INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS

2001 EDITION

EXECUTIVE SUMMARY

ACKNOWLEDGMENTS—*2001 ITRS*

INTERNATIONAL ROADMAP COMMITTEE *Europe*— Wolfgang Arden, Joop Bruines, Gerhard Göltz, Fred van Roosmalen. *Japan*— Toshitaka Fukushima, Seiichiro Kawamura, Toshiaki Masuhara, Makoto Yoshimi. *Korea*— Andy Chung, J. T. Moon, Hee-Koo Yoon. *Taiwan*— Kwan-Yi Chiu, Chenming Hu, Genda Hu, David Lee, C. Y. Lu, Chung Wang, Ping Yang. *U.S.A.*— Bob Doering, Paolo Gargini, Juri Matisoo

TECHNOLOGY WORKING GROUP KEY CONTRIBUTORS

Overall Roadmap Technology Characteristics (including the Chip-Size/Technology Node Study Group)— Joseph Adam, Alan Allan, John Canning, Christopher Case, Chi-Shih Chang, William Chen, Walter Class, Bob Doering, Don Edenfeld, Denis Fandel, Milt Godwin, Randy Goodall, David Jensen, Bill Joyner, Andrew B. Kahng, Seiichiro Kawamura, Harry Levinson, Christopher Long, Ryan Schmidt, Tim Stanley, Dennis Sylvester, Walt Trybula, Peter Zietzoff

Design— Rob Aitken, Lawrence Arledge, Peter Bannon, Ralf Brederlow, Randy Bryant, L. Richard Carley, George Doerre Stephane Donnay, Kwang-Ting Cheng, John Cohn, Donald Cottrell, John Darringer, Jeff Davis, Alfred E. Dunlop, Takashi Fujii, W. Dale Edwards, Yoshiharu Furui, Shinji Furuichi, Hideyuki Hamada, Justin E. Harlow III, Tamotsu Hiwatashi, Mark Horowitz, Alan J. Hu, Noriyuki Itoh, William Joyner, Tomohiko Kadowaki, Andrew B. Kahng, Masaru Kakimoto, Kevin Kolwicz, Grant Martin, Joseph Morrell, Nobuto Ono, Shinji Ono, Jeff Parkhurst, Carl Pixley, Michael J. Rodgers, Resve Saleh, Josef Sauerer, Prashant Saxena, Patrick Schaumont, Herman Schmit, Tsutomu Someya, Dennis Sylvester, Haruyuki Tago, Hirofumi Taguchi, Kunio Uchiyama, Takuzo Uemura, Maarten Vertregt, Ted Vucurevich, Piet Wambacq, Werner Weber, Akihiro Yamada, Seiji Yamaguchi, Ichiro Yamamoto, Takayuki Yamanouchi

Test— Rob Aitken, Richard Antley, Mark Barber, Roger Barth, Jay Bedsole, Phil Burlison, Yi Cai, Don Edenfeld, Kouichi Eguchi, Rudy Garcia, Anne Gattiker, Atul Goel, John Johnson, Bernd Koenemann, Pierre Marilleau, John Matthais, Peter Maxwell, Peter Muhmenthaler, Suresh Nadig, Phil Nigh, Mitsuyasu Ohta, Pete O'Neill, Bill Ortner, Ken Posse, Patrick Quadrel, Rochit Rajsuman, Paul Roddy, Mike Rodgers, Masayuki Sato, Rene Segers, Tim Sheen, Lee Song, Susan Stirrat, Fred Taber, Jody Van Horn, Don Van Overloop, Burnell West, Don Wheater, Tom Williams, Yervant Zorian

Process Integration, Devices, & Structures and Emerging Research Devices— Goncal Badenes, David Blackburn, George Bourinaoff, Joe Brewer, Joop Bruines, Jim Chung, Ramon Compano, Kristin De Meyer, Simon Deleonibus, Theodore A. Dellin, Carlos Diaz, Nobuhiro Endo, Peng Fang, Mike Forshaw, Michel Haond, Steven Hillenius, Toshiro Hiramoto, Yasuaki Hokari, Atsushi Hori, Tadahiko Horiuchi, Margaret Huang, Jim Hutchby, Jiro Ida, Yasuo Inoue, Dae Gwan Kang, Naoki Kasai, Young-Wug Kim, Hideaki Kuroda, Chanho Lee, Ben McKee, Takashi Nakamura, Tak H. Ning, Byong Gook Park, Shizuo Sawada, Dirk Schumann, Kentaro Shibahara, Hyungsoon Shin, Thomas Skotnicki, Toshihiro Sugii, Y. C. (Jack) Sun, Yoshitaka Tadaki, Luan C. Tran, Chung Wang, Jason Woo, Simon Yang, Geoffrey Yeap, Makoto Yoshimi, Peter Zeitzoff, Bin Zhao, Victor Zhirnov

Front End Processes— Mauro Alessandri, Mike Alles, Rick Amos, Scott Becker, Larry Beckwith, Roberto Bez, Joop Bruines, Murray Bullis, Stephanie Butler, Emilio Camerlenghi, Daniele Cantarelli, Paolo Cappelletti, Mark Chang, Walter Class, Phil Crabtree, Michael Current, Simon Deleonibus, Glenn Gale, Gabriella Ghidini, Dinesh Gupta, Mike Jackson, Ukyo Jeong, Kevin Jones, John Hauser, Richard Hockett, Brad Howard, Howard Huff, Sally Iwanski, Amitabh Jain, Ho-Kyu Kang, Seiichiro Kawamura, Adrian Kiermasz, Brian Kirkpatrick, Martin Knotter, Robert Kraft, Michitaka Kubota, Larry Larson, Pat Lysaght, Alfonso Maurelli, Paul Mertens, Alberto Modelli, Anthony Muscat, Dave Myers, Masaaki Niwa, Farad Nouri, Carlton Osborn, Jon Owyang, Conor Rafferty, Karen Reinhardt, Lenny Reuben, Deborah Riley, Mark Sandoval, Tom Seidel, Danny Shum, Yasuhiro Takeda, Georg Tempel, Eric Vogel, Bernd Vollmer, Greg Wilson, Raffaele Zambrano, Peter Zeitzoff

Lithography— Emanuele Baracchi, John Canning, Han Ku Cho, Kevin Cummings, Ralph Dammel, Alain Diebold, Jan-Willem Gemmink, George Gomba, Gerhard Gross, Noreen Harned, Naoya Hayashi, Scott Hector, Y. C. Ku, Harry J. Levinson, Toshi Nishigaki, Hiroshi Ohtsuka, Kurt Ronse, Masaru Sasago, Gil Shelden, Ki Soo Shin, Kazuaki Suzuki, Gary Taylor, Mauro Vasconi, Tony Yen

Interconnect— Riichirou Aoki, Steven Avanzino, Hans-Joachim Barth, Sam Broydo, Steven Burke, Christopher Case, Shinichi Domae, Bob Geffken, Bob Havemann, Harold Hosack, Calvin Hsueh, Jim Hutchby, Marilyn Jones, John Kelly, Ruichen Liu, Mike Mills, Akihiko Ohsaki, Shinichi Ogawa, Hyun Chul Sohn, Joachim Torres, Jeff Wetzel, C. H. Yu

Factory Integration— Jim Ammenheuser, Bob Bachrach, Eddy Bass, Dave Bloss, David Bouldin, Maverick Brown, Ray Bunkofske, Chris Burkhart, Al Chasey, Hugo Chang, Ivan Chou, Eric Christianson, Blaine Crandell, Gino Crispieri, Daren Dance, Klaus Eberhardt, Anne Eldar, John Ellis, Rick Filippuzzi, John Fowler, Karl Gartland, Randy Goodall, Dave Gross, Duane Howard, Michio Honma, Mani Janakiram, Jye Jeng, Carl Johnson, Giichi Inoue, Junji Iwasaki, Shoichi Kodama, Candi Kristoffersen, Leon McGinnis, Dave Miller, Shantha Mohan, Doug Music, Rohan Nageswaran, Seiichi Nakazawa, Roy Odhner, Richard Oechsner, Kunio Oishi, Mikio Otani, Bill O'Grady, Jeff Pettinato, John O'Reilly, Dev Pillai, Lisa Pivin, Joe Reiss, Clift Richardson, Ralph Richardson, Jim Schillawski, Claus Schneider, Doug Scott, Marlin Shopbell, Court Skinner, Atlant Smith, Arnold Steinman, Abol Taghizadeh, Keisuke Tanimoto, Toshi Uchino, Yoshio Watanabe, Hiromi Yajima

Assembly & Packaging— Joe Adam, Xavier Baraton, Mark Bird, Bill Bottoms, Ken Brown, Chi Shih Chang, Bill Chen, Wun-Yan Chen, Chai Tai Chong, Peter Elenius, Ed Fulcher, Ron Gedney, Seiji Hamano, Harold Hosack, Mahadevan Iyer, Bernd Roemer, John Stankus, Henry Utsunomiya, Mike Vanau, Jurgen Wolf

Environment, *Safety*, *& Health*— Jun-ichi Aoyama, Wolfgang Bloch, Ken Clark, Lee Ann English, Terry Francis, Jim Jewett, Ram Mallela, Coleen Miller, Mike Mocella, Mike Scherer, Farhang Shadman, Tim Wooldridge, Walter Worth

Yield Enhancement— Bob Bryant, Scott Buck, Dave Chamness, Jeff Chapman, Darren Dance, Dianne Dougherty, Mel Effron, Warren Evans, William Fil, Lothar Fitzner, Walt Gardner, Milt Godwin, Honey Goel, Christine Gombar, Dan Grelinger Manuela Huber, Rick Jarvis, David Jensen, Paul Jones, Fred Lakhani, Christopher Long, Dan Maynard, James McAndrew, Mike McIntyre, Toshihiko Osada, Michael Patterson, Ralph Richardson, Matt Sbragia, Josh Shenk, Ken Tobin, Hank Walker

Metrology— Peter Borden, Murray Bullis, Alain Deleporte, Alain Diebold, Clive Hayzelden, Dick Hockett, Masahiko Ikeno, Steve Knight, Rudolf Laubmeier, Henry Ma, Ulrich Mantz, Jack Martinez, Cecilia Martner, Fumio Mizuno, Noel Poduje, Michael Postek, Alec Reader, Vincent Vachellerie, Wilfried Vandervorst, Brad van Eck, Mauro Vascone, Andras Vladar, Stefan Zollner

Modeling & Simulation— Herb Bennett, Kevin Cummings, Ed Hall, Harold Hosack, Herve Jaouen, Norihiko Kotani, Gilles LeCarval, Paco Leon, Juergen Lorenz, Mark Lundstom, Wolfgang Molzer, Kenji Nishi, Llanda Richardson, Wim Schoenmaker, Rainer Thoma, Reinout Woltjer

REGIONAL SUPPORT TEAMS

We acknowledge and thank the regional teams for their support:

Regional Liaison— Yumiko Takamori, Japan liaison and senior IRC staff. *Europe*— Isabelle Charrier, Francis Blin, Bernard Baylac, Thomas Repellin. *Japan*— Hiroshi Mayumi, senior IRC staff; *Semiportal*— Naoko Tani, Satoshi Miki, Hiroshi Momose, Yuji Shibata, Satoshi Takasaki. *Korea*— Du-Su An, Joohee Kwan. *Taiwan*— Rachel Huang. *USA*— Alan Allan, senior IRC staff*; International SEMATECH*— Linda Wilson, ITRS information manager / managing editor; Sarah Mangum, ITRS webmaster; Tony Acevedo, Judy Behr, Eduardo Contreras, Judy Curtis, Rolando Duran, Anne Englander, Mario Gonzalez, Mitchell Gonzalez, Kris Light, Jim Perry, Bob Ruliffson, Anne Shelly, David Silva, Dolores Stoner, Terri Theisen, Donna Towery, Michael Turner; *SEMI*— Christina Gretz, Pat Horwath, Christina Torres; *SIA*— Gail Lindus, Molly Marr-Tuttle; *Intel meeting support*— Diane Corona, Barbara Pebworth, Lewis Wentworth, Sally Yanez

FOREWORD

The 2001 edition of *The International Technology Roadmap for Semiconductors (2001 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 *2001 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 *2001 ITRS*.

In 1992 the Semiconductor Industry Association (SIA) coordinated the first efforts of producing what was originally *The National Technology Roadmap for Semiconductors* (*NTRS*). This roadmap of requirements and possible solutions was generated three times—in 1992, 1994, and 1997. 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 (represented by the European Electronics Component Manufacturers Association [EECA]), Korea (Korea Semiconductor Industry Association [KSIA]), Japan (formerly the Electronic Industry Association of Japan [EIAJ] and now the Japan Electronics and Information Technology Industries Association [JEITA]), and Taiwan (Taiwan Semiconductor Industry Association [TSIA]). The initial collaboration of these five organizations 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*.

As the reader will realize by studying this newly created document, the number and the difficulty of the technical challenges facing the semiconductor industry continue to increase as technology moves forward. The red areas signifying "Manufacturable solutions are not known" are, in most cases, shown within a five-year reach.

Traditional scaling, which has been at the basis of the semiconductor industry for the last 30 years, is indeed beginning to show the fundamental limits of the materials constituting the building blocks of the planar CMOS process. However, new materials can be introduced in the basic CMOS structure to replace and/or augment the existing ones to further extend the device scaling approach. Since the assimilation of these new materials into the modified CMOS process gives the device physicist and the circuit designer improved electrical performance similar to the historical trends, this new regime has been often identified as "Equivalent Scaling." Several materials are already under study. It is expected that these new materials will provide a viable solution to extending the limit of the planar CMOS process for the next five–ten years. These subjects are addressed in detail by the *2001 ITRS*.

Despite the use of these new materials, it will be challenging to maintain historical rates of improvement in electrical performance by relying exclusively on improvements in technology. Innovation in the techniques used in circuit and system design will be essential to maintain the historical trends in performance improvement. To achieve this result, it is expected that the integration of multiple silicon technologies on the same chip and a closer integration of package and silicon technology will be necessary. Specifically, given a system cost target, *what technology complexity can be afforded*? In each case it is necessary to evaluate whether a single chip or multiple chips in a single package provide the most cost effective way to obtain the desired performance.

Finally, as the *2001 ITRS* looks at 10–15 years in the future, it becomes evident that most of the known technological capabilities will approach or have reached their limits. In order to provide the computer, communication, consumer, and

other electronics industries with continuously more efficient building blocks, it becomes necessary to investigate new devices that may provide a more cost-effective alternative to planar CMOS in this timeframe. Adequate preparation for this potential transition must include early identification of the possible candidates and then systematically testing their feasibility.

In conclusion, note that the planar CMOS silicon gate technology ultimately resulted from technical investigations initiated in the 1940s. These early studies did not lead to the start of the semiconductor industry, as we know it today, until the late 1960s. It would be difficult for any single company to support the progressively increasing R&D investments necessary to evolve the technology from Traditional Scaling to Equivalent Scaling, and, finally, to investigate and develop a set of new devices usable beyond the limits of CMOS, as indicated in this *2001 ITRS*. Any industry that solely relies on technology evolution is bound to reach, sooner or later, fundamental technical limits. On one hand, the contributors to the *ITRS* agree that much of the R&D needs to be in the shared "*pre-competitive domain*." On the other hand, we recognize that industrial innovation also needs to continue to be fostered and encouraged.

It is the purpose of this *2001 ITRS* to provide a reference document 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. It is hoped that in the future—starting with this document as a common reference and through *cooperative efforts among the various ITRS participants*—the challenge of R&D investments will be cooperatively and more uniformly *shared by the whole industry* while, at the same time, *the fundamental elements that foster innovation will continue to be valued and cultivated by individual companies*.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

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

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

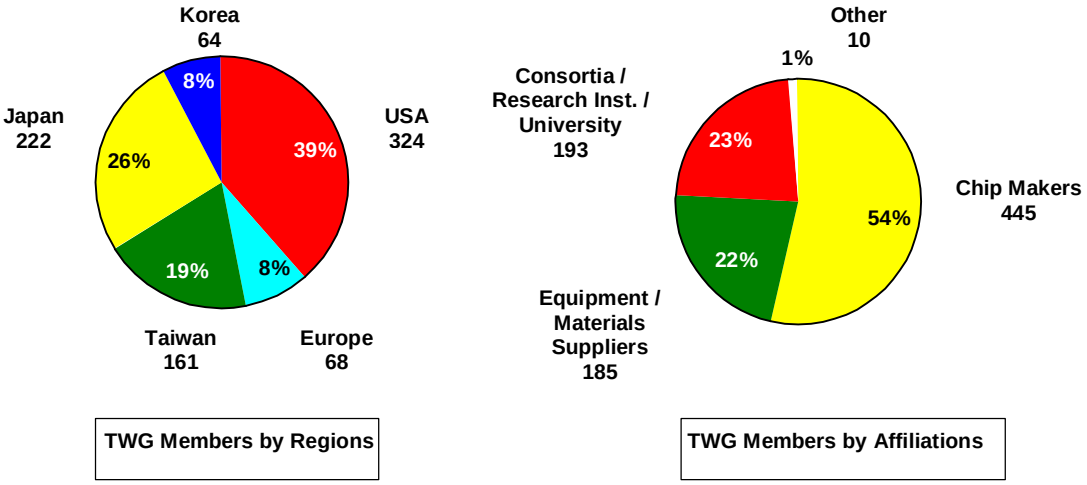


Figure 1 Composition of the Technology Working Group (TWG).

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.

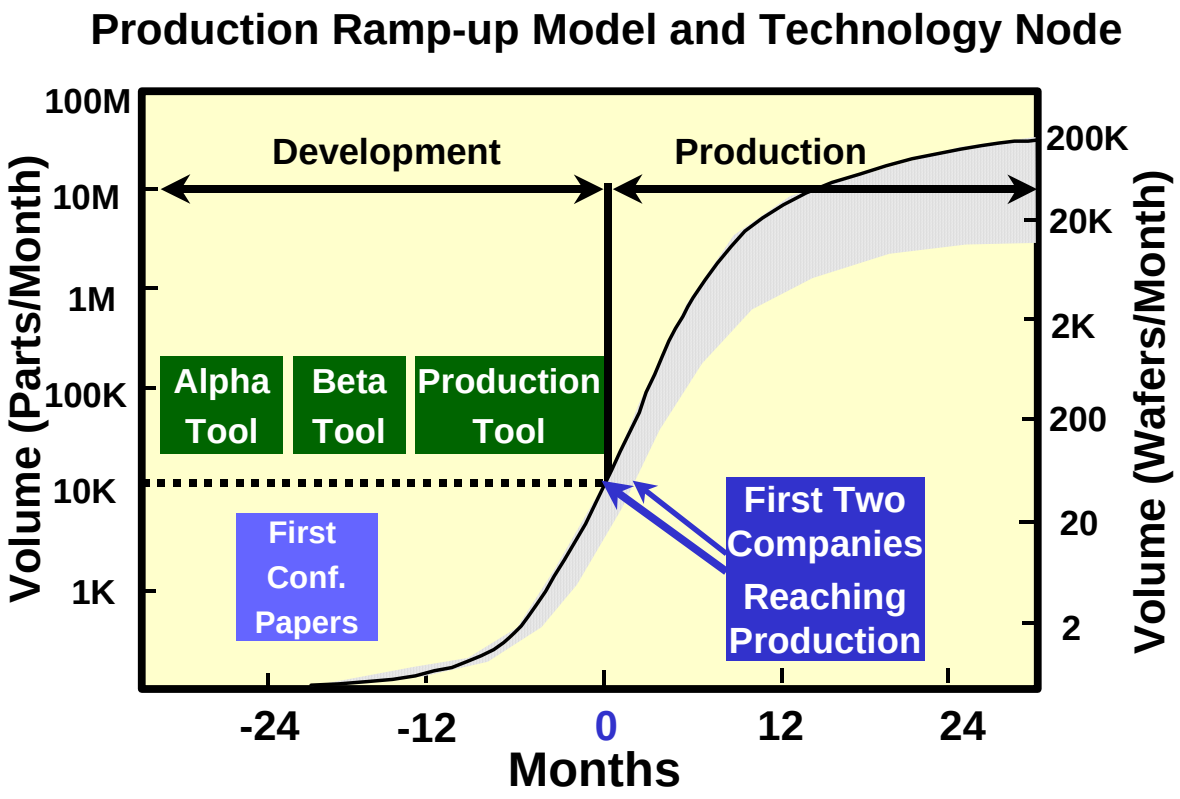


Figure 2 A Typical Production "Ramp" Curve

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.

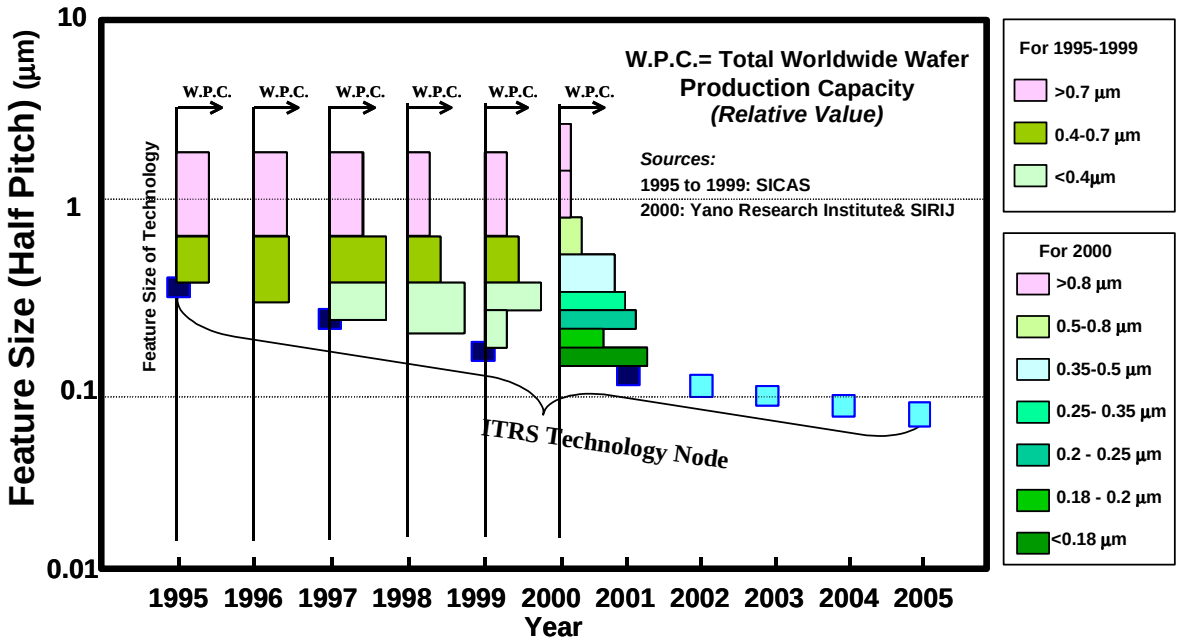
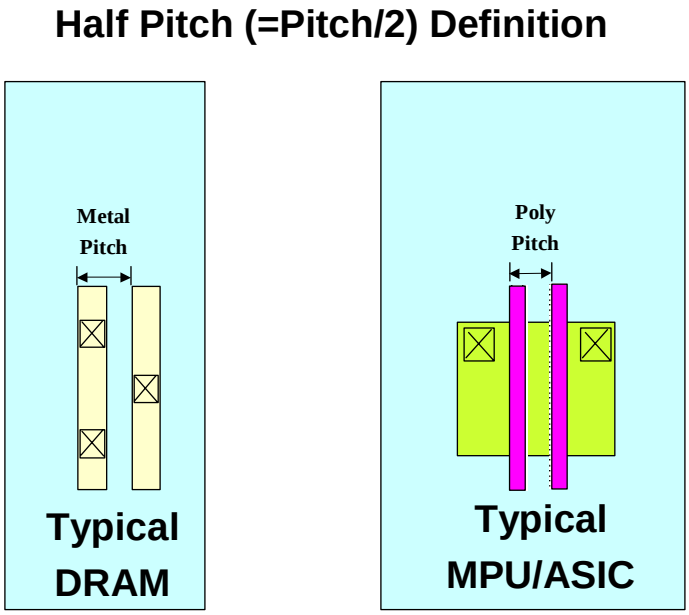


Figure 3 Technology Node Compared to Actual Production Capacity Technology Distribution

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 gross die per 300 mm wafer × 20K wspm × 70% total yield from wafer starts to finished product = 6M 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× increase in bits/chip between generations. For as long as this cycle strictly followed Moore's Law (three-year cycle for 4×), 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*.



** 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.

GRAND CHALLENGES
IN THE NEAR- (THROUGH 2007) AND LONG-TERM (2008 AND BEYOND)

OVERVIEW

This 2001 edition of *The International Technology Roadmap for Semiconductors* identifies the principal technologies expected to be required within the Roadmap timeframe. It should be noted however, that the *ITRS* outlines the efforts of research organizations and research sponsors within industry, government, and universities, and does not limit its scope to new or innovative concepts. *ITRS* is rather intended to encourage creative approaches to advanced microelectronics technologies.

In the working groups chapters' *Difficult Challenges* sections, the top difficult challenges for continuously improving performance and increasing integration, and/or decreasing manufacturing cost in both the near-term (2001 through 2007) and long-term (2008 through 2016) are described by each of International Technology Working Groups.

The *1999 ITRS* warned that there was a wide range of solutions needed but unavailable to meet the technology requirements corresponding to 100 nm technology node. The *1999 ITRS* edition also reported the presence of a potential "Red Brick Wall" or "100 nm Wall" (as indicated by the red cells in the technology requirements) that, by 2005, could block further scaling as predicted by Moore's Law. However, technological progress continues to accelerate. In the process of compiling information for *2001 ITRS*, it was clarified that this "Red Brick Wall" could be reached as early as 2003.

Three difficult challenges corresponding to each of the ITWGs were selected as "Grand Challenges" in order to emphasize their importance. These Grand Challenges were further classified into the two following categories: Enhancing Performance or Cost-effective Manufacturing. In this chapter, these Grand Challenges are summarized according to the "Near Term" and the "Long Term" timeframes of the Roadmap.

IN THE NEAR TERM (THROUGH 2007)

ENHANCING PERFORMANCE

MOS PERFORMANCE AND LEAKAGE [PROCESS INTEGRATION, DEVICES, AND STRUCTURES]

For low power logic (mainly for portable applications), the main issue is low leakage current, which is absolutely necessary in order to extend battery life. Device performance is then maximized according to the low leakage current requirements. Gate leakage current must be controlled, as well as sub-threshold leakage and junction leakage, including band-to-band tunneling. Preliminary analysis indicates that, balancing the gate leakage control requirements against performance requirements, high κ may be required for the gate dielectric by around the year 2005. In any case, one key point is that the requirement for high κ gate dielectric will probably be driven by the needs of low power logic, not by the needs of high performance logic.

PRODUCTION OF NON-CLASSICAL CMOS [PROCESS INTEGRATION, DEVICES, AND STRUCTURES AND FRONT END PROCESSES]

With the rapid scaling of gate length and the relatively slow scaling of the gate equivalent oxide thickness, transistor performance defined by CV/Id will become increasingly difficult to realize with conventional CMOS scaling in subsequent CMOS technology generations. Development of non-conventional CMOS architecture (one that employs double-gate MOS, FinFET, etc.) is required. This architecture may be incorporated into the manufacturing process by about 2007. Timely development, process integration, and qualification are required.

NEW GATE STACK AND MATERIALS [FRONT END PROCESSES]

Continuous reduction of the gate length requires a concomitant reduction in source and drain junction depths and gate dielectric thickness to ensure that scaling benefits are fully realized. One result has been a reduction of gate oxide thickness to less than 2 nm. However, the resulting gate leakage currents make the use of such thin oxides impractical in

many device applications where low standby power consumption is required. For this reason, the present gate oxide dielectric material will eventually be replaced by an alternative material that has a higher dielectric constant (high κ). Furthermore, the gate electrode material will also have to be replaced, since limitations of the traditional doped polysilicon electrode, associated with carrier depletion and boron out-diffusion, would negate the improved device performance that would have resulted from the use of the high κ gate dielectric constant.

CMOS INTEGRATION OF NEW MEMORY MATERIALS AND PROCESSES [FRONT END PROCESSES]

DRAM scaling results in less allowable chip area for DRAM storage capacitors, whereas maintaining reliable electrical performance requires that capacitance not be reduced to less than 25–35 fF. Consequently, process changes that produce thinner capacitor dielectric layers and/or produce a greater plate micro-area per unit allowable chip area have been introduced. These changes require the introduction of dielectric materials that have a higher dielectric constant. Additionally, the capacitor structures are anticipated to migrate from Silicon-Insulator-Silicon (SIS) to Metal-Insulator-Metal (MIM).

In the case of Flash memory devices, the continuous reduction in write voltage requires the use of a thinner tunnel oxide, which degrades data retention. Similarly, the maintenance of an almost constant coupling factor requires scaling of the inter-poly oxide thickness that again results in degraded data retention. This suggests that new tunnel and inter-poly dielectric materials will be introduced into Flash memory process flows.

STARTING MATERIALS ALTERNATE BEYOND 300 MM [FRONT END PROCESSES]

New demand for continued productivity enhancements would dictate the need for a new, large area starting substrate material. Historical trends suggest that the new starting material will have approximately 2.25× the area (1.5× diameter) of current-generation 300 mm substrates. However, it is far from clear whether traditional Czochralski pulling techniques can be scaled to produce cost-effective 450 mm diameter wafers. There is a pressing need for research and engineering to address and come to understand this obstacle.

CD AND L_{EFF} CONTROL [FRONT END PROCESSES AND LITHOGRAPHY]

The control of Critical Dimension (CD) etching especially at gate level has traditionally been one of the more difficult challenges. In recent years it has become common practice to use etch processes that result in a gate dimension (effective gate length: L_{eff}) that is smaller than that printed in the resist. More complex etch processing must be executed while still maintaining the overall gate 3-sigma dimensional tolerance. In addition, the shape of the sidewall profile must be maintained in order to achieve acceptable sidewall oxide coverage and reliability. Another challenge is stopping the etch process at a very thin gate dielectric without cutting a trench into the underlying silicon.

MASK-MAKING [LITHOGRAPHY]

Mask-making capability and cost escalation will continue to be critical to future progress in lithography and will require continued attention. As a consequence of aggressive roadmap acceleration particularly in terms of MPU gate linewidth (post etch) and increased mask error factors (MEFs) associated with low κ1 lithography, mask linewidth controllability fails to meet the requirements of the chipmakers. For example, in the 1997 Roadmap, the 70 nm node device required 4× masks to achieve 9 nm of CD control for isolated lines and 14 nm for contacts. The current CD control requirements are 3.4 nm for isolated lines and 4.3 nm for contact when MEF is assumed to be 1.4 and 3.0, respectively.

Mask equipment and process capabilities for complex optical proximity correction (OPC) and phase shift masks (PSM) are currently available, while mask processes for post-193 nm technologies are still at a research and development stage. Mask damage caused by electrostatic discharge (ESD), which has long been a concern, is likely to be even more problematic as mask feature sizes will shrink further and masks for 157 nm lithography will be kept in atmospheres nearly free of water.

PROCESS CONTROL [LITHOGRAPHY]

Process control, particularly for overlay and linewidths, also represents a major challenge. It is unclear whether metrology, which is fundamental for process control, will be upgraded adequately to meet future requirements. Resist line edge roughness (LER) becomes increasingly significant as gate linewidth control needs to become as precise as size of a polymer unit. Next-generation lithography will require exposure tool of totally new concept. The new tool must be developed and proven to meet reliability and utilization requirements to realize cost-effective production

INTEGRATION OF NEW PROCESSES AND STRUCTURES [INTERCONNECT]

Combinations of new materials, structures, and processes increase integration complexity. In the process of forming contact holes, a barrier metal that prevents interaction between wiring and insulation film, and a related new process technology need to be developed. Also, a technology that fills contact holes with metal so they have a high aspect ratio (A/R) is also required. As the feature size continues to shrink, it becomes more challenging to develop new insulating materials that prevent metal diffusing and a new process technology as well.

COST-EFFECTIVE MANUFACTURING

SCALING OF MAXIMUM QUALITY DESIGN IMPLEMENTATION PRODUCTIVITY [DESIGN]

The number of available transistors scales by a factor of 2 every two years (DRAM) or a factor of 2 at each technology node (MPU), increasing design complexity as well. In order to maintain design quality even after process technologies advance, design implementation productivity must be scaled the same degree that design complexity is scaled.

Improving design productivity and reusing the design are the key considerations for this issue. Namely, overall design productivity of quality- (difficulty-) normalized functions on-chip must scale at the rate of 2× per node.

However, analog and mixed-signal design traditionally suffers from a difficulty in improving design productivity and reusing the design along with process migration. Therefore, there is a strong need to develop a new design methodology to improve those issues by implementing analog and mixed-signal synthesis, verification, and testing.

Embedded software productivity also needs to improve on a similar scale since the on-chip memory size is also growing and some functions are built into such embedded software rather than the hardware.

POWER MANAGEMENT [DESIGN]

Even off-currents in low-power devices increase by a factor of 10 per node, so design technology must maintain constant static power. On the other hand, while power dissipation for high-performance MPU will exceed package limits by 25× in 15 years, design technology must achieve power limits.

As a result, efficient power management requires highly complex controllability across the entire Large-scale Integrated (LSI) circuit. Additionally, any power optimization must simultaneously and fully exploit varying degrees of freedom, for example switching the operating power state of circuits using multi-V_t, multi-T_{ox}, multi-V_{dd} in the LSI core block while guiding the architecture, OS and software.

HIGH-SPEED DEVICE INTERFACES [TEST AND TEST EQUIPMENT]

One major roadblock will be the increasing need for high-frequency, high pincount probes and drastically increased test sockets. In order to overcome this obstacle and find cost effective solutions, urgent research and development is required. The trend towards faster high-speed serial interfaces and an increased port count will continuously drive the need for high-speed analog source/capture and jitter analysis instrument capability during characterization.

Design-For-Test (DFT) / Design-For-Manufacture (DFM) techniques will be a must for manufacturing.

Device interface circuitry must not degrade equipment bandwidth and accuracy, especially in the case of high-frequency differential I/O and analog circuits. Otherwise, it would introduce noise.

HIGHLY INTEGRATED DESIGN AND SOCS [TEST AND TEST EQUIPMENT]

Highly structured DFT approaches are required to enable test access to embedded cores since individual cores require special attention when using DFT and the Built-in Self Test (BIST) to enable testing.

To simplify both test interface requirements and the slow but ever increasing instrument capability trends, analog DFT and BIST techniques must be more fully developed. Also, testing chips containing RF and audio circuits will be a major challenge if they also contain large numbers of noisy digital circuits.

DFT must enable test reuse for reusable design cores. Doing so would totally reduce test development time for highly complex designs.

NEW RELIABILITY SCREENS [TEST AND TEST EQUIPMENT]

In a sense existing methodologies are reaching their limits due to declining capability of IDDQ (Quiescent I_{dd}) testing for devices with a high background current and thermal runaways during burn-in tests. As a result, fundamental research is needed to identify novel infant mortality defects under accelerated stress conditions.

TOOL COST AND R&D COST [LITHOGRAPHY]

While lithography has long accounted for a significant portion of over-all semiconductor manufacturing costs, it is likely to impose even greater concern in aspect of cost control and return on investment (ROI). As throughput of lithography tool drops along with introduction of larger wafers, lithography is likely to become a more dominant factor in total manufacturing costs along with the transition to 300 mm wafers. These issues relating to masks and lithography costs are relevant to optical as well as next-generation lithography. For optical lithography, new resists featuring good pattern fidelity when exposed with short wavelengths (193 nm and 157 nm) must be developed, and new optical materials for lenses, such as CaF_2 will also be needed, after 193 nm (ArF) lithography technology. Inadequacies in performance and supply of resist and CaF_2 have already slowed down advances of lithography technology.

COMPLEXITY IN INTEGRATING FACTORIES [FACTORY INTEGRATION]

Rapid changes in semiconductor technologies, business requirements, and market conditions are making effective and timely factory integration to meet accelerated ramp and yield targets more difficult over time. The factory now must integrate an even larger number of new and different equipment types and software applications in a much shorter time to meet multiple business objectives and customer requirements. To meet this need, equipment and process control systems with standard interfaces that are simple to integrate are critically required. In addition, improvements in factory forecasting and flexible factory information/control systems that can change with business conditions must be developed and implemented.

The number of equipment types and factory software systems used for managing factories is steadily increasing over time. Furthermore, increased integration complexity only serves to accelerate this trend. Therefore, it is necessary to decrease the cycle time to build/ramp new factories, or to convert/ramp existing factories to meet new technology expectations. This is where more complex embedded controller software is needed to enable connections between equipment.

Process technology advances are occurring at ever faster rates especially in the following technology areas: 157 nm lithography, high κ gate stack, low κ dielectrics, silicon-on-insulator (SOI), copper, and others. Nevertheless, factories are expected to ramp up and meet yield targets quicker by using new equipment over the next 2–5 years. This requires integration of yield systems with Work-In- Process (WIP) management.

PRODUCTION EQUIPMENT OEE PERFORMANCE AND EXTENDIBILITY (REUSE) [FACTORY INTEGRATION]

Production equipment is not keeping up with Overall Equipment Efficiency (OEE) and Availability targets, which has an enormous impact on capital and operating costs. The industry is unable to effectively reuse equipment or skills due to the rapid introduction of new equipment (157 nm lithography) and materials (such as copper, high κ gate stack, low κ dielectrics, SOI). Agile manufacturing, advance process control, scheduling and dispatching, e-Diagnostics, and other enablers must be developed to improve equipment OEE and extendibility.

REALIZING EFFICIENT 300 MM CONVERSION [FACTORY INTEGRATION]

The industry must quickly ramp 300 mm factory production to high volumes while achieving the efficiency targets that it had set. Some of these efficiencies include >2.25 more die per wafer than 200 mm, >30% die-cost reduction, 100% AMHS interbay and intrabay systems ergonomics, operational flexibility and cost improvements, and the ability to track and run different recipes for each wafer within a carrier for operational flexibility. Open-standards based production equipment and software must also be implemented per industry priorities for efficient conversion to 300 mm production.

COORDINATED DESIGN TOOLS AND SIMULATORS TO ADDRESS CHIP, PACKAGE, AND SUBSTRATE CO-DESIGN [ASSEMBLY AND PACKAGING]

The continuous migration of semiconductors to smaller features, higher frequency, higher power density, lower voltages, integration of mixed signals, etc., demands a very aggressive packaging technology. Otherwise, packaging will become a limiting factor in the continued evolution of overall semiconductor technology.

To satisfy the performance demands at the system level and to shorten the development cycle, it is necessary to develop integrated design tools and a simulation technology that will simultaneously consider items such as the following: electrical characteristics, thermal dissipation, thermo-mechanical stress, physical requirements and environmental impact. Those tools and technology are used through the design process to mount a chip in a package/module or chip/package on a board. Support from commercial electronic design automation (EDA) suppliers is indispensable. The acceleration of the development of coordinated design tools and simulators is a challenge.

CHEMICAL AND MATERIAL ASSESSMENTS [ESH]

The rapid introduction of new chemicals/materials/processes requires new rapid assessment methodologies to ensure that new chemicals/materials can be utilized in manufacturing without inducing a new hazardous impact on human health, safety, and the environment. Although the methodologies are needed to meet the evaluation and quantification demands on ESH impact, they are currently required to expedite process implementation.

RESOURCE CONSERVATION [ESH]

As the industry grows in size and its technology advances toward finer patterning and larger wafer sizes, the natural tendency is towards increased use of water, energy, chemicals, and materials. Resource conservation especially becomes a major concern with respect to availability, cost reduction, manufacturing location, sustainability, and waste disposal. Thus development of a wide variety of resource effective process equipment becomes necessary.

HIGH ASPECT RATIO INSPECTION [YIELD ENHANCEMENT]

Detecting defects associated with high aspect ratio contacts and combinations of trenches and vias in dual-Damascene structures will continue to be difficult defect detection challenges. More specifically, the detection of via defects near/at the bottom of a Damascene trench will continue to be a Grand Challenge. However, the challenge is complicated by the simultaneous need for high sensitivity and high throughput. High-speed, cost-effective detection tools that satisfy both demands are therefore needed.

NON-VISUAL DEFECT SOURCING AND MANUFACTURE AND TEST ORIENTED DESIGN [YIELD ENHANCEMENT]

Fault isolation complexity is expected to grow exponentially, combining the difficult tasks of defining fault dimensions in the horizontal plane and vertical layers. It is especially difficult to analyze circuit failures that leave no detectable physical remnant. Accordingly, new analysis tools and techniques that can isolate those non-visual failures are needed. Although IC design must be optimized for a given process capability and must be testable/diagnosable, on the other hand many defects that cause electrical faults are not detectable inline. Furthermore, IC design must be optimized for a given process capability and must be testable/diagnosable. Tools are also needed that enable design and process matching so that optimum yields can be achieved.

PROMPT DEVELOPMENT OF METROLOGY TOOLS [METROLOGY]

The continued acceleration of feature size reduction drives metrology solutions for new materials, process, and structures. However, new process control needs are not completely established. As is well known, metrology accelerates yield improvement at every stage of manufacturing and reduces the cost of manufacturing and the time-to-market for new products through better characterization of process tools and processes. Therefore, leading edge developments in measurement tools and technology must be commercialized in a more timely manner.

HIGH ASPECT RATIO INSPECTION [METROLOGY]

Control of high-aspect ratio technologies such as Damascene challenges all metrology methods. Key requirements are void detection in copper lines and pore size distribution in patterned low κ. The need is to have a rapid, in-line observation of a very small number of voids/larger pores. Critical dimension measurements are also required for very high aspect ratio structures that are made from porous dielectric materials and require 3D information for trench and via/contact sidewalls. These measurements will be further complicated by the underlying multi film complexity.

MEASUREMENT OF COMPLEX STRUCTURES [METROLOGY]

Reference materials and a standard measurement methodology are required for new, high κ gate and capacitor dielectrics with interface layers, thin films such as interconnect barriers and low κ dielectric layers, and other processes. Optical measurement of gate and capacitor dielectrics averages over too large an area and needs to characterize interfacial layers.

The same is true for measurement of barrier layers. A measurement methodology is therefore needed for complex material stacks and interfacial properties including physical and electrical properties. Further high frequency dielectric constant measurement advances are required.

IN THE LONG TERM (2008 THROUGH 2016)

ENHANCING PERFORMANCE

NON-CMOS DEVICE AND ARCHITECTURE INCLUDING INTERCONNECT AND MEMORY [PROCESS INTEGRATION, DEVICES, AND STRUCTURES]

It is recognized that research and development on emerging non-CMOS devices is proceeding rapidly toward the end of the Roadmap. Owing to the difficulty of successfully scaling conventional bulk planar CMOS technology to meet the increased performance, density, and reduced power dissipation required for future technology generations, such novel devices will likely be needed eventually. Implementation of non-CMOS device structures and architectures, including interconnect and memory, will drive major changes in process, materials, physics, and design. The emerging non-CMOS devices may coexist with conventional-CMOS integration.

NEXT-GENERATION LITHOGRAPHY [LITHOGRAPHY]

Optical lithography falls short of meeting tough requirements of the 45 nm node and beyond. The long-term Roadmap indicates a need to develop next-generation lithography (NGL) technologies such as extreme ultraviolet lithography (EUV) and electron projection lithography (EPL). As next-generation lithography will require development of a totally new infrastructure, the R & D costs will boost the overall manufacturing costs.

IDENTIFY SOLUTIONS THAT ADDRESS GLOBAL WIRING ISSUES [INTERCONNECT]

Conventional interconnect scaling will no longer satisfy performance requirements. Defining and finding solutions beyond copper and low κ will require material innovation combined with accelerated design, packaging, and unconventional interconnect. New consistent models for transistors, wiring structures, delay calculation, signal length, and target performance are needed.

COST-EFFECTIVE MANUFACTURING

NOISE MANAGEMENT [DESIGN]

Since the operating voltage decreases 20% per technology node, increasing noise sensitivity is becoming a big issue in the design of functional devices (e.g., bits, transistors, gates) and products (such as DRAMs or MPUs). This is becoming more evident due to lower noise headroom especially in low-power devices, coupled interconnects, IR drop and ground bounce in the supply voltage, thermal impact on device off-currents and interconnect resistivities, mutual inductance, substrate coupling, single-event upset (alpha particle), and increased use of dynamic logic families. Consequently, modeling, analysis, and estimation must be performed at all design levels.

ERROR-TOLERANT DESIGN [DESIGN]

The scaling of the design complexity and the increasing transistor count will greatly reduce the potential for failures to occur. In this case, relaxing the requirement for 100% correctness in both transient and permanent failures of signals, logic values, devices, or interconnects may reduce the cost of manufacturing, verification and testing.

Potential solutions are adaptive and self-correcting/self-repairing circuits, and the use of on-chip re-configurability

STARTING MATERIALS ALTERNATE BEYOND 300 MM [FRONT END PROCESSES]

New requirements for continued productivity enhancements will dictate the need for a new, large area starting substrate material. Historical trends suggest that the new starting material will likely have double the area of current-generation 300 mm substrates. However, it is far from clear whether conventional Czochralski pulling techniques can be scaled to produce cost-effective 450 mm diameter wafers. There is an urgent, pressing need for research and engineering to understand and address this obstacle.

POST BULK CMOS AND 450 MM WAFER MANUFACTURING PARADIGMS [FACTORY INTEGRATION]

The conversion to novel devices and 450 mm wafers represent key inflection points for semiconductor manufacturing. The transition from bulk CMOS is expected to occur at the 45 nm technology node (around 2010). Specifics around the types of novel devices and their potential impacts to equipment and the manufacturing facility are not well defined, but are expected to be significant. Further new materials and perhaps even production approaches drastically different than our circular wafer substrate concept may arise.

The conversion to 450 mm wafers is projected to occur in high volume around 2013. This is timing is 12 years beyond the 300 mm conversion and allows three years of lessons learned to occur after the transition from bulk CMOS which should reduce overall industry risk. It also affords time to understand whether 450 mm can be a simple scale up of 300 mm or if more fundamental changes to manufacturing must occur. Cost-effective manufacturability of both novel devices and 450 mm wafers will be important long-term factors to the industry's ability to continue realizing Moore's law.

CHEMICAL AND MATERIAL MANAGEMENT BY ESH DESIGN AND MEASUREMENT METHODS [ESH]

Equipment design engineers and equipment users require timely ESH information regarding ESH characteristics of potential new process chemicals and materials. This information is essential to the proper selection of optimal chemicals and materials for function and ESH impact with respect to reaction product emissions, health and safety properties, materials compatibility with both equipment and other chemical components, flammability, and reactivity. It must be possible to do so while minimizing unnecessary business impact after processes are developed and are in production.

For integrated ESH design and measurement methods, a methodology for determining the lowest ESH impact of materials and processes needs to be developed.

YIELD MODEL DEVELOPMENT FOR NEW MATERIALS AND INTEGRATION [YIELD ENHANCEMENT]

Defect budgets will require frequent revalidation and updates as information about future processing technologies becomes available. Yield models need to consider complex integration issues with respect to random defect-limited yield as well as systematic limited yield (such as parametric yield loss, or circuit yield loss) for future technology nodes. As a result, the models must take into account greater parametric sensitivities, complex integration issues, ultra-thin film integrity, impact of circuit design, and greater transistor packing density.

NON-DESTRUCTIVE PRODUCTION MEASUREMENTS [METROLOGY]

As is well known, surface charging and contamination interfere with electron beam imaging. CD measurements must account for the side wall shape. In addition, CD for the Damascene process may require measurement of trench structures. On the other hand, process control such as focus exposure and etch bias will require greater precision and 3D capability.

As a result, non-destructive (without charging or contaminating the surface) wafer/mask level microscopy for measuring the critical dimensions of 3D structures, overlay, and defect detection are required. Furthermore, analysis of 3D structures such as tapered- or undercut-gate electrodes, trenches, high aspect ratio capacitors, and contacts is needed.

ULTIMATE CMOS SIMULATION CAPABILITY [MODELING AND SIMULATION]

As devices shrink and more materials are introduced into the technology, new analytical techniques that provide the necessary information for the validation of models (i.e., getting the physics and chemistry correct) is critical. In the ultimate CMOS technology, modeling tools for atomistic modeling, nano-scale device modeling, mechanical modeling and integration of simulation modules are needed. Simulation tools that predict the physical limits of the gate materials, the lithography, and the device characteristics especially are required. Furthermore, in the areas where parasitics, reliability effects, and/or statistical variations dominate, full or hybrid quantum descriptions in 2D/3D will be needed.

DIFFICULT CHALLENGES
OF THE INTERNATIONAL TECHNOLOGY WORKING GROUPS

DESIGN

Difficult Challenges	*Summary of Issues*
Productivity	To avoid exponentially increasing design cost, overall productivity of designed functions on chip must scale at > 2× per node. Reuse productivity (including migration) of design, verification and test must also scale at > 2× per node.
Power	Non-ideal scaling of planar CMOS devices, together with the roadmap for interconnect materials and package technologies, presents a variety of challenges related to power management and current delivery.
Manufacturing Integration	"Red bricks"—technology requirements for which no known solutions exist—are increasingly common throughout the *ITRS*. On the other hand, challenges that are impossible to solve within a single technology area of the *ITRS* may be solvable (more cost-effectively) with appropriate intervention from, or partnership with, DT. Feasibility of future technology nodes will come to depend on such "sharing of red bricks."
Interference	Resource-efficient communication and synchronization, already challenged by global interconnect scaling trends, are increasingly hampered by noise and interference. Prevailing signal integrity methodologies in logical, circuit and physical design, while apparently scalable through the 100 nm node, are reaching their limits of practicality.
Error Tolerance	Relaxing the requirement of 100% correctness for devices and interconnects may dramatically reduce costs of manufacturing, verification, and test. Such a paradigm shift is likely forced in any case by technology scaling, which leads to more transient and permanent failures of signals, logic values, devices, and interconnects.

TEST

Five Difficult Challenge ≥65 nm , Through 2007	*Summary of Issues*
High Speed Device Interfaces	A major roadblock will be the need for high-frequency, high pin-count probes and test sockets; research and development is urgently required to enable cost effective solutions with reduced parasitic impedance. High speed serial interface speed and port count trends will continue to drive high speed analog source/capture and jitter analysis instrument capability for characterization. DFT/DFM techniques must be developed for manufacturing. Device interface circuitry must not degrade equipment bandwidth and accuracy, or introduce noise; especially for high-frequency differential I/O and analog circuits.
Highly Integrated Designs	Highly structured DFT approaches are required to enable test access to embedded cores. Individual cores require special attention when using DFT and BIST to enable test. Analog DFT and BIST techniques must mature to simplify test interface requirements and slow ever increasing instrument capability trends. Testing chips containing RF and audio circuits will be a major challenge if they also contain large numbers of noisy digital circuits. DFT must enable test reuse for reusable design cores to reduce test development time for highly complex designs.
Reliability Screens	Existing methodologies are limited (burn-in versus thermal runaway, IDDQ versus background current increases). Research is required to identify novel infant mortality defect acceleration stress conditions
Manufacturing Test Cost	Test cell throughput enhancements are needed to reduce manufacturing test cost. Opportunities include massively parallel test, wafer-level test, wafer-level burn-in, and others. Challenges include device interfacing/contacting, power and thermal management. Device test needs must be managed through DFT to enable low cost manufacturing test solutions; including reduced pin count test, equipment reuse, and reduced test time. Automatic test program generators are needed to reduce test development time. Test standards are required to enable test content reuse and manufacturing agility.
Modeling and Simulation	Logic and timing accurate simulation of the ATE, device interface, and DUT is needed to enable pre-silicon test development and minimize costly post-silicon test content development/debug on expensive ATE. High performance digital and analog I/O and power requirements require significant improvements to test environment simulation capability to ensure signal accuracy and power quality at the die. Equipment suppliers must provide accurate simulation models for pin electronics, power supplies, and device interfaces to enable interface design.

TEST (CONTINUED)

Five Difficult Challenges <65 nm, Beyond 2007	
DUT to ATE interface	Probing capability for optical and other disruptive technologies. Support for massively parallel test - including full wafer contacting. Decreasing die size and increasing circuit density are driving dramatic increases in die thermal density. This problem is further magnified by the desire to enable parallel test to maximize manufacturing throughput. New thermal control techniques will be needed for wafer probe and component test. DFT to enable test of device pins not contacted by the interface and test equipment.
Test Methodologies	New DFT techniques (SCAN and BIST have been the mainstay for over 20 years). New test methods for control and observation are needed. Tests will need to be developed utilizing the design hierarchy. Analog DFT and BIST techniques must mature to simplify test interface requirements and slow ever increasing instrument capability trends. Logic BIST techniques must evolve to support new fault models, failure analysis, and deterministic test. EDA tools for DFT insertion must support DFT selection with considerations for functionality, coverage, cost, circuit performance and ATPG performance.
Defect Analysis	Defect types and behavior will continue to evolve with advances in fabrication process technology. Fundamental research in existing and novel fault models to address emerging defects will be required. Significant advances in EDA tools for ATPG capacity and performance for advanced fault models and DFT insertion are required to improve efficiency and reduce design complexities associated with test.
Failure analysis.	Realtime analysis of defects in multi-layer metal processes are needed. Failure analysis methods analog devices must be developed and automated. Transition from a destructive physical inspection process to a primarily non-destructive diagnostic capability. Characterization capabilities must identify, locate, and distinguish individual defect types.
Disruptive device technologies	Develop new test methods for MEMS and sensors. Develop new fault models for advanced/disruptive transistor structures.

PROCESS INTEGRATION, DEVICES, AND STRUCTURES

Difficult Challenges ≥ 65 nm, Through 2007	*Summary of Issues*
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
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 these transistors will be necessary eventually to control short-channel and other effects in highly scaled devices.
4. Ensuring reliability of new materials and structures in a timely manner.	Accelerated reliability ensurance 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.
Difficult Challenges < 65 nm, Beyond 2007	*Summary of Issues*
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.

FRONT END PROCESSES

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

FRONT END PROCESSES (CONTINUED)

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

LITHOGRAPHY

Five Difficult Challenges ≥65 nm, Through 2007	*Summary of Issues*
Optical mask fabrication with resolution enhancement techniques for ≤ 90 nm and development of post-optical mask fabrication	Development of commercial mask manufacturing processes to meet requirements of Roadmap options (such as registration, CD control, defectivity, and 157 nm films; defect free multi-layer substrates or membranes). Development of equipment infrastructure (writers, inspection, repair, metrology) for a relatively small market.
Cost control and return on investment (ROI)	Achieving constant/improved ratio of tool cost to throughput over time Development of cost-effective resolution enhanced optical masks and post-optical masks including an affordable ASIC solution, such as low cost masks. Achieving ROI for all segments of the industry (chipmakers, equipment and material suppliers, and infrastructure) with sufficient lifetimes for the technologies, especially single node solutions at 90 nm and below.
Process control	Development of processes to control gate linewidths to nearly 3nm, 3 σ Development of new and improved alignment and overlay control methods independent of technology option for < 25 nm overlay.
Resists for ArF and F_2	Outgassing, LER, SEM induced CD changes, etch resistance, and defects as small as 40 nm.
CaF_2	Yield, cost, quality.
Five Difficult Challenges < 65 nm, Beyond 2007	
Mask fabrication and process control	Development of commercial mask manufacturing processes to meet requirements of Roadmap options (defect-free NGL masks, such as EUV multi-layer masks or EPL membranes and stencil masks). Development of equipment infrastructure (writers, inspection of substrates, blanks and patterned masks, repair, metrology) for a relatively small market. Development of mask process control methods to achieve critical dimensions, image placement, and defect density control below the 65 nm node.
Metrology and defect inspection	Capability for measuring critical dimensions down to 9 nm and metrology for overlay down to 9 nm, and patterned wafer defect inspection for defects < 40nm.
Cost control and ROI	Achieving constant/improved ratio of tool cost to throughput over time. Development of cost-effective post-optical masks including an affordable ASIC solution, such as low cost masks. Achieving ROI for industry (chipmakers, equipment and material suppliers, and infrastructure) with sufficient lifetimes for the technologies, especially single node solutions at 45 nm and below.
Gate CD control improvements; process control; resist materials	Development of processes to control gate CDs < 1nm (3 sigma) with appropriate line-edge roughness. Development of new and improved alignment and overlay control methods independent of technology option to < 9 nm overlay.
Tools for mass production	Post optical exposure tools capable of meeting requirements of the Roadmap.

INTERCONNECT

Five Difficult Challenges ≥ 65 nm, Through 2007	*Summary Of Issues*
Introduction of new materials*	The rapid introduction of new materials/processes which are necessary to meet conductivity requirements and reduce the dielectric permittivity create integration and material characterization challenges.
Integration of new processes and structures*	Combinations of materials and processes used to fabricate new structures create integration complexity.
Achieving necessary reliability	New materials, structures and processes create new chip reliability (electrical, thermal and mechanical) exposure. Detecting, testing, modeling and control of failure mechanisms will be key.
Attaining dimensional control	Three-dimensional control (3D CD), with its associated metrology, of interconnect features is necessary for circuit performance and reliability. The multiplicity of levels combined with new materials, reduced feature size, and pattern dependent processes, create this challenge.
Manufacturability and defect management that meet overall cost/performance requirements	As feature sizes shrink, interconnect processes must be compatible with device roadmaps and meet manufacturing targets at the specified wafer size. Plasma damage, contamination, thermal budgets, cleaning of high A/R features, defect tolerant processes, elimination/reduction of control wafers are key concerns. Where appropriate, global wiring and packaging concerns will be addressed in an integrated fashion.
Five Difficult Challenges <65 nm, Beyond 2007	*Summary Of Issues*
Dimensional control and metrology	Multi-dimensional control and metrology of interconnect features is necessary for circuit performance and reliability.
Patterning, cleaning and filling high aspect ratios features	As features shrink, etching, cleaning and filling high aspect ratio structures will be challenging, especially for low κ dual Damascene metal structures and DRAM.
Integration of new processes and structures	Combinations of materials and processes used to fabricate new structures create integration complexity. The increased number of levels exacerbate thermomechanical effects. Novel/active devices may be incorporated into the interconnect.
Continued introductions of new materials and size effects	Further introductions of materials/processes are expected. Microstructure and dimensional effects become important when Cu/low κ interconnect is extended to smaller features.
Identify solutions which address global wiring scaling issues*	Traditional interconnect scaling will no longer satisfy performance requirements. Defining and finding solutions beyond copper and low κ will require material innovation, combined with accelerated design, packaging and unconventional interconnect.

** Top three challenges*

FACTORY INTEGRATION

Difficult Challenges ≥ 65 nm, Through 2007	*Summary of Issues*
Complexity Management	Rapid changes to business needs, demand, and globalization trends Increasing rate of new product and technology introductions Globally disparate factories run as single "virtual factory" Need to meet regulations in different geographical areas Effectively manage complex supply chains Increasing process and product complexity Explosive growth of data collection/analysis requirements Increasing number of processing steps including material movements Multiple lots in a carrier and single wafer control/transport for an equipment group Multiple Products on a wafer Multiple package form factors Larger wafers and carriers driving ergonomic solutions Increasing expectations for integrated material handling and software control systems Increased reliance on factory information and control systems Multiple information and control system interdependencies Standalone and Integrated Reliability required to keep factories operating Co-existence of new factory information and control systems with existing (legacy) Pace of standards definition and implementation is not meeting factory integration needs Addressing complexity while keeping costs in perspective
Factory Optimization	Increased customer expectation to meet on time delivery Balanced throughput and cycle time Reduce time to ramp factories, products, and processes Increased urgency for improved factory effectiveness Ability to measure effectiveness and adjust/optimize factory output High factory yield at startup Reduce effects of parametric variation Reduce wafer and product cost Satisfy all domestic and international regulations
Extendibility, Flexibility, and Scalability	Reuse of building, production and support equipment, and factory information and control systems Across multiple technology nodes and wafer size conversions Factory designs that support rapid process and technology changes and retrofits Understand up-front costs to incorporate EFS Determine which EFS features to include and not to include Minimize downtime to on-going operations Comprehend tighter ESH/Code requirements Comprehend increased purity requirements for process and materials

FACTORY INTEGRATION (CONTINUED)

Difficult Challenges ≥65 nm, Through 2007	*Summary of Issues*
Post Bulk CMOS Manufacturing	New devices beyond traditional bulk CMOS will drive significant changes in process technology and the factory manufacturing capabilities to support it New production equipment must be designed and integrated with the factory Yield and process control capabilities must be modified to new device process Unknown changes to factory operations motivated by new device change Potential for additional process and functional area isolation Need to run CMOS and post CMOS processes within the same factory Rapid technology development and ramp to support high volume manufacturing of new Continued pressure to maintain 0.7× shrink per year efficiency in equivalent die size Development and high volume ramp timeline must meet current technology node Device yield and factory output must meet current roadmap targets Reuse of buildings and equipment to enable new device technology at an affordable cost Potential for additional process and functional area isolation an Need to run CMOS and post CMOS processes within the same factory Comprehend tighter ESH and code requirements
450 mm Wafer Size Conversion	Larger wafers and carriers driving changes to traditional wafer size scale-up strategy Increased requirements for wafer level integration within the factory Significant changes to production equipment and material handling design Pressure to met affordability targets Reuse of buildings and equipment to enable 450 mm at an affordable cost Comprehend tighter ESH and code requirements

ASSEMBLY AND PACKAGING

DIFFICULT CHALLENGES ≥ 65nm / THROUGH 2007	*SUMMARY OF ISSUES*
Improved organic substrates	Tg compatible with Pb free solder processing Increased wireability at low cost Improved impedance control and lower dielectric loss to support higher frequency applications Improved planarity and low warpage at higher process temperatures Low-moisture absorption Low-cost embedded passives
Improved underfills for flip chip on organic substrates	Improve flow, fast dispense/cure, better interface adhesion, lower moisture absorption Higher operating range for automotive in liquid dispense underfills Improved adhesion, small filler size, and improved flow for mold based underfills
Coordinated design tools and simulators to address chip, package, and substrate co-design	Mix signal co-design and simulation environment Faster analysis tools for transient thermal analysis and integrated thermal mechanical analysis Electrical (power disturbs, EMI†, signal integrity associated with higher frequency/current and lower voltage switching) Commercial EDA‡ supplier support
Impact of Cu/low κ on packaging	Direct wirebond and bump to Cu Bump and underfill technology to assure low κ dielectric integrity Improved Mechanical strength of dielectrics Interfacial adhesion
Pb, Sb, and Br free packaging materials	Lower cost materials and processes to meet new requirements, including higher reflow temperatures. Reliability under thermal cycling (stress and moisture)
DIFFICULT CHALLENGES < 65 nm / BEYOND 2007	
Package cost that may greatly exceed die cost	Research investments required for packaging cost reduction are decreasing
Small, high pad count	Array I/O pitches below 80 microns
High Frequency die	Substrate wiring density to support >20 lines/mm Lower loss dielectrics Skin effect above 10GHz
Close gaps between substrate technology and the chip	Interconnect density scaled to silicon (silicon I/O density increasing faster than the printed circuit
System level design capability to integrated chips, passives and substrates	Partitioning of system designs and manufacturing across numerous companies will make required optimization for performance, reliability, and cost of complex systems very difficult. Complex standards for information types and management of information quality along with a structure for moving this information will be required.

** CTE - Coefficient of thermal expansion*

*** UBM - Under bump metallurgy*

†-EMI-Electromagnetic interference

‡ EDA-Electronic design automation

ENVIRONMENT, SAFETY, AND HEALTH

Five Difficult Challenges ≥ 65 nm, Through 2007	*Summary of Issues/Needs*
Chemicals, Materials and Equipment Management	*New Chemical Assessment* Need for quality rapid assessment methodologies to ensure that new chemicals can be utilized in manufacturing, while protecting human health, safety, and the environment without delaying process implementation. Chemicals in existing uses require reassessment when new chemical restrictions are identified. *Chemical Data Collection* Need to document and make available environment, safety, and health characteristics of chemicals. *Chemical Reduction* Need to develop processes that meet technology demands while reducing impact on human health, safety and the environment, both through replacement of hazardous materials with materials that are more benign, and by reducing chemical quantity requirements through more efficient and cost-effective process management. *Environment Management* Need to develop effective management systems to address issues related to disposal of equipment, and hazardous and non-hazardous residue from the manufacturing process.
Resource Conservation	*Natural Resource Conservation (Energy, Water)* Need to design more energy and water efficient processing equipment. *Chemicals and Materials Use* Need more efficient utilization of chemicals and materials. *Resource Recycling* Increase resource reuse and recycling.
Workplace Protection	*Equipment Safety* Need to design ergonomically correct and safe equipment. *Chemical Exposure Protection* Increase knowledge base on health and safety characteristics of chemicals and materials used in the manufacturing and maintenance processes, and of the process byproducts; and implement safeguards to protect the users of the equipment and facility.
Climate Change Mitigation	*Reduce Energy Use of Process Equipment* Need to design energy efficient larger wafer size processing equipment. *Reduce Energy Use of the Manufacturing Facility* Need to design energy efficient facilities to offset the increasing energy requirements of higher class clean rooms. *Reduce High Global Warming Potential (GWP) Chemicals Emission* Need ongoing improvement in methods that reduce emissions from processes using GWP chemicals.
ESH Design and Measurement Methods	*Evaluate and Quantify ESH Impact* Need integrated way to evaluate and quantify ESH impact of process, chemicals, and process equipment, and to make ESH a design parameter in development procedures for new equipment and processes.

YIELD ENHANCEMENT

Difficult Challenge ≥ 65 nm, Through 2007	*Summary of Issues*
Develop and Validate Systematic Yield Models—Process induced defects, equipment generated particles, product/processing measurements, and design/layout sensitivities have to be correlated to yield.	Correlate process-induced defects (PID), particles per wafer pass (PMP), product inspections, and *in situ* measurements. Develop parametric and process-to-design mismatch yield-loss models. Address sampling and statistical issues with ultra-small populations. Increase Yield Model accuracy.
High Aspect Ratio Inspection—High-speed cost-effective tools must be developed that rapidly detect defects associated with high-aspect-ratio contacts/vias/trenches, and particularly defects near/at the bottom of these features.	Poor transmission of energy into bottom of via and back out to detection system Large number of contacts and vias per wafer
Defect/Fault Sourcing for Rapid Yield Learning— Automated, intelligent analysis and reduction algorithms that correlate facility, design, process, test and WIP data must be developed to enable rapid root cause analysis of yield limiting conditions.	Circuit complexity grows exponentially and the ability to rapidly isolate failures on non-arrayed chips is needed. Automated data/image mining and reduction algorithms must be developed to source defects from multiple data sources (facility, design, process and test.)
Correlation of Impurity Level to Yield—Methodology for employment and correlation of fluid/gas types to yield of a standard test structure/product.	Establish an employment methodology for each material type. Define a standard test for yield/parametric effect.
Difficult Challenge < 65 nm, Beyond 2007	*Summary of Issues*
Develop Yield Models that Include New Materials and Integration—Models must comprehend greater parametric sensitivities, complex integration issues, ultra-thin film integrity, impact of circuit design, greater transistor packing, etc.	Develop test structures for new technology nodes. Address complex integration issues. Model ultra-thin film integrity issues. Improve scaling methods for front-end processes including increased transistor packing density.
D*efect Detection*—Detection and simultaneous differentiation of multiple killer defect types is necessary at high capture rates and throughput.	Existing techniques trade-off throughput for sensitivity, but at predicted defect levels, both throughput and sensitivity are necessary for statistical validity. Ability to detect particles at critical size may not exist.
Non-visual Defect Sourcing and Design for Manufacture and Test—Failure analysis tools and techniques are needed to enable localization of defects where no visual defect is detected. Also, IC designs must be optimized for a given process capability and must be testable/diagnosable.	Many defects that cause electrical faults are not detectable inline. Tools are needed that enable design to process matching for optimum yields. Also, testability/diagnose-ability must be designed into the IC for rapid electrical failure sourcing.
Precursors for New Materials—Required purity levels for delivered dielectric pre-cursors are not known or well understood.	Establish methodology for establishing purity standards for new dielectric pre-cursors.

METROLOGY

Five Difficult Challenges ≥65 nm, Through 2007	*Summary of Issues*
Factory level and company wide metrology integration for real time *in situ*, integrated, and inline metrology tools; continued development of robust sensors and process controllers; and data management that allows integration of add-on sensors.	Standards for process controllers and data management must be agreed upon. Conversion of massive quantities of raw data to information useful for enhancing the yield of a semiconductor manufacturing process. Better sensors must be developed for trench etch end point, ion species/energy/dosage (current), and wafer temperature during RTA.
Impurity detection (especially particles) at levels of interest for starting materials and reduced edge exclusion for metrology tools.	Existing capabilities will not meet Roadmap specifications. Very small particles must be detected and properly sized.
Control of high-aspect ratio technologies such as Damascene challenges all metrology methods. Key requirements are void detection in copper lines and pore size distribution in patterned low κ dielectrics.	New process control needs are not yet established. For example, 3-dimensional (CD and depth) measurements will be required for trench structures in new, low κ dielectrics.
Measurement of complex material stacks and interfacial properties including physical and electrical properties.	Reference materials and standard measurement methodology for new, high κ gate and capacitor dielectrics with interface layers, thin films such as interconnect barrier and low κ dielectric layers, and other process needs. Optical measurement of gate and capacitor dielectric averages over too large an area and needs to characterize interfacial layers. The same is true for measurement of barrier layers. High frequency dielectric constant measurements advances need to continue.
Measurement test structures and reference materials.	Scribe lines are shrinking and correlation to variation of chip properties is difficult. Overlay and other test structures are sensitive to process variation, and test structure design must be improved to ensure correlation between scribe line measurement and on chip properties. Standards institutions need rapid access to state of the art development and manufacturing capability to fabricate stable reference materials.
Five Difficult Challenges < 65 nm, Beyond 2007	
Nondestructive, production worthy wafer and mask level microscopy for critical dimension measurement for 3D structures, overlay, defect detection, and analysis	Surface charging and contamination interfere with electron beam imaging. CD measurements must account for side wall shape. CD for Damascene process may require measurement of trench structures. Process control such as focus exposure and etch bias will require greater precision and 3D capability.
Standard electrical test methods for reliability of new materials, such as ultra-thin gate and capacitor dielectric materials, are not available.	The wearout mechanism for new, high κ gate and capacitor dielectric materials is unknown.
Statistical limits of sub-65 nm process control	Controlling processes where the natural stochastic variation limits metrology will be difficult. Examples are low-dose implant, thin gate dielectrics, and edge roughness of very small structures.
3D dopant profiling	The dimensions of the active area approach the spacing between dopant atoms, complicating both process simulation and metrology. Elemental measurement of the dopant concentration at the requested spatial resolution is not possible.
Determination of manufacturing metrology when device and interconnect technology remain undefined.	The replacement devices for the transistor and structure and materials replacement for copper interconnect are being researched.

MODELING AND SIMULATION

Difficult Challenges ≥ 65 nm, Through 2007	*Summary Of Issues*
High Frequency Circuit Modeling (>5 GHz)	Efficient simulation of full-chip interconnect delay. Accurate 3D interconnect model; inductance, transmission line models. High frequency circuit models including non-quasi-static, substrate noise and coupling. Parameter extraction without RF measurements .
Modeling of Ultra Shallow Dopant Distributions, Junctions, and Silicidation	Dopant models and parameters (damage, high- concentration, activation, metastable effects, diffusion, interface and silicide effects). Characterization tools for these ultra shallow geometries and dopant levels.
Modeling Deposition and Etch Variations, Feature Variations across a Wafer	Fundamental physical data (e.g., rate constants, cross sections, surface chemistry); reduced models for complex chemistry. Linked equipment / feature models. CMP (full wafer and chip level, pattern dependent effects). Next generation equipment/wafer models.
Modeling of Lithography Technology	Predictive resist models. Resolution enhancement techniques; mask synthesis (OPC, PSM). 248 nm versus 193 nm versus 157 nm evaluation and tradeoffs. Next-generation lithography system models
Gate Stack Models for Ultra-Thin Dielectrics	Electrical and processing models for alternate gate dielectrics, and alternate gate materials (e.g., MeOx). Model dielectric constant, surface states, reliability, breakdown, and tunneling from process/material conditions.
Difficult Challenges < 65 nm, Beyond 2007	*Summary Of Issues*
Complementing Continuum Tools with Atomistic Ones.	A succession of modeling tools that marry atomistic effects with present day continuum software tools
Ultimate CMOS Simulation Capability	Methods and algorithms that will allow prediction of CMOS limits. Quantum based simulators.
Nano-Scale Device Modeling	New device concepts beyond traditional CMOS based on silicon technology such as vertical MOS, heterostructures, thin-film transistors, single electron transistors.
Thermo-Mechanical Modeling for Reliability	Stress voiding, electromigration, piezoelectric effects, fracture, and adhesion simulation tools
Software Module Integration	Seamless integration of simulation modules with focus on interplay and interfacing of the modules in order to enhance design effectiveness

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

YEAR OF PRODUCTION	1995	1997	1999	2001	2004	2007	2010	2013	2016
Actual Trend Numbers (nm)	**350**	**247.5**	**175**	**123.7**	**87.5**	**61.9**	**43.8**	**30.9**	**21.9**
ITRS Rounded Node Numbers (nm)	**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).

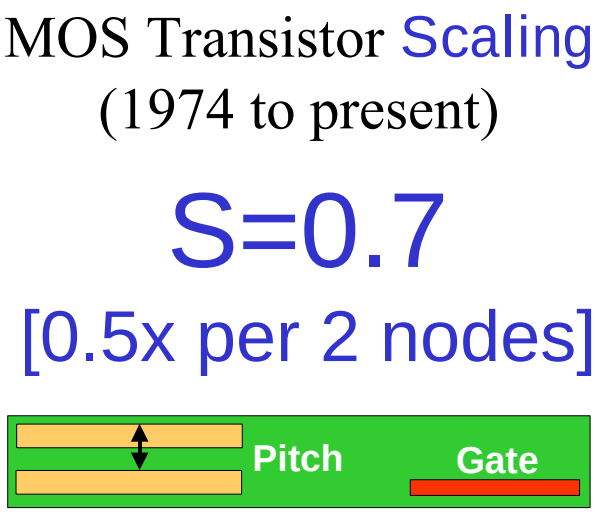


Figure 5 MOS Transistor Scaling—1974 to present

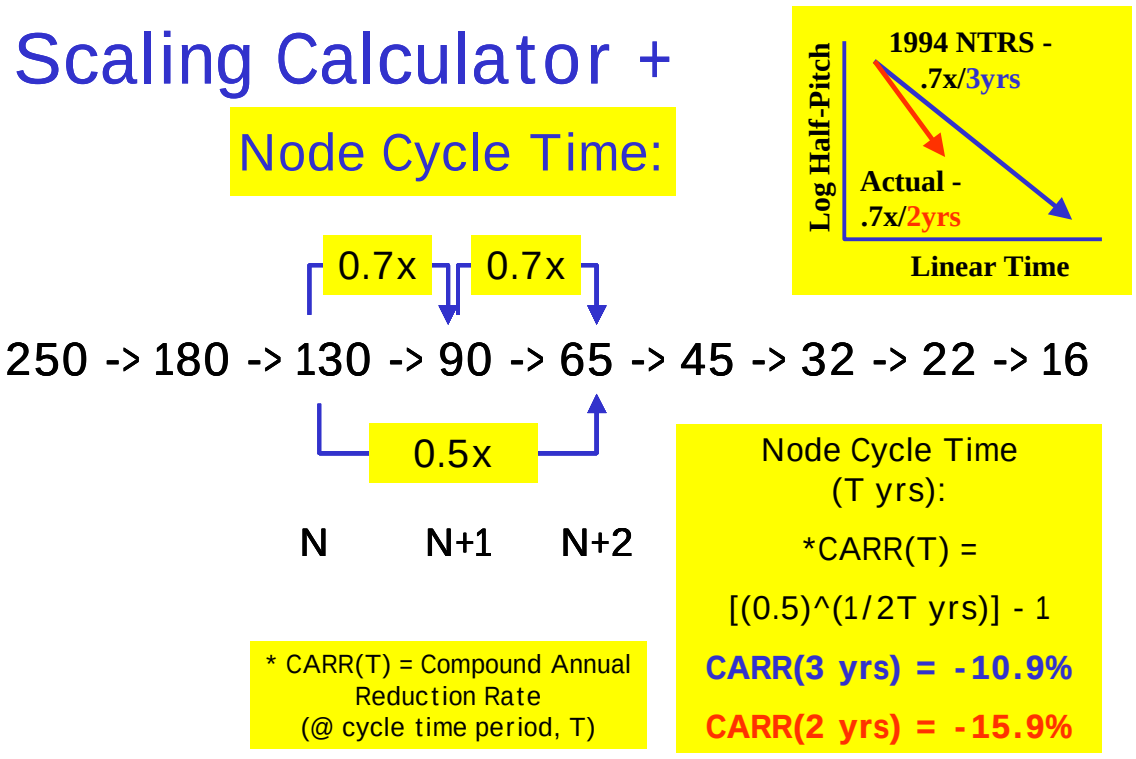


Figure 6 Scaling Calculator

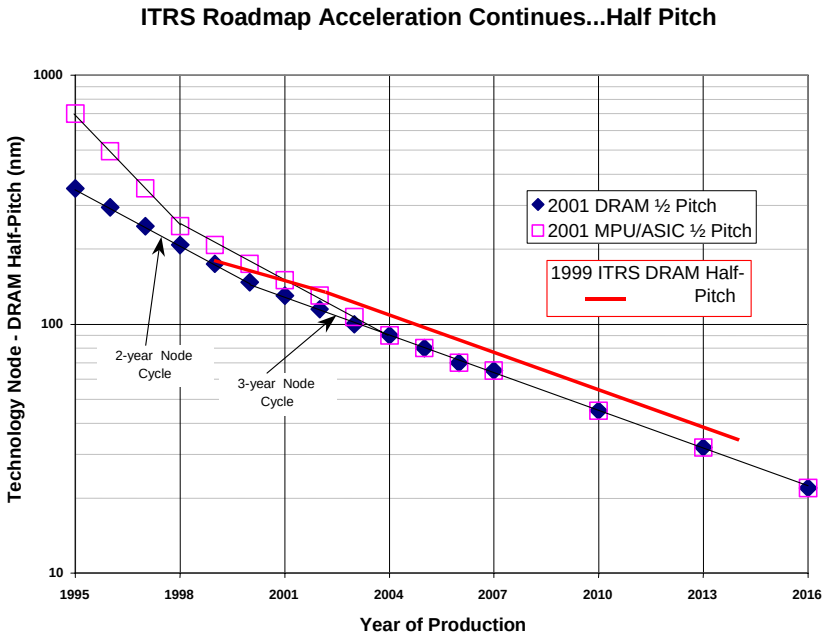


Figure 7 ITRS Roadmap Acceleration Continues—Half Pitch Trends

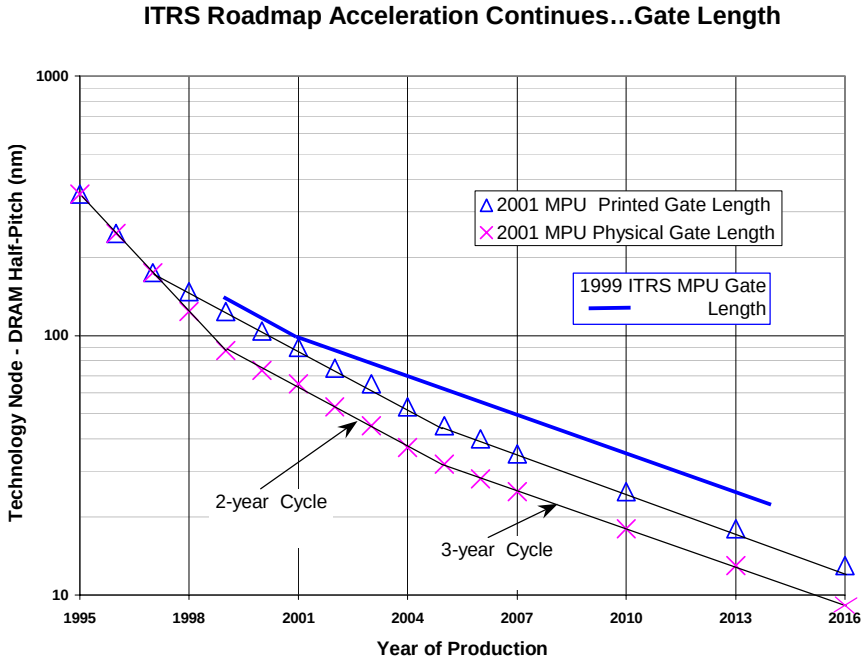


Figure 8 ITRS Roadmap Acceleration Continues—Gate Length Trends

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

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
ASIC/Low Power Printed Gate Length (nm) ††	130	107	90	75	65	53	45
ASIC/Low Power Physical Gate Length) (nm)	90	75	65	53	45	37	32

Table 1b Product Generations and Chip Size Model Technology Nodes—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
ASIC/Low Power Printed Gate Length (nm) ††	32	22	16
ASIC/Low Power Physical Gate Length) (nm)	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:

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 572mm^2 field at Introduction and two chips per 572mm^2 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 572mm^2 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:

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 572mm^2 field at Introduction and two chips per 572mm^2 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 572mm^2 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

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
SRAM Cell (6-transistor) Area factor ++	126.1	123.0	120.3	117.8	115.6	113.7	111.9
Logic Gate (4-transistor) Area factor ++	320.0	320.0	320.0	320.0	320.0	320.0	320.0
SRAM Cell (6-transistor) Area efficiency ++	0.63	0.63	0.63	0.63	0.63	0.63	0.63
Logic Gate (4-transistor) Area efficiency ++	0.50	0.50	0.50	0.50	0.50	0.50	0.50
SRAM Cell (6-transistor) Area w/overhead ++	3.3	2.5	2.0	1.5	1.2	0.93	0.73
Logic Gate (4-transistor) Area w/overhead ++	10.4	8.2	6.5	5.2	4.1	3.3	2.6
Transistor density SRAM (Mtransistors/cm^2)	184	237	305	393	504	646	827
Transistor density logic (Mtransistors/cm^2)	38.6	48.6	61.2	77.2	97.2	122.5	154.3
Generation at introduction *	p04c	—	—	p07c	—	—	p10c
Functions per chip at introduction (million transistors [Mtransistors])	193	243	307	386	487	614	773
Chip size at introduction (mm^2) ‡	280	280	280	280	280	280	280
Cost performance MPU (Mtransistors/cm^2 at introduction) (including on-chip SRAM) ‡	69	87	110	138	174	219	276
Generation at production *	p01c	—	—	p04c	—	—	p07c
Functions per chip at production (million transistors [Mtransistors])	97	122	153	193	243	307	386
Chip size at production (mm^2) §§	140	140	140	140	140	140	140
Cost performance MPU (Mtransistors/cm^2 at production, including on-chip SRAM) ‡	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

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
SRAM Cell (6-transistor) Area factor ++	107.8	106.7	105.7
Logic Gate (4-transistor) Area factor ++	320.0	320.0	320.0
SRAM Cell (6-transistor) Area efficiency ++	0.63	0.63	0.63
Logic Gate (4-transistor) Area efficiency ++	0.50	0.50	0.50
SRAM Cell (6-transistor) Area w/overhead ++	0.22	0.17	0.13
Logic Gate (4-transistor) Area w/overhead ++	0.82	0.65	0.51
Transistor density SRAM (Mtransistors/cm^2)	1718	3532	7208
Transistor density logic (Mtransistors/cm^2)	309	617	1235
Generation at introduction *	p13c	p16c	p19c
Functions per chip at introduction (million transistors [Mtransistors])	1546	3092	6184
Chip size at introduction (mm^2) ‡	280	280	280
Cost performance MPU (Mtransistors/cm^2 at introduction) (including on-chip SRAM) ‡	552	1104	2209
Generation at production *	p10c	p13c	p16c
Functions per chip at production (million transistors [Mtransistors])	773	1546	3092
Chip size at production (mm^2) §§	140	140	140
Cost performance MPU (Mtransistors/cm^2 at production, including on-chip SRAM) ‡	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× 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× 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 ($25\times32 = 800mm^2$) 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

affordable lithography field limit. This accomplished in the present DRAM model by slowing the short-term Moore's Law bits/chip rate from 2×/1.5 years to 2×/ two years. In the later years of the Roadmap the rate is stretched even more to 2×/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× 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

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
Number of Chip I/Os (Number of Total Chip Pads) — Maximum							
Total pads—MPU	3072	3072	3072	3072	3072	3072	3072
Signal I/O—MPU (1/3 of total pads)	1024	1024	1024	1024	1024	1024	1024
Power and ground pads—MPU (2/3 of total pads)	2048	2048	2048	2048	2048	2048	2048
Total pads—ASIC high-performance	3000	3200	3400	3600	4000	4200	4400
Signal I/O pads—ASIC high-performance	1500	1600	1700	1800	2000	2100	2200
Power and ground pads—ASIC high-performance (½ of total pads)	1500	1600	1700	1800	2000	2100	2200
Number of Total Package Pins—Maximum [1]							
Microprocessor/controller, cost-performance	480–1,200	480–1320	500–1452	500–1600	550–1760	550–1936	600–2140
Microprocessor/controller, high-performance	1200	1320	1452	1,600	1,760	1,936	2,140
ASIC (high-performance)	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

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
Number of Chip I/Os (Number of Total Chip Pads) — Maximum			
Total pads—MPU	3840	4224	4416
Signal I/O—MPU (1/3 of total pads)	1280	1408	1472
Power and ground pads—MPU (2/3 of total pads)	2560	2816	2944
Total pads—ASIC high-performance	4800	5400	6000
Signal I/O pads—ASIC high-performance	2400	2700	3000
Power and ground pads—ASIC high-performance (½ of total pads)	2400	2700	3000
Number of Total Package Pins—Maximum [1]			
Microprocessor/controller, cost-performance	780–2782	1014–3616	1318–4702
Microprocessor/controller, high-performance	2782	3616	4702
ASIC (high-performance)	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 cost/pin × 1.10 pins/year = 1.05 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× functionality/chip at flat price every two years = 29%/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 (κ~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

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)							
Pad pitch—ball bond	45	35	30	25	20	20	20
Pad pitch—wedge bond	40	35	30	25	20	20	20
Pad Pitch—area array flip-chip (cost-performance, high-performance)	160	160	150	150	130	130	120
Pad Pitch—peripheral flip-chip (handheld, low-cost, harsh)	150	130	120	110	100	90	80
Cost-Per-Pin							
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
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

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
Pad Pitch—area array (cost-performance, high-performance)	90	80	70
Pad Pitch—peripheral flip-chip (handheld, low-cost, harsh)	60	45	30
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

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
Maximum number wiring levels—maximum	7	8	8	8	9	9	9
Maximum number wiring levels—minimum	7	7	8	8	8	9	9

Table 4d Performance and Package Chips: Frequency, On-Chip Wiring Levels—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
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
Maximum number wiring levels—maximum	10	10	10
Maximum number wiring levels—minimum	9	9	10

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.

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

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

Table 5b Electrical Defects—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
DRAM Random Defect D_0 at production chip size and 89.5% yield (faults/m^2) §	1356	1356	1356
MPU Random Defect D_0 at production chip size and 83% yield (faults/m^2) §§	1464	1116	1134
# Mask Levels – MPU	27	29	29
# Mask Levels – DRAM	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 572mm^2 field at Introduction and two chips per 572mm^2 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 572mm^2 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

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
Power Supply Voltage (V)							
V_{dd} (high performance)	1.1	1.0	1.0	1	0.9	0.9	0.7
V_{dd} (Low Operating Power, high V_{dd} transistors)	1.2	1.2	1.1	1.1	1.0	1.0	0.9
V_{dd} (Low Standby Power, high V_{dd} transistors)	1.2	1.2	1.2	1.2	1.2	1.2	1.1
Allowable Maximum Power [1]							
High-performance with heatsink (W)	130	140	150	160	170	180	190
Cost-performance (W)	61	75	81	85	92	98	104
Battery (W)—(hand-held)	2.4	2.6	2.8	3.2	3.2	3.5	3.5

Table 6b Power Supply and Power Dissipation—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
Power Supply Voltage (V)			
V_{dd} (high performance)	0.6	0.5	0.4
V_{dd} (Low Operating Power, high V_{dd} transistors)	0.8	0.7	0.6
V_{dd} (Low Standby Power, high V_{dd} transistors)	1.0	0.9	0.9
Allowable Maximum Power [1]			
High-performance with heatsink (W)	218	251	288
Cost-performance (W)	120	138	158
Battery (W)—(hand-held)	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× 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.[1] A corresponding analysis conducted from published data for microprocessors yields similar results.[2] 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× linear, 0.5×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.

---

[1] *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.*

[2] *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 572mm^2 field at Introduction and two chips per 572mm^2 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 572mm^2 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

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**
Affordable Cost per Function ++			
DRAM cost/bit (packaged microcents) at samples/introduction	**0.93**	**0.33**	**0.12**
DRAM cost/bit (packaged microcents) at production §	**0.34**	**0.12**	**0.042**
Cost-performance MPU (microcents/transistor) (including on-chip SRAM) at introduction §§	**7.78**	**2.75**	**0.97**
Cost-performance MPU (microcents/transistor) (including on-chip SRAM) at production §§	**4.71**	**1.66**	**0.59**
High-performance MPU (microcents/transistor) (including on-chip SRAM) at production §§	**4.31**	**1.52**	**0.54**
Cost-Per-Pin			
Test Cost			
Volume tester cost per high-frequency signal pin ($K/pin) (high-performance ASIC)—maximum	**4**	**4**	**4**
Volume tester cost per high-frequency signal pin ($K/pin) (high-performance ASIC)—minimum	**2**	**3**	**4**

GLOSSARY

KEY ROADMAP TECHNOLOGY CHARACTERISTICS TERMINOLOGY (WITH OBSERVATIONS AND ANALYSIS)

CHARACTERISTICS OF MAJOR MARKETS

Technology Node—The ground rules of process governed by the smallest feature printed. The half-pitch of first-level interconnect dense lines is most representative of the DRAM technology level required for the smallest economical chip size. For logic, such as microprocessors (MPUs), physical bottom gate length is most representative of the leading-edge technology level required for maximum performance. MPU and ASIC logic interconnect half-pitch processing requirement typically refers to the first polysilicon or metal layer and lags behind DRAM half-pitch, which may also refer either first layer metal or polysilicon. The smallest half-pitch is typically found in the memory cell area of the chip. Each technology node step represents the creation of significant technology progress—approximately 70% of the preceding node, 50% of two preceding nodes. Example: DRAM half pitches of 180, 130, 90, 65, 45, 32 nm, and 22 nm. For cost reasons, high-volume, low-cost ASIC gate-length requirements will typically match DRAM half-pitch targets, but the low-volume leading-edge high-performance ASIC gate-length requirements will track closely with MPUs.

Moore's Law—An historical observation by Intel executive, Gordon Moore, that the market demand (and semiconductor industry response) for functionality per chip (bits, transistors) doubles every 1.5 to 2 years. He also observed that MPU performance [clock frequency (MHz) × instructions per clock = millions of instructions per second (MIPS)] also doubles every 1.5 to 2 years. Although viewed by some as a "self-fulfilling" prophecy, "Moore's Law" has been a consistent macro trend and key indicator of successful leading-edge semiconductor products and companies for the past 30 years.

Cost-per-Function Manufacturing Productivity Improvement Driver—In addition to Moore's Law, there is a historically-based "corollary" to the "law," which suggests that to be competitive manufacturing productivity improvements must also enable the cost-per-function (microcents per bit or transistor) to decrease by -29% per year. Historically, when functionality doubled every 1.5 years, then cost-per-chip (packaged unit) could double every six years and still meet the cost-per-function reduction requirement. If functionality doubles only every two years, as suggested by consensus DRAM and MPU models of the *1999 ITRS*, then the manufacturing cost per chip (packaged unit) must remain flat.

Affordable Packaged Unit Cost/Function—Final cost in microcents of the cost of a tested and packaged chip divided by Functions/Chip. Affordable costs are calculated from historical trends of affordable average selling prices [gross annual revenues of a specific product generation divided by the annual unit shipments] less an estimated gross profit margin of approximately 35% for DRAMs and 60% for MPUs. The affordability per function is a guideline of future market "tops-down" needs, and as such, was generated independently from the chip size and function density. Affordability requirements are expected to be achieved through combinations of—1) increased density and smaller chip sizes from technology and design improvements; 2) increasing wafer diameters; 3) decreasing equipment cost-of-ownership; 4) increasing equipment overall equipment effectiveness; 5) reduced package and test costs; 6) improved design tool productivity; and 7) enhanced product architecture and integration.

DRAM Generation at (product generation life-cycle level)—The anticipated bits/chip of the DRAM product generation introduced in a given year, manufacturing technology capability, and life-cycle maturity (Demonstration-level, Introduction-level, Production-level, Ramp-level, Peak).

MPU Generation at (product generation life-cycle level)—The generic processor generation identifier for the anticipated Microprocessor Unit (MPU) product generation functionality (logic plus SRAM transistors per chip) introduced in a given year, manufacturing technology capability, and life-cycle maturity (Introduction-level, Production-level, Ramp-level, Peak).

Cost-Performance MPU—MPU product optimized for maximum performance and the lowest cost by limiting the amount of on-chip SRAM level-two (L2) cache (example 1Mbytes/2001). Logic functionality and L2 cache typically double every two-year generation.

High-performance MPU—MPU product optimized for maximum system performance by combining a single or multiple CPU cores (example 2@ 25Mt cores in 2001) with a large (example 4Mbyte/2001) level-two (L2) SRAM. Logic functionality and L2 cache typically double every two-year generation by doubling the number of on-chip CPU cores and associated memory.

Product inTER-generation—Product generation-to-generation targets for periodically doubling the on-chip functionality at an affordable chip size. The targets are set to maintain Moore's Law (2×/two years) while preserving economical manufacturability (flat chip size and constant manufacturing cost per unit). This doubling every two years at a constant cost assures that the cost/function reduction rate (inverse productivity improvement) is -29% per year (the target historical rate of reduction). In order to double the on-chip functionality every two years, when technology-node scaling (.7× linear, .5× area) is every three years, an additional device/process design improvement of .8× per two years must be achieved. This requirement represents a design-related (cell-area-factor) area-reduction improvement of at least -11% per year, , and this design-related productivity improvement is in addition to the basic lithography-based area reduction of -21% per year (three-year node cycle). The present *2001 ITRS* consensus target for the rate of increase of DRAM is 2×/chip every two years. However, the *2001 ITRS* forecast of cell-area-factor improvement is only -7% per year on average. This results in an average DRAM inTER-generation chip-size growth of 4.5%/year or about 1.2× every four years. Presently, the MPU transistor area is shrinking only at lithography-based rate (virtually no design-related improvement). Therefore, the *2001 ITRS* MPU inTER-generation functionality model target is 2× transistors/chip every technology node, in order maintain a flat chip size growth throughout the roadmap period.

Product inTRA-generation—Chip size shrink trend within a given constant functions-per-chip product generation. The *2001 ITRS* consensus-based model targets reduce chip size (by shrinks and "cut-downs") utilizing the latest available manufacturing and design technology at every point through the roadmap. The *ITRS* targets for both DRAM and MPU reduce chip size within a generation by minus 50% per technology node.

Year of Demonstration—Year in which the leading chip manufacturer supplies an operational sample of a product as a demonstration of design and/or technology node processing feasibility and prowess. A typical venue for the demonstration is a major semiconductor industry conference, such as the International Solid State Circuits Conference (ISSCC) held by the Institute of Electrical and Electronic Engineers (IEEE). Demonstration samples are typically manufactured with early development or demonstration- level manufacturing tools and processes. Historically, DRAM products have been demonstrated at 4× bits-per-chip every three years at the leading-edge process technology node, typically two–three years in advance of actual market introduction. DRAM demonstration chip sizes have doubled every six years, requiring an increasing number of shrinks and delay before market introduction is economically feasible. Frequently, chip sizes are larger than the field sizes available from lithography equipment, and must be "stitched" together via multiple-exposure techniques that are feasible only for very small quantities of laboratory samples. Example: 1997/ISSCC/1Gb DRAM, versus *ITRS* 1Gb 1999 Introduction-level, 2003 Production-level targets.

Year of INTRODUCTION—Year in which the leading chip manufacturer supplies small quantities of engineering samples (<1K). These are provided to key customers for early evaluation, and are manufactured with qualified production tooling and processes. To balance market timeliness and economical manufacturing, products will be introduced at 2× functionality per chip every two years (every technology node, in the case of MPUs). In addition, manufacturers will delay production until a chip-size shrink or "cut-down" level is achieved which limits the inTER-generation chip-size growth to be flat, or at the most, 1.2× every four years.

Year Of PRODUCTION—Year in which leading chip manufacturers begin shipping volume quantities (10K/month) of product manufactured with qualified production tooling and processes and is followed within three months by a second manufacturer. As demand increases for the leading-edge performance and shrink products, the tooling and processes are being quickly "copied" into multiple modules of manufacturing capacity. For high-demand products, volume production typically continues to ramp to fab design capacity within 12 months. Alpha-level manufacturing tools and research technology papers are typically delivered 24 months prior to volume production ramp. Beta-level tools are typically delivered 12 months prior to ramp, along with papers at industry conferences. The beta-level tools are made production-level in pilot-line fabs, which may also run low volumes of product that is often used for customer sampling and early qualification prior to volume production ramp. Medium-volume production-level DRAMs will be in production concurrently with low-volume introduction-level DRAMs, and also concurrently with very-high-volume, shrunken, previous-generation DRAMs (example: 2003: 1Gb/production, 4G/introduction, plus 512Mb/256Mb/128Mb/64Mb high-volume). Similarly, high-volume cost-performance MPUs are in production concurrently with their lower-volume, large-chip, high-performance MPU counterparts, and also with very-high volume shrinks of previous generations.

Functions/Chip—The number of bits (DRAMs) or logic transistors (MPUs/ASICs) that can be cost-effectively manufactured on a single monolithic chip at the available technology level. Logic functionality (transistors per chip) include both SRAM and gate-function logic transistors. DRAM functionality (bits per chip) is based only on the bits (after repair) on a single monolithic chip.

Chip Size (mm^2)—The typical area of the monolithic memory and logic chip that can be affordably manufactured in a given year based upon the best available leading-edge design and manufacturing process. (Estimates are projected based upon historical data trends and the *ITRS* consensus models).

Functions/cm^2—The density of functions in a given square centimeter = Functions/Chip on a single monolithic chip divided by the Chip Size. This is an average of the density of all of the functionality on the chip, including pad area and wafer scribe area. In the case of DRAM, it includes the average of the high-density cell array and the less-dense peripheral drive circuitry. In the case of the MPU products, it includes the average of the high-density SRAM and the less-dense random logic. In the case of ASIC, it will include high-density embedded memory arrays, averaged with less dense array logic gates and functional cores. In the *2001 ITRS*, the typical high-performance ASIC design is assumed to have the same average density as the high-performance MPUs, which are mostly SRAM transistors.

DRAM Cell Array Area Percentage—The maximum practical percentage of the total DRAM chip area that the cell array can occupy at the various stages of the generation life cycle. At the introduction chip size targets, this percentage must be typically less than 70% to allow space for the peripheral circuitry, pads, and wafer scribe area. Since the pads and scribe area do not scale with lithography, the maximum cell array percentage is reduced in other inTRA-generation shrink levels (typically less than 55% at the production level, and less than 50% at the ramp level).

DRAM Cell Area (μm^2)—The area (C) occupied by the DRAM memory bit cell, expressed as multiplication of a specified *ITRS*-consensus Cell Area Factor target (A) times the square of the minimum half-pitch feature (f) size, that is: $C = Af^2$. To calculate the chip size, the cell area must be divided by the array efficiency, a factor (E) that is statistically derived from historical DRAM chip analysis data. Thus an average cell area (C_{AVE}) can be calculated , which is burdened by the overhead of the drivers, I/O, bus lines, and pad area. The formula is: $C_{AVE} = C/E$. The total chip area can then be calculated by multiplying the total number of bits/chip times the C_{AVE} . Example: 1999: A=8; square of the half-pitch, $f^2 = (180\ nm)^2 = .032\ \mu m^2$; cell area, $C=Af^2=0.26\ \mu m^2$; for 1Gb introduction-level DRAM with a cell efficiency of E=70% of total chip area, the $C_{AVE} = C/E = 0.37\ \mu m^2$; therefore, the 1Gb Chip Size Area=2^{30} bits * 0.37e-6 mm^2/bit = 397 mm^2.

DRAM Cell Area Factor—A number (A) which expresses the DRAM cell area (C) as a multiple of equivalent square half-pitch (f) units. Typically, the cell factor is expressed by equivalent aspect ratios of the half-pitch units (2×4=8, 2×3=6, 2×2=4, 1.6×1.6=2.5, etc.).

SRAM Cell Area Factor—Similar to the DRAM area factor, only applied to a 6-transistor (6t) logic-technology latch-type memory cell. The number expresses the SRAM 6t cell area as a multiple of equivalent square technology-node half-pitch (f) units. Typically, the cell factor of the SRAM 6t cell is 16–25 times greater than a DRAM memory cell area factor.

Logic Gate Cell Area Factor—Similar to the DRAM and SRAM cell area factors, only applied to a typical 4-transistor (4t) logic gate. The number expresses the logic 4t gate area as a multiple of equivalent square technology-node half-pitch (f) units. Typically, the cell factor of the logic 4t gate is 2.5–3 times greater than an SRAM 6t cell area factor, and 40–80 times greater than a DRAM memory cell area factor.

Usable Transistors/cm^2 (High-performance ASIC, Auto Layout)—Number of transistors per cm^2 designed by automated layout tools for highly differentiated applications produced in low volumes. High-performance, leading-edge, embedded-array ASICs include both on-chip array logic cells, as well as dense functional cells (MPU, I/O, SRAM, etc). Density calculations include the connected (useable) transistors of the array logic cells, in addition to all of the transistors in the dense functional cells. The largest high-performance ASIC designs will fill the available production lithography field.

CHIP AND PACKAGE—PHYSICAL AND ELECTRICAL ATTRIBUTES

Number of Chip I/Os–Total (Array) Pads—The maximum number of chip signal I/O pads plus power and ground pads permanently connected to package plane for functional or test purposes, or to provide power/ground contacts (including signal conditioning). These include any direct chip-to-chip interconnections or direct chip attach connections to the board (Package plane is defined as any interconnect plane, leadframe, or other wiring technology inside a package, i.e., any wiring that is not on the chip or on the board.). MPUs typically have a ratio of signal I/O pads to power/ground pads of 1:2, whereas the high-performance ASIC ratio is typically 1:1.

Number of Chip I/Os–Total (Peripheral) Pads—The maximum number of chip signal I/O plus power and ground pads for products with contacts only around the edge of a chip.

Pad Pitch—The distance, center-to-center, between pads, whether on the peripheral edge of a chip, or in an array of pads across the chip.

Number of Package Pins/Balls—The number of pins or solder balls presented by the package for connection to the board (may be fewer than the number of chip-to-package pads because of internal power and ground planes on the package plane or multiple chips per package).

Package Cost (Cost-performance)—Cost of package envelope and external I/O connections (pins/balls) in cents/pin.

CHIP FREQUENCY (MHZ)

On-Chip, Local Clock, High-Performance—On-chip clock frequency of high-performance, lower volume microprocessors in localized portions of the chip.

Chip-To-Board (Off-chip) Speed (High-Performance, Peripheral Buses)—Maximum signal I/O frequency to board peripheral buses of high and low volume logic devices.

OTHER ATTRIBUTES

Lithographic Field Size (mm^2)—Maximum single step or step-and-scan exposure area of a lithographic tool at the given technology node. The specification represents the minimum specification that a semiconductor manufacturer might specify for a given technology node. The maximum field size may be specified higher than the ORTC target values, and the final exposure area may be achieved by various combinations of exposure width and scan length.

Maximum Number of Wiring Levels—On-chip interconnect levels including local interconnect, local and global routing, power and ground connections, and clock distribution.

FABRICATION ATTRIBUTES AND METHODS

Electrical D_0 Defect Density (d/m^{-2})—Number of electrically significant defects per square meter at the given technology node, production life-cycle year, and target probe yield.

Minimum Mask Count—Number of masking levels for mature production process flow with maximum wiring level (Logic).

MAXIMUM SUBSTRATE DIAMETER (MM)

Bulk or Epitaxial or Silicon-on-Insulator Wafer—Silicon wafer diameter used in volume quantities by mainstream IC suppliers. The *ITRS* timing targets, contributed by the Factory Integration ITWG, are based on the first 20K wafer-starts-per-month manufacturing facility.

ELECTRICAL DESIGN AND TEST METRICS

POWER SUPPLY VOLTAGE (V)

Minimum Logic V_{dd}—Nominal operating voltage of chips from power source for operation at design requirements.

Maximum Power High-performance with Heat Sink (W)—Maximum total power dissipated in high-performance chips with an external heat sink.

Battery (W)—Maximum total power/chip dissipated in battery operated chips.

DESIGN AND TEST

Volume Tester Cost/Pin ($K/pin)—Cost of functional (chip sort) test in high volume applications divided by number of package pins.