I and CV^2 performance and power metric calculations. Similar parasitic capacitance values are assumed for the LOP and LSTP devices.
	[12] $\tau = CV/I$ intrinsic delay metric for NMOS device; PMOS CV/I metric assumed to scale proportionally. The CV/I metric provides an indication of the intrinsic switching delay of the device, while $1/\tau$, the reciprocal of CV/I , is a good metric for the intrinsic switching speed of the device. Red/yellow feasibility coloring determined by worst-case saturation current-drive feasibility coloring.
	[13] Improvement in $1/\tau$ NMOS performance metric normalized to the year 2001. Maintains historical approximate 17% per year device performance improvement scaling trend. Red/yellow feasibility coloring determined by worst-case, saturation current-drive feasibility coloring.
Was	[14] CV^2 switching energy metric for an NMOS device with dimensions $W/L_{gate}=3$. The switching energy metric indicates the amount of dynamic power required to switch the device. Maintains approximate 30% per year device switching energy reduction scaling trend. Red/yellow feasibility coloring determined by worst-case saturation current-drive feasibility and mobility/transconductance improvement feasibility coloring.
Is	[14] CV^2 switching energy metric for an NMOS device with dimensions $W/L_{gate}=3$. <u>The power-delay product</u> indicates the amount of dynamic <u>energy</u> required to switch the <u>device through a full (up and down) transition</u> . Maintains approximate 30% per year device switching energy reduction scaling trend. Red/yellow feasibility coloring determined by worst-case saturation current-drive feasibility and mobility/transconductance improvement feasibility coloring.
	[15] Static power dissipation for an NMOS device with dimensions $W/L_{gate}=3$. Assume that the device static power is primarily determined by the sub-threshold current (since the other junction and leakage current components are assumed to be lower). Yellow feasibility coloring by 2007 reflects the potential need for non-classical CMOS or ultra-shallow junction technology in order to control short-channel and/or high-field effects (See Emerging Research Devices section, Non-classical CMOS).

Table 36a Low Operating Power (LOP) Logic Technology Requirements—Near-term

Year of Production		2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)		130	115	100	90	80	70	65
MPU / ASIC ½ Pitch (nm)		150	130	107	90	80	70	65
MPU Printed Gate Length (nm)		90	75	65	53	45	40	35
MPU Physical Gate Length (nm)		65	53	45	37	32	28	25
Physical gate length low-operating power (LOP) (nm) [1]		90	75	65	53	45	37	32
Was	Equivalent physical oxide thickness for LOP T_{ox} (EOT) (nm) [2]	2.0-2.4	1.8-2.2	1.6-2.0	1.4-1.8	1.2-1.6	1.1-1.5	1.0-1.4
Is	<u>EOT: equivalent oxide thickness (physical) for LOP</u> (nm) [2]	2.0-2.4	1.8-2.2	1.6-2.0	1.4-1.8	1.2-1.6	1.1-1.5	1.0-1.4
Electrical thickness adjustment factor (gate depletion and quantum effects) (nm) [3]		0.8	0.8	0.8	0.8	0.8	0.8	0.5
Was	T_{ox} electrical equivalent (nm) [4]	3	2.8	2.6	2.4	2.2	2.1	1.7
Is	<u>Equivalent oxide thickness (electrical)</u> (nm) [4]	3	2.8	2.6	2.4	2.2	2.1	1.7
Nominal LOP power supply voltage (V_{dd}) (V) [5]		1.2	1.2	1.1	1.1	1	1	0.9
Nominal LOP NMOS sub-threshold leakage current, $I_{d leak}$ (@25C) (pA/ μm) [6]		100	100	100	300	300	300	700
Nominal LOP NMOS Saturation drive current, I_{dd} (@ V_{dd} , @25C) ($\mu A/\mu m$) [7]		600	600	600	600	600	600	700
Required percent current-drive "mobility/transconductance improvement" [8]		0%	0%	0%	0%	0%	0%	0%
LOP NMOS Device τ ($C_{gate} * V_{dd} / I_d$ -NMOS) (ps) [9]		2.55	2.45	2.02	1.84	1.58	1.41	1.14
LOP relative device performance [10]		1	1.04	1.3	1.4	1.6	1.8	2.2
Was	Energy per ($W/L_{gate}=3$) device switching transition ($C_{gate} * (3 * L_{gate}) * V_{dd}^2$) (fJ/Device) [11]	0.496	0.424	0.26	0.193	0.128	0.094	0.069
Is	<u>Power-delay product for ($W/L_{gate}=3$) device</u> <u>$[C_{gate} * (3 * L_{gate}) * V_{dd}^2]$</u> (fJ/Device) [14]	0.496	0.424	0.26	0.193	0.128	0.094	0.069
Static power dissipation per ($W/L_{gate}=3$) device (Watts/device) [12]		3.20E-11	2.90E-11	2.10E-11	5.20E-11	4.10E-11	3.30E-11	6.00E-11

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



Table 36b Low Operating Power (LOP) Logic Technology Requirements—Long-term

Year of Production		2010	2013	2016
DRAM ½ Pitch (nm)		45	32	22
MPU / ASIC ½ Pitch (nm)		50	35	25
MPU Printed Gate Length (nm)		25	18	13
MPU Physical Gate Length (nm)		18	13	9
Physical gate length low-operating power (LOP) (nm) [1]		22	16	11
Was	Equivalent physical oxide thickness for LOP T_{ox} (EOT) (nm) [2]	0.8-1.2	0.7-1.1	0.6-1.0
Is	<u>EOT: equivalent oxide thickness (physical) for high-performance</u> (nm) [2]	0.8-1.2	0.7-1.1	0.6-1.0
Electrical thickness adjustment factor (gate depletion and quantum effects) (nm) [3]		0.5	0.5	0.5
Was	T_{ox} electrical equivalent (nm) [4]	1.5	1.4	1.3
Is	<u>Equivalent oxide thickness (electrical)</u> (nm) [4]	1.5	1.4	1.3
Nominal LOP power supply voltage (V_{dd}) (V) [5]		0.8	0.7	0.6
Nominal LOP NMOS sub-threshold leakage current, $I_{sd leak}$ (@25C) (pA/ μm) [6]		1000	3000	10000
Nominal LOP NMOS saturation drive current, I_{dd} (@ V_{dd} , @25C) ($\mu A/\mu m$) [7]		700	800	900
Required percent current-drive "mobility/transconductance improvement" [8]		10%	30%	70%
LOP NMOS device τ ($C_{gate} * V_{dd} / I_{dd}$ -NMOS) (ps) [9]		0.85	0.56	0.35
LOP relative device Performance [10]		3	4.6	7.2
Was	Energy per ($W/L_{gate}=3$) device switching transition ($C_{gate} * (3 * L_{gate}) * V^2$) (fJ/Device) [11]	0.032	0.015	0.006
Is	<u>Power-delay product for ($W/L_{gate}=3$) device [$C_{gate} * (3 * L_{gate}) * V_{dd}^2$]</u> (fJ/Device) [14]	0.032	0.015	0.006
Static power dissipation per ($W/L_{gate}=3$) device (Watts/Device) [12]		5.30E-11	1.00E-10	2.00E-10

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



Table 36c Low Standby Power (LSTP) Technology Requirements—Near-term

Year of Production		2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)		130	115	100	90	80	70	65
MPU / ASIC ½ Pitch (nm)		150	130	107	90	80	70	65
MPU Printed Gate Length (nm)		90	75	65	53	45	40	35
MPU Physical Gate Length (nm)		65	53	45	37	32	28	25
Was	Physical gate length low-standby power (LSTP) (nm) [1]	90	75	65	53	45	37	32
Is	Physical gate length low-standby power (LSTP) (nm) [1]	100	90	75	65	53	45	37
Was	Equivalent physical oxide thickness for LSTP T_{ox} (EOT) (nm) [2]	2.4–2.8	2.2–2.6	2.0–2.4	1.8–2.2	1.6–2.0	1.4–1.8	1.2–1.6
Is	EOT: equivalent oxide thickness (physical) for LOP (nm) [2]	2.4–2.8	2.2–2.6	2.0–2.4	1.8–2.2	1.6–2.0	1.4–1.8	1.2–1.6
	Electrical thickness adjustment factor (gate depletion and quantum effects) (nm) [3]	0.8	0.8	0.8	0.8	0.8	0.8	0.5
Was	T_{ox} electrical equivalent (nm) [4]	3.4	3.2	3	2.8	2.6	2.4	1.9
Is	Equivalent oxide thickness (electrical) (nm) [4]	3.4	3.2	3	2.8	2.6	2.4	1.9
	Nominal LSTP power supply voltage (V_{dd}) (V) [5]	1.2	1.2	1.2	1.2	1.2	1.2	1.1
	Nominal LSTP NMOS sub-threshold current (at 25 °C) ($pA/\mu m$) [6]	1	1	1	1	1	1	1
	Nominal LSTP NMOS saturation current drive (I_{dd}) (at V_{dd} at 25 °C) ($mA/\mu m$) [7]	300	300	400	400	400	400	500
	Required percent current-drive "mobility/transconductance improvement" [8]	0%	0%	0%	0%	0%	0%	0%
Was	LSTP NMOS device τ ($C_{gate} * V_{dd} / I_{d-NMOS}$) (ps) [9]	4.61	4.41	2.96	2.68	2.51	2.32	1.81
Is	LSTP NMOS device τ ($C_{gate} * V_{dd} / I_{d-NMOS}$) (ps) [9]	5.02	4.84	3.31	3.12	2.83	2.66	2.01
Was	LSTP relative device performance [10]	1	1.05	1.6	1.7	1.8	2	2.6
Is	LSTP relative device performance [10]	1	1.04	1.52	1.61	1.77	1.89	2.50
Was	Energy per ($W/L_{gate}=3$) device switching transition ($C_{gate} * (3 * L_{gate}) * V^2$) (fJ/device) [11]	0.448	0.381	0.277	0.204	0.163	0.123	0.095
Is	Power-delay product for ($W/L_{gate}=3$) device ($C_{gate} * (3 * L_{gate}) * V_{dd}^2$) (fJ/Device) [14]	0.542	0.471	0.357	0.292	0.216	0.172	0.122
Was	Static power dissipation per ($W/L_{gate}=3$) device (Watts/device) [12]	3.20E-13	2.90E-13	2.30E-13	1.90E-13	1.60E-13	1.30E-13	1.10E-13
Is	Static power dissipation per ($W/L_{gate}=3$) device (Watts/device) [12]	3.60E-13	3.24E-13	2.70E-13	2.34E-13	1.91E-13	1.62E-13	1.22E-13

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



Table 36d Low Standby Power (LSTP) Technology Requirements—Long-term

Year of Production		2010	2013	2016
DRAM ½ Pitch (nm)		45	32	22
MPU / ASIC ½ Pitch (nm)		50	35	25
MPU Printed Gate Length (nm)		25	18	13
MPU Physical Gate Length (nm)		18	13	9
Was	Physical gate length low-standby power (LSTP) (nm) [1]	22	16	11
Is	Physical gate length low-standby power (LSTP) (nm) [1]	28	20	16
Was	Equivalent physical oxide thickness for LSTP T_{ox} (EOT) (nm) [2]	0.9-1.3	0.8-1.2	0.7-1.1
Is	EOT: equivalent oxide thickness (physical) for high-performance (nm) [2]	0.9-1.3	0.8-1.2	0.7-1.1
Electrical thickness adjustment factor (gate depletion and quantum effects) (nm) [3]		0.5	0.5	0.5
Was	T_{ox} electrical equivalent (nm) [4]	1.6	1.5	1.4
Is	Equivalent oxide thickness (electrical) (nm) [4]	1.6	1.5	1.4
Nominal LSTP power supply voltage (V_{dd}) (V) [5]		1	0.9	0.9
Nominal LSTP NMOS sub-threshold current (at 25 °C) (pA/μm) [6]		3	7	10
Nominal LSTP NMOS saturation current drive (I_{dd}) (at V_{dd} at 25° C) (μA/μm) [7]		500	600	700
Required percent current-drive "mobility/transconductance improvement" [8]		10%	30%	50%
Was	LSTP NMOS device τ ($C_{gate} * V_{dd} / I_{d-NMOS}$) (ps) [9]	1.43	0.91	0.66
Is	LSTP NMOS device τ ($C_{gate} * V_{dd} / I_{d-NMOS}$) (ps) [9]	1.69	1.05	0.82
Was	LSTP relative device performance [10]	3.2	5.1	7
Is	LSTP relative device performance [10]	2.97	4.78	6.15
Was	Energy per ($W/L_{gate}=3$) device switching transition ($C_{gate}*(3*L_{gate})*V^2$) (fJ/device) [11]	0.047	0.024	0.014
Is	Power-delay product for ($W/L_{gate}=3$) device [$C_{gate}*(3*L_{gate})*V_{dd}^2$] (fJ/Device) [14]	0.071	0.034	0.025
Was	Static power dissipation per ($W/L_{gate}=3$) device (Watts/device) [12]	2.00E-13	3.00E-13	3.00E-13
Is	Static power dissipation per ($W/L_{gate}=3$) device (Watts/device) [12]	2.52E-13	3.78E-13	4.32E-13

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



Notes for Table 36a through d

[1] Values set by ORTC.

[1] Values set by ORTC. Assumed to lag high-performance scaling by 2 years. Gate dimensional control is set by the Lithography and FEP (Etch) ITWGs and is assumed to have a three sigma value of $\pm 10\%$. Gate dimension variation is assumed to be the primary factor responsible for driving device parameter variation. Gate length for the LOP and LSTP devices are assumed to be identical.

Was	[2] EOT value set by FEP ITWG in the Front End Processes chapter, Thermal and Thin Film, Doping, and Etching Technology Requirements tables. Yellow/red feasibility coloring set by FEP ITWG projections on gate leakage, thickness control and reliability capability (high κ gate dielectrics will be required around 2005 in order to suppress gate leakage for LSTP). Due to different system applications, the EOT values for the LOP and LSTP devices have been optimized and set independently of each other. Calculations in rest of PIDS table (and the underlying PIDS workbook) based on approximate midpoint EOT values.
Is	[2] EOT value set by FEP ITWG in the Front End Processes chapter, Thermal and Thin Film, Doping, and Etching Technology Requirements tables. Yellow/red feasibility coloring set by FEP ITWG projections on gate leakage, thickness control and reliability capability (high κ gate dielectrics will be required around 2005 in order to suppress gate leakage for LSTP). Due to different system applications, the EOT values for the LOP and LSTP devices have been optimized and set independently of each other. Calculations in rest of PIDS table (and the underlying PIDS workbook) based on approximate midpoint EOT values. <u>For a gate dielectric of thickness T_d and relative dielectric constant κ, EOT is defined by: $EOT = T_d / (\kappa/3.9)$, where 3.9 is the relative dielectric constant of thermal silicon dioxide. The ideal gate capacitance per unit area of the gate dielectric of thickness T_d is the same as that of a gate dielectric made up of thermal silicon dioxide with a thickness of EOT.</u>
	[3] Accounts for gate electrode depletion and inversion-layer quantum effects. Yellow feasibility coloring reflects FEP assessment of polysilicon doping capability; red feasibility coloring reflects the introduction of metal-gate electrodes by 2007 (which reduces the gate depletion value).
	[4] Sum of midpoint EOT and Electrical Thickness Adjustment Factor. Used in CV/I performance metric and CV ² dynamic power metric calculations. Red/yellow feasibility coloring determined by worst-case EOT and Electrical Thickness Adjustment Factor red/yellow feasibility coloring.
Was	[5] Nominal power supply voltage which has been set to smallest value to still maintain sufficient voltage over-drive to allow sufficient circuit switching noise margin (approximately 2.3 times the threshold voltage). Actual power supply voltage values may vary $\pm 10\%$, depending on the particular circuit design application or technology optimization. Due to different system applications, the power-supply voltages for the LOP and LSTP devices have been optimized and set independently of each other. Note, meeting overall system power dissipation requirements will require the use of circuit/system techniques to "turn-off" or "power-down" various circuit blocks.
Is	[5] Nominal power supply voltage ~~which~~ has been set to smallest value to still maintain sufficient ~~voltage over drive to allow sufficient~~ circuit switching noise margin (approximately 2.3 times the threshold voltage). Actual power supply voltage ~~target~~ values may vary $\pm 10\%$, depending on the particular circuit design application or technology optimization. Due to different system applications, the power-supply voltages for the LOP and LSTP devices have been optimized and set independently of each other. Note, meeting overall system power dissipation requirements will require the use of circuit/system techniques to "turn-off" or "power-down" various circuit blocks.
Was	[6] Nominal sub-threshold leakage current is defined as the NMOSFET source current at room temperature with the drain bias set equal to the nominal power-supply voltage and with the gate, source, and substrate biases set to zero volts; all MOSFET device dimensions are assumed to be at their nominal/target values. Total NMOS off-state current is the NMOSFET drain current at room temperature, and is the sum of the NMOS sub-threshold, gate, and junction leakage current components. The sub-threshold leakage current is assumed to be larger than either the gate or junction current components at either room or high-temperature conditions. The increase in sub-threshold current (and the corresponding threshold-voltage value reduction) has been set at a pace that lags the rate of increase of the high-performance device, but that still increases sufficiently to enable continued device performance scaling; power dissipation due to off-state leakage is assumed to not exceed 10% of the total chip power, which is assumed to be 100mW in 2001 and 200mW in 2010 for LOP. For LOP, the yellow feasibility coloring reflects the difficulty of meeting the gate leakage requirements with thin oxy-nitride films and the potential need for non-classical CMOS or ultra-shallow junction technology by 2007 to control short-channel effects and to limit lateral high-field effects. For LSTP, the yellow feasibility coloring in 2003 and 2004 reflects the difficulty of meeting the gate leakage requirements with thin oxy-nitride gate dielectric films, while the red feasibility color from 2005 on reflects the difficulty of implementing a high κ gate dielectric to meet the gate leakage requirements. The above sub-threshold, gate, and junction current scaling scenario also applies to PMOS devices.
Is	[6] Nominal sub-threshold leakage current is defined as the NMOSFET source current at room temperature with the drain bias set equal to the nominal power-supply voltage and with the gate, source, and substrate biases set to zero volts; all MOSFET device dimensions are assumed to be at their nominal/target values. Total NMOS off-state current is the NMOSFET drain current at room temperature, and is the sum of the NMOS sub-threshold, gate, and junction leakage current components. The sub-threshold leakage current is assumed to be larger than either the gate or junction current components at either room or high-temperature conditions. The increase in sub-threshold current (and the corresponding threshold-voltage ~~value~~ reduction) has been set at a pace that lags the rate of increase of the high-performance device, but that still increases sufficiently to enable continued device performance scaling; power dissipation due to off-state leakage is assumed to not exceed 10% of the total chip power, which is assumed to be 100mW in 2001 and 200mW in 2010 for LOP. For LOP, the yellow feasibility coloring reflects the difficulty of meeting the gate leakage requirements with thin oxy-nitride films and the potential need for non-classical CMOS or ultra-shallow junction technology by 2007 to control short-channel effects and to limit lateral high-field effects. For LSTP, the yellow feasibility coloring in 2003 and 2004 reflects the difficulty of meeting the gate leakage requirements with thin oxy-nitride gate dielectric films, while the red feasibility color from 2005 on reflects the difficulty of implementing a high κ gate dielectric to meet the gate leakage requirements. The above sub-threshold, gate, and junction current scaling scenario also applies to PMOS devices.

Notes for Table 36a through d (continued)

[7] Nominal saturation current drive is defined as the MOSFET drain current at room temperature with the gate bias and the drain bias set equal to the nominal power-supply voltage; all MOSFET device dimensions are assumed to be at their nominal/target values. Nominal PMOS saturation current-drive value is assumed to be 40–50% of the nominal NMOS saturation current-drive value. Yellow/red feasibility coloring indicates the projected need for fundamental device current-drive (transconductance/mobility) improvement by 2010 in order to continue approximate 14% per year device performance scaling. NMOS/PMOS current-drive targets are approximate with only 1.5 significant digits of accuracy. The sub-threshold slope, parasitic source/drain resistance, and parasitic gate capacitance scaling have been assumed to be similar to that for the high-performance device. (See the enclosed Excel workbook for the detailed calculations).

[8] Fundamental device mobility/transconductance improvement needed by 2010 in order to continue historical approximate 14% per year device performance scaling. Yellow/red feasibility coloring indicates the difficulty of implementing the fundamental device current-drive (transconductance/mobility) improvement required in 2010 and beyond to continue historical approximate 14% per year device performance scaling. The LOP and LSTP required improvement is projected as lagging that required for the high-performance device.

Was	[9] $\tau = CV/I$ intrinsic delay metric for NMOS device; PMOS CV/I metric assumed to scale proportionally. The CV/I metric provides an indication of the intrinsic switching delay of the device, while $1/\tau$, the reciprocal of CV/I , is a good metric for the intrinsic switching speed of the device. Red/yellow feasibility coloring determined by saturation current-drive feasibility coloring. The C term includes the effect of parasitic gate capacitance, which has been assumed to be equivalent to that for the high-performance device.
Is	[9] $\tau = CV/I$ intrinsic delay metric for NMOS device; PMOS CV/I metric assumed to scale proportionally. The CV/I metric provides an indication of the intrinsic switching delay of the device, while $1/\tau$, the reciprocal of CV/I , is a good metric for the intrinsic switching speed of the device. Red/yellow feasibility coloring determined by saturation current-drive feasibility coloring. <u>The C (gate capacitance) includes the effect of parasitic gate capacitance, which has been assumed to be equivalent to that for the high-performance device. For LSTP, τ has been adjusted to reflect the change in scaling of the physical gate length compared to the 2001 ITRS.</u>
Was	[10] Improvement in $1/\tau$ NMOS performance metric normalized to the year 2001. Maintains approximate 14% per year device performance improvement scaling trend for both LOP and LSTP. Red/yellow feasibility coloring determined by saturation current-drive feasibility coloring.
Is	[10] Improvement in $1/\tau$ NMOS performance metric normalized to the year 2001. Maintains approximate 14% per year device performance improvement scaling trend for both LOP and LSTP. Red/yellow feasibility coloring determined by saturation current-drive feasibility coloring. <u>For LSTP, $1/\tau$ has been adjusted to reflect the change in scaling of the physical gate length compared to the 2001 ITRS.</u>
Was	[11] CV^2 switching energy metric for an NMOS device with dimensions $W/L_{gate}=3$. The switching energy metric indicates the amount of dynamic power required to switch the device. Red/yellow feasibility coloring determined by saturation current-drive feasibility coloring.
Is	[11] CV^2 switching energy metric for an NMOS device with dimensions $W/L_{gate}=3$. The <u>power-delay product</u> indicates the amount of dynamic <u>energy</u> required to switch the device <u>through a full (up and down) transition</u> . Red/yellow feasibility coloring determined by saturation current-drive feasibility coloring. <u>For LSTP, the power-delay product has been adjusted to reflect the change in scaling of the physical gate length compared to the 2001 ITRS.</u>
Was	[12] Static power dissipation for an NMOS device with dimensions $W/L_{gate}=3$. Assume that the device static power is primarily determined by the sub-threshold current (since the other junction and leakage current components are assumed to be lower). Yellow/red feasibility coloring is determined by the sub-threshold leakage current feasibility coloring.
Is	[12] Static power dissipation for an NMOS device with dimensions $W/L_{gate}=3$. Assume that the device static power is primarily determined by the sub-threshold current (since the other junction and leakage current components are assumed to be lower). Yellow/red feasibility coloring is determined by the sub-threshold leakage current feasibility coloring. <u>For LSTP, the static power dissipation per device has been adjusted to reflect the change in scaling of the physical gate length compared to the 2001 ITRS.</u>

Table 37a DRAM Technology Requirements—Near-term

Year of Production	2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm) [1]	130	115	100	90	80	70	65
MPU / ASIC ½ Pitch (nm)	150	130	107	90	80	70	65
MPU Printed Gate Length (nm)	90	75	65	53	45	40	35
MPU Physical Gate Length (nm)	65	53	45	37	32	28	25
DRAM cell size (μm^2) [2]	0.135	0.106	0.06	0.049	0.038	0.029	0.025
DRAM storage cell dielectric equivalent physical oxide thickness, EOT (nm) [3]	2.04	1.8	1.2	1	0.45	0.32	0.22
Was DRAM retention time (ms) [4]	64	64	64	64	64	64	64
Is <u>Minimum</u> DRAM retention time (ms) [4]	64	64	64	64	64	64	64
DRAM soft error rate (fits) [5]	1000	1000	1000	1000	1000	1000	1000

Table 37b DRAM Technology Requirements—Long-term

Year of Production	2010	2013	2016
DRAM ½ Pitch (nm) [1]	45	32	22
MPU / ASIC ½ Pitch (nm)	50	35	25
MPU Printed Gate Length (nm)	25	18	13
MPU Physical Gate Length (nm)	18	13	9
DRAM cell size (μm^2) [2]	0.0122	0.0041	0.0019
DRAM storage cell dielectric equivalent physical oxide thickness, EOT (nm) [3]	0.084	0.028	0.01
Was DRAM retention time (ms) [4]	64	64	64
Is <u>Minimum</u> DRAM retention time (ms) [4]	64	64	64
DRAM soft error rate (fits) [5]	1000	1000	1000

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



42 2002 Update Tables

Notes for Table 37a and b

Was	[1] From ORTC (Overall Roadmap Technology Characteristics) Table 1a and b. These DRAM half pitch numbers are smaller than those in the 2000 ITRS, reflecting a speedup in technology development since 1999.
Is	[1] From ORTC (Overall Roadmap Technology Characteristics) Table 1a and b. These DRAM half pitch numbers are smaller than those in the 2000 ITRS, reflecting a speedup in technology development since 1999. <u>However, there is no clear evidence for a further speed up in the pace of DRAM half pitch scaling during 2001 and the early part of 2002.</u>
Was	[2] The DRAM cell size is driven by the values for DRAM capacity (bits per chip) and chip size, as discussed in more detail in the Front End Process chapter. The capacity and chip size numbers used by FEP are based on the ORTC tables 1a and 1b. Since the FEP DRAM capacity and chip size numbers are quite aggressive, the cell size must also be scaled aggressively. The difficulty will lie in reducing the value of the cell size factor "a", where "a" equals (cell size / F ²), and F is the DRAM half pitch. The required values of "a" are 8 for the 130nm node, 6 for 100nm DRAM half pitch, and 4 for the 32nm node. The "a" value of 8 is probably achievable with current techniques, but the "a" value of 6 will require innovative solutions, as illustrated with yellow zone in this line, while the "a" value of 4 has no known solution for 32nm node and beyond, as illustrated with red zone.
Is	[2] The DRAM cell size is driven by the values for DRAM capacity (bits per chip) and chip size, as discussed in more detail in the Front End Process chapter. The capacity and chip size numbers used by FEP are based on the ORTC tables 1a and 1b. Since the FEP DRAM capacity and chip size numbers are quite aggressive, the cell size must also be scaled aggressively. The difficulty will lie in reducing the value of the cell size factor "a", where "a" equals (cell size / F ²), and F is the DRAM half pitch. The required values of "a" are 8 for the 130nm node, 6 for 100nm DRAM half pitch <u>in 2003</u> , and 4 for the 32nm node. The "a" value of 8 is probably achievable with current techniques, but the "a" value of 6 will require innovative solutions, as illustrated with yellow zone in this line, <u>and may be difficult for the industry to achieve per schedule in 2003.</u> The "a" value of 4 has no known solution for 32nm node and beyond, as illustrated with red zone.
Was	[3] The EOT is driven by the values for DRAM capacity (bits per chip) and chip size, as discussed in more detail in the Front End Process chapter. The capacity and the chip size numbers used by FEP are from ORTC Tables 1a and 1b. Since the values of DRAM capacity and chip size from FEP are quite aggressive, the EOT must also be scaled very aggressively. For the 130nm through the 90 nm nodes, the dielectric material is based on Al ₂ O ₃ or Ta ₂ O ₅ with MIS structure, and hence the color is white. Beyond the 90 nm node, breakthroughs such as MIM structure and higher κ material are needed, so the color is yellow. Finally, for the 65nm node and beyond, there are no known solutions with demonstrated credibility, and hence the color is red. The actual EOT required for each node also depends on the other factors such as cell height and/or 3D structure, film leakage current and contact formation. Trench capacitors have other requirements for the cell dielectric material.
Is	[3] The EOT is driven by the values for DRAM capacity (bits per chip) and chip size, as discussed in more detail in the Front End Process chapter. The capacity and the chip size numbers used by FEP are from ORTC Tables 1a and 1b. Since the values of DRAM capacity and chip size from FEP are quite aggressive, the EOT must also be scaled very aggressively. For the 130nm through the 90 nm nodes, the dielectric material is based on Al ₂ O ₃ or Ta ₂ O ₅ with MIS structure, and hence the color is white. <u>(Note that, notwithstanding the trend to Al₂O₃ or Ta₂O₅, many vendors still used oxide-nitride-oxide (ONO) dielectric in 2001.)</u> Beyond the 90 nm node, breakthroughs such as MIM structure and higher κ material are needed, so the color is yellow. Finally, for the 65nm node and beyond, there are no known solutions with demonstrated credibility, and hence the color is red. The actual EOT required for each node also depends on the other factors such as cell height and/or 3D structure, film leakage current and contact formation. Trench capacitors have other requirements for the cell dielectric material.

[4] Retention time is defined at 85 °C, and is the minimum time during which the data from memory can still be sensed correctly without refreshing a row bit line. The 64 ms specified here is the value needed for PC applications. The retention time depends on the combined interaction of device leakage current, signal strength and signal sensing circuit sensitivity, and also depends on operational frequency and temperature.

[5] This is a typical FIT rate and depends on cycle time and the quality of cell capacitor and sensing circuits.

Table 38a Non-Volatile Memory Technology Requirements—Near-term

Year of Production		2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)		130	115	100	90	80	70	65
MPU / ASIC ½ Pitch (nm)		150	130	107	90	80	70	65
MPU Printed Gate Length (nm)		90	75	65	53	45	40	35
MPU Physical Gate Length (nm)		65	53	45	37	32	28	25
Was	Flash technology node - F (nm) [1]	150	130	115	100	90	80	70
Is	Flash technology node - F (nm) [1]	150	130	107	90	80	70	65
Flash NOR cell size—area factor a in multiples of F ² [2]		10–12	10–12	10–12	11–14	11–14	11–14	11–14
Was	Flash NAND cell size—area factor a in multiples of F ² SLC/MLC [3]	5.5	5.5	4.5	4.5	4.5	4.5/2.3	4.5/2.3
Is	Flash NAND cell size—area factor a in multiples of F ² SLC/MLC [3]	5.5	5.5	5.5	5.5	5.5	5.5	4.5
Was	Flash NOR typical cell size (µm ²) [4]	0.248	0.186	0.145	0.125	0.101	0.08	0.061
Is	Flash NOR typical cell size (µm ²) [4]	0.248	0.186	0.135	0.101	0.08	0.061	0.061
Was	Flash NOR Lg-stack (physical- µm) [5]	0.29–0.31	0.25–0.27	0.22–0.24	0.21–0.23	0.2–0.22	0.2–0.22	0.19–0.21
Is	Flash NOR Lg-stack (physical- µm) [5]	0.29–0.31	0.25–0.27	0.22–0.24	0.2–0.22	0.2–0.22	0.19–0.21	0.19–0.21
Was	Flash NOR highest W/E voltage (V) [6]	8–10	8–10	8–10	8–10	7–9	7–9	7–9
Is	Flash NOR highest W/E voltage (V) [6]	8–10	8–10	8–10	7–9	7–9	7–9	7–9
Was	Flash NAND highest W/E voltage (V) [7]	19–21	18–20	18–20	18–20	18–20	17–19	17–19
Is	Flash NAND highest W/E voltage (V) [7]	19–21	18–20	18–20	17–19	17–19	17–19	15–17
Was	Flash NOR I _{read} (µA) [8]	36–44	35–43	34–42	33–41	31–39	28–36	29–37
Is	Flash NOR I _{read} (µA) [8]	36–44	35–43	34–42	31–39	29–37	28–36	27–35
Was	Flash Coupling Ratio [9]	0.65–0.75	0.65–0.75	0.65–0.75	0.65–0.75	0.65–0.75	0.65–0.75	0.6–0.7
Is	Flash Coupling Ratio [9]	0.65–0.75	0.65–0.75	0.65–0.75	0.65–0.75	0.65–0.75	0.6–0.7	0.6–0.7
Was	Flash NOR tunnel oxide thickness (nm) [10]	9.5–10.5	9.5–10	9–10	9–10	8.5–9.5	8.5–9.5	8.5–9.5
Is	Flash NOR tunnel oxide thickness (nm) [10]	9.5–10.5	9–10	9–10	8.5–9.5	8.5–9.5	8.5–9.5	8–9
Was	Flash NAND tunnel oxide thickness (nm) [11]	8.5–9.5	8.5–9	8–9	8–9	8–9	7.5–8	7.5–8
Is	Flash NAND tunnel oxide thickness (nm) [11]	8.5–9.5	7–8	7–8	7–8	7–8	7–8	6–7
Was	Flash NOR interpoly dielectric thickness (nm) [12]	13–15	12–14	11–13	11–13	10–12	9–11	9–11
Is	Flash NOR interpoly dielectric thickness (nm) [12]	13–15	12–14	11–13	10–12	9–11	9–11	8.5–10.5
Was	Flash NAND interpoly dielectric thickness (nm) [13]	14–16	13–15	12–14	12–14	12–14	11–13	10–12
Is	Flash NAND interpoly dielectric thickness (nm) [13]	14–16	13–15	13–15	13–15	13–15	13–15	10–13
Flash endurance (erase/write cycles) [14]		1.00E+05	1.00E+05	1.00E+05	1.00E+05	1.00E+05	1.00E+05	1.00E+05
Flash nonvolatile data retention (years) [15]		10	10–20	10–20	10–20	10–20	10–20	10–20
Flash maximum number of bits per cell (MLC) [16]		2	2	4	4	4	4	4
FeRAM technology node - F (nm) [17]		500	350	250	220	180	150	130
FeRAM cell size—area factor a in multiples of F ² [18]		60	40	24	16	10	10	10
FeRAM cell size (µm ²) [19]		15	4.9	1.5	0.518	0.324	0.225	0.169
FeRAM cell structure [20]		2T2C	1T1C	1T1C	1T1C	1T1C	1T1C	1T1C
FeRAM capacitor structure [21]		planar	planar	stack	stack	stack	stack	3D
Ferro capacitor voltage (V) [22]		3	3	2.5	1.8	1.5	1.3	1.2
FeRAM endurance (read/write cycles) [23]		1.00E+12	1.00E+13	1.00E+14	1.00E+15	>1E16	>1E16	>1E16
FeRAM nonvolatile data retention (years) [24]		10	10	10	10	10	10	10

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



Table 38b Non-Volatile Memory Technology Requirements—Long-term

Year of Production		2010	2013	2016
DRAM ½ Pitch (nm)		45	32	22
MPU / ASIC ½ Pitch (nm)		50	35	25
MPU Printed Gate Length (nm)		25	18	13
MPU Physical Gate Length (nm)		18	13	9
Flash technology node - F (nm) [1]		50	35	25
Flash NOR cell size —area factor a in multiples of F^2 [2]		12–15	13–16	14–17
Flash NAND cell size —area factor a in multiples of F^2 SLC/MLC [3]		4.5/2.3	4.5/2.3	4.5/2.3
Flash NOR typical cell size (μm^2) [4]		0.034	0.018	0.01
Flash NOR Lg-stack (physical- μm) [5]		0.17–0.19	0.14–0.16	0.12–0.14
Flash NOR highest W/E voltage (V) [6]		7–9	7–9	7–9
Was	Flash NAND highest W/E voltage (V) [7]	17–19	16–18	16–18
Is	Flash NAND highest W/E voltage (V) [7]	15–17	15–17	15–17
Flash NOR I_{read} (μA) [8]		27–33	25–31	22–28
Flash Coupling Ratio [9]		0.6–0.7	0.6–0.7	0.6–0.7
Flash NOR tunnel oxide thickness (nm) [10]		8–9	8	8
Flash NAND tunnel oxide thickness (nm) [11]		6–7	6–7	6–7
Flash NOR interpoly dielectric thickness (nm) [12]		8–10	6–8	4–6
Was	Flash NAND interpoly dielectric thickness (nm) [13]	10–12	9–11	9–11
Is	Flash NAND interpoly dielectric thickness (nm) [13]	10–13	9–10	9–10
Flash endurance (erase/write cycles) [14]		1.00E+06	1.00E+06	1.00E+07
Flash nonvolatile data retention (years) [15]		10–20	20	20
Flash maximum number of bits per cell (MLC) [16]		8	8	8
FeRAM technology node - F (nm) [17]		100	70	50
FeRAM cell size —area factor a in multiples of F^2 [18]		8	8	8
FeRAM cell size (μm^2) [19]		0.08	0.039	0.02
FeRAM cell structure [20]		1T1C	1T1C	1T1C
FeRAM capacitor structure [21]		3D	3D	3D
Ferro capacitor voltage (V) [22]		1	0.7	0.7
FeRAM endurance (read/write cycles) [23]		>1E16	>1E16	>1E16
FeRAM nonvolatile data retention (years) [24]		10	10	10

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known



OVERALL ROADMAP TECHNOLOGY CHARACTERISTICS

BACKGROUND

The Overall Roadmap Technology Characteristics (ORTC) tables are created early in the Roadmap process and are used as the basis for initiating the activities of the International Technology Working Groups (ITWGs) in producing their detailed chapters. These tables are also used throughout the renewal effort of the Roadmap as a means of providing synchronization among the TWGs by highlighting inconsistencies between the specific tables. The process to revise the tables include increasing levels of cross-TWG and international coordination and consensus building to develop underlying models of trends and to reach agreement on target metrics. As a result, the ORTC tables went through several iterations and reviews.

The metric values of the ORTC tables can be found throughout the Roadmap in greater detail in each Technology Working Group chapter. The information in this section is intended to highlight the current rapid pace of advancement in semiconductor technology. It represents a completion of the revision update and renewal work that began in 2000. Additionally, an ORTC Glossary is provided as an appendix.

OVERVIEW OF 2001 REVISIONS

DEFINITIONS

As noted above, the Overall Roadmap Technology Characteristics tables provide a consolidated summary of the key technology metrics. Please note that the year header on the tables may refer to different points in the development/life cycle of integrated circuits (ICs), depending on the individual line item metric. However, unless otherwise specified for a particular line item, the default year header still refers (as in previous Roadmaps) to the year when product shipment first exceeds 10,000 units per month of ICs from a manufacturing site using “production tooling.” Additional clarification was provided this year by the *ITRS* executive International Roadmap Committee (IRC), requiring a second company to start production within three months. To satisfy this definition, ASIC production may represent the cumulative volume of many individual product line items processed through the facility. Please see the Glossary section for additional details on “Technology Node” and “Production” timing definitions.

Furthermore, new IRC guidelines clarified the definition of a technology node as the achievement of significant advancement in the process technology. To be explicit, a technology node was defined as the achievement of an approximate 0.7× reduction per node (0.5× per two nodes) (Figure 5). The period of time in which a new technology node is reached is called a “technology-node cycle” (Figure 6). It is acknowledged that continuous improvement occurs between the technology nodes, and this is reflected by including data between nodes in the annual columns of the “Near-term years” tables. The “Long-term years” table columns are three-year increments of the 2001 *ITRS* timeframe.

ROADMAP TIMELINE

The 2001 edition of the Roadmap maintains a 15-year projection, from 2001 as a reference year and through 2016. However, the timing of future technology nodes has changed from the 1999 edition.

By international consensus, the 130 nm node was pulled in an additional year, continuing the historical precedent for two-year technology-node cycles since 350 nm/1995 (250 nm node in 1997, 180 nm node in 1999). Although there is the possibility of a continuation of the two-year-node cycle trend, the present consensus projects a three-year cycle for DRAM interconnect half-pitch nodes throughout the 2001–2016 Roadmap period (Figure 7).

DRAM interconnect (metal or polysilicon—refer to the Glossary for additional detail) half-pitch will continue to be used as the most representative feature of leading-edge semiconductor manufacturing technology for defining the achievement of a technology node. However, additional data analysis indicates an aggressive trend for the lagging MPU and ASIC interconnect half-pitches to catch up to DRAM half-pitch by 2004 (Figure 7).

The 2001 *ITRS* includes a correction of the past “rounding” convention for the technology node labels. The actual mathematical trend reduces the nodes by 50% every other node, resulting in an actual versus rounded node number targets, starting from 350 nm in 1995 as follows:

Table C Rounded versus Actual Trend Numbers

<i>YEAR OF PRODUCTION</i>	<i>1995</i>	<i>1997</i>	<i>1999</i>	<i>2001</i>	<i>2004</i>	<i>2007</i>	<i>2010</i>	<i>2013</i>	<i>2016</i>
<i>Actual Trend Numbers (nm)</i>	350	247.5	175	123.7	87.5	61.9	43.8	30.9	21.9
<i>ITRS Rounded Node Numbers (nm)</i>	350	250	180	130	90	65	45	32	22

Note the new “rounding” corrections become more critical as the industry moves into the double-digit technology nodes. Some regions, for past publication consistency, will continue to track the previous technology nodes beginning with 100 nm/2003, resulting in milestones that are placed one year earlier than the present 2001 roadmap convention (70 nm/2006; 50 nm/2009; 35nm/2012; 25 nm/2015).

The *printed* MPU gate length received a major correction to more an aggressive starting point in 2001. In addition, a new *physical* gate length is being tracked that further reduces the bottom gate length dimension of a fully-processed transistor. Both the printed and physical gate length trends are forecast to continue scaling by about 70% per two-year cycle through the 32nm physical MPU gate length in 2005, but are expected to return to a three-year cycle trend thereafter, consistent with the present DRAM half-pitch trend forecast (Figure 8).

The ORTC metrics, which guide the Roadmap, are often used by semiconductor companies as a set of targets that need to be achieved ahead of schedule to achieve industry leadership. Thus, the highly competitive environment of the semiconductor industry quickly tends to make obsolete many portions of the ORTC metrics and, consequently, the Roadmap. Hopefully, our annual update process will provide sufficiently close tracking of the evolving international consensus on technology directions to maintain the usefulness of the *ITRS* to the industry.

For example, there is some anticipation that DRAM half-pitch nodes could undergo an additional one-year pull-in. This possibility will be re-evaluated during the year 2002 *ITRS* Update process, along with the possibility of using a two-year node cycle as a longer-term trend. To reflect the variety of cycles and to allow for closer monitoring of future roadmap shifts, it was agreed to continue the practice of publishing annual technology requirements from 2001 through 2007, called the “Near-Term Years,” and at three-year (node) intervals thereafter, called the “Long-Term Years” (2010, 2013, 2016).

MOS Transistor *Scaling* (1974 to present)

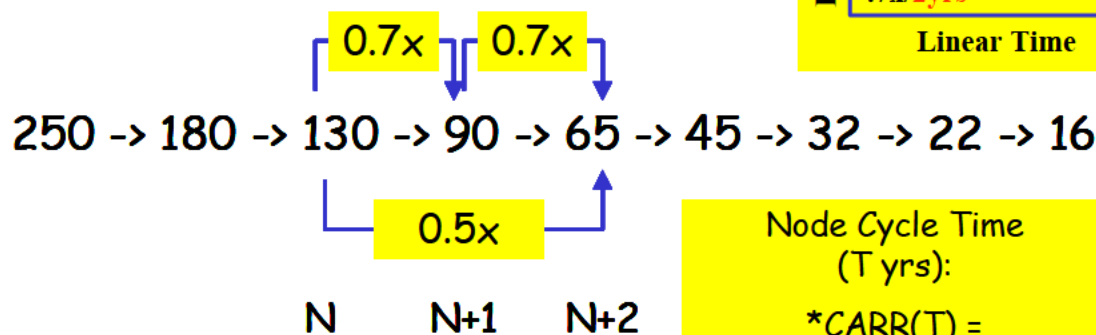
$$S=0.7$$

$$[0.5x \text{ per } 2 \text{ nodes}]$$

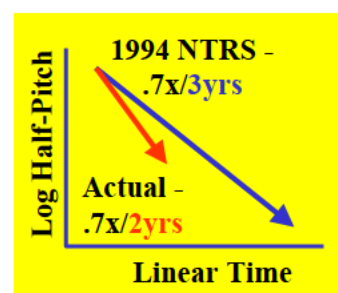


Figure 5 MOS Transistor Scaling—1974 to present

Scaling Calculator + Node Cycle Time:



* CARR(T) = Compound Annual
Reduction Rate
(@ cycle time period, T)



Node Cycle Time
(T yrs):

*CARR(T) =

$$[(0.5)^{(1/2T \text{ yrs})}] - 1$$

CARR(3 yrs) = -10.9%

CARR(2 yrs) = -15.9%

Figure 6 Scaling Calculator

ITRS Roadmap Acceleration Continues...Half Pitch

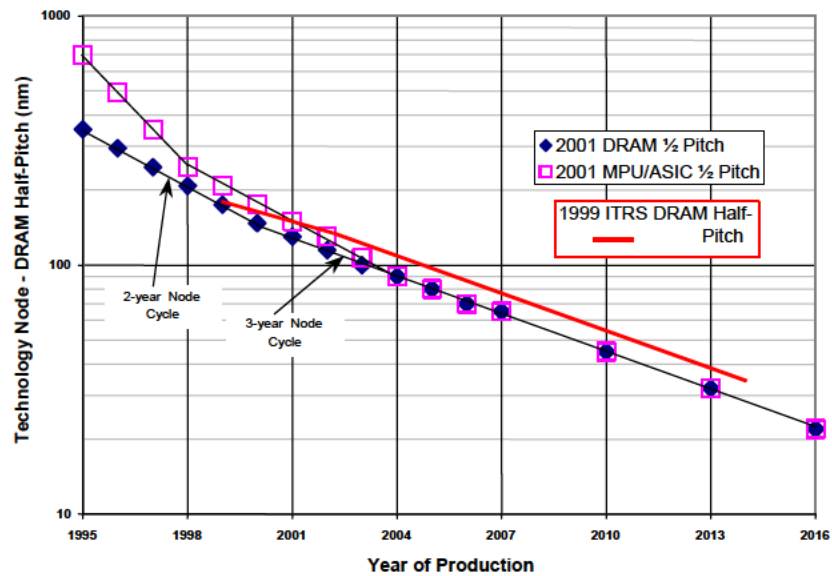


Figure 7 ITRS Roadmap Acceleration Continues—Half Pitch Trends

ITRS Roadmap Acceleration Continues...Gate Length

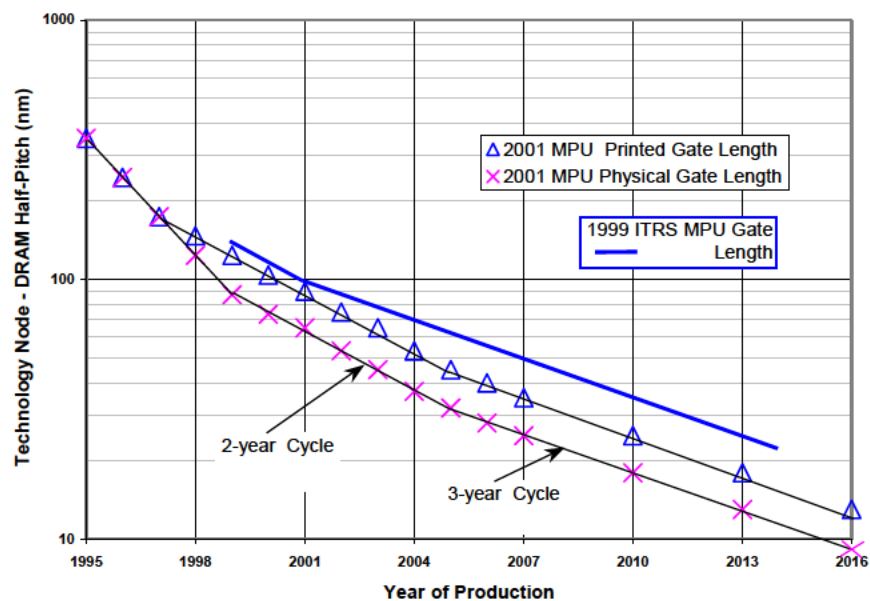


Figure 8 ITRS Roadmap Acceleration Continues—Gate Length Trends

PRODUCT GENERATIONS AND CHIP-SIZE MODEL

In this section, we will discuss “product generations” and their relationship to the technology nodes, since, in the past, these terms have often been used interchangeably. However, the historically simple picture of a new DRAM product generation every three years (at 4× the previous density and based on an essentially new set of technology features) has become obsolete as a way to define technology nodes. For this 2001 edition, the “technology node” is still linked to an anticipated DRAM feature size (minimum metal or polysilicon half-pitch). However, implications of this connection are diminishing as the product evolution/shrink path becomes more complex.

Historically, DRAM products have been recognized as the technology drivers for the entire semiconductor industry. Prior to the early 1990s, logic (as exemplified by MPU) technology was developed at a slower pace than DRAM technology. During the last few years, the development rate of new technologies used to manufacture microprocessors has accelerated. Microprocessor products are closing the half-pitch technology gap with DRAM, and are now driving the most leading-edge lithography tools and processes—especially for the capability to process the isolated-line feature of the printed and physical gate length. With this 2001 Roadmap it is recognized that DRAM and microprocessor products share the technology leadership role.

However, several fundamental differences exist between the two families of products. Due to strong commodity market economic pressure to reduce cost and increase fab output productivity, DRAM product emphasizes the minimization of the chip size. Therefore development of DRAM technology focuses mainly on minimization of the area occupied by the memory cell. However, this pressure to minimize cell size is in conflict with the requirement to maximize the capacitance of the cell for charge storage performance, which puts pressure on memory cell designers to find creative ways through design and materials to meet minimum capacitance requirements while reducing cell size. In addition, to closely pack the highest number of DRAM cells in the smallest area requires minimization of cell pitch.

Microprocessors have also come under strong market pressure to reduce costs while maximizing performance. Performance is enabled primarily by the length of the transistor gate and by the number of interconnect layers. The 2001 *ITRS* teams have reached consensus on models for the required functionality, chip size, cell area, and density for the ORTC tables. Additional line items were added to communicate the model consensus, and the underlying model assumptions are included in notations. Table 1a and 1b summarize the near and long technology node metrics. As agreed, the key *ITRS* technology node identifier would continue to be the DRAM half-pitch, but also included are the aggressive MPU gate-length performance-driven feature sizes. For completeness, the MPU/ASIC product metal half-pitch are also tracked and that will trail slightly behind or equal to the DRAM half-pitch. The ASIC/low power gate lengths are also included, and lag behind the leading-edge MPU in order to maximize standby and operating current drain. See the Glossary section for additional detail on the definition of the half-pitch and gate-length features. For each product generation, both the leading-edge (“at introduction”) and the high-volume (“at production”) DRAM products are indicated

It should be noted that the long-term average annualized reduction rate in feature size is projected to continue at approximately 11%/year (~30% reduction/three years), even though this rate accelerated to approximately 16%/year (~30% reduction/two years) in the time interval 1995–2001 (refer to Figure 5). As mentioned above, the overall schedule for introduction of a new product generation has been accelerated by one additional year.

Table 1a Product Generations and Chip Size Model Technology Nodes—Near-term Years

<i>Year of Production</i>	<i>2001</i>	<i>2002</i>	<i>2003</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>
<i>DRAM ½ Pitch (nm)</i>	130	115	100	90	80	70	65
<i>MPU/ASIC ½ Pitch (nm)</i>	150	130	107	90	80	70	65
<i>MPU Printed Gate Length (nm) ††</i>	90	75	65	53	45	40	35
<i>MPU Physical Gate Length (nm)</i>	65	53	45	37	32	28	25
<i>ASIC/Low Power Printed Gate Length (nm) ††</i>	130	107	90	75	65	53	45
<i>ASIC/Low Power Physical Gate Length (nm)</i>	90	75	65	53	45	37	32

Table 1b Product Generations and Chip Size Model Technology Nodes—Long-term years

<i>Year of Production</i>	<i>2010</i>	<i>2013</i>	<i>2016</i>
<i>DRAM ½ Pitch (nm)</i>	45	32	22
<i>MPU/ASIC ½ Pitch (nm)</i>	45	32	22
<i>MPU Printed Gate Length (nm) ††</i>	25	18	13
<i>MPU Physical Gate Length (nm)</i>	18	13	9
<i>ASIC/Low Power Printed Gate Length (nm) ††</i>	32	22	16
<i>ASIC/Low Power Physical Gate Length (nm)</i>	22	16	11

Notes for Tables 1a and 1b

†† MPU and ASIC gate-length (in resist) node targets refer to the most aggressive requirements, as printed in photoresist (which was by definition also “as etched in polysilicon,” in the 1999 ITRS).

However, during the 2000/2001 ITRS development, trends were identified, in which the MPU and ASIC “Physical” gate lengths may be reduced from the “as-printed” dimension. These “Physical” gate-length targets are driven by the need for maximum speed performance in logic Microprocessor (MPU) products, and are included in the Front End Processes (FEP), Process Integration, Devices, and Structures (PIDs), and Design ITWG Tables as needs that drive device design and process technology requirements.

Table 1c DRAM Production Product Generations and Chip Size Model—Near-term Years

Year of Production	2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)	130	115	100	90	80	70	65
MPU/ASIC ½ Pitch (nm)	150	130	107	90	80	70	65
MPU Printed Gate Length (nm)	90	75	65	53	45	40	35
MPU Physical Gate Length (nm)	65	53	45	37	32	28	25
Cell area factor [a]	8	8	6	6	6	6	6
Cell area [$Ca = af^2$] (μm^2)	0.130	0.103	0.061	0.049	0.039	0.031	0.024
Cell array area at production (% of chip size) §	54.8%	55.3%	55.7%	56.1%	56.4%	56.7%	57.0%
Generation at production §	512M	512M	1G	1G	2G	2G	4G
Functions per chip (Gbits)	0.54	0.54	1.07	1.07	2.15	2.15	4.29
Chip size at production (mm^2)§	127	100	118	93	147	116	183
Gbits/cm ² at production §	0.42	0.54	0.91	1.15	1.46	1.85	2.35

Table 1d DRAM Production Product Generations and Chip Size Model—Long-term Years

Year of Production	2010	2013	2016
DRAM ½ Pitch (nm)	45	32	22
MPU/ASIC ½ Pitch (nm)	45	32	22
MPU Printed Gate Length (nm)	25	18	13
MPU Physical Gate Length (nm)	18	13	9
Cell area factor [a]	6	4	4
Cell area [$Ca = af^2$] (μm^2)	0.012	0.004	0.002
Cell array area at production (% of chip size) §	57.7%	58.1%	58.4%
Generation at production §	8G	32G	64G
Functions per chip (Gbits)	8.59	34.36	68.72
Chip size at production (mm^2)§	181	239	238
Gbits/cm ² at production §	4.75	14.35	28.85

Notes for Tables 1c and 1d

§ DRAM Model—Cell Factor (design/process improvement) targets are as follows

1999–2002/8×; 2003–2010/6×; 2011–2016/4×

DRAM product generations are usually increased by 4×bits/chip every four years with interim 2×bits/chip generations, except

- at the Introduction phase, after the 8Gbit interim generation, the introduction rate is 4×/five years (2×/two–three years); and
- at the Production phase, after the interim 32Gbit generation, the introduction rate is 4×/five years (2×/two–three years).

InTER-generation chip size growth rate varies to maintain one chip per 572mm² field at Introduction and two chips per 572mm² field at Production. The more aggressive “best case opportunity” technology node trends allow the Production-phase products to remain at 2×bits/chip every two years and still fit within the target of two DRAM chips per 572mm² field size, through the 32Gbit interim generation. The InTRA-generation chip size shrink model is 0.5× every technology node in-between cell factor reductions.

Refer to the Glossary for definitions of Introduction, Production, InTERgeneration, and InTRAgeneration terms.

Table 1e DRAM Introduction Product Generations and Chip Size Model—Near-term Years

Year of Production	2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)	130	115	100	90	80	70	65
MPU/ASIC ½ Pitch (nm)	150	130	107	90	80	70	65
MPU Printed Gate Length (nm)	90	75	65	53	45	40	35
MPU Physical Gate Length (nm)	65	53	45	37	32	28	25
Cell area factor [a]	8	8	6	6	6	6	6
Cell area [$Ca = af^2$] (μm^2)	0.130	0.103	0.061	0.049	0.039	0.031	0.024
Cell array area at introduction (% of chip size) §	71.3%	71.8%	72.2%	72.6%	72.9%	73.2%	73.5%
Generation at introduction §	2G	2G	4G	4G	8G	8G	16G
Functions per chip (Gbits)	2.15	2.15	4.29	4.29	8.59	8.59	17.18
Chip size at introduction (mm^2) §	390	308	364	287	454	359	568
Gbits/cm ² at introduction §	0.55	0.70	1.18	1.49	1.89	2.39	3.03

Table 1f DRAM Introduction Product Generations and Chip Size Model—Long-term Years

Year of Production	2010	2013	2016
DRAM ½ Pitch (nm)	45	32	22
MPU/ASIC ½ Pitch (nm)	45	32	22
MPU Printed Gate Length (nm)	25	18	13
MPU Physical Gate Length (nm)	18	13	9
Cell area factor [a]	6	4	4
Cell area [$Ca = af^2$] (μm^2)	0.012	0.004	0.002
Cell array area at introduction (% of chip size) §	74.2%	74.6%	74.9%
Generation at introduction §	32G	64G	64G
Functions per chip (Gbits)	34.36	68.72	68.72
Chip size at introduction (mm^2) §	563	373	186
Gbits/cm ² at introduction §	6.10	18.42	37.00

Notes for Tables 1e and 1f

§ DRAM Model—Cell Factor (design/process improvement) targets are as follows

1999–2002/8×; 2003–2010/6×; 2011–2016/4×

DRAM product generations are usually increased by 4×bits/chip every four years with interim 2×bits/chip generations, except

5. at the Introduction phase, after the 8Gbit interim generation, the introduction rate is 4×/five years (2×/two–three years); and
6. at the Production phase, after the interim 32Gbit generation, the introduction rate is 4×/five years (2×/two–three years).

InTER-generation chip size growth rate varies to maintain one chip per 572mm² field at Introduction and two chips per 572mm² field at Production. The more aggressive “best case opportunity” technology node trends allow the Production-phase products to remain at 2×bits/chip every two years and still fit within the target of two DRAM chips per 572mm² field size, through the 32Gbit interim generation. The InTRA-generation chip size shrink model is 0.5×every technology node in-between cell factor reductions.

Refer to the Glossary for definitions of Introduction, Production, InTERgeneration, and InTRAgeneration terms.

Table 1g MPU (High-volume Microprocessor) Cost-Performance Product Generations and Chip Size Model—Near-term Years

<i>Year of Production</i>	<i>2001</i>	<i>2002</i>	<i>2003</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>
<i>DRAM ½ Pitch (nm)</i>	130	115	100	90	80	70	65
<i>MPU/ASIC ½ Pitch (nm)</i>	150	130	107	90	80	70	65
<i>MPU Printed Gate Length (nm)</i>	90	75	65	53	45	40	35
<i>MPU Physical Gate Length (nm)</i>	65	53	45	37	32	28	25
<i>SRAM Cell (6-transistor) Area factor ++</i>	126.1	123.0	120.3	117.8	115.6	113.7	111.9
<i>Logic Gate (4-transistor) Area factor ++</i>	320.0	320.0	320.0	320.0	320.0	320.0	320.0
<i>SRAM Cell (6-transistor) Area efficiency ++</i>	0.63	0.63	0.63	0.63	0.63	0.63	0.63
<i>Logic Gate (4-transistor) Area efficiency ++</i>	0.50	0.50	0.50	0.50	0.50	0.50	0.50
<i>SRAM Cell (6-transistor) Area w/overhead ++</i>	3.3	2.5	2.0	1.5	1.2	0.93	0.73
<i>Logic Gate (4-transistor) Area w/overhead ++</i>	10.4	8.2	6.5	5.2	4.1	3.3	2.6
<i>Transistor density SRAM (Mtransistors/cm²)</i>	184	237	305	393	504	646	827
<i>Transistor density logic (Mtransistors/cm²)</i>	38.6	48.6	61.2	77.2	97.2	122.5	154.3
<i>Generation at introduction *</i>	p04c	—	—	p07c	—	—	p10c
<i>Functions per chip at introduction (million transistors [Mtransistors])</i>	193	243	307	386	487	614	773
<i>Chip size at introduction (mm²) ‡</i>	280	280	280	280	280	280	280
<i>Cost performance MPU (Mtransistors/cm² at introduction) (including on-chip SRAM) ‡</i>	69	87	110	138	174	219	276
<i>Generation at production *</i>	p01c	—	—	p04c	—	—	p07c
<i>Functions per chip at production (million transistors [Mtransistors])</i>	97	122	153	193	243	307	386
<i>Chip size at production (mm²) §§</i>	140	140	140	140	140	140	140
<i>Cost performance MPU (Mtransistors/cm² at production, including on-chip SRAM) ‡</i>	69.0	87.0	109.6	138.0	173.9	219.1	276.1

Notes for Tables 1g and 1h

++ The MPU area factors are analogous to the "cell area factor" for DRAMs. The reduction of area factors has been achieved historically through a combination of many factors, for example—use of additional interconnect levels, self-alignment techniques, and more efficient circuit layout. However, recent data has indicated that the improvement (reduction) of the area factors is slowing, and is virtually flat for the logic gate area factor.

* p is processor; numerals reflect year of production; c indicates cost-performance product. Examples—the cost-performance processor, p01c, was introduced in 1999, but not ramped into volume production until 2001; similarly, the p04c, is introduced in 2001, but is targeted for volume production in 2004.

‡ MPU Cost-performance Model—Cost-performance MPU includes Level 2 (L2) on-chip SRAM (512Kbyte/1999), and the combination of both SRAM and logic transistor functionality doubles every technology node cycle.

§§ MPU Chip Size Model—Both the cost-performance and high-performance MPUs InTER-generation chip size growth rates are targeted to be flat through 2016, made possible by doubling the on-chip functionality every technology node cycle. The InTRA-generation chip size shrink model is 0.5× every two-year technology node through 2001, then 0.5× every three-year technology node after 2001.

Table 1h MPU (High-volume Microprocessor) Cost-Performance Product Generations and Chip Size Model—Long-term Years

<i>Year of Production</i>	<i>2010</i>	<i>2013</i>	<i>2016</i>
<i>DRAM ½ Pitch (nm)</i>	45	32	22
<i>MPU/ASIC ½ Pitch (nm)</i>	45	32	22
<i>MPU Printed Gate Length (nm)</i>	25	18	13
<i>MPU Physical Gate Length (nm)</i>	18	13	9
<i>SRAM Cell (6-transistor) Area factor ++</i>	107.8	106.7	105.7
<i>Logic Gate (4-transistor) Area factor ++</i>	320.0	320.0	320.0
<i>SRAM Cell (6-transistor) Area efficiency ++</i>	0.63	0.63	0.63
<i>Logic Gate (4-transistor) Area efficiency ++</i>	0.50	0.50	0.50
<i>SRAM Cell (6-transistor) Area w/overhead ++</i>	0.22	0.17	0.13
<i>Logic Gate (4-transistor) Area w/overhead ++</i>	0.82	0.65	0.51
<i>Transistor density SRAM (Mtransistors/cm²)</i>	1718	3532	7208
<i>Transistor density logic (Mtransistors/cm²)</i>	309	617	1235
<i>Generation at introduction *</i>	p13c	p16c	p19c
<i>Functions per chip at introduction (million transistors [Mtransistors])</i>	1546	3092	6184
<i>Chip size at introduction (mm²) ‡</i>	280	280	280
<i>Cost performance MPU (Mtransistors/cm² at introduction) (including on-chip SRAM) ‡</i>	552	1104	2209
<i>Generation at production *</i>	p10c	p13c	p16c
<i>Functions per chip at production (million transistors [Mtransistors])</i>	773	1546	3092
<i>Chip size at production (mm²) §§</i>	140	140	140
<i>Cost performance MPU (Mtransistors/cm² at production, including on-chip SRAM) ‡</i>	552	1104	2209

Table 1i High-Performance MPU and ASIC Product Generations and Chip Size Model—Near-term Years

Year of Production	2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)	130	115	100	90	80	70	65
MPU/ASIC ½ Pitch (nm)	150	130	107	90	80	70	65
MPU Printed Gate Length (nm)	90	75	65	53	45	40	35
MPU Physical Gate Length (nm)	65	53	45	37	32	28	25
Logic (Low-volume Microprocessor) High-performance ‡							
Generation at production **	p01h	—	p03h	—	p05h	—	p07h
Functions per chip (million transistors)	276	348	439	553	697	878	1106
Chip size at production (mm ²) §§	310	310	310	310	310	310	310
High-performance MPU Mtransistors/cm ² at production (including on-chip SRAM) ‡	89	112	142	178	225	283	357
ASIC							
ASIC usable Mtransistors/cm ² (auto layout)	89	112	142	178	225	283	357
ASIC max chip size at production (mm ²) (maximum lithographic field size)	800	800	572	572	572	572	572
ASIC maximum functions per chip at production (Mtransistors/chip) (fit in maximum lithographic field size)	714	899	810	1020	1286	1620	2041

Table 1j High-Performance MPU and ASIC Product Generations and Chip Size Model—Long-term Years

Year of Production	2010	2013	2016
DRAM ½ Pitch (nm)	45	32	22
MPU/ASIC ½ Pitch (nm)	45	32	22
MPU Printed Gate Length (nm)	25	18	13
MPU Physical Gate Length (nm)	18	13	9
Logic (Low-volume Microprocessor) High-performance ‡			
Generation at production **	—	p13h	—
Functions per chip (million transistors)	2212	4424	8848
Chip size at production (mm ²) §§	310	310	310
High-performance MPU Mtransistors/cm ² at production (including on-chip SRAM) ‡	714	1427	2854
ASIC			
ASIC usable Mtransistors/cm ² (auto layout)	714	1427	2854
ASIC maximum chip size at production (mm ²) (maximum lithographic field size)	572	572	572
ASIC maximum functions per chip at ramp (Mtransistors/chip) (fit in maximum lithographic field size)	4081	8163	16326

Notes for Tables 1i and 1j

** p is processor, numerals reflect year of production; h indicates high-performance product. Examples—the high-performance processor, p99h, was ramped into volume production in 1999; similarly, the p01h, is introduced in 2001.

‡ MPU High-performance Model—High-performance MPU includes large L2 and L3 on-chip SRAM (2MByte/1999) plus a larger logic core (P99h core = 25M transistor (Mtransistors) both SRAM and Logic functionality doubles every technology node cycle.

§§ MPU Chip Size Model—Both the cost-performance and high-performance MPUs InTER-generation chip size growth rates are targeted to be flat through 2016, made possible by doubling the on-chip functionality every technology node cycle. The InTRA-generation chip size shrink model is 0.5x every two-year technology node through 2001, then 0.5x every three-year technology node after 2001.

CHIP-SIZE, LITHOGRAPHIC-FIELD, AND WAFER-SIZE TRENDS

Despite the continuous reduction in feature size of about 30% every three years, the size of first DRAM product demonstration in technical forums such as the IEEE International Solid State Circuits Conference (ISSCC) has continued to double every six years (an increase of about 12%/year). This increase in chip area has been necessary to accommodate 59% more bits/capacitors/transistors per year in accordance with Moore's Law (historically doubling functions per chip every 1.5–2 years). However, to maintain the historical trend of reducing cost/function by ~25–30%/year, it is necessary to continuously enhance equipment productivity, increase manufacturing yields, use the largest wafer size available, and, most of all, increase the number of chips available on a wafer.

The increase in the gross number of chips available on a wafer is primarily obtained by reducing the area of the chip by means of a combination of smaller feature size (shrink/scaling) and product/process redesign (compaction). For instance, using the latest models, it is forecast that the introduction chip area of a cost-effective product generation [which doubles the inter-generation (generation-to-generation) functionality every two years] must either remain flat or grow no faster than 20% every four years. Furthermore, the area must be shrunk at an intra-generation (within a generation) annual reduction rate of 50% (the square of the $.7\times$ lithography reduction rate) during every technology node period.

Affordable DRAM products must also achieve virtually flat intra-generation chip-sizes, and they must also maintain a cell area array efficiency ratio of less than 70% of total chip area. Therefore, DRAM products require aggressive cell area factors (cell area in units of minimum-feature-size-squared). The Front End Processes International Technology Working Group has provided the cell area factors and detailed the challenges and needs for solutions to meet the aggressive cell area goals in the Front End Processes chapter. Due to the importance of tracking/coordinating these new challenges, the DRAM cell area factor, the target cell sizes, and the cell array area percentage of total chip-size line items will continue to be tracked in ORTC Tables 1c, d, e, and f. (Refer to the Glossary for additional details.)

In 2001 the Design ITWG improved the MPU chip size model to update with the latest transistor densities, large on-chip SRAM, and smaller target chip sizes. The Design ITWG has also added additional detail to the model, including transistor design improvement factors. The Design ITWG notes that design improvements occur at a slow rate in SRAM transistors and very little in logic gate transistors. Almost all the “shrink” and density improvement comes from lithography-enabled interconnect half-pitch scaling alone.

The present MPU chip size model reflects additional competitive requirements for affordability and power management by targeting flat chip size trends for both high-performance MPUs (310mm^2) and cost-performance MPUs (140mm^2). Due to the MPU two-year-cycle half-pitch “catch-up phase” through the year 2004, the MPU products may be able to maintain flat chip sizes due to lithography improvements alone. However, after 2004, the intra-generation chip size of MPUs can remain flat only by slowing the rate of on-chip transistors to double every technology node.

Due to the forecasted return to a three-year technology node cycle, the present MPU chip-size model slows the Moore's Law rate of on-chip transistors to $2\times$ every three years. In order to maintain a flat chip size target and also return to the historical doubling every two years of on-chip functionality (transistors), MPU chip and process designers must add additional design/process improvements to the fundamental lithography-based scaling trends. The new target metrics of the MPU model are summarized in Tables 1g, h, i, and j.

To improve productivity, it is necessary to increase the output of good chips at each step in the fabrication process. The ability of printing multiple chips in a single exposure is a key productivity driver and is determined by the field size of the lithographic tool and the size and aspect ratio of the chips being printed on the wafer. In the past, lithography exposure field sizes doubled every other technology node to meet the demand for increasing chip sizes. The result was the achievement of very large step-and-scan fields ($25\times 32 = 800\text{mm}^2$) by 1999. However, the Lithography ITWG indicates that maintaining the large field size under continued reduction of exposure features is increasing costs dramatically. Therefore, the ITWG forecasts a requirement for the economically affordable lithography field to be reduced to a 572mm^2 level (22×26) by the 90 nm node. That trend is shown in Tables 2a and b.

DRAM chip sizes were deemed to be the most appropriate driver of affordable lithography field sizes. In the present *ITRS* chip-size model for DRAMs, the introduction-level chip size is targeted to be smaller than the new affordable 572mm^2 lithography field size, fitting at least one introduction-level chip size within the field. The production-level DRAM model fits at least two die within the affordable field. The combination of technology-node scaling and cell design improvements (A-factor reduction) accomplishes that goal, while also maintaining a goal of doubling on-chip bits every two years. However, the slowing of DRAM design improvements causes a requirement to add fewer on-chip bits to stay under the

162 Overall Roadmap Technology Characteristics

affordable lithography field limit. This accomplished in the present DRAM model by slowing the short-term Moore's Law bits/chip rate from $2\times/1.5$ years to $2\times/$ two years. In the later years of the Roadmap the rate is stretched even more to $2\times/2.5$ to three years, as required. The data targets for the DRAM model are included in Tables 1c, d, e, and f.

Both the DRAM and MPU models depend upon achieving the aggressive DRAM and MPU design and process improvement targets. If those targets slip, then pressure will increase to print chip sizes larger than the present roadmap, or further slow the rate of "Moore's-Law" on-chip functionality. Either of these consequences will result in a negative impact upon cost-per-function reduction rates—the classical measure of our industries productivity-improvement and competitiveness.

With increasing cost reduction pressures, the need for the 300 mm productivity boost will also increase in urgency, especially for leading-edge manufacturers, but the poor economy will create challenges and limit capital investment. The 2001 Wafer-Diameter Generation roadmap (see Tables 2a and b) is consistent with the ramp of 300 mm capacity beginning 2001. Also, the first manufacturing capability for the next $1.5\times$ wafer size conversion to 450 mm diameter is not anticipated to be required until 2013 in the present roadmap. However, should the other productivity-improvement drivers (lithography and design/process improvements) fail to stay on schedule, there would be a need to accelerate the use of increased wafer diameter as a productivity improvement.

The affects of future technology acceleration/deceleration and the timing of the next wafer generation conversion requires the development and application of comprehensive long-range factory productivity and industry economic models. Such industry economic modeling (IEM) work is being sponsored and carried out jointly by SEMI and International SEMATECH.

Table 2a Lithographic-Field and Wafer-Size Trends—Near-term Years

(Note: 2001 Lithographic field sizes represent current capability)

Year of Production	2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)	130	115	100	90	80	70	65
MPU/ASIC ½ Pitch (nm)	150	130	107	90	80	70	65
MPU Printed Gate Length (nm)	90	75	65	53	45	40	35
MPU Physical Gate Length(nm)	65	53	45	37	32	28	25
Lithography Field Size							
Lithography Field Size—area (mm ²)	800	800	800	800	800	800	572
Lithographic field size — length (mm)	32	32	32	32	32	32	26
Lithographic field size — width (mm)	25	25	25	25	25	25	22
Maximum Substrate Diameter (mm) — High-volume Production (>20K wafer starts per month)							
Bulk or epitaxial or SOI wafer	300	300	300	300	300	300	300

Table 2b Lithographic-Field and Wafer Size Trends—Long-term Years

Year of Production	2010	2013	2016
DRAM ½ Pitch (nm)	45	32	22
MPU/ASIC ½ Pitch (nm)	45	32	22
MPU Printed Gate Length (nm)	25	18	13
MPU Physical Gate Length (nm)	18	13	9
Lithography Field Size			
Maximum lithographic field size—area (mm ²)			
Lithography Field Size—area (mm ²)	572	572	572
Maximum lithographic field size—length (mm)	26	26	26
Maximum lithographic field size—width (mm)	22	22	22
Maximum Substrate Diameter (mm)—High-volume Production (>20K wafer starts per month)			
Bulk or epitaxial or SOI wafer	300	450	450

PERFORMANCE OF PACKAGED CHIPS

NUMBER OF PADS AND PINS / PAD PITCH, COST PER PIN, FREQUENCY

The demand for a higher number of functions on a single chip requires the integration of an increased number of transistors or bits (memory cells) for each product generation. Typically, the number of pads and pins necessary to allow Input/Output (I/O) signals to flow to and from an integrated circuit increases as the number of transistors on a chip increases. (Refer to Tables 3a and b)

Additional power and ground connections to the chip are also necessary to optimize power management and to increase noise immunity. Based upon chip pad-count numbers supplied by the Test ITWG, logic products (MPUs and high-performance ASICs) both approach 4-6K pads over the *ITRS* period. The MPU products are forecast to increase the total number of pads through this period by nearly 50%, and the ASICs double the maximum number of pads per chip. The two product types also differ significantly in the ratio of power/ground pads. The MPU product pad counts typically have 1:3 signal I/O pads and 2:3 power and ground pads, or two power/ground pads for every signal I/O pad. Unlike MPUs, high-performance ASIC product pad counts typically include one power/ground pad for each signal I/O pad.

Table 3a Performance of Packaged Chips: Number of Pads and Pins—Near-term Years

<i>Year of Production</i>	<i>2001</i>	<i>2002</i>	<i>2003</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>
<i>DRAM ½ Pitch (nm)</i>	130	115	100	90	80	70	65
<i>MPU/ASIC ½ Pitch (nm)</i>	150	130	107	90	80	70	65
<i>MPU Printed Gate Length (nm)</i>	90	75	65	53	45	40	35
<i>MPU Physical Gate Length (nm)</i>	65	53	45	37	32	28	25
<i>Number of Chip I/Os (Number of Total Chip Pads) — Maximum</i>							
<i>Total pads—MPU</i>	3072	3072	3072	3072	3072	3072	3072
<i>Signal I/O—MPU (1/3 of total pads)</i>	1024	1024	1024	1024	1024	1024	1024
<i>Power and ground pads—MPU (2/3 of total pads)</i>	2048	2048	2048	2048	2048	2048	2048
<i>Total pads—ASIC high-performance</i>	3000	3200	3400	3600	4000	4200	4400
<i>Signal I/O pads—ASIC high-performance</i>	1500	1600	1700	1800	2000	2100	2200
<i>Power and ground pads—ASIC high-performance (½ of total pads)</i>	1500	1600	1700	1800	2000	2100	2200
<i>Number of Total Package Pins—Maximum [1]</i>							
<i>Microprocessor/controller, cost-performance</i>	480–1,200	480–1320	500–1452	500–1600	550–1760	550–1936	600–2140
<i>Microprocessor/controller, high-performance</i>	1200	1320	1452	1,600	1,760	1,936	2,140
<i>ASIC (high-performance)</i>	1700	1870	2057	2263	2489	2738	3012

Notes for Tables 3a and 3b

[1] Pin counts will be limited for some applications where fine pitch array interconnect is used by PWB technology and system cost.

The highest pin count applications will as a result use larger pitches and larger package sizes.

The reference to signal pin ratio will also vary greatly dependent on applications with an expected range from 2:1 to 1:4

Table 3b Performance of Packaged Chips: Number of Pads and Pins—Long-term Years

<i>Year of Production</i>	<i>2010</i>	<i>2013</i>	<i>2016</i>
<i>DRAM ½ Pitch (nm)</i>	45	32	22
<i>MPU/ASIC ½ Pitch (nm)</i>	45	32	22
<i>MPU Printed Gate Length (nm)</i>	25	18	13
<i>MPU Physical Gate Length (nm)</i>	18	13	9
<i>Number of Chip I/Os (Number of Total Chip Pads) — Maximum</i>			
<i>Total pads—MPU</i>	3840	4224	4416
<i>Signal I/O—MPU (1/3 of total pads)</i>	1280	1408	1472
<i>Power and ground pads—MPU (2/3 of total pads)</i>	2560	2816	2944
<i>Total pads—ASIC high-performance</i>	4800	5400	6000
<i>Signal I/O pads—ASIC high-performance</i>	2400	2700	3000
<i>Power and ground pads—ASIC high-performance (½ of total pads)</i>	2400	2700	3000
<i>Number of Total Package Pins—Maximum [1]</i>			
<i>Microprocessor/controller, cost-performance</i>	780–2782	1014–3616	1318–4702
<i>Microprocessor/controller, high-performance</i>	2782	3616	4702
<i>ASIC (high-performance)</i>	4009	5335	7100

Package pin count (Tables 3 a and b) and cost-per-pin (Tables 4 a and b), provided by the Assembly and Packaging ITWG, point out challenges to future manufacturing economics. Based upon the projected growth in the number of transistors/chip, it is forecast that the number of package pin/balls will continue to grow at an annual rate of approximately 10%, while the cost/pin decreases at 5%/year. These trends make it more challenging for suppliers of packaging technologies to deliver cost-effective solutions, because the overall average cost of packaging will increase annually at 5%/year ($.95 \text{ cost/pin} \times 1.10 \text{ pins/year} = 1.05 \text{ cost/year}$).

In the very competitive consumer electronics product environment, prices for high-volume, high-tech products such as PCs and cell phones tend to remain flat or even decrease. These same high-tech products typically also deliver twice the performance every two years. This is the end-use market environment of the leading-edge semiconductor manufacturer, and it is the fundamental economic driver behind the *ITRS* economic requirement to reduce cost per function (bits, transistors) at an annual 30% or faster rate ($2 \times \text{functionality/chip at flat price every two years} = 29\%/\text{year}$).

If future semiconductor component products must be targeted to maintain constant or decreasing prices and the average number of pins per unit increases at 10% while the average cost per pin decreases at only 5%, then the following will occur:

1. the average packaging share of total product cost will double over the 15-year roadmap period, and
2. the ultimate result will be greatly reduced gross profit margins and limited ability to invest in R&D and factory capacity.

This conclusion is one of the drivers behind the industry trends to reduce the overall system pin requirements by combining functionality into Systems-on-Chip (SoC) and through the use of multi-chip modules, bumped chip-on-board (COB), and other creative solutions.

In addition to the need to increase functionality while exponentially decreasing cost per function, there is also a market demand for higher-performance, cost-effective products. Just as Moore's Law predicts that functions-per-chip will double every 1.5–2 years to keep up with consumer demand, there is a corresponding demand for processing electrical signals at progressively higher rates. In the case of MPUs, processor instructions/second have also historically doubled every 1.5–2 years. For MPU products, increased processing power, measured in millions of instructions per second (MIPs), is accomplished through a combination of "raw technology performance" (clock frequency) multiplied by "architectural performance" (instructions per clock cycle). The need for a progressively higher operational frequency associated with an

increasing average chip size will continue to demand the development of novel process, design, and packaging techniques.

These considerations are reflected in Tables 4c and d, which includes line items contributed by the Design and Assembly and Packaging ITWGs to forecast the maximum on-chip and chip-to-board frequency trends. The highest frequency obtainable in each product generation is directly related to the intrinsic transistor performance (on-chip, local clock). The difference between this “local” frequency and the frequency of signals traveling across the chip increases due to degradation of signal propagation delay caused by line-to-line and line-to-substrate capacitive coupling. Additional signal degradation is associated with the inductance of wire bonds and package leads. Direct chip attachment may eventually be the only viable way to eliminate any parasitic effect introduced by the package. To optimize signal and power distribution across the chip, it is expected that the number of layers of interconnect will continue to increase. As size downscaling of interconnect also continues, wider use of copper (low resistivity) and various inter-metal insulating materials of progressively lower dielectric constant ($\kappa \sim 2-3$) will be adopted in the chip fabrication process. Multiplexing techniques will also be used to increase the chip-to-board operating frequency (off-chip).

Table 4a Performance and Package Chips: Pads, Cost—Near-term Years [Update]

Year of Production		2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)		130	115	100	90	80	70	65
MPU/ASIC ½ Pitch (nm)		150	130	107	90	80	70	65
MPU Printed Gate Length (nm)		90	75	65	53	45	40	35
MPU Physical Gate Length (nm)		65	53	45	37	32	28	25
Chip Pad Pitch (micron)								
Was	Pad pitch—ball bond	45	35	30	25	20	20	20
Is	Pad pitch—ball bond	45	<u>40</u>	<u>35</u>	<u>30</u>	<u>25</u>	20	20
Was	Pad pitch—wedge bond	40	35	30	25	20	20	20
Is	Pad pitch—wedge bond	<u>50</u>	<u>50</u>	<u>40</u>	<u>40</u>	<u>35</u>	<u>35</u>	<u>30</u>
Was	Pad Pitch—area array flip-chip (cost-performance, high-performance)	160	160	150	150	130	130	120
Is	Pad Pitch—area array flip-chip (cost-performance, high-performance)	<u>200</u>	<u>180</u>	150	150	<u>100</u>	<u>100</u>	<u>80</u>
Was	Pad Pitch—peripheral flip-chip (handheld, low-cost, harsh)	150	130	120	110	100	90	80
Is	Pad Pitch—peripheral flip-chip (handheld, low-cost, harsh)	<u>80</u>	<u>80</u>	<u>60</u>	<u>60</u>	<u>40</u>	<u>40</u>	<u>30</u>
Cost-Per-Pin								
Was	Package cost (cents/pin) (cost-performance)—minimum—maximum	0.80–1.60	0.75–1.44	0.70–1.30	0.66–1.17	0.61–1.06	0.56–1.03	0.64–1.00
Is	Package cost (cents/pin) (cost-performance)—minimum—maximum	0.80–1.60	0.75– <u>1.30</u>	0.70– <u>1.24</u>	0.66–1.17	0.61– <u>1.11</u>	0.56– <u>1.06</u>	0.64– <u>1.01</u>
Package cost (cents/pin) (Memory)—minimum—maximum		0.36–1.54	0.34–1.39	0.32–1.26	0.30–1.14	0.28–1.03	0.27–0.93	0.27–0.84

Table 4b Performance and Package Chips: Pads, Cost—Long-term Years [Update]

Year of Production		2010	2013	2016
DRAM ½ Pitch (nm)		45	32	22
MPU/ASIC ½ Pitch (nm)		45	32	22
MPU Printed Gate Length (nm)		25	18	13
MPU Physical Gate Length (nm)		18	13	9
Chip Pad Pitch (micron)				
Pad pitch—ball bond		20	20	20
Pad Pitch—wedge bond		20	20	20
Was	Pad Pitch—area array (cost-performance, high-performance)	90	80	70
Is	Pad Pitch—area array (cost-performance, high-performance)	<u>70</u>	<u>70</u>	<u>50</u>
Was	Pad Pitch—peripheral flip-chip (handheld, low-cost, harsh)	60	45	30
Is	Pad Pitch—peripheral flip-chip (handheld, low-cost, harsh)	<u>20</u>	<u>20</u>	<u>15</u>
Cost-Per-Pin				
Package cost (cents/pin) (cost-performance)— minimum—maximum		0.49–0.98	0.42–0.93	0.36–0.79
Package cost (cents/pin) (Memory)— minimum—maximum		0.22–0.54	0.19–0.39	0.19–0.33

Table 4c Performance and Package Chips: Frequency On-Chip Wiring Levels—Near-term Years [Update]

YEAR OF PRODUCTION		2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)		130	115	100	90	80	70	65
MPU/ASIC ½ Pitch (nm)		150	130	107	90	80	70	65
MPU Printed Gate Length (nm)		90	75	65	53	45	40	35
MPU Physical Gate Length (nm)		65	53	45	37	32	28	25
Chip Frequency (MHz)								
On-chip local clock		1,684	2,317	3,088	3,990	5,173	5,631	6,739
Chip-to-board (off-chip) speed (high-performance, for peripheral buses)[1]		1,684	2,317	3,088	3,990	5,173	5,631	6,739
Was	Maximum number wiring levels—maximum	7	8	8	8	9	9	9
Is	Maximum number wiring levels—maximum	<u>10</u>	<u>10</u>	<u>12</u>	<u>13</u>	<u>14</u>	<u>14</u>	<u>14</u>
Was	Maximum number wiring levels—minimum	7	7	8	8	8	9	9
Is	Maximum number wiring levels—maximum	<u>10</u>	<u>10</u>	<u>12</u>	<u>9</u>	<u>10</u>	<u>10</u>	<u>10</u>

Table 4d Performance and Package Chips: Frequency, On-Chip Wiring Levels—Long-term Years [Update]

Year of Production		2010	2013	2016
DRAM ½ Pitch (nm)		45	32	22
MPU/ASIC ½ Pitch (nm)		45	32	22
MPU Printed Gate Length (nm)		25	18	13
MPU Physical Gate Length (nm)		18	13	9
Chip Frequency (MHz)				
On-chip local clock		11,511	19,348	28,751
Chip-to-board (off-chip) speed (high-performance, for peripheral buses)[1]		11,511	19,348	28,751
Was	Maximum number wiring levels—maximum	10	10	10
Is	Maximum number wiring levels—maximum	14	15	15
Was	Maximum number wiring levels—minimum	9	9	10
Is	Maximum number wiring levels—minimum	10	11	11

Note for Tables 4c and 4d

[1] The off chip frequency is expected to increase for a small number of high speed pins which will be used in combination with a large number of lower speed pins

[2] In 2001, high-speed serial communications transceiver devices are achieving chip-board frequencies of 3.125 GHz using CMOS, and 10 GHz using SiGe. In 2002 it is expected that 10 GHz transceivers will be fabricated using CMOS. 40 GHz SiGe devices are expected in 2003. The roadmap for higher levels of integration with wider bus widths, is shown in the High Frequency Serial Communications section in the Test chapter.

Add

[3] The Minimum number of wiring levels represents the Interconnect metal levels, and the Maximum number of Interconnect wiring levels includes the Minimum number of wiring levels plus additional optional levels required for power, ground, signal conditioning, and integrated passives (i.e., capacitors).

ELECTRICAL DEFECT DENSITY

The latest targets for electrical defect density of DRAM, MPU, and ASIC (necessary to achieve 83–89.5 % chip yield in the year of volume production) are shown in Tables 5a and b. The allowable number of defects is calculated by taking into account the different chip sizes based on the latest chip size model forecasts, as reported in Table 1 for DRAM and microprocessors. In addition, the data in the table are now reported only at the production-level of the product life-cycle. Other defect densities may be calculated at different chip sizes at the same technology node by using the formula found in the *Yield Enhancement* chapter. The approximate number of masks for logic devices is included as an indicator of the ever-increasing process complexity.

Table 5a Electrical Defects—Near-term Years

<i>Year of Production</i>	<i>2001</i>	<i>2002</i>	<i>2003</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>
<i>DRAM ½ Pitch (nm)</i>	130	115	100	90	80	70	65
<i>MPU/ASIC ½ Pitch (nm)</i>	150	130	107	90	80	70	65
<i>MPU Printed Gate Length (nm)</i>	90	75	65	53	45	40	35
<i>MPU Physical Gate Length (nm)</i>	65	53	45	37	32	28	25
<i>DRAM Random Defect D_0 at production chip size and 89.5% yield (faults/m²) §</i>	1,963	2,493	2,148	2,748	1,752	2236	1426
<i>MPU Random Defect D_0 at production chip size and 83% yield (faults/m²) §§</i>	1,356	1,356	1,356	1,356	1,356	1,356	1,356
<i># Mask Levels – MPU</i>	25	25	25	25	25	27	27
<i># Mask Levels – DRAM</i>	21	22	24	24	24	24	24

Table 5b Electrical Defects—Long-term Years

<i>Year of Production</i>	<i>2010</i>	<i>2013</i>	<i>2016</i>
<i>DRAM ½ Pitch (nm)</i>	45	32	22
<i>MPU/ASIC ½ Pitch (nm)</i>	45	32	22
<i>MPU Printed Gate Length (nm)</i>	25	18	13
<i>MPU Physical Gate Length (nm)</i>	18	13	9
<i>DRAM Random Defect D_0 at production chip size and 89.5% yield (faults/m²) §</i>	1356	1356	1356
<i>MPU Random Defect D_0 at production chip size and 83% yield (faults/m²) §§</i>	1464	1116	1134
<i># Mask Levels – MPU</i>	27	29	29
<i># Mask Levels – DRAM</i>	26	26	26

Notes for Tables 5a and 5b

D_0 —defect density

§ DRAM Model—Cell Factor (design/process improvement) targets are as follows

1999–2002/8×; 2003–2010/6×; 2011–2016/4×

DRAM product generations are usually increased by 4×bits/chip every four years with interim 2×bits/chip generations, except

1. at the Introduction phase, after the 8Gbit interim generation, the introduction rate is 4×/five years (2×/two–three years); and
2. at the Production phase, after the interim 32Gbit generation, the introduction rate is 4×/five years (2×/two–three years).

InTER-generation chip size growth rate varies to maintain one chip per 572mm² field at Introduction and two chips per 572mm² field at Production. The more aggressive “best case opportunity” technology node trends allow the Production-phase products to remain at 2×bits/chip every two years and still fit within the target of two DRAM chips per 572mm² field size, through the 32Gbit interim generation. The InTRA-generation chip size shrink model is 0.5×every technology node in-between cell factor reductions.

Refer to the Glossary for definitions of Introduction, Production, InTERgeneration, and InTRAgeneration terms.

§§ MPU Chip Size Model—Both the cost-performance and high-performance MPUs InTER-generation chip size growth rates are targeted to be flat through 2016, made possible by doubling the on-chip functionality every technology node cycle. The InTRA-generation chip size shrink model is 0.5×every two-year technology node through 2001, then 0.5×every three-year technology node after 2001.

POWER SUPPLY AND POWER DISSIPATION

Reduction of power supply voltage (see Tables 6a and b) is driven by several factors—reduction of power dissipation, reduced transistor channel length, and reliability of gate dielectrics. As seen in Tables 6a and b, the value of the power supply voltage is now given as a range.

Selection of a specific V_{dd} value continues to be a part of the analysis undertaken to simultaneously optimize speed and power for an IC, leading to a range of usable power supply voltages in each product generation. Values of V_{dd} as low as 0.5 volts are expected to be achieved by 2013, but the lowest target is now 0.4V by 2016 (versus 0.3V by 2014 in the 1999 ITRS).

Maximum power trends (e.g., for MPUs) are presented in three categories—1) high-performance desktop applications, for which a heat sink on the package is permitted; 2) cost-performance, where economical power management solutions of the highest performance are most important; and 3) portable battery operations. In all cases, total power consumption continues to increase, despite the use of a lower supply voltage. The increased power consumption is driven by higher chip operating frequencies, the higher interconnect overall capacitance and resistance; and the increasing gate leakage of exponentially-growing and scaled on-chip transistors.

Table 6a Power Supply and Power Dissipation—Near-term Years

<i>Year of Production</i>	<i>2001</i>	<i>2002</i>	<i>2003</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>
<i>DRAM ½ Pitch (nm)</i>	130	115	100	90	80	70	65
<i>MPU/ASIC ½ Pitch (nm)</i>	150	130	107	90	80	70	65
<i>MPU Printed Gate Length (nm)</i>	90	75	65	53	45	40	35
<i>MPU Physical Gate Length (nm)</i>	65	53	45	37	32	28	25
<i>Power Supply Voltage (V)</i>							
<i>V_{dd} (high performance)</i>	1.1	1.0	1.0	1	0.9	0.9	0.7
<i>V_{dd} (Low Operating Power, high V_{dd} transistors)</i>	1.2	1.2	1.1	1.1	1.0	1.0	0.9
<i>V_{dd} (Low Standby Power, high V_{dd} transistors)</i>	1.2	1.2	1.2	1.2	1.2	1.2	1.1
<i>Allowable Maximum Power [1]</i>							
<i>High-performance with heatsink (W)</i>	130	140	150	160	170	180	190
<i>Cost-performance (W)</i>	61	75	81	85	92	98	104
<i>Battery (W)—(hand-held)</i>	2.4	2.6	2.8	3.2	3.2	3.5	3.5

Table 6b Power Supply and Power Dissipation—Long-term Years

<i>Year of Production</i>	<i>2010</i>	<i>2013</i>	<i>2016</i>
<i>DRAM ½ Pitch (nm)</i>	45	32	22
<i>MPU/ASIC ½ Pitch (nm)</i>	45	32	22
<i>MPU Printed Gate Length (nm)</i>	25	18	13
<i>MPU Physical Gate Length (nm)</i>	18	13	9
<i>Power Supply Voltage (V)</i>			
<i>V_{dd} (high performance)</i>	0.6	0.5	0.4
<i>V_{dd} (Low Operating Power, high V_{dd} transistors)</i>	0.8	0.7	0.6
<i>V_{dd} (Low Standby Power, high V_{dd} transistors)</i>	1.0	0.9	0.9
<i>Allowable Maximum Power [1]</i>			
<i>High-performance with heatsink (W)</i>	218	251	288
<i>Cost-performance (W)</i>	120	138	158
<i>Battery (W)—(hand-held)</i>	3.0	3.0	3.0

Note for Table 6a and 6b

[1] Power will be limited more by system level cooling and test constraints than packaging

COST

Tables 7a and b are dedicated to cost trends. The ability to reduce the cost per function by an average 25–30% each year represents one of the unique features of the semiconductor industry and is a direct consequence of the market pressure to continue to deliver twice the functionality on-chip every 1.5–2 years in an environment of constant or reducing prices. In support of this cost reduction, a continuously increasing amount of investment is needed for R&D and manufacturing capital. Even on a per-factory basis, the capital cost of manufacturing continues to escalate. Yet, the semiconductor industry has historically delivered two times as many functions per chip every 1.5–2 years with no or only a moderate increase in chip size and cost (approximately constant cost per cm^2 of silicon). This technological and economic performance is the fundamental engine behind the growth of the semiconductor industry.

However, the customers in today's competitive market environment are resistant to even "moderate" increases in cost and the rate of doubling functions per chip (Moore's Law) is slowing. Therefore, the semiconductor manufacturers must seek a new model to deliver the same cost-per-function reduction requirements that have fueled industry growth. Consequently, the 1999 ITRS proposed a new model for achieving the required reduction: provide the customer twice the functionality every two years at a constant cost and average selling price (ASP) per chip. The 2001 ITRS uses the model and results in the same 29% cost reduction of a function (bit, transistor, etc.) that has been achieved historically by delivering four times the functionality per chip every three years at $1.4\times$ increase in cost per unit. The DRAM and MPU cost models continue to use the need for that 29% cost-per-function productivity reduction rate as an economic driver of the industry. Therefore, that core cost-per-function trend has been used to set the INTRA-generation trends for the affordable cost/bit and cost/transistor for DRAM and microprocessors, respectively. Extrapolation of historical trends would indicate an "at introduction" affordable cost/bit of approximately 21 microcents for 2-Gbit DRAMs in 2001. In addition, the historical trends indicate that, within a DRAM generation, a 45%/year reduction in cost/bit should be expected.² A corresponding analysis conducted from published data for microprocessors yields similar results.³ As a result the 29%/year target for reduction in affordable cost/transistor from generation to generation is also being used in the MPU model, along with the 45%/year reduction rate within the same generation.

However, the Design ITWG has updated the MPU model, based upon recent data. The new data indicates that logic transistor size is improving only at the rate of the lithography ($0.7\times$ linear, $0.5\times$ area reduction every technology node). Therefore in order to keep the MPU chip sizes flat, the number of transistors can be doubled only every technology node. The technology node rate is projected to return to a three-year cycle after 2001, therefore the transistors per MPU chip can double only every three years after 2001. DRAM memory bit cell design improvements are also slowing down, and the rate of bits per chip will also be slowing in the future to keep chip sizes under control. To compensate for slowing DRAM and MPU functions-per-chip, there will be increasing pressure to find alternative productivity enhancements from the equivalent productivity scaling benefits of chip and system-level architecture and designs.

Even though the rate of increase of on-chip functionality could slow in the future, the amount of functions/chip is still growing exponentially. As the number of functions/chip continues to increase, it becomes increasingly difficult and, therefore, costly to test the final products. This is reflected in the escalating cost of testers. Even though the cost/pin of testers is forecast to decline between 0% and 10% per year (Tables 7 a and b), the number of pins grows at 10%/year (Tables 4 a and b). Therefore, the need for accelerated implementation of Built-In-Self-Test (BIST) and Design-For-Testability (DFT) techniques will continue within the time frame of the 2001 International Technology Roadmap for Semiconductors. Further discussion is detailed in the *Test* chapter.

² McClean, William J., ed. *Mid-Term 1994: Status and Forecast of the IC Industry*. Scottsdale: Integrated Circuit Engineering Corporation, 1994.

McClean, William J., ed. *Mid-Term 1995: Status and Forecast of the IC Industry*. Scottsdale: Integrated Circuit Engineering Corporation, 1995.

³ a) Dataquest Incorporated. *x86 Market: Detailed Forecast, Assumptions, and Trends*. MCRO-WW-MT-9501. San Jose: Dataquest Incorporated, January 16, 1995.

b) Port, Otis; Reinhardt, Andy; McWilliams, Gary; and Brull, Steven V. "The Silicon Age? It's Just Dawning," Table 1. *Business Week*, December 9, 1996, 148–152.

Table 7a Cost—Near-term Years

Year of Production	2001	2002	2003	2004	2005	2006	2007
DRAM ½ Pitch (nm)	130	115	100	90	80	70	65
MPU/ASIC ½ Pitch (nm)	150	130	107	90	80	70	65
MPU Printed Gate Length (nm)	90	75	65	53	45	40	35
MPU Physical Gate Length (nm)	65	53	45	37	32	28	25
Affordable Cost per Function ++							
DRAM cost/bit at (packaged microcents) at samples/introduction	21	14.8	10.5	7.4	5.3	3.7	2.6
DRAM cost/bit at (packaged microcents) at production §	7.7	5.4	3.8	2.7	1.9	1.4	0.96
Cost-performance MPU (microcents/transistor) (including on-chip SRAM) at introduction §§	176	124	88	62	44	31	22
Cost-performance MPU (microcents/transistor) (including on-chip SRAM) at production §§	107	75	53	38	27	19	13.3
High-performance MPU (microcents/transistor) (including on-chip SRAM) at production §§	97	69	49	34	24	17	12
Cost-Per-Pin							
Test Cost							
Volume tester cost per high-frequency signal pin (\$K/pin) (high-performance ASIC)—maximum	4.0	3.0	3.0	3.0	3.0	3.0	3.0
Volume tester cost per high-frequency signal pin (\$K/pin) (high-performance ASIC)—minimum	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Notes for Tables 7a and 7b

++ Affordable packaged unit cost per function based upon Average Selling Prices (ASPs) available from various analyst reports less Gross Profit Margins (GPMs); 35% GPM used for commodity DRAMs and 60% GPM used for MPUs; 0.5×/two years inTER-generation reduction rate model used; .55×/year inTRA-generation reduction rate model used; DRAM unit volume life-cycle peak occurs when inTRA-generation cost per function is crossed by next generation, typically seven–eight years after introduction; MPU unit volume life-cycle peak occurs typically after four–six years, when the next generation processor enters its ramp phase (typically two–four years after introduction).

§ DRAM Model—Cell Factor (design/process improvement) targets are as follows

1999–2002/8×; 2003–2010/6×; 2011–2016/4×.

DRAM product generations are usually increased by 4×bits/chip every four years with interim 2×bits/chip generations, except

- 1) at the Introduction phase, after the 8Gbit interim generation, the introduction rate is 4×/five years (2×/two–three years); and
- 2) at the Production phase, after the interim 32Gbit generation, the introduction rate is 4×/five years (2×/two–three years).

InTER-generation chip size growth rate varies to maintain one chip per 572mm² field at Introduction and two chips per 572mm² field at Production. The more aggressive “best case opportunity” technology node trends allow the Production-phase products to remain at 2×bits/chip every two years and still fit within the target of two DRAM chips per 572mm² field size, through the 32Gbit interim generation. The InTRA-generation chip size shrink model is 0.5× every technology node in-between cell factor reductions.

Refer to the Glossary for definitions of Introduction, Production, InTERgeneration, and InTRAgeneration terms.

§§ MPU Chip Size Model—Both the cost-performance and high-performance MPUs InTER-generation chip size growth rates are targeted to be flat through 2016, made possible by doubling the on-chip functionality every technology node cycle. The InTRA-generation chip size shrink model is 0.5× every two-year technology node through 2001, then 0.5× every three-year technology node after 2001.

Table 7b Cost—Long-term Years

<i>Year of Production</i>	<i>2010</i>	<i>2013</i>	<i>2016</i>
<i>DRAM ½ Pitch (nm)</i>	45	32	22
<i>MPU/ASIC ½ Pitch (nm)</i>	45	32	22
<i>MPU Printed Gate Length (nm)</i>	25	18	13
<i>MPU Physical Gate Length (nm)</i>	18	13	9
<i>Affordable Cost per Function ++</i>			
<i>DRAM cost/bit (packaged microcents) at samples/introduction</i>	0.93	0.33	0.12
<i>DRAM cost/bit (packaged microcents) at production §</i>	0.34	0.12	0.042
<i>Cost-performance MPU (microcents/transistor) (including on-chip SRAM) at introduction §§</i>	7.78	2.75	0.97
<i>Cost-performance MPU (microcents/transistor) (including on-chip SRAM) at production §§</i>	4.71	1.66	0.59
<i>High-performance MPU (microcents/transistor) (including on-chip SRAM) at production §§</i>	4.31	1.52	0.54
<i>Cost-Per-Pin</i>			
<i>Test Cost</i>			
<i>Volume tester cost per high-frequency signal pin (\$K/pin) (high-performance ASIC)—maximum</i>	4	4	4
<i>Volume tester cost per high-frequency signal pin (\$K/pin) (high-performance ASIC)—minimum</i>	2	3	4