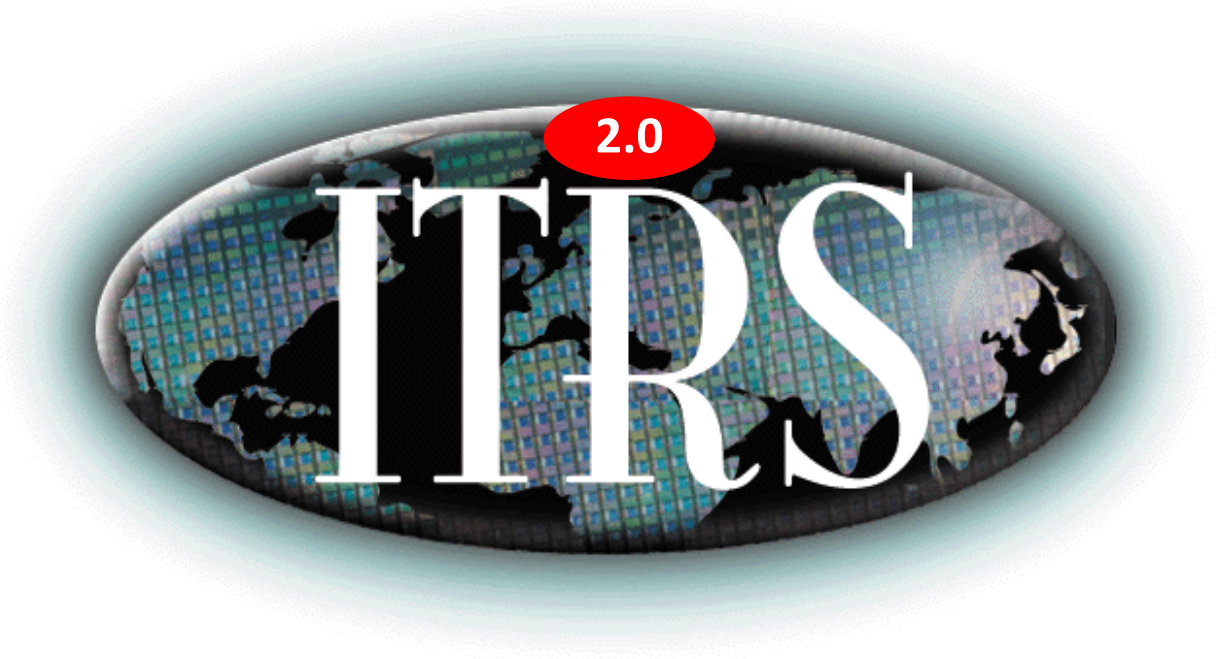


INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS 2.0

2015 EDITION

EXECUTIVE REPORT

THE ITRS 2.0 IS DEVISED AND INTENDED FOR TECHNOLOGY ASSESSMENT ONLY AND IS WITHOUT REGARD TO ANY COMMERCIAL CONSIDERATIONS PERTAINING TO INDIVIDUAL PRODUCTS OR EQUIPMENT.

The ITRS is Jointly Sponsored
by

European Semiconductor Industry Association

Japan Electronics and Information Technology Industries Association

Korea Semiconductor Industry Association

Taiwan Semiconductor Industry Association

Semiconductor Industry Association

2015 ITRS 2.0 ACKNOWLEDGMENTS

INTERNATIONAL ROADMAP COMMITTEE

◈*Europe*—Mart Graef, Bert Huizing, Reinhard Mahnkopf
◈*Japan*— Hidemi Ishiuchi, Yoshihiro Hayashi, Nobuyuki Ikumi, Hiroyuki Miyakawa
◈*Korea*—Siyoung Choi, Jae Hoon Choi,
◈*Taiwan*—Sam Pam, Wilman Tsai
◈*U.S.A.*—Paolo Gargini, Taffy Kingscott, Linda Wilson

TECHNOLOGY WORKING GROUP KEY CONTRIBUTORS

2015 Cross TWG Study Group (Technology Pacing)—Alan Allan, Dave Armstrong, An Chen, Mustafa Badaroglu, Joel Barnett, Roger Barth, Herbert Bennett, Bill Bottoms, Juan-antonio Carballo, Carlos Diaz, Alain Diebold, Paul Feeney, Mike Gaitan, Paolo Gargini, Mike Garner, Hidemi Ishiuchi, Dan Herr, Hirofumi Inoue, Scott Jones, Andrew Kahng, Leo Kenny, Rich Liu, Jürgen Lorenz, Steve Moffat, James Moyne, Mark Neisser, Kwok Ng, George Orji, Lothar Pfitzner, Gopal Rao, Thomas Skotnicki, Hitoshi Wakabayashi, Mike Walden, Linda Wilson, Osamu [Sam] Yamazaki, Victor Zhirnov, Paul Zimmerman

2015 Cross TWG Study Group (More than Moore)—Herbert Bennett, Bill Bottoms, Michel Brillouët, Juan-Antonio Carballo, Patrick Cogez, Erik DeBenedictis, Michael Gaitan, Mart Graef, Bert Huizing, Andrew Kahng, Reinhard Mahnkopf, Grace O' Malley, Chuck Richardson

System Integration—Juan-Antonio Carballo, Wei-Ting Jonas Chan, Andrew B. Kahng, Masaru Kakimoto, Siddhartha Nath, Toshitada Saito, Kenshu Seto, Gary Smith, Ichiro Yamamoto

Test and Test Equipment—Kenichi Anzou, Dave Armstrong, John Aslanian, Roger Barth, Yi Cai, Krishnendu Chakrabarty, Tapan Chakraborty, Sreejit Chakravarty, Wendy Chen, William Chui, Steve Comen, Zoe Conroy, Adam Cron, Al Crouch, Ted Eaton, Stefan Eichenberger, Bill Eklow, Paul Emmett, Ira Feldman, Francois-Fabien Ferhani, Shawn Fetterolf, Paul Franzon, Piergiorgio Galletta, Anne Gattiker, Sandeep Goel, Kazumi Hatayama, Hirokazu Hirayama, Hisao Horibe, Shuichi Ito, Hongshin Jun, Masahiro Kanase, Rohit Kapur, Toshiaki Kato, Brion Keller, Ajay Khoche, Satoru Kitagawa, Marc Knox, Masashi Kondo, Ken Lanier, Lenny Leon, Marc Loranger, Anthony Lum, Erik Jan Marinissen, Peter Maxwell, Cedric Mayor, Teresa McLaurin, Milanjan Mukherjee, Takeshi Nagasaka, Masaaki Namba, Naoaki Narumi, Phil Nigh, Akitoshi Nishimura, Hermann Obermeir, Jayson Park, Mike Peng-Li, Frank Poehl, Chris Portelli-Hale, Bill Price, Jeff Rearick, Herb Reiter, Mike Ricchetti, Michael Rodgers, Mark Roos, Tomonori Sasaki, Keno Sato, Masayuki Sato, Yasuo Sato, Ryuji Shimizu, Yoichi Shimizu, Fumio Sonoda, Hiroyoshi Suzuki, Tetsuo Tada, Satoru Takeda, Steven Tilden, Erik Volkerink, Adam Wright, Yervant Zorian

Process Integration, Devices, and Structures—Yasushi Akasaka, Hiroaki Ammo, Dimitri Antoniadis, Mustafa Badaroglu, Gennadi Bersuker, Azeez Bhavnagarwala, Frederic Boeuf, Alex Burenkov, Chorng-Ping Chang, Kin (Charles) Cheung, Kristin DeMeyer, James Fonseca, Yuzo Fukuzaki, Bill Gallagher, Christopher Henderson, Herve Jaouen, Gaspard Hiblot, Toshiro Hiramoto, Digh Hisamoto, Jim Hutchby, Jiro Ida, Hirofumi Inoue, Kunihiko Iwamoro, Moon-Young Jeong, Malgorzata Jurczak, Naoki Kasai, SungGeun Kim, Gerhard Klimeck, Fred Kuper, Hajime Kurata, Chung Lam, Robert Lander, Rich Liu, Witek Maszara, Toru Mogami, Kwok Ng, Tak Ning, Masaaki Niwa, Tony Oates, Hidekazu Oda, Sang Hyun Oh, Tatsuya Ohguro, Sam Pan, Jongwoo Park, Thierry Poiroux, Siddharth Potbhare, Kirk Prall, Xiang Qi, Mehdi Salmani-Jelodar, Thomas Schulz, Saurabh Sinha, Thomas Skotnicki, James Stathis, Toshihiro Sugii, Shinichi Takagi, Akira Tanabe, Tetsu Tanaka, Cheng-Tzung Tsai, Wilman Tsai, Hitoshi Wakabayashi, Philip Wong, Yanzhong Xu, Geoffrey Yeap, Makoto Yoshimi, Scott Yu, Peter Zeitzoff

Outside System Connectivity—RF & AMS- Hiroaki Ammo, James Bateman, Herbert Bennett, Pascal Chevalier, Hassan Farooq, David Harame, Anthony Immorlica, Jay John, James Li, Peter Magnee, Tatsuya Ooguro, Mark Rocchi, Michael Schroter, Sorin Voinigescu, *Photonic Interconnects*- Hiroaki Ammo, Carlos Augusto, Keren Bergman, Frederic Boeuf, Bill Bottoms, Mark Brongersma, William Chen, Patrick Cogez, Jason Coder, Jennifer Dionne, Shanhui Fan, Michael Gaitan, C. Michael Garner, Dan Gorcea, Andrew B. Kahng, Ashok Kapoor, Lionel Kimerling, Ashok Krishnamoorthy, David A.B. Miller III, Dick Otte, Chuck Richardson, Bernard Sautreuil, Jelena Vuckovic, Jean-Baptiste David, Alexei Tchelnokov

Heterogeneous Components—Steve Bart, Raji Baskaran, Steve Briet, Asif Chowdhury, Alissa Fitzgerald, Koji Fukumoto, Michael Gaitan, Buzz Hardy, Brian Jamieson, Akihiro Koga, Ron Lawes, Karen Lightman, Pete Loeppert, Erik Jan Lous, Mary Ann Maher, Jianmin Miao, Takashi Mihara, Arthur Morris, Yasutaka Nakashiba., Goro Nakatani, Fumihiko Nakazawa, Rob O'Reilly, Mervi Paulasto-Kröckel, Sasha Revel, Philippe Robert, John Rychcik, Tetsu Tanaka, Steve Tilden, Hiroshi Toshiyoshi, Joost van Beek, Chris van Hoof, Randy Wagner, Marcie Weinstein

Emerging Research Devices—Hiro Akinaga, Fabien Alibart, Joerg Appenzeller, Tetsuya Asai, Yuji Awano, Manuel Bibes, George Bourianoff, Bill Bottoms, Rod Bowman, Michel Brillouet, Geoffrey Burr, John Carruthers, Ralph Cavin, Chorn-Ping Chang, An Chen, Christophe Chevallier, U-In Chung, Byung Jin Cho, Sung Woong Chung, Luigi Colombo, Shamik Das, Erik DeBenedictis, Simon Deleonibus, Tetsuo Endoh, Paul Franzon, Akira Fujiwara, Mike Garner, Chakku Goplan, Bogdan Govoreanu, Wilfried Haensch, Dan Hammerstrom, Tsuyoshi Hasegawa, Shigenori Hayashi, Dan Herr, Marc Heyns, Mutsuo Hidaka, Toshiro Hiramoto, Jim Hutchby, Adrian Ionescu, Kiyoshi Kawabata, Seiichiro Kawamura, Masashi Kawasaki, Alex Khitun, Antoine Khoueir, Suhwan Kim, Hyoungjoon Kim, Young-Bae Kim, Atsuhiro Kinoshita, Dae-Hong Ko, Gwan-hyeob Koh, Hiroshi Kotaki, Michael Kozicki, Ilya Krivorotov, Franz Kreupl, Mark Kryder, Zoran Krivokapic, Kee-Won Kwon, Jangeun Lee, Jung-Ho Lee, Chenhsin Lien, Yun Fook Liew, Tsu-Jae King Liu, Louise Lome, Wei Lu, T.P. Ma, Reinhard Mahnkopf, Matthew Marinella, Stefan Meskers, Rene Meyer, Yasuyuki Miyamoto, Hiroshi Mizuta, Kotaki Murali, Kwok Ng, Michael Niemier, Fumiyuki Nihei, Dmitri Nikonov, Kei Noda, Byung-Gook Park, Ferdinand Peper, Er-Xuan Ping, Yaw Obeng, Yutaka Ohno, Frank Register, Heike Riel, Shriram Ramanathan, Dave Roberts, Kaushik Roy, Sayeef Salahuddin, Yoshiaki Saito, Akihito Sawa, Shintaro Sato, Barry Schechtman, Frank Schwierz, Sadas Shankar, Takahiro Shinada, Masayuki Shirane, Kaushal Singh, Thomas Skotnicki, Satoshi Sugahara, Shin-ichi Takagi, Tsutomu Teduka, Ming-Jinn Tsai, Ken Uchida, Thomas Vogelsang, Yasuo Wada, Rainer Waser, Jeff Welser, Frans Widershoven, Philip Wong, Dirk Wouters, Kojiro Yagami, David Yeh, In-Seok Yeo, Hiroaki Yoda, In-K Yoo, Victor Zhirnov

Emerging Research Materials—Hiro Akinaga, Scott Anderson, Nobuo Aoi, Bernd Appelt, Koji Arita , Hidefumi Asano, Koyu Asai, Yuji Awano, Rama Ayothi, Agnès Barthélémy, Joel Barnett, Chris Bencher, Daniel-Camille Bensahel, Kris Bertness, Stacey Bent, Manuel Bibes, Idriss Blakey, Mikael Björk, Serge Blonkowski, Kirill Bolotin, Lars Bomholt, August Bosse, Bill Bottoms, Rod Bowman, Jennifer Braggin, Alex Bratkovski, Robert Bristol, Ahmed Busnaina, David. A. Cahill, Jeff Calvert, , Bernie Capraro, Gustau Catalan, Ian Cayrefourcq, Alek Chen, An Chen, Xiangyu Chen, Zhihong Chen, Joy, Cheng, Toyohiro Chikyow, Byung Jin Cho, Luigi Colombo, Michael Current, Ralph Dammel, Shamik Das, Juan DePablo, Anton Devilliers, Geraud Dubois, Catherine Dubourdieu, Glenn Fredrickson, Nathan Fritz, Roman Gafiteanu, Martin Giles, C. Michael Garner, Emmanuel Giannelis, Valeriy Ginzburg, Michael Goldstein, Padma Gopolan, Andrew Grenville, Roel Gronheid, Wilfried Haensch, Cliff Henderson, Daniel Herr, Hiro Hibino, Marc Hillmyer, Toshiro Hiramoto, ChiaHua Ho, Sung Eun Hong, Phil Hustad, Jim Hutchby, Harold Hwang, Diane Hymes, Koji Inoue, Kohei Ito, Hatsuo Ishida, Taisuke Iwai, Ajey Jacob, David Jamieson, Ali Javey, Berry Jonker, Yeon Sik Jung, Andrew Kahng, Ted Kamins, Masashi Kawaski, Leo Kenny, Choong-Un Kim, Philip Kim, Sang Ouk Kim, Joseph Kopanski, Michael Kozicki, Mark Kryder, Sean King, Atsuhiro Kinoshita, Paul Kohl, Blanka Magyari-Kope, Victor Krivokapic, Mark Kryder, Yi-Sha Ku, Hiroshi Kumigashira, Nabil Laachi, Jang Ein Lee, Ming-Hsiu (Eric) Lee, Yao-Jen Lee, Yi Jun Lee, Francois Leonard, Harry Levinson, Chenhsin Lien, Liew Yun Fook, Shawna Liff, Lloyd Litt, Ratchana Limary, Guanyang Lin, Chi Chun Liu, Cong (Colin) Liu, Timothy E. Long,Arantxa Maestre-Caro, Blanka Magyari-Kope, Prashant Majhi, Jochen Mannhart, Matthew Marinella, Francois Martin, Jennifer McKenna, Fumihiro Matsukura, Nobuyuki Matsuzawa, Dan Millward, Subhasish Mitra, Yoshiyuki Miyamoto, Steve Moffatt, Parvaneh Mokarian, Daniel Moraru, Mick Morris, Daniel Murray, Azad Naeemi, Christophe Navarro, Boris Naydenov, C. Gomez-Navarro, Paul Nealey, Mark Neisser, Kwok Ng, Quoc Ngo, Yoshio Nishi, Kei Noda, Yaw Obeng, Chris Ober, Katsumi Ohmori, Yutaka Ohno, Laurent Pain, Tomás Palacios, Samuel Pan, Jason Parker, Er-Xuan Ping , Jed Pitera, Joel Plawsky, Enrico Prati, Moshe Preil, Steve Putna, Ramamoorthy Ramesh, Nachiket Raravikar, Ben Rathsack, Heike Riel, Dave Roberts, Ricardo Ruiz, Thomas Russell, Tadashi Sakai, Dan Sanders, Sayeed Salahuddin, Gurtej Sandhu, Krishna Saraswat, Chandra Sarma, Hideyuki Sasaki, Shintaro Sato, Akihito Sawa, Mizuki Sekiya, Sadasivan Shankar, Matthew Shaw, Michael Sheehan, Takahiro Shinada, David Shykind, Michelle Simmons, Kaushal K. Singh, Bruce Smith, Tom Smith, Mark Somervell, Satofumi Souma, Naoyuki Sugiyama, Raja Swaminathan, Masahiro Takemura, Koki Tamura, Anna Tchikoulaeva, Serge Tedesco, Jim Thackeray, Raluca Tiron, Yoshihiro Todokoro, Yasuhide Tomioka, Luan Tran, Peter Trefonas, Jean-Marc Triscone, Ming-Jinn Tsai, Wilman Tsai, Emanuel Tutuc, Ken Uchida, Yasuo Wada, Kang Wang, Jeffrey Weinhold, Andrew Whittaker, C.P. Wong, H.S. Philip Wong, Karen Wooley, Dirk Wouters, Wen-Li Wu, Shigeru Yamada, Fu-Liang Yang, Linda He Yi, Hiroaki Yoda, Bin Yu, Todd Younkin, Victor Zhirnov, Paul Zimmerman, Alfred Zinn, Ehrenfried Zschech, Ajoy Zutshi

Front End Processes—John C. Arnold, Souvik Banerjee, Joel Barnett, Meredith Beebe, Skip Berry, Ian Brown, Jeff Butterbaugh, Luke Chang, Glenn Gale, Michael Goss, Chris Gottschalk, Aomar Halimaoui, Yukinobu Hikosaka, Chris Hobbs, Eric A. Joseph, Andreas Kadavanich, Sivananda Kanakasabapathy, Deoksin Kil, Simon Kirk, Martin Knotter, Yong-Sun Koh, Kaushik Kumar, Cathy Labelle, Yannick Le Tiec, Kee Jeung Lee, Wei-Yip Loh, Prashant Majhi, Kevin McLaughlin, Paul Mertens, Ichiro Mizushima, Dan Mocuta, Bich-Yen Nguyen, Jin-Goo Park, Jagdish Prasad, Srini Raghavan, Maud Rao, Abbas Rastegar, Rick Reidy, Stefan Schmitz, Akshey Sehgal, Wolfgang Sievert, Chris Sparks, Allan Upham, Steven Verhaverbeke, M. Watanabe, Richard Wise, Klaus Wolke

Lithography—Tsukasa Azuma, Chris Bencher, Tatsuo Chijimatsu, Brian Cha, Will Conley, Ralph Dammel, Greg Denbeaux, Anton deVilliers, Masayuki Endo, Nigel Farrar, Ted Fedynyshyn, Heiko Feldmann, Emily Gallagher, C. Michael Garner, Reiner Garreis, Cesar Garza, TS Gau, Bob Gleason, Frank Goodwin, Roel Gronheid, Naoya Hayashi, Long He, Yoshiaki Ikuta, Rik Jonckeere, Hyungsang Joo, Franklin Kalk, Kunihiko Kasama, Patrick Kearney, Insung Kim, Sachiko Kobayashi, Christof Krautschik, Y.C. Ku, Keishiro Kurihara, Jongwook Kye, David Kyser, Michael Lercel, ChangMoon Lim, Shy-Jay Lin, Lloyd Litt, Greg McIntyre, Dan Millward, Matt Malloy, Pawitter Mangat, Hiroaki Morimoto, Venkat Nagaswami, Hideo Nakashima, Patrick Naulleau Mark Neisser, Katsumi Ohmori, Yasushi Okubo, Masahiko Okumura, Kazuya Ota, Laurent Pain, Eric Panning, Moshe Preil, Doug Resnick, Morty Rothschild, Yoshitake Shusuke, Mark Slezak, Osamu Suga, Kazuhiro Takahashi, Takao Tamura, Serge Tedesco, Raluca Tiron, Walt Trybula, Plamen Tzviatkov, Rick Uchida, Fumikatsu Uesawa, Mauro Vasconi, Keiji Wada, Phil Ware, John Wiesner, Jim Wiley, Grant Willson, Obert Wood, Stefan Wurm, Jiro Yamamoto, Tetsuo Yamaguchi, Pei-Yang Yan, Anthony Yen, JeongHo Yeo, John Zimmerman

Interconnect—Nobuo Aoi, Lucile Arnaud, Koji Ban, Hans-Joachim Barth, Eric Beyne, Boyan Boyanov, Jon Candelaria, Chung-Liang Chang, Hsien-Wei Chen, Wen-Chih Chiou, Gilheyun Choi, Jinn-P. Chu, Mike Corbett, Alexis Farcy, Paul Feeney, Takashi Hayakawa, Paul Ho, Cheng-Chieh Hsieh, Masayoshi Imai, Atsunobu Isobayashi, Raymond Jao, Shin-Puu Jeng, Morihiro Kada, Sibum Kim, Nobuyoshi Kobayashi, Kaushik Kumar, Noh-jung Kwak, Hyeon Deok Lee, Anderson Liu, James Lu, Toshiro Maekawa, David Maloney, Akira Matsumoto, Azad Naeemi, Mehul Naik, Tomoji Nakamura, Yuichi Nakao, Akira Ouchi, Sesh Ramaswami, Hideki Shibata, Michele Stucchi, Zsolt Tokei, Thomas Toms, Manabu Tsujimura, Kazuyoshi Ueno, Osamu Yamazaki, Paul Zimmerman

Factory Integration—Daniel Babbs, Jonathan Chang, Gino Crispieri, Peter Csatary, Mike Czerniak, Chih-Wei (David) Huang, Slava Libman, Les Marshall, Supika Mashoro, Rick McKee, Steve Moffatt, James Moyne, Andreas Neuber, Kevin Pate, Dave Roberts, Markus Pfeffer, Dan Stevens, Ines Thurner, Makoto Yamamoto

Heterogeneous Integration—Amit Agrawal, Bernd Appelt, Muhannad S. Bakir, Dale Becker, Steve Bezuk, W. R. Bottoms, Yi-jen Chan, William Chen, Dan Evans, Michel Garnier, Steve Greathouse, Tom Gregorich, Richard Grzybowski, George Harman, Mike Hung, Ph.D, John Hunt, Rong-Shen Lee, Li Li, Sebastian Liau, Weichung Lo, Debendra Mallik, Keith Newman, Gary Nicholls, John Osenbach, Richard F. Otte, Bob Pfahl, Gilles Poupon, Gamal Refai-Ahmed, Charles Richardson, Theresa Sze, Coen Tak, Kripesh Vaidyanathan, James Wilcox, Jie Xue,

Environment, Safety, and Health—Laurie Beau, Mario Chen, Paul Connor, Reed Content, Terry Francis, Mike Garner, F. M. Hsu, Tom Huang, Francesca Illuzzi, Deb Kaiser, Leo Kenny, Alan Knapp, S. J. Ko, Hsi-An Kwong, Chris Lee, Slava Libman, Joey Lu, Mary Majors, Supika Mashiro, Toshi Matsuda, Rick McKee, Steve Moffatt, James Moyne, Andreas Neuber, Alain Pardon, Srini Raghavan, Dave Speed, Masahiro Takemura, Steve Tisdale, Pete Trefonas, Nausikaa van Hoornick, Kurt Werner, Dan Wilcox, Walter Worth, Don Yeaman

Yield Enhancement—Scott Anderson, Sabrina Anger, Joel Barnett, Dwight Beal, David Blackford, Yannick Bordel, Jennifer Braggin, Richard Brulls, Marc Camezind, Jan Cavelaars, Hubert Chu, Mike Clarke, Robert Clark, Jeff Covington, John DeGenova, Gil Delgado, Charley Dobson, Arnaud Favre, Guiseppe Fazio, Francois Finck, Herve Fontaine, Dan Fuchs, Takashi Futatsuki, Guillaume Gallet, Astrid Gettel, Rick Godec, Milton Goldwin, Barry Gotlinsky, Mathias Haeuser, Asad Haider, Jeff Hanson, Teruyuki Hayashi, Christoph Hocke, Masahiko Ikeno, Francesca Illuzzi, Hans Jansen, Jost Kames, Barry Kennedy, Keith Kerwin, Suhas Ketkar, Y.J. Kim, Katsunobu Kitami, Kaoru Kondoh, Naoki Kotani, John Kurowski, Sumio Kuwabara, Bob Latimer, Ravi Laxman, Slava Libman, Rick Lindblom, Juergen Lobert, Rushikesh Matkar, Yasuhiko Matsumoto, Tim Miller, Fumio Mizuno, William Moore, Jens Mueller, Chris Muller, Hiroshi Nagaishi, Andreas Neuber, Kazuo Nishihagi, Andreas Nutsch, Jim Ohlsen, Kevin Pate, Dilip Patel, Ruben Pessina, Lothar Pfitzner, Larry Rabellino, Abbas Rastegar, Dieter Rathei, Rich Riley, David Roberts, Dan Rodier, Biswanath Roy, Philippe Rychen, Koichiro Saga, Hideyuki Sakaizawa, Tony Schleisman, Akshey Sehgal, Yoshimi Shiramizu, Drew Sinha, Jim Snow, Terry Stange, Isamu Sugiyama, Paul Tan, Yoshitaka Tatsumoto, Ines Thurner, Hiroshi Tomita, Takahiro Tsuchiya, Allan Upham, Gerald Weineck, Dan Wilcox, Hubert Winzig, Bernie Zerfas

Metrology—Carlos Beitia; Mark Berry; Benjamin Bunday; Umberto Celano; SooBok Chin; Hyun Mo Cho; Alain Diebold; Brendan Foran, Michael Gaitan; Michael Garner; Christina Hacker; Karey Holland Masahiko Ikeno Yun Jung Jee; Eiichi Kawamura; Adrain Kiermasz; Joseph Kline; Stephen Knight; Delphine Le-Cunff; Scott List; Seong-Min Ma; Philippe Maillot; Jack Martinez; Niwa Masaaki; Yaw Obeng; George Orji; Brennan Peterson; Narenda Rana; David Seiler, Victor Vartanian; Andras Vladar; Yuichiro Yamazaki.

Modeling and Simulation—Jean-Charles Barbe, Augusto Benvenuti, Alex Burenkov, Mauro Ciappa, Wolfgang Demmerle, Andreas Erdmann, Wladek Grabinski, Tibor Grasser, Vincent Huard, Bert Huizing, Herve Jaouen, Andre Juge, An De Keersgieter, Gerhard Klimeck, Wilfried Lerch, Jürgen Lorenz, Wolfgang Molzer, Victor Moroz, Chandra Mouli, Jürgen Niess, Yon-Sup Pang, Hwasik Park, Paul Pfäffli, Peter Pichler, Yi-Ming Sheu, I.C. Yang

Table of Contents

List of Figures

List of Tables

EXECUTIVE SUMMARY

For the past 50 years the Semiconductor Industry has marched at the pace of Moore's Law. Transistor scaling associated with doubling the number of transistors every two years has been and continues to be the unique feature of the semiconductor industry. As a consequence, as transistors became smaller they could also be switched from the off to the on state at faster rates while simultaneously became cheaper to manufacture. System integrators assembled new products utilizing the building block provided by the semiconductor industry but system integrators were barely able to complete assembling a new system when a yet new more powerful IC was becoming available. Any new technology generation enabled multiple new products with better performance than the previous ones. Integrated device manufacturers (IDM) were in full control of the pace at which the electronics industry was progressing by setting the marching pace at which the whole electronics industry ecosystem was moving forward. Therefore past editions of technology roadmaps concentrated on forecasting the rate of transistor scaling and how transistor density and performance affected the evolution of integrated circuits (IC).

In the past 15 years the advent of the Internet, the extensive deployment of Wi-Fi base stations, consumer acceptance of a broad variety of wireless mobile appliances plus the successful combination of fabless companies working in conjunction with foundries has completely changed the electronics industry. System integrators are nowadays able to conceive, design and realize any integrated circuit they wish without having to recur to integrate device manufacturers. System Integrators can nowadays integrate multiple functionalities in a single chip called System on Chip (SOC) or by means of integrating multiple dice in a single package called System in Package (SIP) as opposed to connecting multiple specialized ICs on a board. It is clear that these methods of integration are more efficient and less costly than acquiring several separate ICs (e.g., microprocessor, graphic processor, multiple memory types, usb etc.) and assembling them on a board. System integrators are by and large setting the pace of innovation for the electronics industry.

Each new technology generation produces faster transistors that can switch faster than those produced with the previous technology generation. In the past this electrical feature of transistors enabled microprocessors to operate at higher frequencies and therefore computer performance as measured by industry benchmarks, like measured by million of instruction executed per second (MIPS), continued to improve at very fast rates. Computer architecture has not changed since Von Neumann introduced its concept on how to perform computing in 1945. Through the years nobody paid too much attention to the fact that power consumption of integrated circuits kept on increasing with any new technology generation since the industry motto was: "***performance at any cost!"*** This approach came to an end at the beginning of the past decade when fundamental thermal limits were reached by some ICs. Even though the transistor count has kept on increasing now and then at Moore's Law pace and transistors are able to operate with each new technology generation at higher frequency then before it has become practically impossible to keep on conjunctly increasing both of these factors due to physical limitations on power dissipation; one of the two feature (i.e., either number of transistor or frequency) had to level off in order to make the ICs capable to operate under practical thermal conditions. Frequency was selected as the sacrificial victim and it has stalled in the few GHz since the middle of the previous decade.

These limitations on maximum useable frequency have impacted the rate of progress of the computer industry that has been compelled to develop such methods as complex software algorithms and clever instruction management to improve performance to partially compensate for the aforementioned conditions. The architecture of the microprocessors has changed from single core to multi-core. In this arrangement the processor can run in the few GHz range while the output rate is increased multifold by combining the output of multiple cores to produce the output signal. Unfortunately this parallel type of solution cannot be used in all cases since some problems can only be solved in a serial way. However, these performance limitations did not impact the development and expansion of the mobile Internet society. Consumers began accessing the Internet via desktop appliances and then progressively got used to access it via mobile multipurpose appliances. Reaching any source of information via the Internet takes tens of milliseconds due to the speed at which signals can travel on any interconnect lines so microprocessors operating in the few GHz frequency range are more than adequate to handle the communication traffic. Cell phones began operation in the 90s using frequencies in the 800-900MHz ranges in accordance with specifications of the Global System for Mobile Communications (GSM). These operational frequencies evolved with the use of 4G and LTE to the 2,500-2,700MHz ranges. Operation in these frequency ranges is still well within the capabilities of ICs. In the past 10 years, cell phones and mobile appliances in general have become a viable means of accessing the Internet. Cell phone power consumption is typically below 5 watts so this value is well within the thermal limits of ICs operation. Most recently, access to the Internet via Wi-Fi has been continuously increasing since the areas of coverage are continuously extending; mobile appliances have become the most convenient means of communication and access to any source of information anywhere at any time.

The IC industry has also contributed to provide valuable technology building blocks to other industries and by so doing enabling new devices like micro-mechanical systems (MEMS), flat panel displays, multiple sensors and so on. All of these somewhat dissimilar technologies have been readily included into mobile appliances by means of heterogeneous integration.

The insatiable demand for information has led to the creation of gigantic clusters of servers and memory banks named Data Centers. In this environment performance is still the name of the game and using complex cooling systems can mitigate power issues. Power consumption of Data Centers is rapidly escalating into the hundreds of Megawatts range. Communications within the Data Centers and for long distances is handled via fiber optics because of their stellar low rate of attenuation.

All of the above systems specifications dictate the requirements for the semiconductor industry.

Minimal power consumption of transistor operation has become the main requirement for the semiconductor industry dictated by the new ecosystem of the electronics industry. On the other hand, the requirement for a continuously increasing number of transistors according to Moore's Law continues unabated. As an example this can be demonstrated by observing that the number of transistors in the latest cell phone Application Processors has kept on increased from the 1Billion transistors of the A5 to 2Billions for A6 to 3Biliion for A6X. It is projected that the upcoming A9 will range from 2.7 to 4.5 Billions depending upon the application.

To satisfy the product demand for higher transistor counts the semiconductor industry is approaching a new era of scaling.

Geometrical scaling characterized the 70s, 80s and 90s. This was the first generation of transistor scaling. Major material and structural limitations were identified in the mid-90s and the research community initiated the foundation of a new scaling approach that was heralded by the ITRS in 1998. This was named Equivalent Scaling. Strained silicon, high-k/metal gate, FinFET and use of other semiconductor material (e.g., Germanium) represent the main features of this scaling approach.

As features approach the 10nm range and below it becomes clear that the semiconductor industry is running out of horizontal space.

Memory products have always been the leaders in transistor density and so it is not surprising that the solution to this problem is coming from Flash memories. Multiple companies have announced that future products will fully utilize the vertical dimension. This is not too dissimilar from the approach taken in Manhattan, Tokyo, Hong Kong or similarly highly crowded places to deal with space limitations: skyscraper have become the standard approach to maximize "packing density".

For the reasons exposed above the new scaling method is called "3D Power Scaling".

In the Beyond CMOS section the reader will find multiple new and exciting devices that after 10 years of research are demonstrating the possibility of becoming key players in the next decade.

Finally, all these exciting technologies need to be realized in a High Volume Manufacturing (HVM) Factory. Controlling nanometer features across 300mm wafers remain an amazing challenge that the industry is managing across the globe but will this be possible on 450mm wafers? Real time data collection, analysis and consequent disposition of wafers via a fully automated material handing system will remain the goals of the HVM Factory in the next decade until "full light out" operation is reached.

1. INTRODUCTION: THE EVER CHANGING ENVIRONMENT

The Semiconductor Industry was born in the 70s as an electronics component industry with three main business drivers. The first driver consisted in providing cost effective memory devices to the computer industry. Memories soon became the undisputed volume leaders.

The second driver consisted in production of application specific integrated circuits (ASIC). Any company that required new and unique functionalities to realize novel products developed a set of functional and electrical targets that were communicated to one or more semiconductor manufacturers. The latters turned these targets into specifications for one or more Integrated Circuits (IC) but all the Intellectual Properties (IP) belonged to the companies that commercialized the final produces and not to the IC manufactures.

Thirdly, ICs allowed cost effective integration of simple building blocks used in the design of many systems. These ICs typically consisted in small logic gates (RTL, DTL, TTL, NAND, J-K Flip-Flop etc.) and operational amplifiers (μA 709, μA 741 etc.) just to mention a few. Customers demanded pin-out and functionality standardization for memory devices while ASIC products were customer specific. Logic devices reprogrammable by software (e.g., microprocessors) were developed to minimize design cycle time and to pave the way for subsequent high volume ASIC devices once the product functionality was demonstrated.

Phase 1

In the 70s and 80s ***system specifications were solidly in the hands of system integrators***. New DRAM memory technologies were introduced every three years to keep up with the introduction of new computer systems. Typically a new generation of computers required a four-fold increase in memory and that it is why the famous "4X/3Years" memory density and pace for introduction of new memory technologies was established. This 4x increase in memory size was realized by means of decreasing transistor features but also by substantially increasing the die size and this had a very negative impact on product cost as fewer dice were made available by this approach on each wafer. In the long rum the latter resulted in a limiting factor for DRAM memory manufactures that had to default to doubling the number of transistors every two years. Leading producers of memory technologies generated requirements for development of new building blocks for the manufacturing process that required better equipment; makers of logic devices were operating on a 4 year cycle trailing behind memory manufactures and therefore adopted a subset of similar process blocks and most of the same equipment developed for memory technologies.

Phase 2

In the early 90s the very profitable Personal Computer market required more complex and faster microprocessors. To feed the growth of this market the makers of logic ICs were compelled to accelerate the introduction of new technologies from 4-years cycle to a more aggressive 2-year cycle to support faster introduction of new products. In the late 90s memory makers were compelled to follow the 2x/2-year introduction cycle when eventually the die size of new memory products became too large to be economic. The strong correlation between technology scaling and electrical performance of most ICs, as for instance higher operational frequency, made the introduction of new higher performance products, and PC in particular, completely subordinate to enhanced performance of microprocessors. This strong scaling-product correlation made those IC makers that controlled the introduction of leading technologies the real controllers of the whole electronics industry. As a direct result of this trend, a ***substantial part of system specifications, performance and profits shifted in the hands of IC manufacturers***. After all this ecosystem was beneficial to everybody in the electronics industry and as long as performance of microprocessors and computers built accordingly to Von Neumann architecture was benefitting form higher operational frequency, the entire semiconductor industry revenue continued to grew. The IC industry grew at an average pace of 17%/year during this period. IDM manufacturers of logic devices were able to capitalize the most from this new shift in the balance of power.

Like it happens very often, the IC industry believed in the 90s that this market model (i.e., faster processors supporting larger amounts of memory supporting continuous growth for the semiconductor and electronics industries) was going to last forever but things were going to change in a very dramatic way with the advent of the new century.

Phase 3

A completely new ecosystem emerged during the past decade:

- First of all, the relatively moderate incremental cost of technology development from one generation to the next (10% or below) fueled the aggressive bi-annual introduction of new semiconductor technologies that allowed ICs, consisting of hundreds of million of transistors, to be produced in a very cost effective manner. This made it possible to integrate extremely complex systems constituted of logic, memory, graphics and other functionalities on a single die at a very attractive cost. Furthermore, progress in packaging technology enabled the placement of multiple dice within a single package. These categories of devices were defined as system on chip (SOC) and system in package (SIP), respectively.
- Second, manufacturers of integrated circuits offering foundry services were able to provide the capability of realizing complex systems either on a single die or in a single package (the "New ASICs") at very attractive costs. This led to the emergence of a very profitable business model whereby fabless companies could generate IC designs, while actual wafer production was done elsewhere by a foundry company.
- Third, sophisticated process equipment, developed for the production of advanced integrated logic and memory circuits, proliferated to adjacent technology fields making the realization of flat panel displays (FPD), MEMS

sensors, radio components and passives, etc., possible at very reasonable costs. All these different technologies were then integrated in a single revolutionary system. This new way of making products by integrating multiple heterogeneous technologies successfully emerged around the middle of the previous decade and was named "More than Moore" (MtM).

Under these conditions ***system integrators were once again in the position to fully control system design and product integration.***

In the same timeframe, the successful adoption of the Internet and the rapid rise of world-wide adoption of mobile phones promoted the extensive deployment of fiber optic cables and the proliferation of multiple wireless technologies ranging from communication satellites to tens of thousands of "base stations" which enabled an unprecedented level of global and mobile connectivity.

This ecosystem facilitated also the creation of completely new and unexpected markets of which the many "Social Networks" represents one of the latest examples.

Intense research on how to enhance the functionality of mobile devices continues to bear fruits and has made them the ultimate instrument that allows customers to enjoy ubiquitous access to any type of communication and information material. Furthermore, new additional capabilities enabled by the integration of computer and communication technologies are fueling research on yet broader markets. Substantial in road is being made in making available to customers multiple sensorial inputs that are remotely generated. Among other things this technology is opening the way to enabling remote medical interaction between doctors and patience. Multiple companies also herald the arrival in the next few years of self-driving automobiles as another example of the integration of powerful on-board computers connected via Wi-Fi to navigation information and capable of simultaneously collecting information from multiple sensors operating from optical to RADAR frequencies.

All of the above elements have been referred to as the "Internet of Things" (IoT) or recently even more generally as the "Internet of Everything" (IoE) since any aspect of society ranging from human to objects to everything you can think of is becoming an "element" connected to and by the Internet. Innovative products introduced by telecommunication companies, by companies operating data centers and by content providers are battling for dominant positions in this newly created market. It is clear that all of these innovations could not have occurred without the support of the semiconductor industry that has provided the building blocks for all the above applications.

It was simple to understand the role of ICs in the product chain up to the 90s but with this completely different ecosystem we must ask ourselves an important and fundamental question:

"What is the role of the semiconductor industry in this new ecosystem?"

2. INTERNET OF EVERYTHING (IOE)

The ITRS already emphasized the importance of the Internet in previous reports but in the past 5 years IoT has evolved into IoE and it is important to dedicate a paragraph to this subject since the IoE is shaping the whole society and creating multiple industries.

The US Department of Defense awarded contracts as early as the 1960s for packet network systems, including the development of the ARPANET (which would become the first network to use the Internet Protocol.)

Access to the ARPANET was expanded in 1981 when the National Science Foundation (NSF) funded the Computer Science Network (CSNET). Since the mid-1990s, the Internet has had a revolutionary impact on culture and commerce, including the rise of near-instant communication by electronic mail, instant messaging, voice over Internet Protocol (VoIP) telephone calls, two-way interactive video calls, and the World Wide Web. This worldwide connectivity has created new phenomena like social networking and online shopping. Increasing amounts of data are transmitted at higher and higher speeds over fiber optic networks operating at 1-Gbit/s, 10-Gbit/s, soon 40-Gbit/s and more. The Internet's takeover of the global communication landscape was almost instant in historical terms: it only communicated 1% of the information flowing through two-way telecommunications networks in the year 1993, already 51% by 2000, and more than 97% of the telecommunicated information by 2007. Today the Internet continues to grow, driven by ever-greater amounts of online information, commerce, entertainment, and social networking. Access to the Internet was

originally done via desktop computers but the introduction of smart phones in 2007 and tablets in 2010 has revolutionized the way people interact via the Internet. The world of communications has truly become a wireless, ubiquitous and continuously interconnected world.

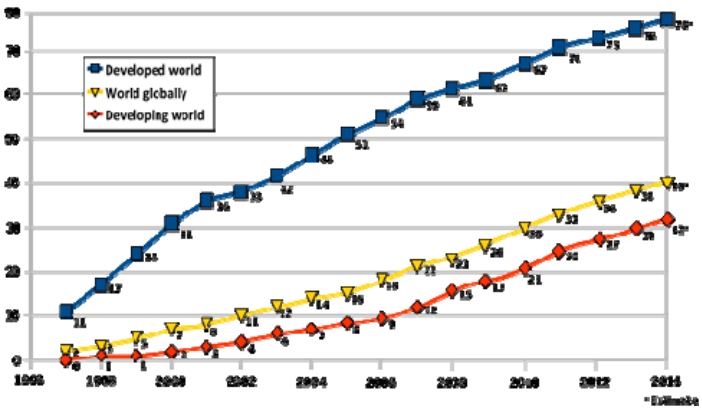


Fig. 2.1 Internet Penetration

Until 10 years ago the Internet was mostly used as a means of collecting information and communications but progressively it evolved into a commercial instrument through which people could book services, acquire goods, pay bills and so on. This however was just the beginning since different types of cameras and sensors began to be connected to the Internet and by so doing opening new exciting uses and applications. At present the Internet of Things allows objects to be sensed and controlled remotely across existing network infrastructure, creating opportunities for more direct integration between the physical world and computer-based systems, and resulting in improved efficiency, accuracy and economic benefit

As the list of elements that can be connected to the Internet keeps on increasing a new term has been coined out: The Internet of Everything (IoE). The term IoE expands on the concept of the "Internet of Things" in that it connects not just physical devices but quite literally ***everything*** by getting them all on the network. IoE works to connect more devices onto the network, stretching out the edges of the network and expanding the roster of what can be connected. IoE has a major play in all industries, from retail to telecommunications to banking and financial services. It also opens the door to remote medicine in which a doctor can analyze results of tests remotely taken by patients and evaluates them. Of course placing all this personal information on the web or accessible via the web poses a severe challenge for security. Almost everyday we can hear news of hackers stealing valuable information and it is imperative that appropriate measures are taken to prevent these problems but despite this more and more people are willing to "venture" in the IoE since the benefits offered by the many capabilities offered by the IoE are overwhelming.

With this all said and done it is important to remember that the IoE could not happen if semiconductors did not power communication devices, data centers, routers and sensors. The advent of Foundries and Fabless companies enabled the customization of semiconductor products that now cover all the aspect of IoE and it would be a mistake to assume that the semiconductor industry is by now a mature industry and it has not much more to offer. The advent of third phase of device integration (i.e., 3D Power Scaling) plus the many new capabilities associated with the introduction of revolutionary materials in the semiconductor industry will revolutionize how computers are built. New Computers built with revolutionary Architectures enabled by New Devices will be surrounded from top to bottom with a variety of new sensorial capabilities that will offer new and exciting options to system designers (see Rebooting Computing section for more details).

But we should not be overwhelmed by the endless applications that we can dream of since in the end, from a system point of view, the IoE is nothing more than a distributed computer capable of accessing via the Internet (Routers and base stations) remote memories (Data Centers), connected to a variety of sensors and actuators. Software and protocols then tie it all together.

In the near future the IoE will allow customers to retrieve, analyze, assemble, summarize information and provide also actionable recommendations. Be careful, the IoT will soon think for you!

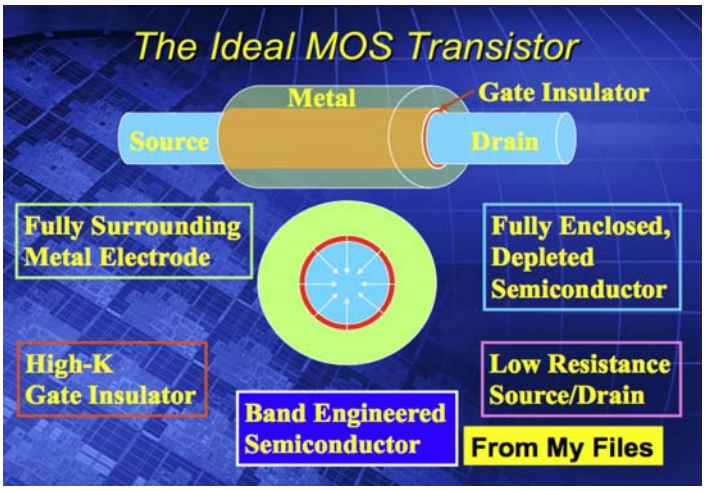
3. 2015 ITRS 2.0 EDITION AND MOVING FORWARD…

"Nothing is New, but never is the same"

The foundations of IC industry were laid out with the invention of the self-aligned silicon gate planar process in the late 1960s. Moore's predictions of the doubling of transistor generated on an annual and then bi-annual pace formulated in 1965 and in 1975 in conjunction with Dennard's scaling guidelines led to the growth of the semiconductor industry until the beginning of the last decade.

This was the (First) Era of Geometrical (classical) Scaling. This type of scaling was the foundation of the National Technology Roadmap for Semiconductors (NTRS) initiated in 1991.

The ITRS laid out the foundations of the (Second) Era: Equivalent Scaling (e.g., strained silicon, high-K/metal gate, Multigate transistors and use of non-silicon semiconductors in general) between 1998 and 2000.



1998 ITRS: Equivalent Scaling vision of required transistor innovation Fig. 1

The implementation of these technologies successfully supported the growth of the semiconductor industry in the past decade and it will continue to do so until the end of the present decade and beyond.

In the next decade ITRS 2.0 predicts that the advent of the third phase of scaling ***"3D Power Scaling"*** will become the driver of the rejuvenated semiconductor industry and this ***answers the question*** posed before about the future of the semiconductor industry: "***Yes the semiconductor industry will continue to be a key enabler of the IoE***"

Moore's Law is now entering into a third phase characterized by vertical integration and performance specifications driven towards reduction of power in either the active or the stand-by modes.

Fig. 3.1 The Ideal MOS Transistor

It is important to realize that despite the many announcements of the end of Moore's Law it is expected that the transistor density of future IC will continue to increase in the foreseeable future at historical rates.

Even though producers of logic integrated circuits have been the most prominent champions of Moore's Law it is also true that ***producers of memory devices have been the leaders in production of Integrated Circuits with the tightest tolerances and the smallest dimensions. In few words, they have been the real drivers of Moore 'Law and will continue to do so in the future.***

DRAM producers already dealt with space problems as far back as the "Megabit" memory generation. In order to build capacitors back then with enough storage capacity it would have required more silicon area than the area needed to build the transistors and therefore DRAM producers adopted stack capacitor or trench capacitor solutions to take advantage of the vertical dimension either above or below the surface of the silicon.

Nowadays, Flash memory producers are facing a similar problem since they are running out of horizontal space, cost of producing integrated memory circuits of small dimensions keeps on rising while the number of stored electrons in the floating gate keeps on decreasing [Fig.3.2]. To eliminate these problems Flash memory producers have already demonstrated and announced several new products that stack multiple layers of memory on top of each other in a single integrated circuit [Fig. 3.3]. As many as 32 and 48 layers of Flash memory have been reported. Flash memory devices constituted by more than 100 layers have been predicted. Flash memory producers promise a transistor density acceleration of memory ICs realized with 3D technology ***far exceeding the historical 2x/2year rate***. Moore's Law is once again on the verge of a substantial acceleration.

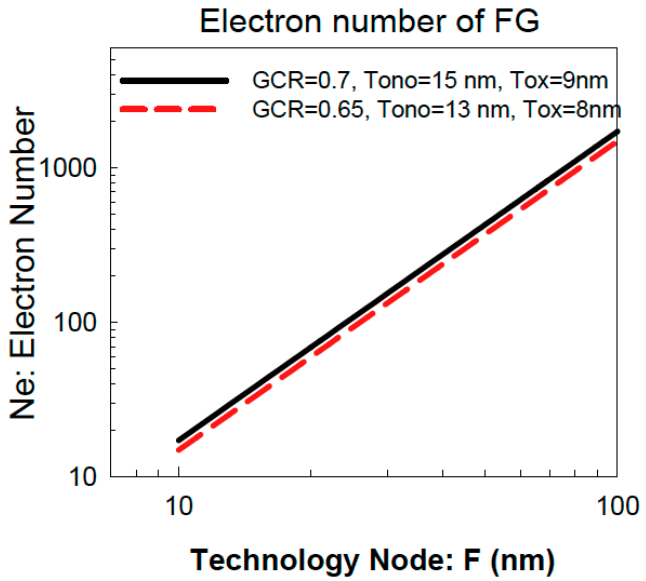


Fig. 3.2 The numbers of electrons on a floating gate will be reduced to 10 when design rules reach the 10nm

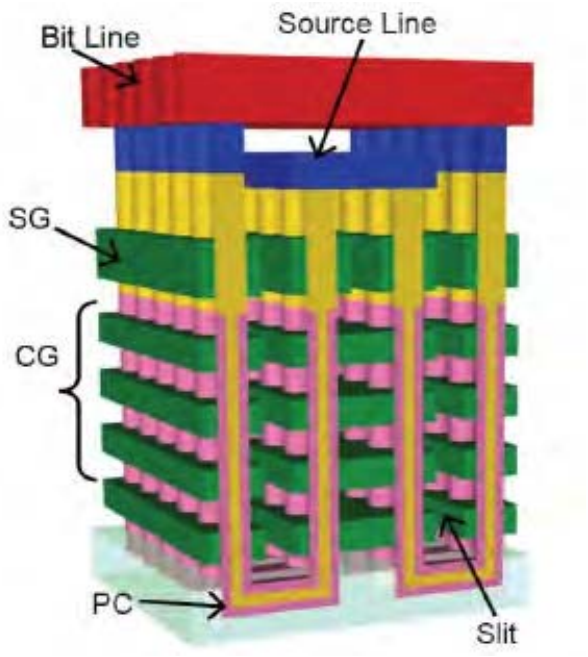


Fig. 3.3 Toshiba's BiCS (Bit Cost Scalable)

Once again it is relevant to notice that, as a method of manufacturing the densest IC reaches physical limits a new, more powerful one, takes over the role of carrying the industry forward for at least one or two decades. Despite this recurrent "demise and resurgent" sequence of events many prophesies about the end of Moore's Law have been published every 10 years or so. May be the best way to eliminate these recurrent misunderstandings consists in capturing in a single phrase

the essence of these repeating events. This goal could be accomplished by using a well-known and familiar saying and morphing it to epitomize the most relevant historical trend of the semiconductor industry as follow:

"Moore's Law is dead, long live Moore's Law!"

4. Overall 2015 2.0 Process and Structure

4.1. Background

As outlined before the advent of Fabless Design Houses and Foundries has revolutionized the way in which business is done in the new semiconductor industry, System Integrators have regained full control of the business model. This implies that system requirements are set at the beginning of any new product design cycle and step by step related requirements percolate down through the manufacturing production chain to the semiconductor manufactures. No longer a faster microprocessor triggers the design of a new PC but on the contrary the design of a new smart phone generates the requirements for new ICs and other related components. Under these conditions it became clear in 2012 that the ITRS needed to adapt and morph to the new ecosystem [Fig 4.1]. It was anticipated that this transformation process would take sometime and it was decided that the 2013 ITRS was going to be the last of its kind. Next, 2014 and 2015 were going to be dedicated to the construction of a new roadmap that was named ITRS 2.0.

The new ecosystem needed to be viewed though a top-down viewer and for this purpose the elements outlined in Fig. 4 led to the formation of seven Focus Teams [Fig.4.2]

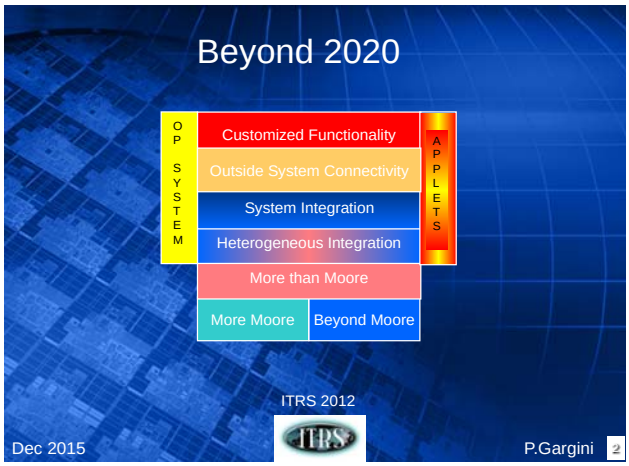


Fig. 4.1 The New Ecosystem

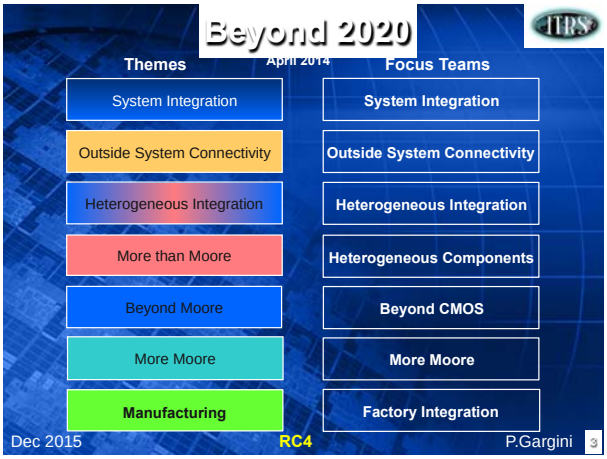


Fig. 4.2 The New Focus teams

4.2. Roadmapping Process

The ITRS process and content has kept on evolving to match the needs of the semiconductor industry. Collaboration of industry and research continues to be essential to understand the technical challenges and future needs in the various spectra of micro- and Nano-electronics. Over the past few years, the ITRS teams have worked closely to assess emerging technologies. They have been preparing to support the technical needs of both the new world of interconnectedness between humans and hardware (IoE) as well as how to handle the complex world of information processing known as Big Data. These examples are only part of the new frontiers of inventions and discoveries. As these new focus topics continue to emerge, the ITRS teams will continue to determine what this means for the global electronics industry. The industry must define the new drivers to help it stay on a path of productivity and profitability, while promoting environmental health and encouraging areas of innovation for new scientists and technologists

The most relevant subjects of the ITRS were originally divided among eleven International Technology Working Groups (ITWGs). As time went by the industry became more complex and the number of ITWGs increased to 17 in 2013. During the 2014-15 timeframe ITRS 2.0 was reorganized into 7 Focus Teams.

For the 2015 ITRS 2.0 the Focus Teams are the following:

- System Drivers
- Heterogeneous Integration
- Heterogeneous Components
- Process Integration, Devices, and Structures
- Outside System Connectivity
- More Moore Beyond CMOS
- Factory Integration

The existing 17 ITWGs were redistributed according to this new organization [Fig 4.3]. All the ITWGs are still operational but most of their output is channeled into the Focus team. An additional section ***"For Expert Only"*** includes additional detailed information provided by the ITWGs that was not completely included into the Focus Teams. The Focus Teams and the ITWGs are composed of experts from industry (chip-makers as well as their equipment and materials suppliers), government research organizations, and universities.

		ITRS 2.0						
		1	2	3	4	5	6	7
2013 ITRS ITWGs		SI	OSC	HI	HC	MM	BC	FI
1 System Drivers		X						
2 Design		X						
3 Test and Test Equipment				X				
4 Process Integration, Devices and Structures (PIDS)						X		
5 Emerging Research Devices (ERD)							X	
6 Emerging Research Materials (ERM)							X	
7 Front End Process						X		
8 Lithography						X		
9 Interconnect						X		
10 RF and A/MS Techonologies			X					
11 Yield Enhancement								X
12 Factory Integration								X
13 Assembly and Packaging				X				
14 Enviroment, Safety and Health								X
15 Metrology						X		
16 Modeling and Simulation						X		
17 MEMs					X			

Fig. 4.3 Transition from ITWGs to FTs

4.3. 2015 ITRS 2.0 FOCUS TEAMS

System Integration (SI)

The mission of the System Integration (SI) chapter in ITRS2.0 is to establish a top-down, system-driven roadmapping framework for key market drivers of the semiconductor industry drivers in the 2015-2030 period. The SI chapter is currently developing and constructing roadmaps of relevant system metrics for mobile, datacenter and Internet of Things (IoT) drivers.

Heterogeneous Integration (HI)

Heterogeneous Integration refers to the integration of separately manufactured components into a higher-level assembly that in the aggregate provides enhanced functionality and improved operating characteristics. In this definition components should be taken to mean any unit whether individual die, MEMS device, passive component and assembled

package or sub-system that are integrated into a single package. The operating characteristics should also be taken in its broadest meaning including characteristics such as system level cost-of-ownership.

The mission of Heterogeneous Integration Focus Team is to provide guidance to industry, academia and government to identify key technical challenges with sufficient lead-time that they do not become roadblocks preventing the continued progress in electronics that is essential to the future growth of the industry and the realization of the promise of continued positive impact on mankind. The approach is to identify the requirements for heterogeneous integration in the electronics industry through 2030, determine the difficult challenges that must be overcome to meet these requirements and, where possible, identify potential solutions.

Heterogeneous Components (HC)

Heterogeneous Components are electronic devices that are manufactured using micro and Nano fabrication technologies and assembled together into the heterogeneous systems. The Heterogeneous Components Chapter identifies trends in high-market-growth-potential applications, determines the device performance requirements that are needed to enable those applications, and uses this information to discover technology gaps that must be solved in order to produce the future products that are envisioned by the roadmap.

The Mission of Heterogeneous Components Focus Team is to provide industry, academia and government to identify key technical challenges related to new heterogeneous components needed by the IoE ecosystem.

Outside System Connectivity (OSC)

The mission of OSC consist in identify and assess capabilities needed to connect most elements of the Internet of Everything (IoE) and highlight technology needs and gaps. This includes supporting connection of a broad range of sensors, devices, and products and to support information communication, processing and analysis for many applications (i.e. mobility, energy, health, and others) with wire line, wireless and optical interconnect technologies.

More Moore (MM)

MM Focus Team in ITRS provides physical, electrical and reliability requirements for logic and memory technologies to sustain More Moore (PPAC: power, performance, area, cost) scaling for big data, mobility, and cloud (IoT and server) applications and forecast logic and memory technologies (15 years) in main-stream/high-volume manufacturing (HVM)

Beyond CMOS (BC)

The goal of this Focus Team is to survey, assess and catalog viable new information processing devices and system architectures due to their relevance on technological choices. It is also important to identify the scientific/technological challenges gating their acceptance by the semiconductor industry as having acceptable risk for further development. Another goal is to pursue long-term alternative solutions to technologies addressed in More-than-Moore (MtM) ITRS entries.

This is accomplished by addressing two technology-defining domains: 1) extending the functionality of the CMOS platform via heterogeneous integration of new technologies, and 2) stimulating invention of new information processing paradigms.

Factory Integration (FI)

The Factory Integration (FI) focus area of ITRS 2.0 is dedicated to ensuring that the semiconductor-manufacturing infrastructure contains the necessary components to produce items at affordable cost and high volume. Realizing the potential of Moore's Law requires taking full advantage of device feature size reductions, new materials, yield improvement to near 100%, wafer size increases, and other manufacturing productivity improvements. This in turn requires a factory system that can fully integrate additional factory components and utilize these components collectively to deliver items that meet specifications determined by other ITRS 2.0 focus areas as well as cost, volume and yield targets.

5. SYSTEM INTEGRATION: ITRS 2.0 VIEW OF THE NEW ECOSYSTEM

In providing a top-down view of the new ecosystem it is necessary to have some simple depiction of what is under study including some basic definitions

ITRS 2.0 has developed also several tables to capture all of these elements and the reader will be able to find detailed information in the chapters addressing these subjects.

The executive summary presents a succinct overview of the above elements progressing from system requirement all the way to device specifications.

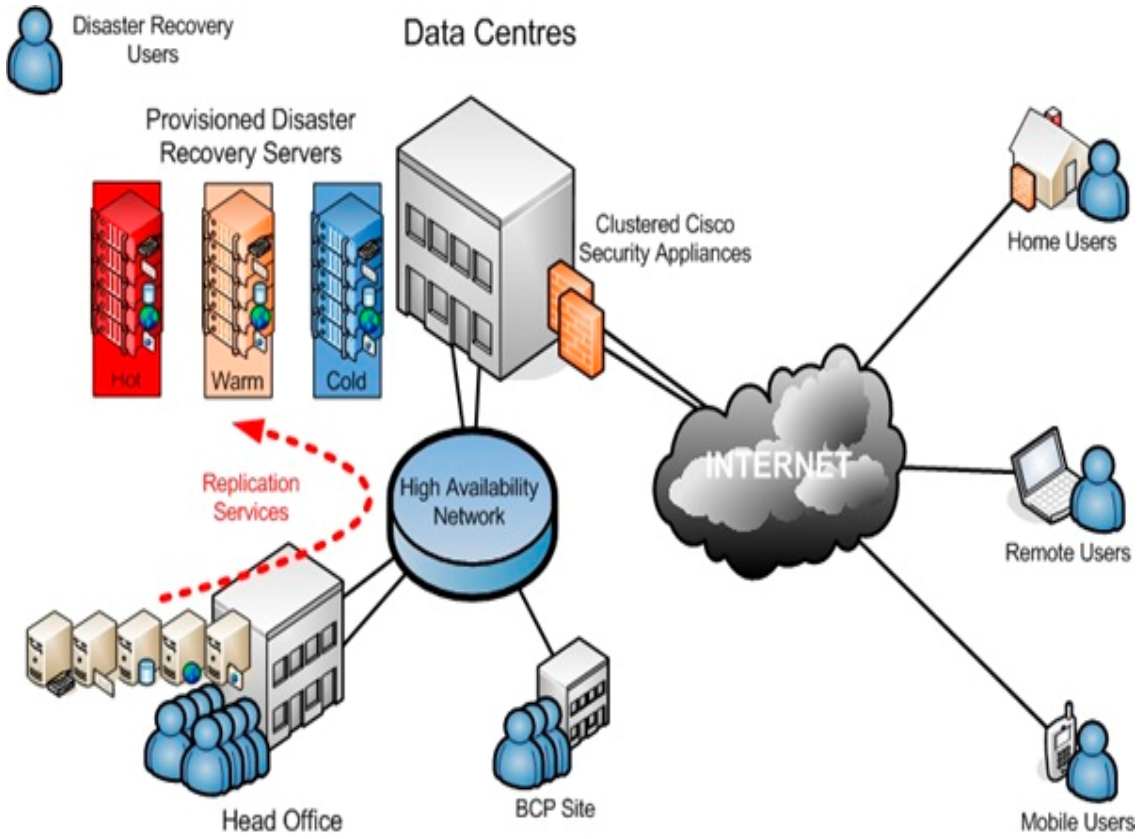


Fig.5.1 Internet: Example of Data Centers and Connectivity

5.1. DATA CENTERS

Data centers are physical or virtual infrastructure [Fig. 5.1] used by enterprises to house computer, server and networking systems and components for the company's information technology (IT) needs, which typically involve storing, processing and serving large amounts of mission-critical data to clients in a client/server architecture.

In order to improve processing performance the IC industry has adopted a multi-core architecture for any type of Processors. This technical approach derives from the power limitations imposed by operating with a single core. Multiple cores, however operating at a lower frequency than a single core can provide enhanced output by processing information in a parallel way.

The increased demand for information processing in a data center will drive a 30-fold increase in the number of cores.

Data Centers

Categories	Year	2015	2017	2019	2021	2023	2025	2027	2029
Data Center Global Indicators	Number of Cores (K)	360	1044	3008	4935	5825	7578	8967	10602
	Total Memory Storage (PB)	300	1559	4676	14029	42088	126264	378792	1136377
	Area of One Server Building (MSF)	0.5	0.9	1.6	2.2	2.2	2.42	2.42	2.42
	Total Power Consumed by Datacenter (MkWh)	779.8	839	1004.7	1137.2	1226.1	1380.7	1635.6	2044.3
	Total Switching Capability of Datacenter BW (Tb/s)	1000	2512	6309	10000	15849	25119	39811	63096
	Data Center Efficiency (GFLOPS/W)	2.4	4.9	10	17	24	33.9	47.9	67.8
Servers	Server Units/Rack	40	40	40	40	40	40	40	40
	Number of Cores/socket	18	29	47	59	74	93	117	147
	Power /Single Server Unit (W)	700	700	700	700	700	700	700	700
	Main Memory / Single Server Unit (GB)	32	45	64	76	91	108	129	154
	Power Consumed by Cores/single socket (W)	165	159	153	149	145	141	137	133
Switching	Network Bandwidth/Unit (Gb/s)	40	40	100	100	100	400	400	400
	Network Bandwidth/Rack (Gb/s)	1600	1600	4000	4000	4000	16000	16000	16000
	Rack Switch Capacity BW (Gb/s)	1200	1200	3000	3000	3000	12000	12000	12000
Power	Power efficiency (Grid delivery/Data Center Use)	0.55	0.57	0.59	0.61	0.63	0.65	0.67	0.69
	Power Consumed by Networking and Switching (MkWh)	21.91	55.05	138.26	87.66	138.93	220.19	348.97	553.08
	Power Consumed by Storage (MkWh)	0.0657	0.259	0.449	0.778	1.347	2.334	4.043	7.004
	Power Consumed for facility cooling (MkWh)	38.99	55.02	86.13	237.31	264.26	307.03	374.89	482.56

Data Centers operate as repository of data and given the insatiable need expressed by the user for more and more data it is expected that the requirement for memory storage will be the fastest growing item in the ITRS 2.0 Horizons an almost a 4000-fold increase in memory is forecast.

Under these conditions the area of the building will keep on increasing by 5x. These requirements will more than double the power requirements as measured in GWh (or MKWh). In order to control the rate of increase in power it is mandatory that the power of each server and the number of servers/rack be kept constant across the next 15 years.

Moving information within the Data Center and moving data to the network and receiving data from the network is of fundamental importance for the efficient operation of the Data Center and its connection to the Internet as an efficient means of communication to the users. Switching bandwidth per unit is presently operating at 40Gb/sec but it will need to reach the 400 Gbit/sec by 2025. This will require that the Network Bandwidth /rack will need to reach the 16,000 Gb/sec range to feed the insatiable need for data requested by the network.

A severe concern is constituted by the amount of energy needed to cool the facility that will escalate to almost 500 GWh in the time horizon. It is not surprising that Data centers need to come with an associated power plant. This escalating power consumption explains also while many Data Centers are located in proximity of large water sources, like lakes, and in predominantly cold regions.

5.2. IoT AND IoE

The **IoE** has four main components inextricably connected to each other: physical things, data, processes, and people.

1. The Internet of Things (IoT) is the **global network of physical objects** accessed through the Internet and incorporating the infrastructure for the internet-connected world of devices-objects-things. As it was stated before, the IoT is growing into the globally interconnected world where everything is connected. World. The IoE includes the following ITRS 2.0 application areas: mobile products, big data systems, the Cloud 2.0, biomedical products, transportation components and subsystems and interconnect.

2. **Data** that are generated by all of the devices-objects-things in item 1 above.

3. **Smart applications for processing** in a timely manner the data generated by the IoT to **deliver the right information to the right machine or person** at the correct time and thereby for solving problems (societal, economic, environmental, and the like) for industry and governments.

4. **Application software/programming** interfaces that connect people in more relevant and valuable ways

IoE

Categories	Year	2015	2017	2019	2021	2023	2025	2027	2029
Power	Energy source (B = battery; H = energy harvesting)	B	B	B + H	B + H	B + H	B + H	B + H	B + H
	Lowest VDD Used By Components (V)	0.8	0.75	0.7	0.65	0.65	0.55	0.45	0.45
	Deep suspend current of MCU (nA)	100	72	52	38	27	20	14	10
	Conversion efficiency of DC-to-DC Conversion (%)	80%	82%	86%	88%	89%	91%	93%	95%
	Spatial Power Density of DC Converter (W/mm^2)	1	1.17	1.36	1.59	1.85	2.16	2.52	2.94
	Peak Current Consumed by Connectivity Interface (mA)	50	19.28	7.44	2.87	1.11	0.43	0.16	0.06
	Transmission Power per bit (µW/bit)	2.48	0.972	0.381	0.149	0.058	0.023	0.009	0.004
Form factor	Module footprint (mm2)	500	500	280	179	115	73	47	30
Performance	MCU Number of Cores	1	1	1	1	1	1	1	1
	MCU Current / Operation frequency (mA/MHz)	30	21.7	15.7	11.3	8.9	7.7	6.7	5.8
	Max MCU Frequency (MHz)	200	235	277	306	316	327	338	350
	MCU Flash Size (KB)	1024	1024	2048	4096	4096	8192	8192	8192
	MCU Dhrystone MIPS (DMIPS)	200	242	293	354	429	519	628	759
Peripheral	Number of Sensors Integrated to System	4	8	10	12	12	13	13	13
	Max Sensor Power (µW)	2850	1397	1009	729	617	522	442	374

Devices utilized in the IoE environment need to consume the list amount of power in order to insure extended operability.

This requires operation at progressively lower supply voltage whether this is provided by battery sources or by scavenging energy from the environment. It is very challenging to operate CMOS at 0.5V to 0.4V. This requirement is driving device research towards new transistor like Tunnel transistors (TFET) that are expected to operate at voltages as low as 300mV or even lower.

This requirement for lower power consumption is driving completely new device specifications. This is one of the fundamental changes in methodology introduced by ITRS 2.0.

For similar reasons, transmission power for bit needs to be reduced by about 3 orders of magnitude in the time horizon. Module footprint remains a major challenge, as more ICs or other devices need to be assembled in a limited space (e.g., smart phone). MCU frequency is expected to only moderately increase due to the limitation imposed on power consumption both during operation and in standby. The number of sensors [Fig. 5.2] will easily increase by 3X while the sensor power will have to be reduced to about 13% of nowadays power consumption for mobility sake.

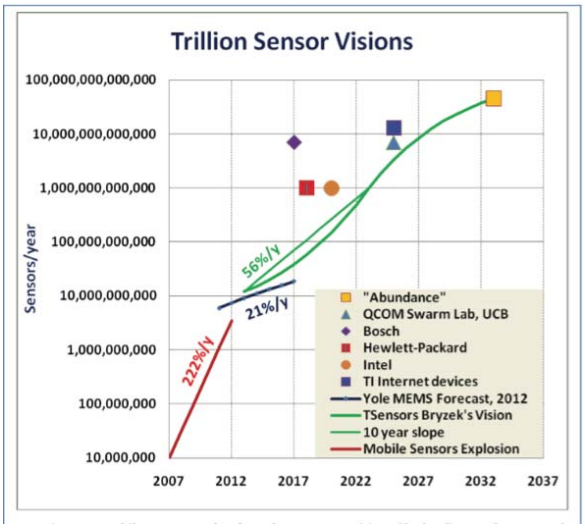


Fig. 5.2 Sensors will populate the world of the IoE

5.3. MOBILITY

In recent years, mobile devices [Fig. 5.3], notably smartphones, have shown significant expansion of computing capabilities. Since smartphone systems are built with multiple heterogeneous ICs (e.g., logic, memory, Microelectromechanical systems (MEMS), and radio-frequency (RF)), it is imperative that the tradeoffs at the system level be fully understood. Beyond the current ITRS SOC-CP roadmap, ITRS 2.0 introduces a new mobile driver to comprehend and roadmap metrics at a higher, system level for mobility applications. The number of cores for Application Processors (AP) will moderately increase by 4x in the time horizon. The increase in the number of cores in mobile is much smaller than what is forecasted to occur in the Data center but the extreme restriction on power consumption is responsible for this slow rate of increase. It is however relevant to notice that total number of transistors continues to increase (A7~1Billion, A8~2Billion) while the die size is actually decreasing (A7 to A8, 13% die size reduction) to accommodate for the limited space available in the phone platform [Fig. 5.4, Fig. 5.5].

Fig. 5.3 The wireless world

Mobile

	Year	2015	2017	2019	2021	2023	2025	2027	2029
Input Metrics	Number of AP cores	4	9	18	18	28	36	30	25
	Number of GPU cores	6	19	49	69	141	247	273	303
	Max frequency of any Component in System (GHz)	2.7	2.9	3.2	3.4	3.7	4	4.3	4.7
	Number of Mega pixels in Display	2.1	2.1	3.7	8.8	8.8	33.2	33.2	33.2
	Band Width between AP and Main memory (Gb/s)	25.6	34.8	52.6	57.3	61.9	61.9	61.9	61.9
	Number of Sensors	14	16	20	20	21	21	22	22
	Number of Antennas	11	13	13	14	15	15	15	15
	Number of ICs	7-10	7-10	7-10	7-10	7-10	7-10	7-10	7-10
	Cellular data rate growing ~1.3x/year (MB/s)	12.5	12.5	21.63	40.75	40.75	40.75	40.75	40.75
	Wi-Fi data rate evolving with standards (Mb/s)	867	867	867	7000	7000	28000	28000	28000
Output Metrics	PCB area of main Components (cm2)	62	69	76	84	93	103	103	103
	Board power averaged at ~7%/year (W)	4.2	4.64	5.12	5.64	6.22	6.86	7.56	8.48

On the other hand, the number of cores in Graphics Processing Units (GPU) will increase by 50x in this time horizon. This increase in GPU processing capacity is necessary to keep up with the increasingly growing number of megapixel offered by the displays. The number of megapixels will in fact increase by 15x in the future. It is expected that the traffic between AP and memory will have to correspondingly increase more than 2x.

Sensors associated with mobile devices will about double in the time horizon but the total number of major ICs will need to remain almost constant at around 10 units or below due to the limitations of space and power.

Cellular data will continue to grow at about 1.3x/year. This due to the fact that users are becoming more and more familiar to utilizing the phone for data acquisition and manipulation.

According to the announced standards, ranging from the present 802.11ac to the future Wireless HD 1.1, the Wi-Fi data rate will grow at ~1.4x/year. Power growth will have to be limited but some growth cannot be avoided.

The whole cell phone mobility relies on the existence of a cell within which the phone can operate. The size of the cell depends on

1. The Application Processor (AP) power level
2. The protocol used (e.g., 802.11a/b/g/n/ac etc.)
3. The Data rates that are allowed

As an example 802.11g standard works in the 2.4 GHz band (like 802.11b) and it operates at a maximum physical layer bit rate of 54 Mbit/s exclusive of forward error correction codes, or about 22 Mbit/s average throughput. An access point compliant with802.11g using the stock antenna might have a range of 100 m (330 ft.).

The maximum power depends on country regulations and on the availability of channels. As an example depending on the specific regulatory class the power limits at 5GHz may range from 40mW to 1000mW according to Federal Communication Commission (FCC). In essence, the mobile devices will have to rely very heavily on the surrounding infrastructure to be able to communicate with the IoE.

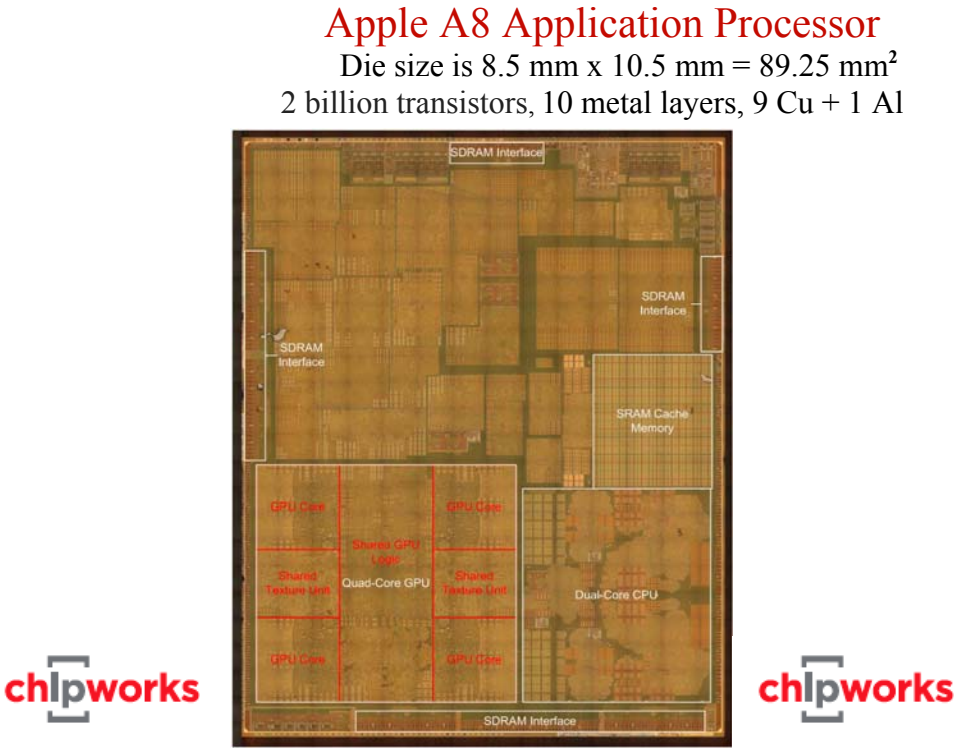


Fig. 5.4 A8 key features and die layout

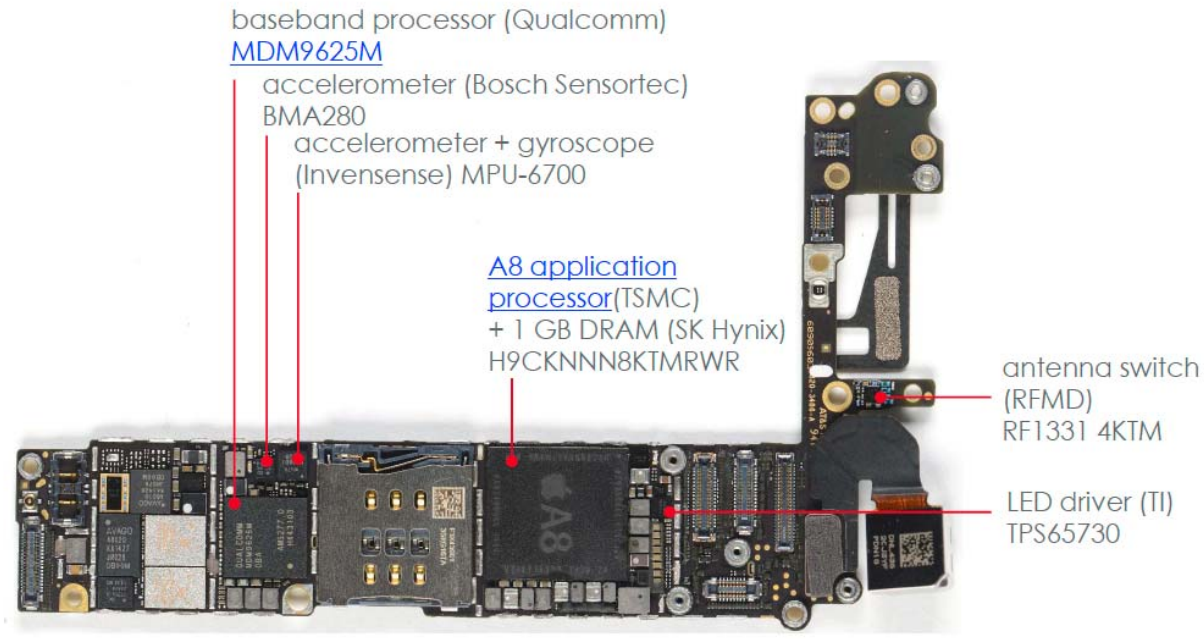


Fig. 5.5 iPhone 6 Application Processor and phone layout

6. HETEROGENEOUS INTEGRATION

Heterogeneous Integration refers to the integration of separately manufactured components into a higher-level assembly that, in the aggregate, provides enhanced functionality and improved operating characteristics.

The mission of the Heterogeneous Integration (HI) Focus Team is to provide guidance to industry, academia and government to identify key technical challenges with sufficient lead-time that they do not become roadblocks preventing the continued progress in electronics that is essential to the future growth of the electronics industry and the realization of the promise of continued positive impact on mankind. The approach is to identify the requirements for heterogeneous integration in the electronics industry through 2030, determine the difficult challenges that must be overcome to meet these requirements and, where possible, identify potential solutions.

The focus for Heterogeneous Integration (HI) is identification of the difficult challenges and the potential solutions for meeting those demands for the next 15 years. The primary integration technology for the potential solutions will be complex, 3D System in Package (SiP) architectures.

System in Package (SiP) is a combination of multiple active electronic components of different functionality, assembled in a single unit, which provides multiple functions associated with a system or sub-system [Figure 6.1]. A SiP may optionally contain passives, MEMS, optical components, and other packages and devices.

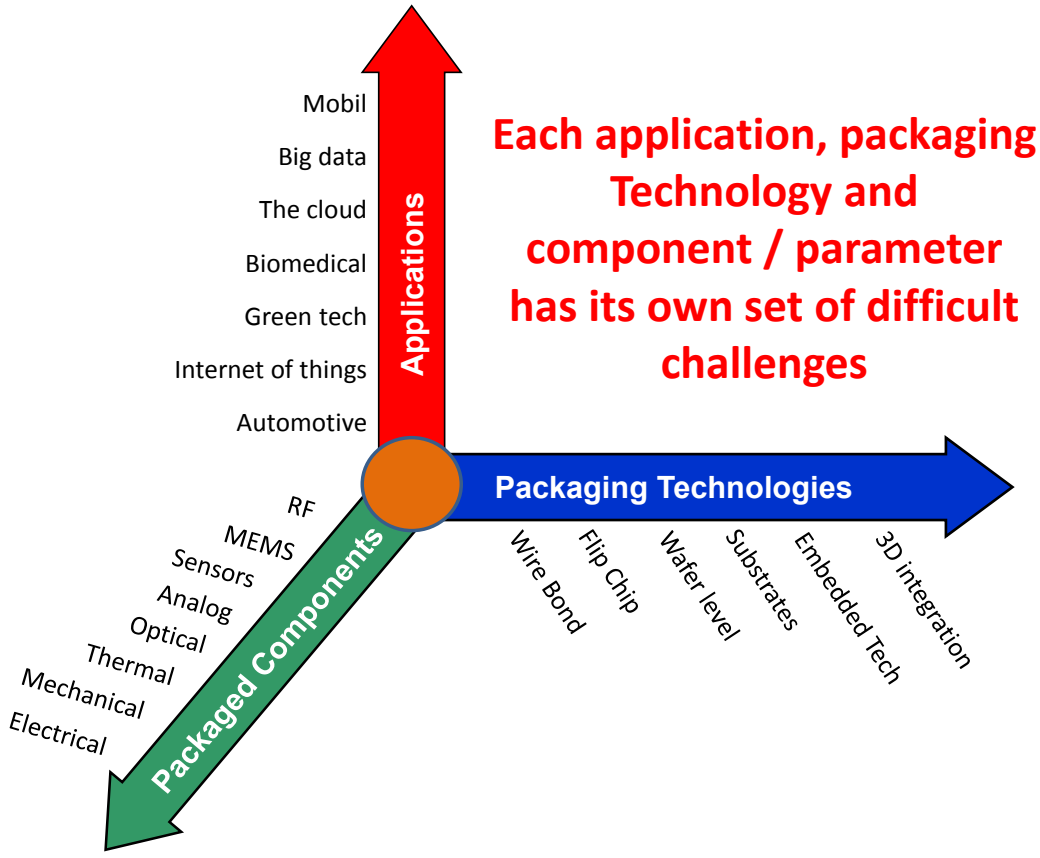


Fig. 6.1 Elements incorporated into complex SiP packages through heterogeneous integration

There are several types of SiP solutions as shown in Figure 6.2. Each of these solutions applies to a different application. For instance combining stacking and side-by-side multichip module Ball Grid Arrays (BGAs) are used across the PC, consumer and communication applications for their low cost, design flexibility and proven technology. Particularly important would be the integration of several different semiconductor technologies, from different foundries onto a single package.

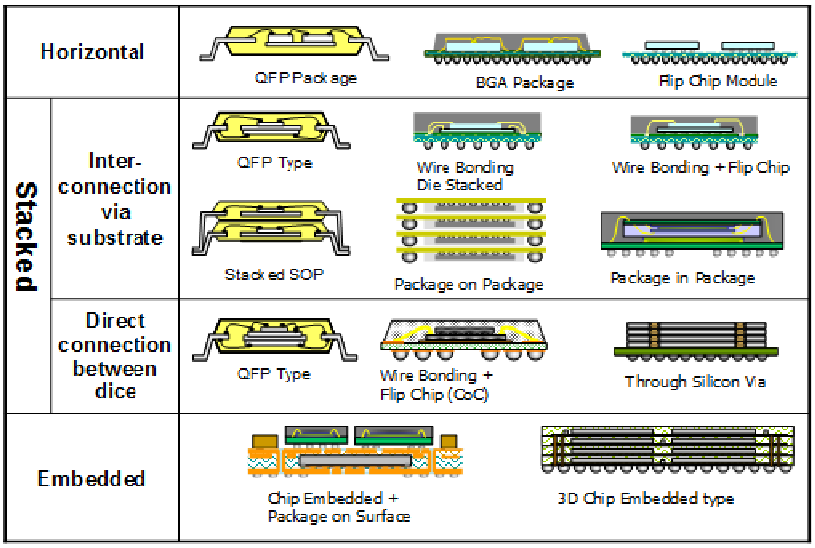


Fig. 6.2 Categories of SiP

The state of the art today replaces the wire bond in the staggered die stack with stacked die using TSV interconnects between die. The specification for the device using the TSV based package has higher speed, decreased package volume, decreased latency and has reduced operating power 50% (Toshiba data). The specification for this SiP is in the Table 1 below.

Package Type		NAND Dual x8 BGA-152	
Storage Capacity (GB)		128	256
Number of Stacks		8	16
External Dimension (mm)	**W**	14	14
	D	18	18
	H	1.35	1.90
Interface		Toggle DDR	

Table 1 Specification for SiP

Besides the comparison with SoC, SiP is often compared with multiple monolithic packages for a new subsystem or system level product package. SiP has often been considered more costly than the corresponding combination of monolithic packages. It is sometimes true when only the packaging costs are compared. But the total cost of the system building and the accompanying advantages must be taken into account. SiP offers customers benefits of smaller footprint of devices and fewer numbers of connecting traces between devices, which results in lower number of PCB layers.

The heat generation from an IC is highly non-uniform with areas of very high local heat fluxes at few locations on the die. Future trends show an increase in thermal design power and an increase in both average power density and local power density (also known as “hot spots”). Hot spot thermal management limits the thermal solution for the package. Even when total power of the component remains within design specification, the hot spot power density increase could limit the device performance and reliability.

2.5D and 3D Package Integration

The technical challenges of 2D scaling in CMOS technology are becoming evident as we reach the practical limits of planar features and new solution utilizing in a variety of ways the third dimension are becoming available at both the device and the package level.

6.1. 2.5 D INTEGRATION

The use of silicon interposers for advanced packaging, first introduced as a product by Xilinx, is now in use for 2.5D integration [Fig. 6.3]. The introduction of product with 2.5D integration used a TSV interposer to connect multiple chips side by side and provide high bandwidth connections between the die, RDL to map the die geometries to the printed circuit board geometries and TSVs to connect the top and bottoms layers of the interposer. The benefits of higher bandwidth, lower power and reduced latency are compelling and several other products are in development using this technology.

The initial selection of this packaging technology was driven by the need for improved performance and it was very successful in achieving that goal. Nevertheless the widespread adoption of this technology has not followed. The primary limitation was the high cost of the interposer incorporating TSVs. Driving the cost down for the 2.5D technology is a critical challenge since it not only delivers major performance advantages, it also drive the industry down the learning curve for the cost effective adoption for 3D-TSV.

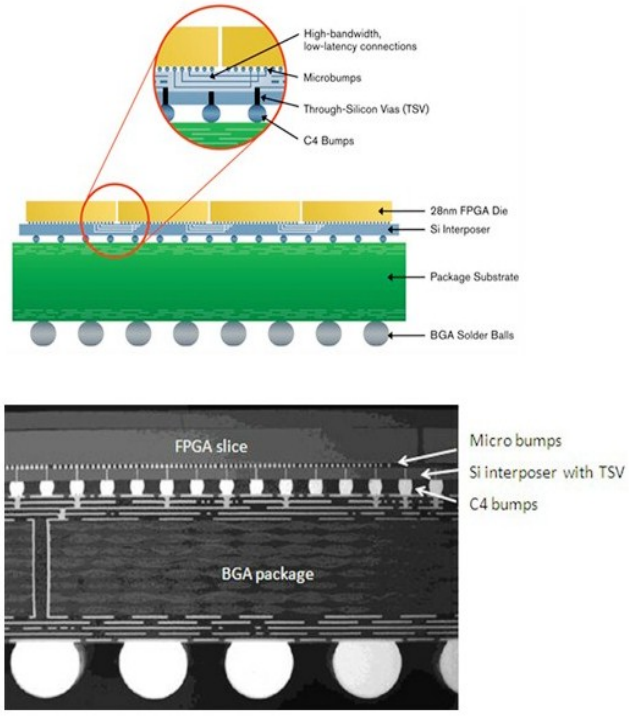


Fig. 6.3 2.5D Interposer for Xilinx FPGA

6.2. 3D INTEGRATION

The effort to continue the pace of progress for the semiconductor industry has been focused on two technical directions. Functional diversification is referred to as “more than Moore” and it is one technical direction. It is the driving force behind System-In-Package (SiP) architecture. The second is the use of the third dimension. This allows increase in functional density for a given node and also improves performance and reduces power requirements. There are several driving forces for 3D integration [Fig. 6.4]. The potential benefits include higher performance, reduced power requirement, reduced latency, smaller size and eventually lower cost compared to 2.5D and conventional 2D packaging.

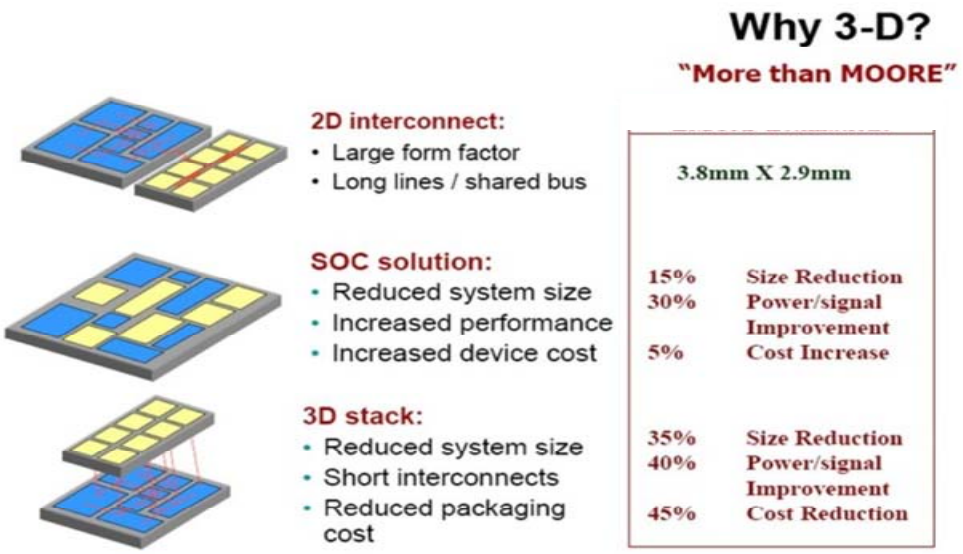


Fig. 6.4 Driving Forces for 3D Integration

6.3. WAFER LEVEL PACKAGING

Wafer Level Packaging (WLP) has been defined as a technology in which all of the IC packaging process steps are performed at the wafer level. The original WLP definition required that all package IO terminals be continuously located within the chip outline (fan-in design) producing a true chip size package.

This definition described a Wafer Level Chip Scale Package, or WLCSP, with the processing of a complete silicon wafer. From a systems perspective, under this definition, the limitation on WLP was how many I/O could be placed under the chip and still have a board design that can be routed.

However, there are products coming to market that do not fall under this earlier definition of WLP. These new packages have been described as "Fan-out" WLP. They are constructed by placing individual sawn die into a polymer matrix or silicon carrier that has the same form factor as the original silicon wafer. These "Reconstituted" artificial wafers are then processed through all of the same processes that are used for "real" silicon wafers, and finally sawn into separate packages.

WLCSP are mainly being used in portable consumer markets where small size, thickness, weight, and electrical performance are additional advantages to cost. Major trends include work for cost efficient rerouting with multi-layer RDL and improved design and simulation tools for WLP technologies.

With the introduction of TSVs, IPDs, Fan-out, and MEMS packaging technologies, WLP products can be used in a much broader range of applications, with higher I/O counts, and greater functional complexity. These packaging technologies open new opportunities for WLPs in the packaging field.

WLP now incorporates many different structures to meet specific application targets. A variety of WLP types are shown in the Figure 6.5

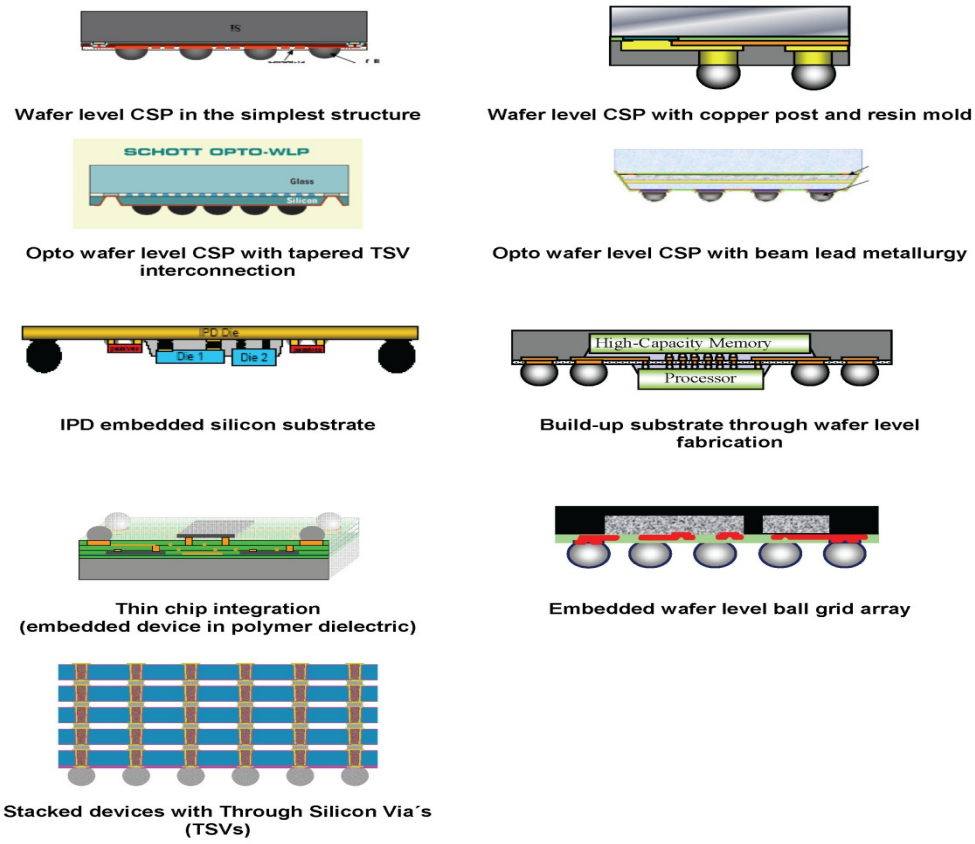


Fig. 6.5 Examples of existing Wafer Level Packaging Types

6.4. DIFFICULT CHALLENGES FOR SIP

The major challenges for digital systems today are power requirement and integrity, thermal management and the limitations in physical density of bandwidth.

The total power requirement can be reduced by dropping the operating voltage, reducing the interconnect distance, reducing the capacitance and reducing the operating frequency. Stacking the circuits in layers will reduce the interconnect length by approximately the square root of the number of layers.

In 3D integration the thermal density is increased and the surface area to exhaust the heat is reduced. Several solutions are described in the HI Chapter.

As the density and performance of the logic circuits are increased and the physical area available to address the resulting increase in demand for bandwidth is reduced new approaches are necessary. There have been proposals to bring photonic data transmission to the IC. These proposals may address the bandwidth but the wavelength employed for photonic data transmission is many times larger than the transistors.

The solution is to bring photonics to a direct bandgap semiconductor circuit in a SiP package and convert to a very wide electrical bus on the package. The photonic connection must us wavelength multiplexing in a single mode fiber to reach the physical density of bandwidth required for future SiPs.

6.5. DIFFICULT CHALLENGES FOR HETEROGENEOUS INTEGRATION

Several key limitations faced by the Semiconductor industry in the near term will involve most, if not all, of the Technical Working Groups (TWGs) and Focus Teams. These include:

- Manage the power and thermal dissipation requirements through both reduction in power requirements and improving the heat dissipation capability of packages.
- Increasing the physical density of bandwidth in order to make use of the enormous gains in the physical density of processor power.
- Support the growing functional diversity requirements driven by "More than Moore" technologies.
- Support the reliability, power integrity and thermal management challenges of 3D integration.
- Drive down cost in assembly and packaging to reduce the impact of packaging cost not scaling to match device cost.
- Reduce time to market and by co-design and simulation that includes electrical, thermal, mechanical and, in some cases chemical, requirements for the device, package and system.
- Support reliability in the face of transistors that will wear out.
- Support the global network changes required to handle the coming yottabyte (10^{24}) level data traffic requirement.
- Developing designs, equipment, materials and processes to support the introduction of 450mm wafer production.

6.6. HETEROGENEOUS COMPONENTS

The advent and fast growing market demand of hand-held mobile devices, such as smartphones and tablet computers, has driven components manufacturers to produce miniaturized components with ever increasing performance and functionality. Their performance requirements are driven by systems evolution or new "killer apps" that can emerge unexpectedly into the scene. While the market for smartphones and tablets is expected to continue to grow with continuing advances in product performance and functionality, new markets are also emerging including wearables and many other applications related to the Internet of Everything. These applications fuel requirements, development, and production of Heterogeneous Components [Fig.6.6]. Heterogeneous components can have use in multiple applications, such as automotive and portable consumer devices, but each of the applications often has different device performance requirements. For example, automotive applications require accelerometers with high reliability and use in extremes in environmental conditions yet accelerometers in mobile devices require low cost and low power consumption. Thus a main requirement for low cost could be traded off with other requirements such as long-term reliability or accuracy. Device performance requirements will need to be developed for each application area that is considered in this roadmap. In summary, the development of a roadmap for Heterogeneous Components is intrinsically much broader and more fragmented than most of the traditional roadmaps outlined in other paragraphs.

The ITRS established the MEMS Technology Working Group in 2010 and published the first edition of the MEMS chapter of the ITRS Roadmap in 2011. The MEMS Chapter included accelerometers, gyroscopes, inertial measurement unites, microphones, and RF MEMS resonators, galvanic switches, and varactors. As the ITRS transitions its roadmapping activities to ITRS 2.0, the MEMS TWG has transitioned its efforts to support the Heterogeneous Components Focus Area (HC FT) of ITRS2.0. This provides an opportunity to expand the range of applications and device technologies that were originally addressed in the MEMS Roadmap.

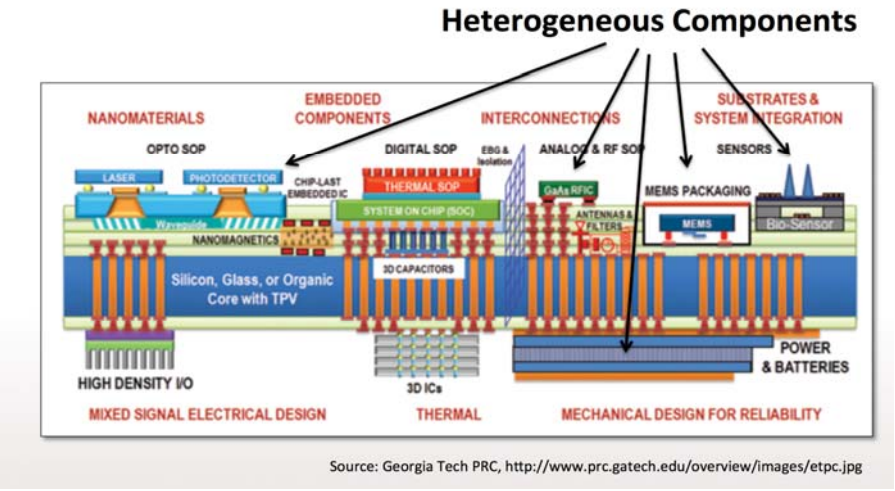


Fig. 6.6 Illustration of Heterogeneous Components in a Heterogeneous System.

A major challenge to roadmapping MEMS technology has been the diversity of applications for MEMS. For example, MEMS devices include pressure sensors, ink jet printers cartridge, accelerators, digital projectors, bolometers, gas sensors, surgical tools, microphones, portable medical diagnostic systems, and more. Furthermore, there has been a long history of MEMS of a one-device one-application paradigm, where each device had a unique manufacturing process. In order to address these challenges, the working group chose to set focus on applications related portable, wereable, and automotive.

6.7. AUTOMOBILES

As automobiles become more complex, signal processing technologies are increasingly being used to create the automobile of the future. This future car will need to adapt to changes in driving conditions, provide driving directions, keep in touch with the office and family members, and provide quality audio and video entertainment—all while providing more safety and running more efficiently than ever before. No easy task.

The most common application beyond airbag systems is the almost ubiquitous Automatic Braking System (ABS). Until very recently, most ABS systems did not use an inertial sensor. They simply read wheel speed and apply pulsed braking if the wheels are thought to be skidding. However, most all-wheel-drive systems and some newer high performance ABS systems, look at longitudinal acceleration to determine if the chassis is still moving. This is particularly important for all-wheel-drive equipped vehicles where all four wheels may have lost traction due to the application of drive torque. Electronic Stability Control is another application for MEMS sensors that assists the driver in regaining control of the automobile just as it is starting to skid.

MEMS Gyros used for roll over sensing do not require the same resolution as those used in ESC systems, but they must have excellent rejection of external shock and vibration and have a larger dynamic range. MEMS gyroscopes are now commonly used in this application because of their insensitivity to external shock and vibration.

6.8. CONSUMER PORTABLE

The release of the Nintendo Wii in November 2006 can be considered to mark the beginning of consumer portable MEMS device applications. The Wii wireless controller incorporated a 3-axis MEMS accelerometer that determined motion and position of the controller, bringing a new dimension to game playing applications. The remote allows the user to interact with the console using gestures and by pointing at the screen. The accelerometer and optical sensor that are built into the remote enable this functionality.

Table 2 Examples of heterogeneous components with their key attributes for automotive, smart phone, wearable, and medical applications, exemplifying how their requirements may differ with application.

	Automotive	Smart Phone	Wearable	Medical
Inertial Sensors (accelerometers, gyroscopes and IMUs)	Crash Sensors, Inertial Positioning: reliability, accuracy, large ambient temperature range	Tilt, Movement, Inertial Positioning: low cost, low power.	Movement sensor, heart rate – low cost, low power.	Movement, tilt – reliability, low power, FDA approval
Microphones	Hands free communications	Voice, noise cancellation, ultrasonic communications	Voice, heart rate, ultrasonic communications – low cost, low power	Voice, heart rate - reliability, low power, FDA approval
Pressure Sensors	Tire, manifold, altitude, vapor pressure	Altitude, barometric pressure	Blood pressure – low cost, low power	Blood pressure – reliability, low power, FDA approval
Micro Speakers		Sound reproduction – small size, low power	Sound reproduction – small size, low power	
Conductivity			Perspiration	
Humidity			Ambient environmental conditions	
eNose		Environmental and Health monitoring	Environmental and Health monitoring	

Apple introduced the iPhone in June 2007. Analogous to how the Wii remote revolutionized gaming, the iPhone can be considered to have revolutionized mobile phones. The iPhone advanced the functionality of mobile telephones by providing a more advanced graphical user interface with Internet browsing and email, among other things. iPhone has continued to evolve and utilize a broad variety of MEMS [Fig. 6.7]. For instance, the MEMS accelerometer technology detected the direction of gravity, which enabled the display to rotate so that it was always kept upright, and also provided an interface to game applications that could be purchased from the "app store."

The MEMS technologies that supported these consumer portable applications did not require the same levels of accuracy and reliability as the automotive applications from which the technologies evolved. The primary drivers for these applications were cost, size and low power dissipation.

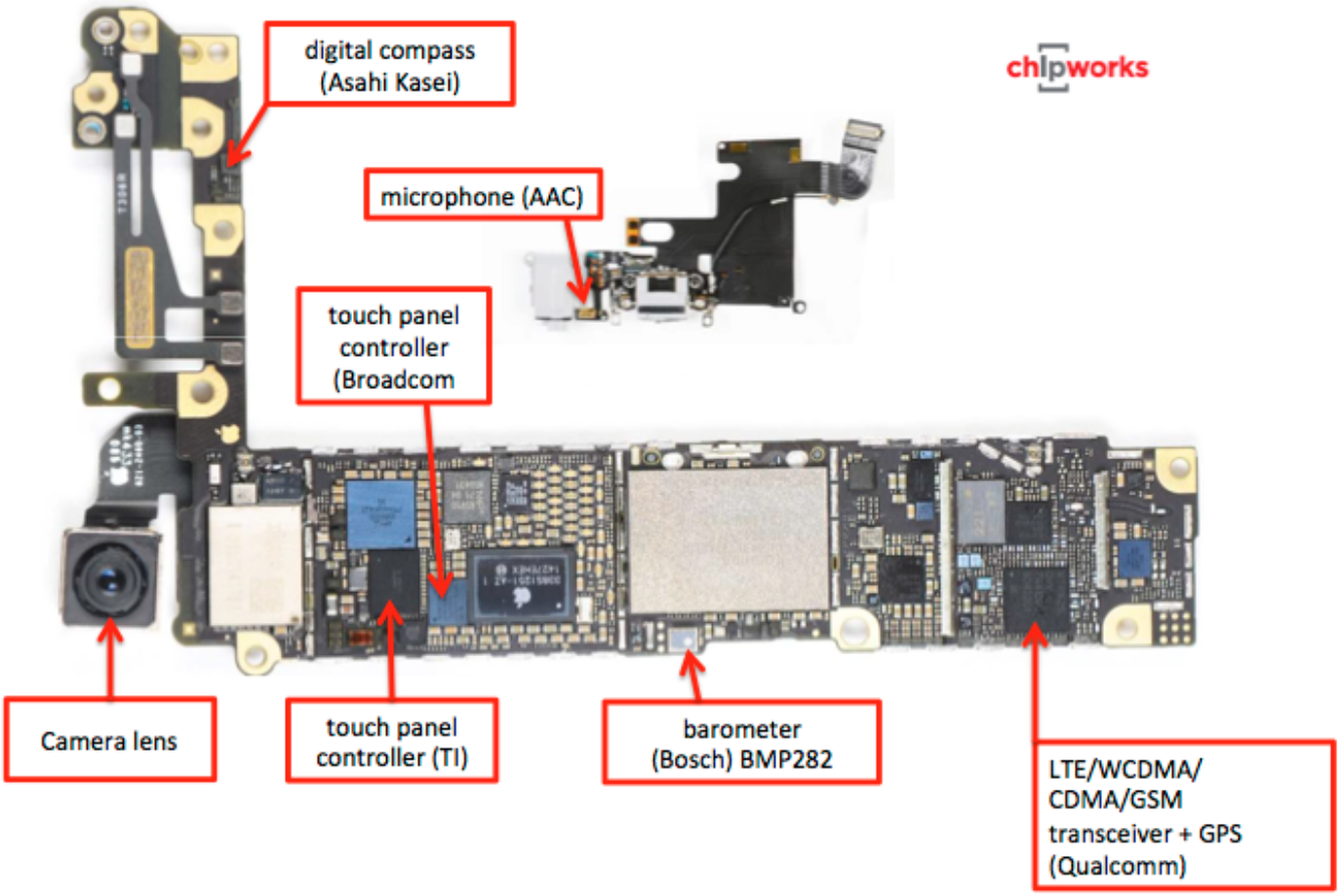


Fig. 6.7 Sensors bridge the world of Electronics with the world of humans

Another significant component of the consumer MEMS evolution was in the area of packaging. Previous MEMS products, mostly automotive sensors, had a cost structure that allowed the use of mechanically robust, open cavity packages, such as ceramic DIP packages. The benefit of such packages was that the MEMS device could be mounted in an elastically isolated way, decoupling it from package induced stresses. However, such package technologies were much too expensive for the consumer market. Methods for using low cost, plastic packages were required to meet cost targets. This lead to technologies for capping the MEMS device, to allow plastic over-molding and the development of sophisticated die attach and stress relief methods.

6.9. CONSUMER WEARABLE AND HEALTHCARE

The rapid increase of digital data and connected technologies is revolutionizing healthcare. The healthcare system used to be highly centralized, disease oriented and focused on acute care. It is changing today towards keeping people healthy, raising each individual's awareness of their health and inducing efficient behavioral changes. As they become empowered to maintain a healthy lifestyle, a large portion of them is eager to collect data about their health, track trends over time and share their health performance on their social network. Others could have a mild condition that can benefit from continuous monitoring. This growing part of the population with the desire to monitor lifestyle and health is often called the Worried-well. Worried-wells are information seekers. In the last decade, worried-wells were searching the web for symptoms they would experience, looking to correlate them with possible health conditions. In the last few years, websites have been introduced to connect people with similar conditions or symptoms (e.g. patients-like-me). In the future, the worried-wells will have a new army of technologies at their disposal to monitor and improve their health. Silicon and MEMS technologies are making that revolution possible.

Today the mobile phone can already provide a great deal of health information. Accelerometers can track activity and sleep. Built-in optical sensors are available that can sense heart rate when the user is touching the phone. The camera in the phone can be used for purposes as diverse as checking the calorie content of a food item, or identifying your emotions

based on facial expression recognition. A broad spectrum of mobile phone apps has been developed to analyze this data, and deliver it to the consumer in an intelligible and actionable manner. A recent survey reported over 16,000 consumer health apps available on the Apple store as of August 2013 and that 66% of Americans would use mobile health apps to manage their health.

In summary, MEMs have found a key role by means of heterogeneous integration in a variety of applications in the new ecosystem of the electronics industry [Fig. 6.8].

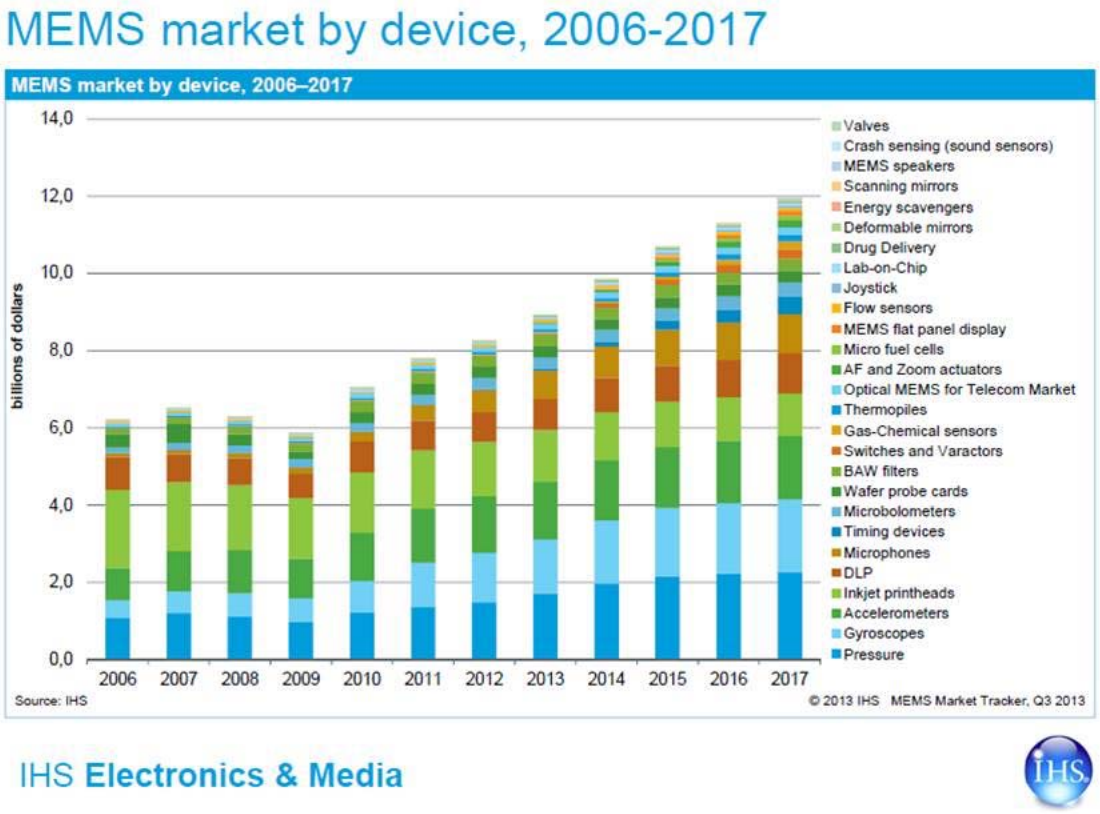


Fig. 6.8 The multi-faceted market of MEMS

7. CONNECTIVITY

Mobile devices and systems are connected to the internet primarily through RF wireless communications such as WiFi. Communication between fixed systems is achieved through a combination of RF wireless, copper wiring and photonic interconnects. Information processing systems are physically limited by both energy dissipation and communication of information, especially in dense, high-speed systems. Overall, the majority of energy dissipation in systems is in the interconnections and communication parts of any system. Photonic interconnects, consisting mainly of lasers and fibers, offers solutions to both of these problems, enabling lower energy communication and higher information densities, especially for all longer distance connections beyond the chip itself.

7.1. RF & AMS WIRELESS

Radio frequency (RF), high frequency (HF), and analog/mixed-signal (AMS) technologies serve the rapidly growing communications markets that include many of the physical components for the Internet of Everything (IoE) and represent essential and critical technologies for the success of many semiconductor manufacturers. RF wireless technology is one of the simplest ways to connect IoT devices to the Internet through WiFi, Cellular or proprietary wireless connections. Thus, the number of devices employing wireless RF is expected to increase dramatically over the life of the roadmap. This will require high volumes of energy efficient devices and ICs with RF and AMS capabilities. Furthermore, more sophisticated antennas will be needed to connect to multiple networks in compact systems, such as mobile phones, etc.

Communications products and emerging products with functionalities enabled by more-than-Moore (MtM) RF, HF, and AMS technologies are becoming key drivers for volume manufacturing. Consumer products account for over half of the demand for semiconductors. Fourth generation (4G) cellular phones and tablets now have a much higher RF and AMS semiconductor content and now are a very large fraction of the mobile market compared to only 5 % of the market a few years ago. With different technologies capable of meeting technical requirements, time to market and overall system cost will govern technology selection.

The requirements for transceiver integrated circuits (ICs) are technology drivers that contribute substantially to the recent ITRS-defined More-than-Moore (MtM) thrust. This 2016 ITRS RF and AMS Section is divided into the four analog-carrier frequency bands – low frequency (LF) less than 0.4 GHz, radio frequency (RF) 0.4 GHz to 30 GHz, millimeter-wave (mm-wave) 30 GHz to 300 GHz, and terahertz (THz) greater than 300 GHz. The table below lists a few examples of applications for each of these bands.

Analog - Carrier Frequency Bands			
LF Analog (0.0 GHz-0.4 GHz)	**RF (0.4 GHz-30 GHz)**	**Millimeter-Wave (30 GHz-300 GHz)**	**THz (> 300 GHz)**
Example Applications for Each Frequency Band			
Automotive controls	**Cellular**	**60 GHz point-to-point**	
On-chip regulators	**WLAN**	**Touchless gesture control, Imaging, sensing**	
Power management	**Serializer/ Deserializer**	**Automotive radar**	
	ADC,DAC	**Wireless backhaul**	

The IoT-enabled applications listed above are part of the scope for the RF and AMS ITWG and currently lag technology and processing capabilities for reliable manufacturing similar to digital CMOS. Although no applications are currently identified for the THz band, capabilities are being developed and potential applications are being investigated in research.

7.2. PHOTONIC TRANSMITTERS

The most commonly used optical transmitters are semiconductor devices such as **light emitting diodes** (LEDs) and **laser diodes**. The difference between LEDs and laser diodes is that LEDs produce incoherent light, while laser diodes produce coherent light. For use in optical communications, semiconductor optical transmitters must be designed to be compact, efficient, and reliable, while operating in an optimal wavelength range, and must be capable of being directly modulated at high frequencies.

7.3. FIBER-OPTIC COMMUNICATION

Fiber-optic communication is a method of transmitting information from one place to another by sending pulses of light through an optical fiber. The light forms an electromagnetic carrier wave that is modulated to carry information. First developed in the 1970s, fiber-optics revolutionized the telecommunications industry and played a major role in the advent of the Information Age. Because of its advantages over electrical transmission, optical fibers have largely replaced copper wire communications in core networks in the developed world. Fiber-optics are used by many

telecommunications companies to transmit telephone signals, Internet communication, and cable television signals. Researchers at Bell Labs have reached Internet speeds of over 100 petabit×kilometer per second using fiber-optic communication.

The process of communicating using fiber-optics involves the following basic steps: Creating the optical signal involving the use of a transmitter, relaying the signal along the fiber, ensuring that the signal does not become too distorted or weak, receiving the optical signal, and converting it into an electrical signal.

The two main advantages associated with the use of optics are:

1. **Low power loss over distance**. Attenuation of single mode optical signals is **< 0.3db/Km** over hundreds of Km.
2. **Very high information density**. A single mode 0.125 mm diameter optical fiber can transmit information **>100 Exabits/s/mm^2**. (This implies: A bandwidth of ~200Terahertz, spectral efficiency of 10 bits/Hz and a fiber diameter of 125 microns.)

In addition, in the future, optical technology could offer new possibilities for information processing at low or even essentially zero power; however, significant research is still needed for both materials and devices capable of optical amplification and logic functions

Office & Factory LAN

Year	2015	2017	2019	2021	2023	2025	2027	2029
LAN to ~10Km Max Data Rate/ Wavelength (Gb/s)	40	100	400	400	1000	1000	2000	2000
LAN Mode	Multi /single-mode	Multi /single-mode	Multi /single-mode	Multi /single-mode	Single-mode / multimode	Single-mode / multimode	Single-mode / multimode	Single-mode / multimode
LAN Wavelengths	4	4	4	4	8	8	8	8
Physical Modulation Method (direct or external)	Direct / external	Direct / external	Direct / external	Direct / external	Direct / external	Direct / external	Direct / external	Direct / external
Laser	VCSEL/ edge emitters	VCSEL/ edge emitters	VCSEL/ edge emitters	VCSEL/ edge emitters	VCSEL/ edge emitters	VCSEL/ edge emitters	VCSEL/ edge emitters	VCSEL/ edge emitters
LAN wavelength spacing (nm)	20	20	20	20	10	10	10	10
LAN Energy Consumption (pJ/bit)	45	25	20	20	10	10	10	10

7.4. Local Area Networks

Local Area Networks (LANs) that transmit data 100 meters to a few kilometers utilize much of the basic technology originally developed by the telecommunication industry. LANs generally require data rates of at least 1Gb/s and rarely use more than a few wavelengths in each fiber. While copper has worked well to 1Gb/s, optical methods are more in demand to implement 10Gb/s & 100 Gb/s Ethernet. The 100 Gb/s Ethernet Standard recommends fiber for distances greater than 7 meters. In many cases, the cost to raise capacity is minimized by adding more fiber rather than implementing multiplexing or changing to higher data rate transceivers.

7.5. Long Distance Connectivity

The first transatlantic telephone cable to use optical fiber went into operation in 1988.

Third-generation fiber-optic systems operated at 1.55 μm and had losses of about 0.2 dB/km. Engineers overcame earlier difficulties using conventional InGaAsP semiconductor lasers. Scientists used fiber designed to have minimal dispersion at 1.55 μm or by limiting the laser spectrum to a single mode. These developments eventually allowed third-generation systems to operate commercially at 2.5 Gbit/s with repeater spacing in excess of 100 km.

The fourth generation of fiber-optic communication systems used optical amplification and wavelength-division multiplexing to increase data capacity. These two improvements caused a revolution that resulted in the doubling of system capacity every 6 months starting in 1992 until a bit rate of 10 Tb/s was reached by 2001. In 2006 a bit-rate of 14 Tbit/s was reached over a single 160 km line using optical amplifiers.

The focus of development for the fifth generation of fiber-optic communications is on extending the wavelength range over which a WDM system can operate. The conventional wavelength window, known as the C band, covers the wavelength range 1.53-1.57 µm, and *dry fiber* has a low-loss window promising an extension of that range to 1.30-1.65 µm.

7.6. WAVELENGTH-DIVISION MULTIPLEXING (WDM)

In fiber-optic communications, **wavelength-division multiplexing** (**WDM**) is a technology which multiplexes a number of **optical carrier** signals onto a single optical fiber by using different wavelengths (i.e., colors) of laser light. This technique enables bidirectional communications over one strand of fiber, as well as multiplication of capacity.

The term ***wavelength-division multiplexing*** is commonly applied to an **optical carrier** (which is typically described by its wavelength), whereas **frequency-division multiplexing** typically applies to a **radio carrier** (which is more often described by frequency). Since *wavelength* and *frequency* are tied together through a simple directly inverse relationship, in which the product of frequency and wavelength equals ***c*** (the propagation speed of light), the two terms actually describe the same concept.

Optical Carrier transmission rates are a standardized set of specifications of transmission bandwidth for digital signals that can be carried on Synchronous Optical Networking (SONET) fiber optic networks. Transmission rates are defined by rate of the bitstream of the digital signal and are designated by hyphenation of the acronym **OC** and an integer value of the multiple of the basic unit of rate, e.g., OC-48. The base unit is 51.84 Mbit/s. Thus, the speed of optical-carrier-classified lines labeled as OC-*n* is *n* × 51.84 Mbit/s.

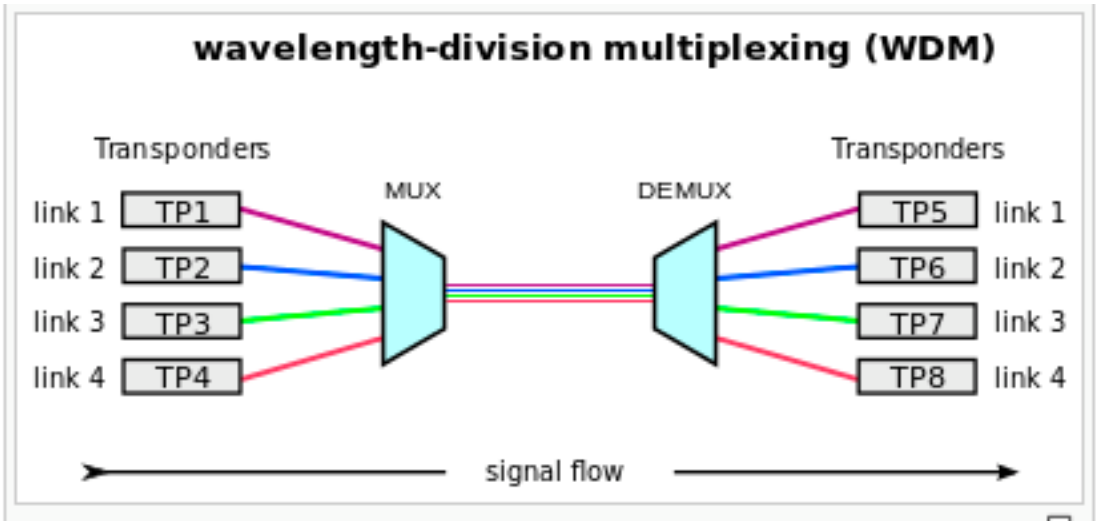


Fig.7.1 WDM schematics

A WDM system uses a multiplexer at the transmitter to join the several signals together, and a demultiplexer at the receiver to split them apart [Figure 7.1]. With the right type of fiber it is possible to have a device that does both simultaneously, and can function as an optical add-drop multiplexer. The optical filtering devices used have conventionally been etalons (stable solid-state single-frequency Fabry–Pérot interferometers in the form of thin-film-coated optical glass).

WDM Module Performance	CWDM4							
Year	2015	2017	2019	2021	2023	2025	2027	2029
System structure								
Data Rate/ Lane (Gbit/s)	25	25	50	50	50	100	100	200
Single Channel GBaud/s	25	25	100	200	200	400	400	800
Distance (km)	<2km	<2km	<10km	<80km	<80km	<80km	<80km	<80km
Number of Wavelengths	1	4	8	8	8	8	16	16
Number of Bits per symbol (HOM)	1	1	2	4	4	4	8	8
Additional link penalty due to HOM (dB)	0	0	5	8	8	8	11	11
Device Performance								
TX I/O loss (dB)	2.5	2.5	2	1.5	1.5	1.5	1	1
RX I/O loss (dB)	4	4	3.5	3	2.5	2	2	2
MUX IL (dB)	0	2	1.5	1.5	1.5	1.5	1.5	1.5
DEMUX IL (dB)	0	2	1.5	1.5	1.5	1.5	1.5	1.5
Target RX Sensitivity (dBm)	-12	-12	-12	-12	-12	-12	-15	-15
Noise Penalty at the receiver side (dB) (25G as reference)	0	0	1.5	3	4.5	6	7.5	9
Laser Output Power (dBm) [3]	13	14	16	16	16	16	16	16
Photo detector BW (GHz)	17.5	17.5	35	35	35	70	70	140
Modulator BW (GHz)	17.5	17.5	35	35	35	70	70	140

HOM: higher order modulation (i.e. PAM Qpsk, etc.)
TX: transmitter path
RX: receiver
TIA: trans impedance amplifier
PD: photodetector
FEC: error code correction
PAM: Phase amplitude modulation
Qpsk: Quadrature Phase-Shift Keying

7.7. TELECOMMUNICATIONS (LONG RANGE COMMUNICATIONS)

Telecommunications will continue to support increasing data volumes both to Data Centers, Offices, and FTTX to support home users. To support this increased volume of data, the data rate per wavelength will need to increase from the current 200Gb/sec-wavelength to 1000Gb/s-wavelength in 2029. The maximum data rate per fiber will increase from 50Tb/s to 250Tb/s in 2023 as shown in Table OSC7, which requires over 100 wavelengths per fiber and modulation of polarization and higher order modulation. The telecommunication industry is leading the development of these technologies.

Fiber to the *x* (FTTX) is a generic term for any broadband network architecture using optical fiber to provide all or part of the local loop used for last mile telecommunications. As fiber optic cables are able to carry much more data than copper cables, especially over long distances, copper telephone networks built in the20th Century are being replaced by fiber. **FTTX** is a generalization for several configurations of fiber deployment, arranged into two groups: FTTP/FTTH/FTTB (Fiber laid all the way to the premises/home/building) and FTTC/N (fiber laid to the cabinet/node, with copper wires completing the connection).

Telecommunications Optical Interconnect Requirements

Year	2015	2017	2019	2021	2023	2025	2027	2029
Data rate/ wavelength single transceiver (Gb/s)	200	400	400	400	1000	1000	1000	1000
Bandwidth efficiency, Bits/ symbol	8	12	16	16	20	20	20	20
Modulation Method	DP-QPSKx 2	DP-16QAM	DP-16QAM	DP-16QAM	DP-64QAM	DP-64QAM	DP-64QAM	DP-64QAM
Max data rate/fiber	50Tb/s	50Tb/s	100Tb/s	100Tb/s	250 Tb/s	250 Tb/s	250 Tb/s	250 Tb/s
Distance between repeaters	80km to 1000km	80km to 1000km	80km to 1000km	80km to 1000km	80km to 1000km	80km to 1000km	80km to 1000km	80km to 1000km
Optical wavelength	1310, 1550/1610	1310, 1550/1610	1310, 1550/1610	1310, 1550/1610	1310,1 550/1610	1310, 1550/1610	1310, 1550/1610	1310, 1550/1610
Single channel optical power	16dbm	16dbm	16dbm	16dbm	16dbm	16dbm	16dbm	16dbm
Total optical power	20dbm	20dbm	20dbm	20dbm	20dbm	20dbm	20dbm	20dbm
# wavelengths/ waveguide	<168	<168	<168	<168	<168	<168	<168	<168
Wavelength spacing, GHz	50/100	50/100	50/100	50/100	50/100	50/100	50/100	50/100
Physical Modulation Method (direct or secondary)	Secondary modulator, PIC	Secondary modulator, PIC	Secondary modulator, PIC	Secondary modulator, PIC	Secondary modulator, PIC	Secondary modulator, PIC	Secondary modulator, PIC	Secondary modulator, PIC
Optical mode; multi/single	single	single	single	single	single	single	single	single

8. MORE MOORE

8.1. OVERALL TRANSISTOR TRENDS

Transistor cost and performance will continue to be strongly correlated to dimensional and functional scaling of CMOS as information processing technology is driving the semiconductor industry into a broadening spectrum of new applications according to 2015 ITRS 2.0.

Strained silicon, high- κ /metal-gate and Multigate transistors are now widely used in IC manufacturing. A significant part of the research to further improve device performance is presently concentrated on III-V materials and Ge. These materials promise higher nobilities than Si devices.

In order to take advantage of the well-established Si platform, it is anticipated that the new high-mobility materials will be epitaxially grown on Si substrate. Beyond implementation of these new materials, the Emerging Research Device (ERD) section reports completely new transistors, operating on new principles like tunneling (e.g. TFET) or spin that offer the possibility of operating at very low power. A lot of progress has been reported on these devices and their introduction into manufacturing appears consistent with the advent of the next decade.

Furthermore, a large variety (like never before) of new memory devices operating on completely new principles are extensively reported and some of them are approaching the pre-manufacturing phase.

FIN FET devices can be simply characterized by observing that these transistors are nothing more than planar transistors rotated by 90 degrees along the source-gate-drain side. It has been already demonstrate that by means of a combination of horizontal scaling and by reducing the number of FINs it is possible to scale linear dimensions of transistors by more than historical 70%. Elimination of one FIN does not have any deleterious effect on performance since making the remaining FINs proportionally taller can compensate the lack of current carrying capability of the missing FIN. On the other hand, FIN FET transistors were primarily introduced to reduce leakage current and for being capable of operating at lower voltages than classical planar CMOS. Low power operation is now the most sought after attribute. Despite this enhanced scaling capability 2D scaling will eventually approach fundamental limits towards the end of this decade and therefore both logic and memory devices are exploring the use of the vertical dimension (3D) as a means of continuing packing more transistors in each cm^2 of silicon.

The combination of 3D device architecture and low power device will usher the (Third) Era of Scaling, identified in short as "3D Power Scaling". Increase in the number of transistors per unit area will eventually be accomplished by stacking multiple layers of transistors

No new breakthroughs are reported for classical interconnections since no viable materials with resistivity below copper exist. However, progress in manipulation of edgeless wrapped materials (e.g., carbon nanotubes, graphene combinations etc.) offer the promise of "ballistic conductors," which may emerge in the next decade.

3D integration of multiple dice offers possible avenues towards reducing interconnect resistance by increasing the conductor cross-section (vertical) and by reducing the length of each interconnect path. For instance, integrating memory device (die) immediately above logic device (die) and connecting them by means of wide through silicon vias (TSV) can accomplish this result.

8.2. MEMORY TRENDS.

Memory technologies will continue to drive the most aggressive pitch scaling and the highest transistor count. ***Memory technologies have been and will always be the drivers of Moore's Law.***

DRAM products are approaching fundamental limitations as scaling DRAM capacitors is becoming very difficult in 2D structures. It is expected that these limits will be reached by 2024 and after this year DRAM technology will saturate at the 32Gbit level unless some major breakthrough will occur.

Flash memory on the other hand will lead the semiconductor industry towards the next revolution in transistor density. By 2020 the 2D Flash topological method will reach a practical limit with respect to cost effective realization of pitch dimensions. Due to the small volume of the floating gate charge amount and charge retention will become fundamental problems as fewer and fewer electrons will be available for the memory function. However, Flash producers are turning the transistors completely vertical in order to increase density while relaxing the dimensions of the storage gate to higher values. By so doing it will be possible to store more charge in the floating gate. Already prototypes of 32 and 48 layers of Flash transistors planes have been announced. This integration technology will become dominant in the next few years and will spread to logic circuits as well. However, the concept of "increased functionality" in the era of IoE is not longer exclusively equivalent to either increased transistor density or performance. New ways of integrating a system have become essential to producing system of higher functionality. The first method consists in extending the functionality of the CMOS platform via heterogeneous integration of new technologies, and the second method consists in stimulating inventions of devices that support new information-processing paradigms

DRAM TECHNOLOGY							
YEAR OF PRODUCTION	2015	2017	2019	2021	2024	2027	2030
Half Pitch (Calculated Half pitch) (nm)	24	20	17	14	11	8.4	7.7
DRAM cell size (μm^2)	0.00346	0.00240	0.00116	0.00078	0.00048	0.00028	0.00024
DRAM cell FET structure	RCAT+Fin	RCAT+Fin	VCT	VCT	VCT	VCT	VCT
Cell Size Factor: a	6	6	4	4	4	4	4
Array Area Efficiency	0.55	0.55	0.5	0.5	0.5	0.5	0.5
V_{int} (support FET voltage) [V]	1.1	1.1	1.1	1.1	0.95	0.95	0.95
Support min. V_{tn} (25C, $G_{m,max}$, V_d=55mV)	0.40	0.40	0.40	0.40	0.37	0.37	0.37
Minimum DRAM retention time (ms)	64	64	64	64	64	64	64
DRAM soft error rate (fits)	1000	1000	1000	1000	1000	1000	1000
Gb/1chip target	8G	8G	16G	16G	32G	32G	32G

NAND Flash							
Year of Production	2015	2016	2020	2022	2024	2028	2030
2D NAND Flash uncontacted poly 1/2 pitch – F (nm)	15	14	12	12	12	12	12
3D NAND minimum array 1/2 pitch -F (nm)	80nm	80nm	80nm	80nm	80nm	80nm	80nm
Number of word lines in one 3D NAND string	32	32-48	64-96	96-128	128-192	256-384	384-512
Dominant Cell type (FG, CT, 3D, etc.)	FG/CT/3D	FG/CT/3D	FG/CT/3D	FG/CT/3D	FG/CT/3D	FG/CT/3D	FG/CT/3D
Product highest density (2D or 3D)	256G	384G	768G	1T	1.5T	3T	4T
3D NAND number of memory layers	32	32-48	64-96	96-128	128-192	256-384	384-512
Maximum number of bits per cell for 2D NAND	3	3	3	3	3	3	3
Maximum number of bits per cell for 3D NAND	3	3	3	3	3	3	3

8.3. LOGIC TRENDS

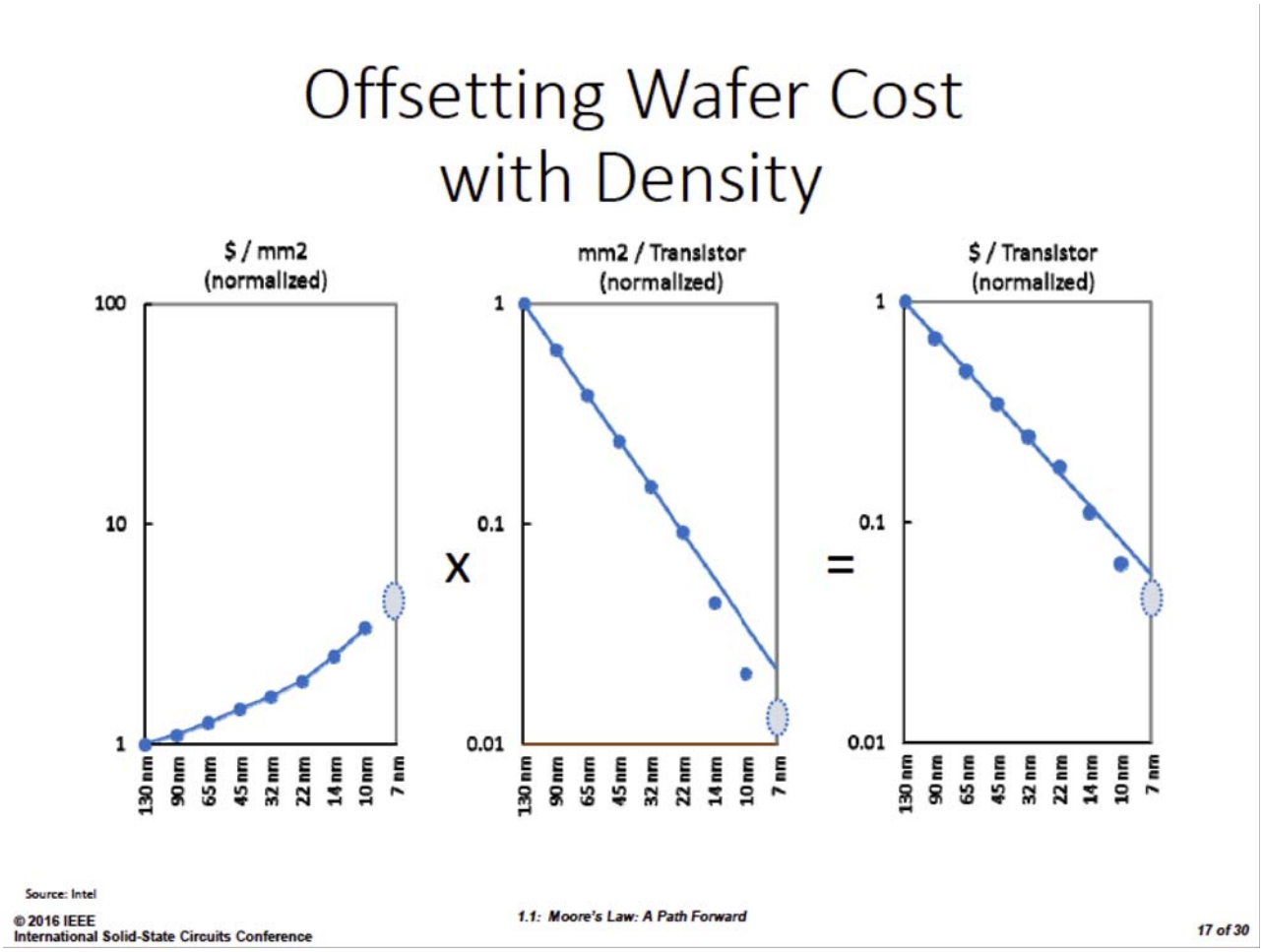
The advent of ubiquitous access to the Internet by means of Wi-Fi communication technologies has led to a vast deployment of mobile products. The need for extended time of operation by means of energy provided by a battery has changed the driver of transistor performance from maximum speed to reduced power consumption. This is accomplished by reducing power supply voltage and by designing transistors with steeper sub threshold slope. However, the need for improved performance is reflected in the drive for Vt reduction of more than 100% in the time horizon for Low (Power) Performance transistors to partially compensate for the reduction if power supply voltage.

Electrical Properties of Transistors

YEAR OF PRODUCTION	2015	2017	2019	2021	2024	2027	2030
Logic device technology naming	P70M56	P48M36	P42M24	P32M20	P24M12G1	P24M12G2	P24M12G3
Logic industry "Node Range" Labeling (nm)	"16/14"	"11/10"	"8/7"	"6/5"	"4/3"	"3/2.5"	"2/1.5"
Logic device structure options	FinFET FDSOI	FinFET FDSOI	FinFET LGAA	FinFET LGAA VGAA	VGAA, M3D	VGAA, M3D	VGAA, M3D
DEVICE ELECTRICAL SPECS							
Power Supply Voltage - Vdd (V)	0.80	0.75	0.70	0.65	0.55	0.45	0.40
Subthreshold slope - [mV/dec]	75	70	68	65	40	25	25
Inversion layer thickness - [nm]	1.10	1.00	0.90	0.85	0.80	0.80	0.80
Vt sat (mV) at Ioff=100nA/μm - HP Logic	129	129	133	136	84	52	52
Vt sat (mV) at Ioff=100pA/μm - LP Logic	351	336	333	326	201	125	125
Effective mobility (cm2/V.s)	200	150	120	100	100	100	100
Rext (Ohms.μm) - HP Logic	280	238	202	172	146	124	106
Ballisticity. Injection velocity (cm/s)	1.20E-07	1.32E-07	1.45E-07	1.60E-07	1.76E-07	1.93E-07	2.13E-07
Vdsat (V) - HP Logic	0.115	0.127	0.136	0.128	0.141	0.155	0.170
Vdsat (V) - LP Logic	0.125	0.141	0.155	0.153	0.169	0.186	0.204
Ion (μA/um) at Ioff=100nA/μm - HP logic w/ Rext=0	2311	2541	2782	2917	3001	2670	2408
Ion (μA/μm) at Ioff=100nA/μm - HP logic, after Rext	1177	1287	1397	1476	1546	1456	1391
Ion (μA/μm) at Ioff=100pA/μm - LP logic w/ Rext=0	1455	1567	1614	1603	2008	1933	1582
Ion (μA/μm) at Ioff=100pA/μm - LP logic, after Rext	596	637	637	629	890	956	821
Cch, total (fF/μm2) - HP/LP Logic	31.38	34.52	38.35	40.61	43.14	43.14	43.14
Cgate, total (fF/μm) - HP Logic	1.81	1.49	1.29	0.97	1.04	1.04	1.04
Cgate, total (fF/μm) - LP Logic	1.96	1.66	1.47	1.17	1.24	1.24	1.24
CV/I (ps) - FO3 load, HP Logic	3.69	2.61	1.94	1.29	1.11	0.96	0.89
I/(CV) (1/ps) - FO3 load, HP Logic	0.27	0.38	0.52	0.78	0.90	1.04	1.12
Energy per switching [CV2] (fj/switching) - FO3 load, HP Logic	3.47	2.52	1.89	1.24	0.94	0.63	0.50

Requirements for power reduction apply also to the High Performance transistors. Similarly, despite these reductions in operating voltage it is expected that Ion for both types of transistors will remain almost constant across the time horizon and this in not an easy result to accomplish. Total channel capacitance will increase by about 30% event though gate capacitance will continue to decrease. Finally transistor speed will continue to increase by a 4X factor for HP transistors enabling 1THz intrinsic functionality. Overall energy per switching will decrease to about 15% from nowadays values.

The introduction of FinFET transistors into manufacturing in 2011 revolutionized the way transistors are built and it resulted essential not only to extending Moore's Law but to also to accelerate the pace of reduction of transistor and pitch dimensions in this decade [Fig.8.1].



From: 2016 ISSCC. "Moore's Law: A Path Forward". Bill Holt, Intel

Fig. 8.1 Increased transistor density enabled by FinFET

However, it is expected that new types of transistors will be required in the next decade to accomplish the indicated reduction in Subthreshod slope. Beginning with 2005 an extensive search for a new "switch" was initiated across the world under very broad guidelines. Multiple new types of switches were identified by 2010 and substantial experimental results were reported in 2015. Early on the new desirable properties of the new "switches" were classified in three categories:

1. ***The new devices were required to function in al least two separate logic states but the more the better***
2. ***It was desirable that the devices could operate with minimal power consumption (i.e., consuming less power than CMOS devices) and with no power consumption at all in their stand-by mode***
3. ***It was also desirable that the devices should have the ability to retain memory information with essentially no power consumption (i.e., no power supply)***

Based on past experience few fundamental modes of physical operation were identified. In one case the devices relied on flow of electric charge, in another case the devices relied on magnetic properties and finally some devices would operate in a completely new way. While the first mode of operation could still be associated with charges flowing from one location to another the second mode of operation was associated with stationary magnetic dipoles that did not consume any power in their stand by condition. It was quickly realized that even though it was well known that electrons carried a negative charge it was also known that they also carried a magnetic dipole (spin) associated with them even though this latter property had never been used for construction of commercial Integrated Circuits.

In 2010 an extensive review of progress on novel devices and possible applications was published in the Proceeding of IEEE (Vol.98, No.12, December 2010). Among these devices the ***tunnel transistor*** appeared as a leading candidate with respect to operation at very low voltage. The operation can be explained in very simple terms. In a typical NMOS

transistor the charges in the conduction band of the source region are prevented from flowing to the drain region by the potential barrier generated by the NP junction existing between the source and the body of the transistor. Application of positive voltage to the gate lowers this barrier and lets electrons flow to the drain. However, the electrons in the source region have an energy distribution that spreads to values exceeding the potential barrier even when there is no voltage applied to the gate. Under these conditions some level of leakage cannot be eliminated. Conceptually, it is possible to completely eliminate the effect of the tail of the electron energy distribution by deciding to operate via the electrons located in the valence band of the source region. Under these conditions charges would not be able to overcome the combination of the band gap and junction barrier voltage when operating in a stand-by mode! While it is true that any leakage from source to drain could be eliminated under these conditions it is also true that in order to allow any current at all to flow from the source to the drain region it would be necessary to somehow apply a voltage that would overcome the combination of band-gap and PN junction potentials. So, no leakage but also no current could flow under normal voltage operations.

However, Prof. Leo Esaki had demonstrated that when two energy bands are brought into very close proximity, charges could flow from one band to the other by tunneling though this band-to-band potential barrier. He received a Nobel Prize in 1973 for this invention.

Further improvements in the practical implementation of this concept to transistor fabrication has led to very promising results in recent years. It has been demonstrated that, under proper conditions, flow of current can occur at very low gate voltage and leakage current is practically negligible in the off state. Subthreshod slopes below 30mV/dec have been demonstrated. The research effort is now concentrated on selecting the appropriate materials. However, while essentially steep subthreshold transition from off to on state have been observed when TFETs operate at low temperature (i.e., 28mv/Dec at 77K) at present leakage increases as operation approaches the 300K temperatures. This is due to unwanted electronic states that allow current to flow between the two bands. The problem to be solved is not different from the problem of eliminating surface states in early MOS that scientists had to deal with for 20 years. This time the "states" are buried somewhere between the two bands across which the charges have to tunnel. Among the many proposed solution the use of 2D material seems very promising since their very own structure eliminates one set of undesirable "dangling bonds".

Several devices utilizing the spin property of electrons have also shown promising results.

In particular, the concept of operating a device with current (dynamic mode) but then storing the result by means of a magnetic state (static mode) has been demonstrated. Spin-transfer torque (STTM) is an effect in which the orientation of a magnetic layer in a magnetic tunnel junction or spin valve can be modified using a spin-polarized current. Once the magnet is polarized no current is required to keep the magnet in this state. In essence this device embodies both the current carrying properties of electrons as well the ability of transferring magnetic information to create a permanent state in a magnetic layer. This effect represents the ability to electrically program a magnetic memory that can store information without using any energy.

These few are just few examples indicating that by the year 2020 several new devices will be available to work in conjunction or better that CMOS on some specific applications.

"Moore's Law is dead, long live Moore's Law"

The question of how long will Moore's Law last has been posed an infinite number of times since the 80s and every 5-10 years publications claiming the end of Moore's Law have appeared from the most unthinkable and yet "reputedly qualified" sources. Despite these alarmist publications the trend predicted by Moore's Law has continued unabated for the past 50 years by morphing from one scaling method to another, where one method ended the next one took over. This concept has completely eluded the comprehension of casual observes that have mistakenly interpreted the end of one scaling method as the end of Moore's Law. As stated before, bipolar transistors were replaced by PMOS that were replaced by NMOS that were also replaced by CMOS. Equivalent Scaling succeeded Geometrical Scaling when this could not longer operate and now 3D Power Scaling is taking off.

By 2020-25 device features will be reduces to a few nanometers and it will become practically impossible to reduce device dimensions any further. At first sight this consideration seems to prelude to the unavoidable end of the integrated circuit era but once again the creativity of scientists and engineers has devised a method

"To snatch victory from the jaws of defeat"

The basic concept of this solution is actually rather obvious if we only observe places like Manhattan, Tokyo, Seoul or Hong Kong; once real estate space was fully utilized people discovered the unexplored space of the vertical dimension and began building skyscrapers.

Nowadays, Flash memory producers are facing a similar problem as they are running out of space to further scale Flash transistors. Furthermore, cost of producing integrated memory circuits of small dimensions keeps on rising while the number of stored electrons in the floating gate keeps on decreasing. To eliminate this problem Flash memory producers have already demonstrated and announced several new products that stack multiple layers of memory transistors on top of each other in a single integrated circuit. As many as 32 and 48 layers of Flash memory have been reported. Flash memory devices constituted by more than 100 layers have been predicted. The basic approach consists in building transistors that are 100% vertically oriented. Arrays of columns of 32 or 48 transistors have been built to increase transistor density as measured by the number of transistors/cm^2 and reduce cost since this approach requires fewer mask layers that traditional 2D manufacturing methods.

This vertical trend is led by Flash memories but it is expected that it will become an industry wide trend in the next decade. FinFET structures were built to control as much as possible the potential in the transistor channel. It was however anticipated by the ITRS in the 1998-2000 timeframe that eventually transistors would be fabricated with gates completely surrounding the semiconductor. Orienting the transistor substrate vertically and then completely surrounding it with a sequence of dielectric and metal layers deposited by means of deposition to fabricate the composite gate structure can more easily accomplished if the transistor is vertically oriented. It is clear that this method reduces the transistor footprint and in conjunction with creating multiple layers of transistors one on top of the other will accelerate the level of transistor density beyond Moore's Law traditional trends.

Physical Properties of Transistors

YEAR OF PRODUCTION	2015	2017	2019	2021	2024	2027	2030
Logic device technology naming	P70M56	P48M36	P42M24	P32M20	P24M12G1	P24M12G2	P24M12G3
Logic industry "Node Range" Labeling (nm)	"16/14"	"11/10"	"8/7"	"6/5"	"4/3"	"3/2.5"	"2/1.5"
Logic device structure options	FinFET FDSOI	FinFET FDSOI	FinFET LGAA	FinFET LGAA VGAA	VGAA, M3D	VGAA, M3D	VGAA, M3D
LOGIC DEVICE GROUND RULES							
MPU/SoC Metalx ½ Pitch (nm)[1,2]	28.0	18.0	12.0	10.0	6.0	6.0	6.0
MPU/SoC Metal0/1 ½ Pitch (nm)	28.0	18.0	12.0	10.0	6.0	6.0	6.0
Contacted poly half pitch (nm)	35.0	24.0	21.0	16.0	12.0	12.0	12.0
L_g: Physical Gate Length for HP Logic (nm) [3]	24	18	14	10	10	10	10
L_g: Physical Gate Length for LP Logic (nm)	26	20	16	12	12	12	12
FinFET Fin Half-pitch (new) =0.75 or 1.0 M0/M1 (nm)	21.0	18.0	12.0				
FinFET Fin Width (nm)	8.0	6.0	6.0				
FinFET Fin Height (nm)	42.0	42.0	42.0		END OF	2D	DOMAIN
Footprint drive efficiency - FinFET	2.19	2.50	3.75				
Lateral GAA Lateral Half-pitch (nm)			12.0	10.0			
Lateral GAA Vertical Half-pitch (nm)			12.0	9.0			
Lateral GAA Diameter (nm)			6.0	6.0			
Footprint drive efficiency - lateral GAA, 3x NWs stacked			2.4	2.8	START	OF 3D	DOMAIN
Vertical GAA Lateral Half-pitch (nm)				10.0	6.0	6.0	6.0
Vertical GAA Diameter (nm)				6.0	5.0	5.0	5.0
Footprint drive efficiency - vertical GAA, 3x NWs stacked				2.8	3.9	3.9	3.9
Device effective width - [nm]	92.0	90.0	56.5	56.5	56.5	56.5	56.5
Device lateral half pitch (nm)	21.0	18.0	12.0	10.0	6.0	6.0	6.0
Device width or diameter (nm)	8.0	6.0	6.0	6.0	5.0	5.0	5.0

To emphasize the transition to 3D Power Scaling the columns describing 2D transistor dimensions have been left FinFET cells completely blank from 2021 on.

9. DRIVING THE IC INDUSTRY TO THE LIMITS OF CMOS AND BEYOND

9.1. FUTURE DEVICES CATEGORIES

Continued dimensional and functional scaling of CMOS is driving information processing technology into a broadening spectrum of new applications. Many of these applications are enabled by performance gains and/or increased complexity realized by scaling. Because dimensional scaling of CMOS eventually will approach fundamental limits, several new alternative information processing devices and microarchitectures for existing or new functions are being explored to extend the historical integrated circuit scaling cadence and sustain performance gain beyond CMOS scaling. This is driving interest in new devices for information processing and memory, new technologies for heterogeneous integration of multiple functions (a.k.a. "More than Moore"), and new paradigms for system architecture.

The semiconductor industry is facing three classes of difficult challenges related to extending integrated circuit technology to new applications and to beyond the end of CMOS dimensional scaling.

1. One class relates to propelling CMOS beyond its ultimate density and functionality by integrating a new high-speed, high-density, and low-power memory technology onto the CMOS platform.
2. Another class is to extend information processing substantially beyond that attainable by CMOS using an innovative combination of new devices, interconnect and architectural approaches for extending CMOS and eventually inventing a new information processing platform technology.
3. The third class is to invent and reduce to practice long term alternative solutions to technologies that address existing More than Moore (MtM) ITRS topical entries.

These difficult challenges need to be are addressed in the period ranging from 2020 to 2030.

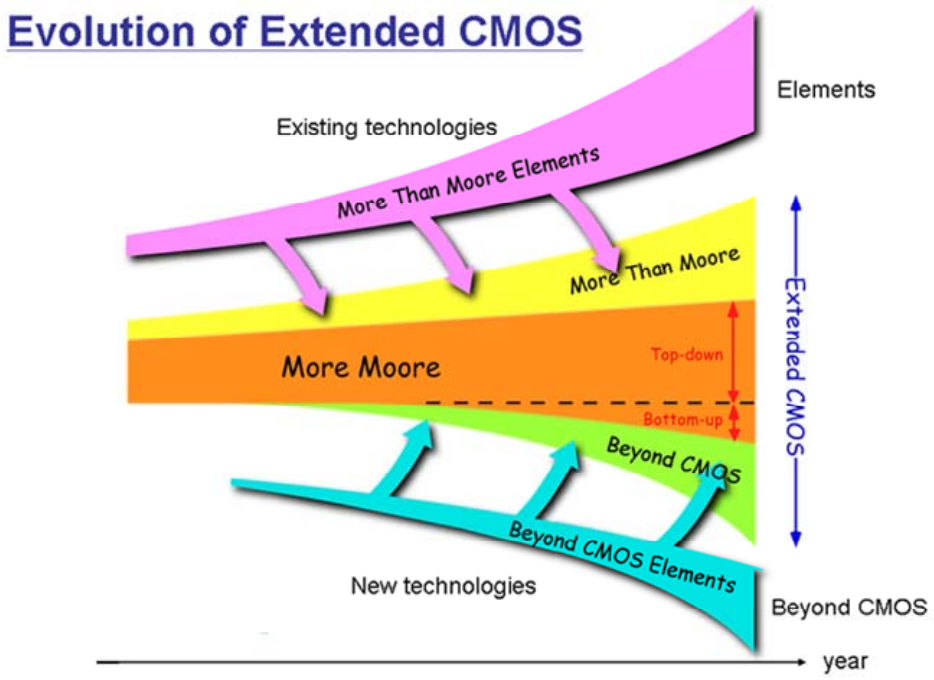


Fig. 9.1 Relationships of More Moore, More-than-Moore, and Beyond CMOS devices (Courtesy of Japan ERD).

Difficult challenges gating development of emerging research devices are divided into three parts: those related to memory technologies, those related to information processing or logic devices, and those related to heterogeneous integration of multi-functional components, *a.k.a.* More-than-Moore (MtM) or Functional Diversification.

One challenge is the need of a new memory technology that combines the best features of current memories in a fabrication technology compatible with CMOS process flow and can be scaled beyond the present limits of SRAM and FLASH. This would provide a memory device fabrication technology required for both stand-alone and embedded memory applications. The ability of an MPU to execute programs is limited by interaction between the processor and the memory, and scaling does not automatically solve this problem. The current evolutionary solution is to increase MPU cache memory, thereby increasing the floor space that SRAM occupies on an MPU chip. This trend eventually leads to a decrease of the net information throughput. In addition to auxiliary circuitry to maintain stored data, volatility of semiconductor memory requires external storage media with slow access (e.g., magnetic hard drives, optical CD, etc.). Therefore, development of electrically accessible non-volatile memory with high speed and high density would initiate a revolution in computer architecture. This development would provide a significant increase in information throughput beyond the traditional benefits of scaling when fully realized for nanoscale CMOS devices.

A related challenge is to sustain scaling of CMOS logic technology beyond 2020. One approach to continuing performance gains as CMOS scaling matures in the next decade is to replace the strained silicon MOSFET channel (and the source/drain) with an alternate material offering a higher potential quasi-ballistic-carrier velocity and higher mobility than strained silicon. Candidate materials include strained Ge, SiGe, a variety of III-V compound semiconductors, and carbon materials. Introduction of non-silicon materials into the channel and source/drain regions of otherwise silicon MOSFET (i.e., onto a silicon substrate) is fraught with several very difficult challenges. These challenges include heterogeneous fabrication of high-quality (*i.e.*, defect free) channel and source/drain materials on non-lattice matched silicon, minimization of band-to-band tunneling in narrow bandgap channel materials, elimination of Fermi level pinning in the channel/gate dielectric interface, and fabrication of high-κ gate dielectrics on the passivated channel materials. Additional challenges are to sustain the required reduction in leakage currents and power dissipation in these ultimately scaled CMOS gates and to introduce these new materials into the MOSFET while simultaneously minimizing the increasing variations in critical dimensions and statistical fluctuations in the channel (source/drain) doping concentrations.

The industry is now addressing the increasing importance of a new trend, “More than Moore” (MtM), where added value to devices is provided by incorporating functionalities that do not necessarily scale according to "Moore's Law“. An MtM section was first introduced into the ERD chapter in 2011 with the initial coverage on wireless technologies. The 2013 version added devices with learning capabilities. This version further expands the coverage to emerging devices for hardware security. Traditionally, the ITRS has taken a “technology push” approach for roadmapping “More Moore”, assuming the validity of Moore’s law. In the absence of such a law, different methodologies are needed to identify and guide roadmap efforts in the MtM-domain.

A longer-term challenge is invention and reduction to practice of a manufacturable information processing technology addressing “beyond CMOS” applications. For example, emerging research devices might be used to realize special purpose processor cores that could be integrated with multiple CMOS CPU cores to obtain performance advantages. These new special purpose cores may provide a particular system function much more efficiently than a digital CMOS block, or they may offer a uniquely new function not available in a CMOS-based approach. Solutions to this challenge beyond the end of CMOS scaling may also lead to new opportunities for such an emerging research device technology to eventually replace the CMOS gate as new information processing primitive element. A new information processing technology must also be compatible with a system architecture that can fully utilize the new device. A non-binary data representation and non-Boolean logic may be required to employ a new device for information processing. These requirements will drive the need for new system architectures.

9.2. FUTURE MEMORIES

Research on new memory devices has been the most productive field of research for the past 10 years [Fig.9.2]. Multiple innovative ways of storing information have been proposed and are intensively undergoing evaluation. The field of Non-volatile memory has been the most productive.

The purpose of many memory systems is to store massive amounts of data, and therefore *memory capacity* (or *memory density*) is one of the most important system parameters. In a typical memory system, the memory cells are connected to form a two-dimensional array, and it is essential to consider the performance of memory cells in the context of this array architecture. A memory cell in such an array can be viewed as being composed of two fundamental components: the ‘*storage node’* and the *‘select device’,* the latter of which allows a given memory cell in an array to be addressed for read

or write. Both components impact scaling limits for memory. For several emerging resistance-based memories, the storage node can, in principle, be scaled down below 10 nm, and the select device will limit the memory density.

Planar transistors (e.g. FET or BJT) are typically used as select devices. In a two-dimensional layout using in-plane select FETs the cell layout area is A_{cell}=(6-8) $F2$. In order to reach the highest possible 2-D memory density of $4F^2$, a vertical select transistor can be used.

At the highest level, memory technologies are separated by the ability to retain data without power. Nonvolatile memory offers essential use advantages, and the degree to which non-volatility exists is measured in terms of the length of time that data can be retained. Volatile memories also have a characteristic retention time that can vary from milliseconds to (for practical purposes) the length of time that power remains on. Nonvolatile memory technologies are further categorized by their maturity. Flash memory is considered the baseline nonvolatile memory because it is highly mature, well optimized, and has a significant commercial presence. Flash memory is the benchmark against which prototypical and emerging nonvolatile memory technologies are measured. Prototypical memory technologies are at a point of maturity where they are commercially available (generally for niche applications), and have a large scientific, technological, and systems knowledge base available in the literature.

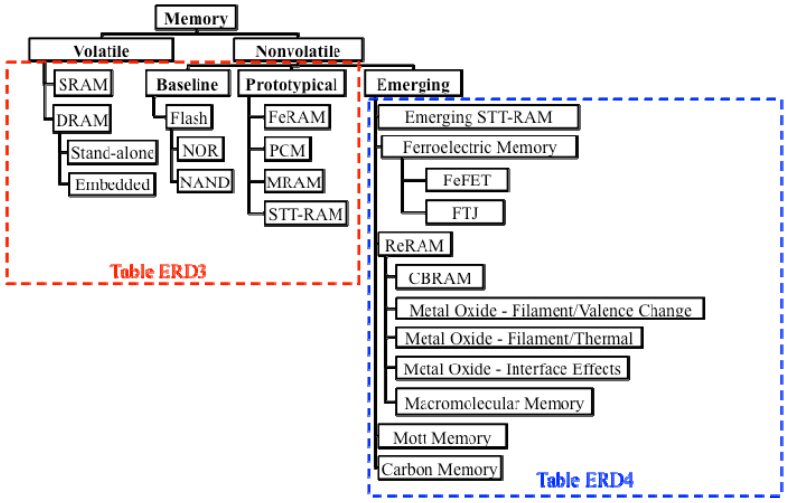


Fig. 9.2 Taxonomy of emerging memory devices

Among the many emerging memory devices it is important to mention a few:

The redox-based nanoionic memory operation is based on a *change in resistance* of a MIM structure caused by ion (cation or anion) migration combined with redox processes involving the electrode material or the insulator material, or both . Three classes of electrically induced phenomena have been identified that involve chemical effects that effects which relate to redox processes in the MIM cell. In these three ReRAM classes, there is a competition between thermal and electrochemical driving forces involved in the switching mechanism.

Ferroelectric memory. Coding digital memory states by the electrically alterable polarization direction of ferroelectrics has been successfully implemented and commercialized in capacitor-based Ferroelectric Random Access Memory (Table ERD3). However, in this technology the identification of the memory state requires a destructive read operation and largely depends on the total polarization charge on a ferroelectric capacitor, which in terms of lateral dimensions is expected to shrink with every new technology node. In contrast to that, alternative device concepts, such as the ferroelectric field effect transistor (FeFET) and the ferroelectric tunnel junction (FTJ), allow for a non-destructive detection of the memory state and promise improved scalability of the memory cell.

Ferroelectric FET. The ferroelectric tunnel junction, a ferroelectric ultra-thin film commonly sandwiched by asymmetric electrodes and/or interfaces, exhibits ferroelectric polarization induced resistive switching by a non-volatile modulation of barrier height. With the tunneling current depending exponentially on the barrier height, the ferroelectric dipole orientation either codes for a high or a low resistance state in the FTJ, which can be read out non-destructively. The

resulting Tunneling Electro Resistance (TER) effect of FTJs, the ratio between HRS and LRS, is usually in the range of 10 to 100. However, giant TER of > 10^4 has most recently been reported in a super-tetragonal $BiFeO_3$ based FTJ by Yamada et al.

Carbon Based memory. Recently, various allotropes of carbon, such as amorphous carbon (α-C) and diamond-like (DLC) carbon, carbon nanotubes and graphenic carbon were explored for memory applications. The device structure consists of an electrode/carbon/electrode capacitor structure, whose resistance can be changed by appropriate current pulses. One of the proposed switching mechanism is creation and destruction of conductive sp^2-bonds in an otherwise insulating carbon matrix consisting of sp^3-bonds.

STT RAM. Recently, some novel mechanisms have been discovered to switch spin-transfer-torque RAM (STT-RAM). Among them, voltage-included magnetization switching and spin Hall effects are the most promising candidates. Voltage pulse induction of magnetization switching has emerged as a candidate for a high efficiency nonvolatile memory. The voltage driven nature of the technology allows us to use high resistance magnetic tunneling junctions and fits to the existing CMOS design. It has the potential to offer ultra-low power, high speed and long endurance memory cells beyond conventional STT-MRAM technology, which is based on a spin-polarized current injection.

The voltage writing technique is based on a magnetic anisotropy change because of application of an electric field at the interface between magnetic metal and an insulator. The change in magnetic anisotropy switches magnetization from perpendicular to in-plane for a quasi-static process. With high-speed pulse application, however, the magnetization undertakes a precessional switching within around 1 nsec. Since the precessional switching is a toggle switching, a read procedure is required before a writing pulse. Alternative method to use voltage effect is a combination with spin-transfer effect. For this case a precise control of device parameters or a special pulse shape required. Voltage torque MRAM (VT-MRAM) is a very attractive candidate as for a storage class memory, DRAM and/or cash memory. For magnetic cells less than 20 nm in diameter, however, development of a material with a higher voltage effect is essential.

9.3. Storage Class Memory Devices

Storage-class memory (SCM) describes a device category that combines the benefits of solid-state memory, such as high performance and robustness, with the archival capabilities and low cost of conventional hard-disk magnetic storage. Such a device requires a nonvolatile memory (NVM) technology that could be manufactured at a very low cost per bit.

A number of suitable NVM candidate technologies have long received research attention, originally under the motivation of readying a "replacement" for NAND Flash, should that prove necessary. Yet the scaling roadmap for NAND Flash has progressed steadily so far, without needing any replacement by such technologies. So long as the established commodity continues to scale successfully, there would seem to be little need to gamble on implementing an unproven replacement technology instead.

In July 2015, Intel and Micron jointly announced a new nonvolatile memory technology, called "3D-Xpoint." This technology is said to offer 1000x lower latency and 1000x higher endurance than NAND Flash, at a density that is 10x higher than DRAM. (Note that it is most likely that the latency referred to here is write latency rather than read latency, since NAND write latency is much slower than its read latency.) 3D-Xpoint technology, said to have been implemented at the 128Gbit chip level, is based on a two-layer stacked crossbar array, with each intersection point containing a nonvolatile memory device and a nonlinear access device [Micron3DXpoint]. The particular nonvolatile memory device was not specified, other than that it depends on bulk changes of resistance, nor was the nonlinear access device described. Speculation based on patent searches and job solicitations suggest that the technology may be a combination of some variant of phase change memory and an Ovonic Threshold Switching access device.

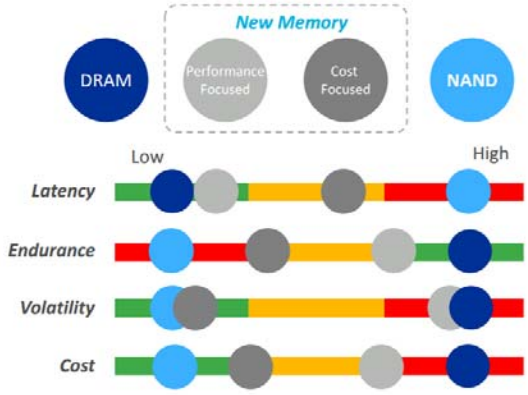


Fig. 9.3 Comparison of performance of different memory technologies

9.4. EMERGING LOGIC AND ALTERNATIVE INFORMATION PROCESSING DEVICES

One of the central objectives of recent research in devices beyond silicon transistors consists in the development of novel logic switches that might replace the silicon transistor as the device driving technological development within the semiconductor industry. Such a replacement is thought potentially to be viable if one or more of the following capabilities is afforded by a novel device: (1) an increase in device density (and corresponding decrease in cost) beyond that achievable by ultimately scaled CMOS; (2) an increase beyond CMOS in switching speed, e.g., through improvements in the normalized drive current or reduction in switched capacitance; (3) a reduction beyond CMOS in switching energy, associated with a reduction in overall circuit energy consumption; or (4) the enabling of novel information processing functions that cannot be performed as efficiently using conventional CMOS.

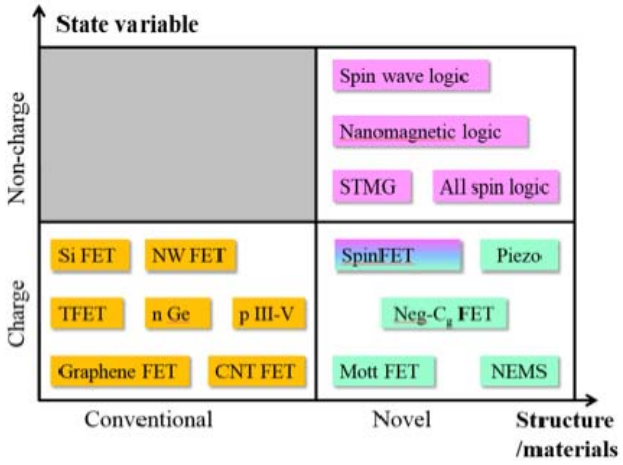


Fig. 9.4 Taxonomy of options for emerging logic devices.

The devices examined in this chapter are differentiated according to (1) whether the structure and/or materials are conventional or novel, and (2) whether the information carrier is electron charge or some non-charge entity. Since a conventional FET structure and material imply a charge-based device, this classification results in a three-part taxonomy.

Many devices have been proposed in the past 10 years. This section present a selection of some of the most promising new devices

9.5. DEVICES FOR CMOS EXTENSION

Carbon Nanotube FETs. For many researchers, the search for an ideal semiconductor to be used in FETs succeeded when single-walled carbon nanotubes (CNTs) were first shown to yield promising devices more than 15 years ago. Owing to their naturally ultrathin body (~1 nm diameter cylinders of hexagonally bonded carbon atoms), superb electron and hole transport properties, and reasonable energy gap of ~0.6 – 0.8 eV, CNTs offer solutions in most of the areas that other semiconductors fundamentally fail when scaled to the sub-10 nm dimensional scale.

The most prominent advantages of CNT FETs over other options for aggressively scaled devices are the room temperature ballistic transport of charge carriers, the reasonable energy gap, the demonstrated potential to yield high performance at low operating voltage, and scalability to sub-10 nm dimensions with minimal short channel effects.

Nanowire FETs Nanowire field-effect transistors are structures in which the conventional planar MOSFET channel is replaced with a semiconducting nanowire. Such nanowires have been demonstrated with diameters as small as 0.5 nm. They may be composed of a wide variety of materials, including silicon, germanium, various III-V compound semiconductors (GaN, AlN, InN, GaP, InP, GaAs, InAs), II-VI materials (CdSe, ZnSe, CdS, ZnS), as well as semiconducting oxides (In_2O_3, ZnO, TiO_2), etc. Importantly, at low diameters, these nanowires exhibit quantum

confinement behavior (i.e., 1-D ballistic conduction) and match well with the gate-all-around structure that may permit the reduction of short channel effects and other limitations to the scaling of planar MOSFETs.

Circuit and system functionality of nanowire devices have been demonstrated, including vertical InAs MOSFETs with 103 GHz switching speed, a down-conversion mixer based on vertical InAs transistors that showed cut-off frequency of 2 GHz, and extended, programmable arrays ("tiles") of nonvolatile nanowire-based flash memory that are used to build circuits such as full-adder, full-subsratcor, multiplexer, demultiplexer, clocked D-latch and finite state machines.

III-V Channel Replacement Devices III-V compounds semiconductors as replacements for n-type channel materials have attracted considerable attention because of their excellent bulk electron motilities. To realize III-V p-FETs on a Si substrate, some methods are reported.

Tunnel FETs Tunneling Field Effect Transistors (TFETs) have the potential to achieve a low operating voltage by overcoming the thermally limited subthreshold swing voltage of 60mV/decade. In its simplest form, a TFET is a gated reverse-biased p-i-n junction. There are two mechanisms that can be used to achieve a low voltage turn on. The gate voltage can be used to modulate the thickness of the tunneling barrier at the source channel junction and thus modulate the tunneling probability. Changing the electric field in the tunneling junction controls the thickness of the tunneling barrier.

9.6. BEYOND-CMOS DEVICES

Spin FET. Spin-transistors are classified as "non-conventional charge-based extended CMOS devices", and can be further divided into two categories: the spin-FETs proposed by Datta and Das and spin-MOSFETs proposed by Sugahara and Tanaka. The structures of both types of spin transistors consist of a ferromagnetic source and a ferromagnetic drain that act as a spin injector and a detector, respectively. Although the devices have similar structures, they have quite different operating principles.

In spin-MOSFETs, the gate has the same current switching function as in ordinary transistors, whereas in the spin-FETs, the gate acts to control the spin direction by utilizing the Rashba spin-orbit interaction. Both types of devices behave as a transistor and function as a magnetoresistive device. The important features of spin transistors are that they allow a variable current to be controlled by the magnetization configuration of the ferromagnetic electrodes (spin-MOSFETs) or the spin direction of the carriers (spin-FETs), and they offer the capability for nonvolatile information storage using the magnetization configurations.

Negative Gate-Capacitance FET. Based on the energy landscapes of ferroelectric capacitors, it has been suggested that by replacing the standard insulator of a MOSFET gate stack with a ferroelectric insulator of appropriate thickness, it should be possible to implement a step-up voltage transformer that will amplify the gate voltage. Such a device is called a negative gate capacitance FET. The gate operation in this device would lead to subthreshold swing (STS) lower than 60 mV/decade and might enable low voltage/low power operation. The main advantage of such a device is that it is a relatively straightforward replacement of conventional FET. Thus, high Ion levels similar to advanced CMOS would be achievable with lower voltages. An early experimental attempt to demonstrate a low-STS NCFET, based on a P (VDF-TrFE)/SiO_2 organic ferroelectric gate stack, was reported in 2008 and subsequently in 2010 in a more controlled structure. These experiments thus established the proof of concept of <60mV/decade operation using the principle of negative capacitance.

NEMS Switch. Micro/Nano-Electro-Mechanical (M/NEM) switches are devices that employ electrostatic force to actuate a movable structure to make a conductive path between two electrodes. Mechanical switches feature two fundamental properties that are unattainable in MOSFETs: *zero off-state leakage* and *zero subthreshold swing*. The first property provides for zero standby power dissipation, while the second property suggests the potential to operate at very low voltage for low dynamic power dissipation as well. Another advantage is that a mechanical switch can be operated with either positive or negative voltage polarity due to the ambipolar nature of the electrostatic force, so that an electrostatically actuated relay can be configured as a pull-down or pull-up device, respectively. M/NEM switches can be fabricated using conventional planar processing techniques (thin-film deposition, lithography and etch steps), with a final release step in which a sacrificial material such as silicon dioxide, photoresist, polyimide or silicon is selectively removed to form the actuation and contact air-gaps.

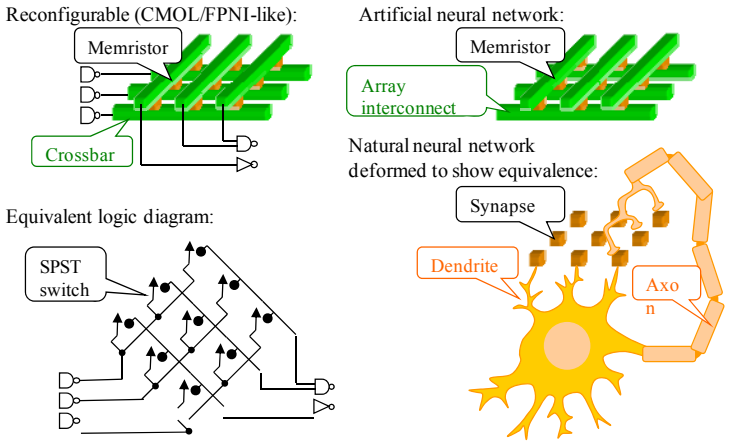
All-Spin Logic Devices. The recently proposed concept of all spin logic (ASL) uses magnets to represent non-volatile binary data while the communication between magnets is achieved using spin currents in spin coherent channels with the energy coming from the power supply. The ASL concept is based on key scientific advancements of the last

decade. These advancements have blurred the distinction between spintronics and magnetics, creating the possibility of a device capable of providing a low power alternative to charge-based information processing. In particular, the two key recent advances are (1) the demonstration of spin injection into metals and semiconductors from magnetic contacts and (2) the switching of a second magnet by the injected spins. These demonstrations suggest an all-spin approach to information processing. Magnets inject spins and spins turn magnets (digital bits) forming a closed "ecosystem" which takes advantage of both analog (spin currents) and digital (bistable magnets) properties without the need to convert to charge. It has been shown that ASL can potentially reduce the switching energy-delay product by a significant amount, but there are major challenges to be overcome.

9.7. DEVICES WITH LEARNING CAPABILITIES

The ITRS traditionally views computer performance as the product of device performance multiplied by a factor related to computer architecture – but it is possible to raise computer performance in a different way. An information processing function can be realized by the placement of gates and wires on a custom ASIC or by coding the function in software and running it on a microprocessor. The ASIC implementation will be several orders of magnitude faster and more power efficient because moving a bit from one gate or functional block to another via a wire is much more efficient than having a microprocessor move the bit in software, incurring as overhead the movement of hundreds of bits of instruction, opcode, cache access, etc. to interpret the machine's instruction set for the software.

The Field Programmable Gate Array (FPGA) achieved the flexibility objective of the last paragraph long ago, but memristor-class devices show the possibility meeting the efficiency objective as well. As illustrated in Figure9.5 , there are several approaches in the literature where memristors are used as essentially Single Pole Single Throw (SPST) switches that connect some form of general wiring array into specific wiring interconnections. By altering the pattern of switches, a series of gates in a base CMOS layer can be reconnected into an arbitrary configuration. This is essentially the same architectural concept as an FPGA, but the traditional FPGA implements a non-volatile switch with a sizeable CMOS circuit.



Reconfigurable-memristor functional mapping. Memristors fill the role of a SPST switch in a crossbar wiring network. The logic circuit can be "rewired" by appropriate switch/memristor settings

Artificial-natural neural network mapping. Memristors fill the role of synapses with row/column wires filling the role of dendrites and axons.

Fig. 9.5 Two variants of learning devices for configuration

Devices That Learn by Examples

Brain-inspired, neural systems are capable of performing tasks close to the level of human intellect and there is strong indication that these systems can be made more efficient using physical devices that support learning. Computers that learn by examples have been successfully applied to activities that have previously been beyond the reach of computers, such as the Watson computer playing Jeopardy. To the extent a new type of computer can address formerly human-only tasks, the proper cost comparison becomes computer cost versus the cost of a human being paid a salary, working in an office, and so forth. The Jeopardy-playing computer demonstration shows that computer methods are known for performing at least some formerly human-only tasks, but the Watson computer was actually a 2500-core microprocessor cluster. If we assume the artificial intelligence community has correctly identified the artificial neural network as the

necessary computational primitive for this new type of computing, it would be important to find a way to build an artificial neural network that is more efficient than a microprocessor cluster.

9.8. EMERGING COMPUTING ARCHITECTURES

Emerging devices are creating opportunities for new computing machines, including machines with non-conventional architectures. At the same time, designers are developing emerging architectural concepts, such as neuro-inspired architectures that incorporate conventional computing devices. This sub-chapter includes a survey of a number of these technologies, together with a summary of open challenges [Fig. 9.6].

Horizontally, the chart is first organized with "Program-Centric" architectures on the left and "Data-Centric" architectures on the right. "Program centric" refers to processors wherein the designer dictates the detailed design, including every step of the program flow. Program-centric processors are further split into conventional stored program Von Neumann machines and non-Von Neumann processors that are not driven from a stored software program. "Data-centric" refers to processors wherein some of the details of the final "design" are driven by the data that is passed through the system during a training step, either by the system manufacturer at the factory, by a systems integrator before shipping to customers, or even by customers in the field. For example, the weights in a neural network are learned for a recognition task by showing the network many data-examples. The final horizontal dimension is the split of each column into CMOS and non-CMOS device enabled categories.

Vertically, a dashed red line between "deterministic and non-deterministic" computing divides the table. "Deterministic" computing refers to fixed-precision (whether floating or fixed point) and low bit-error-rate compute models. "Non-deterministic" computing refers to computing that does not rely on these paradigms, and the compute solution has some sort of probabilistic nature to it.

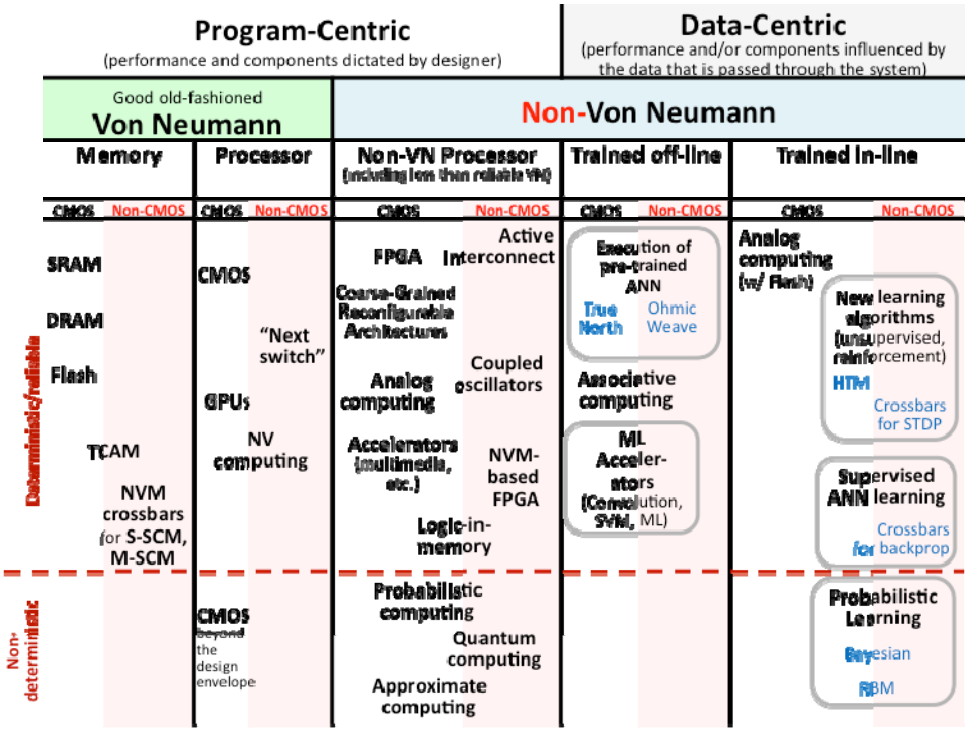


Fig. 9.6 Categories of computing architecture

9.9. LOGIC DEVICES BENCHMARKING

Preliminary analyses sponsored by SRC/NRI and the Intel study surveyed the potential logic opportunities afforded by major emerging switches using a variety of information tokens and communication transport mechanisms. Specifically, the projected effectiveness of these devices used in a number of logic gate configurations was evaluated, and normalized to CMOS at the 15nm generation as captured by the ITRS. The initial work has focused on "standard" Boolean logic architecture, since the CMOS equivalent is readily available for comparison. It should be noted that the majority of devices are evaluated via simulations, since many of them have not yet been built, so it should be considered only a "snapshot in time" of the potential of any given device [Figure 9.7].

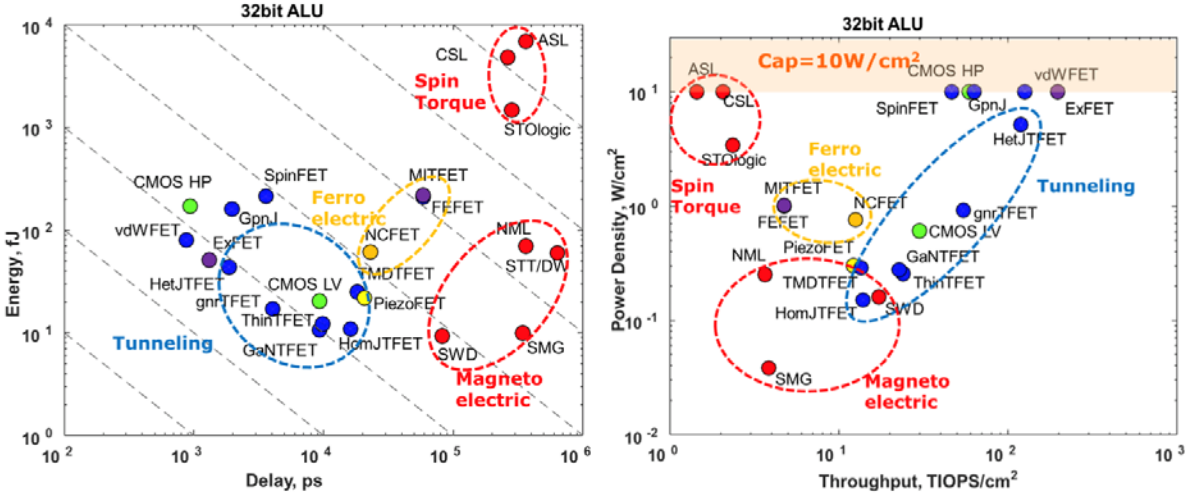


Fig. 9.7 (a) Energy vs. delay plot of 32bit ALU built from benchmarked devices; (b) power vs. throughput of 32bit ALU built from these devices, reflecting power-constrained (< 10 W/cm^2) throughput

Based on the current data and observations, it is clear that CMOS will remain the primary basis for IC chips for the coming decade. While it is unlikely that any of the current emerging devices could entirely replace CMOS, several do seem to offer advantages, such as ultra-low power or nonvolatility, which could be utilized to augment CMOS or to enable better performance in specific application spaces. One potential area for entry is that of special purpose cores or accelerators that could off-load specific computations from the primary general purpose processor and provide overall improvement in system performance. This is particularly attractive given the move to multi-core chips: while most are homogeneous today, if scaling slows in delivering the historically expected performance improvements in future generations, heterogeneous multi-core chips may be a more attractive option. These would include specific, custom-designed cores dedicated to accelerate high-value functions, such as accelerators already widely used today in CMOS (e.g. Encryption/Decryption, Compression/Decompression, Floating Point Units, Digital Signal Processors, etc.), as well as potentially new, higher-level functions (e.g. voice recognition). While integrating dissimilar technologies and materials is a big challenge, advances in packaging and 3D integration may make this more feasible over time, but the performance improvement would need to be large to balance this effort.

10. FACTORY INTEGRATION

The Factory Integration (FI) focus area of the ITRS is dedicated to ensuring that the semiconductor-manufacturing infrastructure contains the necessary components to produce items at affordable cost and high volume [Figure 10.1]. Realizing the potential of Moore's Law requires taking full advantage of device feature size reductions, new materials, yield improvement to near 100%, wafer size increases, and other manufacturing productivity improvements. This in turn requires a factory system that can fully integrate additional factory components and utilize these components collectively to deliver items that meet specifications determined by other ITRS focus areas as well as cost, volume and yield targets. Preserving the decades-long trend of 30% per year reduction in cost per function also requires capturing all possible cost reduction opportunities.

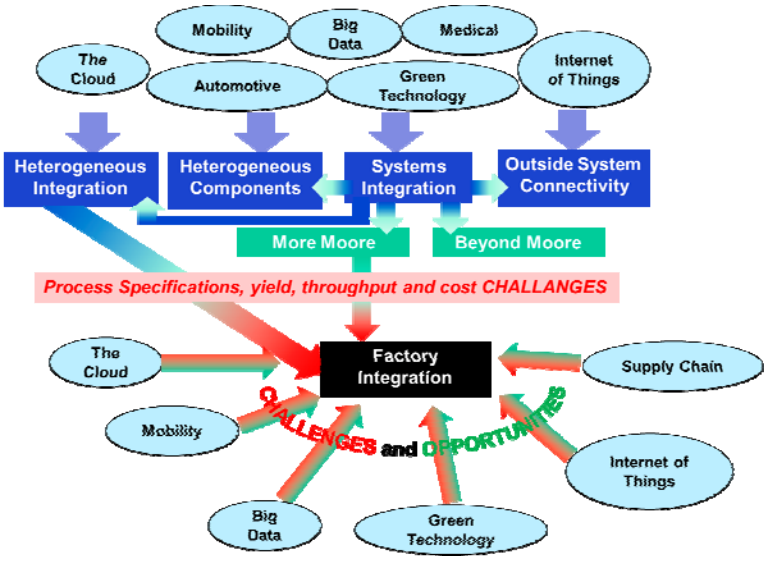


Fig.10.1 The Role of Factory Integration

10.1. FACTORY REQUIREMENTS

In order to clearly understand the integrated factory requirements and at the same time define measurable and actionable metrics, the factory integration chapter defines four functional areas, corresponding to physical boundaries that are required to perform semiconductor manufacturing. They are Production Equipment (PE), Material Handling Systems (MHS), Factory Information & Control Systems (FICS), and, Factory Facilities (Facilities).

Production Equipment (PE) covers process and metrology equipment and their interfaces to other factory elements. It also focuses on addressing equipment related productivity losses.

The Production Equipment (PE) specification has been enhanced with a new section focusing on communication between host and PE in support of PE energy and utility savings. In the future, this specification may extend to the communication with subsystems inside the equipment. The specification is aligned with work being done in SEMI standards.

Advanced Process Control (APC), which includes run-to-run process control, fault detection, fault classification, fault prediction and statistical process control technologies, and often leverages integrated metrology for process control) has evolved past the state of being an add-on capability to being a design-in requirement both at the equipment and fab-wide level. This does not mean the equipment will necessarily have embedded APC, but it does mean that the equipment will need to leverage APC capabilities (either internally or externally) to meet productivity and cost reduction targets. This is further emphasized in 2015 with the addition of a Control Systems Architecture (CSA) sub-chapter in FI. *Cross-leveraging 300 mm and 450 mm factory challenges*—we have addressed several 300 mm challenges, but it is still necessary to continue to address these challenges as we migrate to 450mm. We need to provide solutions that can be used in both domains as much as possible so as to leverage economy of scale and resource pooling. FI issues such as: 1) cycle time improvement, 2) yield improvement, 3) productivity waste reduction, 4) higher process controllability, and, 5) reduction in utilities, power consumption and emission with even more progressive targets, should have very similar solutions and roadmaps in 300mm and 450mm.

Production Equipment Technology Requirements

Year of Production	*2015*	*2017*	*2019*	*2021*	*2023*	*2025*	*2027*	*2029*
DRAM ½ Pitch (nm) (contacted)	24	22	18	15	12	10	9	8
Wafer Diameter (mm)	*300*	*300*	*300*	*450*	*450*	*450*	*450*	*450*
Process equipment availability (A80)--300mm	>95%	>95%	>96%	>96%	>96%	>96%	>96%	>96%
Process equipment availability (A80)--450mm			93%	> 93%	>93%	>93%	>93%	>93%
Metrology equipment availability (A80)--300mm	>98%	>98%	>98%	>98%	>98%	>98%	>98%	>98%
Metrology equipment availability (A80)--450mm			> 95%	> 96%	>96%	>96%	>96%	>96%
Maximum allowed electrostatic field on wafer and mask surfaces (V/m) for ESD prevention	2,600	2,000	1,550	1,300	1,000	775	650	TBD
Maximum recommended electrostatic field at chrome mask surfaces (V/m) for EFM	500	500	500	500	500	500	500	500
Minimum equipment data output rates (Hz) from a tool	10Hz	100Hz	100Hz	1kHz	1kHz	>1kHz	>1kHz	TBD
Pervasiveness of APC as an integral part of equipment design and operation	Partial	All	All	All	All	All	All	All
Pervasiveness of predictive technologies such as virtual metrology PdM, yield prediction and predictive scheduling in certain equipment components (e.g., vacuum, abatement, gas supply systems) feeding into overall equipment predictive solution, to support improvements such as reduction in unscheduled downtime and improved yield.	Partial	Partial	Partial	All	All	All	All	All
Pervasiveness of Equipment Health Monitoring capability as a common health indication capability across tools	Partial	All	All	All	All	All	All	All

10.2. FACILITIES REQUIREMENTS

Facilities include the overall physical buildings, cleanroom and facility infrastructure systems, including tool hook up. The ITRS Facilities scope does not include adjacent general office spaces and corporate functional areas. It is important to note that the following requirements will affect the facility and support facility infrastructure system with respect to their complexity and costs:

- Production equipment
- Manufacturing goals
- Management philosophies
- Environmental, safety, and health (ESH) goals
- Building codes and standards
- Defect-reduction and wafer cost reduction targets
- Disruptive manufacturing technology migration

The industry continues to demand facilities that are increasingly flexible, environmentally benign, extendable and reliable, services that come online more quickly, and are more cost-effective. However, production equipment requirements, ESH compliance and factory operational flexibility continue to drive increased facility capital and operating costs. Production and support equipment are becoming more complex, larger, and heavier, thereby driving the need for a continuous increase in factory size and tool packing density.

Meeting production equipment requirements (such as vibration and air, gas, and liquid purity levels) at the point-of-use may be a more cost-effective approach to meeting future requirements without increasing facility costs or sacrificing flexibility. For example, reducing facility vibration requirements and then working with production equipment manufacturers to ensure proper vibration control at the tool could reduce overall costs without decreasing the facility's flexibility. Reduction of air, gas, and chemical purity and piping installation specifications on central supply systems and

introducing localized purification systems to the specific equipment or areas requiring such measures can also help control costs, improve flexibility and enhance operating reliability.

Electrostatic discharge (ESD) sensitivity trends will have larger impact on manufacturing process yields as the device feature size decreases. Companies will need to increase their efforts to verify that the installed ESD controls are capable of handling these devices and to make any necessary improvement in ESD control methods. This could include changes in the ESD control item limits, change in the frequency of compliance verification, and other forms of ESD monitoring, such as ESD event detection. (ESD) Technology Roadmap.

Electromagnetic Interference (EMI) (see the standard SEMI E33 for definition) causes variety of problems for semiconductor manufacturing, including, but not limited to, equipment lockup and malfunction, sensor misreading, metrology errors, sensitive component damage and others. There are many sources of EMI in semiconductor environment that include electromagnetic emission from ESD, operation of equipment, especially high-energy tools, motors and actuators, wireless communication and alike. Co-location of sensitive equipment with high-energy tools, cabling, ground problems, improper maintenance of equipment and others further aggravate EMI problems.

Despite the continuous device feature size shrinkage and increase of process complexity in process technology according to Moore's Law, the drive towards the reduction in manufacturing cost will result in the introduction of larger wafer sizes, such as the pending use of 450mm wafers. Such a change will also have implications on the design and construction of a wafer manufacturing facility due to increases in overall size, height, and weight of process equipment, their utility consumption, and other process-driven facility requirements such as AMC, EMI, ESD and acoustic controls.

Facilities Technology Requirements

Year of Production	*2015*	*2017*	*2019*	*2021*	*2023*	*2025*	*2027*	*2029*
DRAM ½ Pitch (nm) (un-contacted Poly) (f) [2]	24	22	18	15	12	10	9	*8*
Wafer Diameter (mm)	*300*	*300*	*300*	*300*	*300*	*300*	*300*	*300*
Wafer Diameter (mm)				*450*	*450*	*450*	*450*	*450*
# Mask Levels—DRAM	*39*	*39*	*39*	*39*	*39*	*39*	*39*	*39*
Manufacturing (clean room) area m2/wafer starts per month/number / of mask layers (300mm)	0.0058	0.0058	0.0058	0.0058	0.0058	0.0058	0.0058	0.0058
Manufacturing (clean room) area m2/wafer starts per month/number of mask layers (450mm)			0.011	0.0075	0.0075	0.0075	0.0065	0.006
Design Criteria for Facility critical vibration areas (lithography, metrology, other) (□m/sec)	6.25 (VC D)	6.25 (VC D)	6.25 (VC D)	6.25 (VC D)	6.25 (VC D)	6.25 (VC D)	6.25 (VC D)	6.25 (VC D)
Design Criteria for Facility non-critical vibration areas (□m/sec)	50 (VC A)	50 (VC A)	50 (VC A)	50 (VC A)	50 (VC A)	50 (VC A)	50 (VC A)	50 (VC A)
Maximum allowable electrostatic field on facility surfaces (V/m) for ESD prevention	2700	2200	1850	1550	1350	1100	900	800
Continuous Radiated emission Limit for Facility Allowable Low Frequency (0-30kHz) Magnetic Field (nT/) for EMI sensitive area	80	60	40	20	10	8	6	5
Continuous Radiated emission Limit for Facility Allowable Low Frequency (0-30kHz) Magnetic Field (nT/) for very EMI sensitive area	8	6	4	2	1	1	1	1
Continuous Radiated emission Limit for Facility Allowable High Frequency (30MHz - 3GHz) Electric/Magnetic Field (V/m) for EMI sensitive area (Far Field)	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.2
Continuous Radiated emission Limit for Facility Allowable High Frequency (30MHz - 3GHz) Electric/Magnetic Field (V/m) for EMI sensitive area (Near Field)	1.0	0.8	0.7	0.7	0.5	0.5	0.5	0.5
Transient Radiated emission Limit for Facility Allowable Spectrum (30MHz - 3GHz) Electric/Magnetic Field (V/m) for EMI sensitive area (Far Field)	1	1	0.8	0.8	0.7	0.7	0.5	0.5
Transient Radiated emission Limit for Facility Allowable Spectrum (30MHz - 3GHz) Electric/Magnetic Field (V/m) for EMI sensitive area (Near Field)	2	2	1.5	1.5	1.0	1.0	0.8	0.7
Continuous Conducted Emission Limit Allowable Continuous Noise (9kHz - 30MHz) Level (dBuV)	90	90	80	80	70	70	70	70
Transient Conducted Emission Limit Allowable transient signal Level (V) for regular environment	0.5	0.5	0.4	0.4	0.3	0.3	0.3	0.3
Transient Conducted Emission Limit Allowable transient signal Level (V) for sensitive environment	0.3	0.3	0.2	0.2	0.1	0.1	0.1	0.1

Ratio of Tool Idle vs. Processing Energy Consumption (kW/kW)	**0.60**	**0.60**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**

10.3. IMPACT OF BIG DATA OF FACTORY OPERATIONS

To improve factory operations and traceability companies must invest in solutions to effectively manage their data growth. Data generation, storage and usage have increased in the factory because of the improvements of semiconductor equipment computer interfaces that provide higher rates for data collection and additional equipment parameter data availability. In addition to the increase of equipment generated data, manufacturing data analysis requires more complex data integration because the needed data comes from multiple sources and databases. Transactional volumes, velocity responsiveness, quantity, variety are ***exceeding traditional database and file systems processing capabilities***, and veracity of data created. This explosion of data growth in manufacturing has created a set of requirements that are commonly referred to as "Big Data". As a result there are significant efforts across industry to define Big Data and the Big Data problem. A consolidated effort is being headed by NIST (National Institute of Standards and Technology). Big Data is characterized by an increase in: data volume, velocity of generation (as well as variability in collection and storage rates), variety of data sources, difficulty in verifying the veracity, or "quality", of the data, and difficulty in obtaining maximum value from the data through efficient analytics and processing. From an information technology perspective, Big Data represents data sets whose size, type, speed of creation, or data quality make them impractical to process and effectively analyze with traditional database technologies and related tools in a cost- or time-effective way.

Moving to big data solutions involves addressing any number five Vs at various levels. Currently this is often accomplished by enhancing existing systems, e.g., to support larger data volumes or improved data quality. However over the longer term it is anticipated that all of manufacturing will move to include more big-data friendly solutions such as those that contain Hadoop Ecosystem components. Initially these solutions will be used primarily for off-line, non-real-time applications such as off-line data mining to support generation and maintenance of prediction models. In these areas, the move to big data-friendly solutions will be motivated by reduced cost of ownership with respect to data volumes, improved analysis processing speeds, and increased analysis capabilities resulting largely from the parallel processing capabilities of the ecosystem. Over the longer term some of these solutions will likely be used for on-line non-real-time applications; the development to support this capability will likely come from outside of the semiconductor industry.
The infrastructure needs for cloud computing for the factory can be borrowed from commercial computer clouds.

Making data available for advanced analytics will likely be challenging because of multiple levels of user data accessibility needs. Determination of standardized policies will be applied to Big Data to make sure internal and external users have access to the data. Empowerment in the organization to explore and discover uncovered patterns and internal resources will likely perform trends in the factory. Big Data will need to be secured and managed by the factory but access to it may be limited by security policies or firewalls inherent to the computer or server infrastructure

11. CHALLENGES AND POSSIBLE SOLUTIONS

11.1. SYSTEM INTEGRATION

11.1.1 Table SYSINT1: Summary of system integration challenges.

Near term (within 3 years)	*Sub-challenges*	*Relation between drivers*
Design productivity	System integration, AMS/MEMS co-design and design automation SIP and 3D (TSV-based) planning and implementation flows Heterogeneous integration (optical, mechanical, chemical, biomedical, etc.)	Mobile/IoT: Beneficial for system dimension scaling, performance improvement, and cost. Datacenter: Beneficial for system-wide bandwidth and power efficiency
Power management	Dynamic and static, system- and circuit-level power optimization	Mobile/ IoT: Beneficial for battery life Datacenter: Beneficial for cooling cost and energy fee
Manufacturability	Performance/power variability, device parameter variability, lithography limitations impact on design, mask cost, quality of (process) models	Mobile/datacenter/IoT: Beneficial for cost reduction and reliability improvement
Bandwidth / service latency	High performance memory / NVM interfaces, memory / processor stacking	Mobile: Beneficial for improving display capacity and develop more sophisticated services Datacenter: Beneficial for faster responses
Cooling	Temperature-constrained physical implementation, 3D integration/packaging	Mobile/datacenter: avoid heating issues

Long term (> 3 years)	*Sub-challenges*	*Relation between drivers*
Design productivity	System-Level Design Automation (SDA) Executable Specification	Mobile/datacenter/IoT: Beneficial for faster design turn-around-time and less design effort
Power management	On-die power sensors, silicon photonics, novel transistors and memory	Mobile/ IoT: Beneficial for battery life Datacenter: Beneficial for cooling cost and energy fee
Manufacturability	Sequential 3D integration 3D transistors (LGAA, VGAA, CNT) Novel memory technologies	Mobile/datacenter/IoT: Beneficial for cost reduction and reliability improvement
Bandwidth / service latency	High radix network, interfaces with novel memory devices	Mobile: Beneficial for improving display capacity and develop more sophisticated services Datacenter: Beneficial for faster responses
Cooling	Microfluidic cooling (single-phase / two-phase)	Mobile/datacenter: Avoid heating issues

AMS—analog/mixed signal MEMS—micro-electro-mechanical systems TSV—through silicon via NVM—non-volatile memory LGAA/VGAA—lateral/vertical gate-all-around CNT—carbon nanotube

11.1.2 Table SYSINT3: Key challenges and potential solutions of the mobile driver.

Challenges	Metrics (Description)	Roadblocks	Potential solutions
Form Factor Challenge	#Sensors, #ICs, #Antennas, (#Components ↑) Memory bandwidth (PCB routing complexity ↑, #connectors ↑)	Increasing PCB footprint occupied by connectors and components	1. Package-level integration 2. Through-silicon via (TSV)-based 3D integration 3. Sequential 3D integration [1]
		Integration of heterogeneous components	1. 3D integration [2] 2. Unified technology (CMOS RF/CMOS MEMS)
		Die area explosion due to more functionalities	1. Technology scaling
System-Level Power Management	Max freq., #AP cores, #GPU cores, Memory bandwidth (Power consumption ↑)	High-speed off-processor memory buses	1. Logic/memory stacking 2. Embedded DRAM [14]
		Increasing #sensors and #IC components	1. 3D integration 2. Highly-integrated RF/AP/peripheral controllers
System- wide Bandwidth Scaling	Memory bandwidth, #MPixel, Cellular data rate, WiFi data rate (Bandwidth requirement ↑)	High-speed off-processor memory buses	1. Logic/memory stacking 2. Embedded DRAM
		Increasing inter-component bandwidth requirement	1. 3D integration 2. Highly-integrated RF/AP/peripheral controllers
Communication Bandwidth Scaling	Cellular data rate, WiFi data rate (Bandwidth requirement ↑)	Increasing communication modes/bandwidth requirement (2015) for cellular phone and WiFi	1. Technology scaling for high performance RF devices 2. High-performance RF/baseband 3. Integrated multi-standard communication circuits (2016)
Sensor Pixel Scaling	#MPixel (Pixel density ↑, Optical design complexity ↑)	Pixel dimension scaling limited by optical performance	1. 3D integration (back-side illumination) [12] 2. Computational sensing [11]

11.1.3 Table SYSINT7: Key challenges and potential solutions of the datacenter and microserver drivers.

Challenges	Metrics	Roadblocks	Potential solutions
Service Latency Challenge	#MPU cores/rack unit, DRAM cap./ rack unit, Max Freq., DRAM bandwidth, Off-MPU bandwidth (Performance requirement ↑)	Low #hop connections	High-radix network [3]
		Low bit/J transmission	Silicon photonics [3] which can deliver higher switch BW and lower pJ/bit
		High performance memory architecture	NVMs to replace hard drives [3]
		High storage bandwidth	Distributed storage nodes [3]
		Encrypted data processing	Distributed compression and encryption engines [3]
		NVM reliability	Novel memory devices [3]
			Control algorithm [4]
Node Density/Cooling/ Power Management Challenge	#MPU cores/rack unit, DRAM cap./ rack unit, Max Freq., DRAM bandwidth, Off-MPU bandwidth (Power consumption ↑)	Die areas increase due to more functionalities	Moore's Law scaling
		Low power processor architecture	64-bit ARM core [5]
		Lack of one-fits-all processor architecture	Modularized processor 3D stacks [5]
		Power management for different application context	Integrated on-die power sensors [6]
Electro-Optical Integration Challenge	DRAM bandwidth, Off-MPU bandwidth (Bandwidth requirement ↑)	On-chip light source	Compound semiconductor [7] [13] Silicon Raman laser [8]
		On-chip detector / modulator	Compound semiconductor [9]
		Bendable fibers for limited server space	Material innovation [10]

[1] *Sequential 3D integration refers to 3D integration with fine-pitch TSV (close to gate pitch) while TSV-based refers to 3D integration with coarse-pitch TSV at function block level.*
[2] *3D integration refers to the super set of sequential 3D, TSV-based 3D, memory/logic stacking, sensor/logic stacking, etc.*

11.1.4 Table SYSINT10: Key challenges and potential solutions of the IoT driver.

Challenges	Metrics	Roadblocks	Potential solutions
Transistor device design and scaling	MCU #Cores, MCU ION / Operation frequency (μA/MHz), MCU Flash Size (KB), Deep suspend current (nA)	Leakage current management	Device: FinFET, LGAA, VGAA, and CNT
		Reliability issues due to logic transistor scaling	
		Threshold voltage scaling	
		Data corruption due to NVM device scaling	Device: Emerging memory devices (e.g., RRAM) Design technology: 3D stacking
IP/Sensor integration and scaling (More than Moore) challenge	#Sensors, Max Sensor Power (μW), DC-DC efficiency (%), DC-DC power density (W/mm^2)	Exclusive technology (analog, MEMS, logic… etc.)	Heterogeneous 3D integration
		On-die voltage regulation, scaling of passive components [42], conversion efficiency between different input/output of regulators	On-die magnetics [43], configurable switching capacitor [Alon11]
Supply voltage scaling challenge	Lowest VDD (V), Battery Power Density (Watt-Hr/Liter), Peak Tx/Rx current (mA), Tx/Rx power per bit (μW/bit)	Threshold voltage scaling and performance requirement	Near-threshold computing [45], asynchronous computing [46], stochastic computing [47], approximate computing [48]

11.2. HETEROGENEOUS INTEGRATION

11.2.1 Table HI-1 Heterogeneous Integration Difficult Challenges > 16nm

Table HI-1 Heterogeneous Integration Difficult Challenges

Difficult Challenges ≥16 nm	***Summary of Issues***
Impact of BEOL including Cu/low κ on packaging	– Improved fracture toughness of dielectrics
	– Interfacial adhesion
	– Probe damage for copper/ultra low κ
Wafer level Packaging	– 3D integration in WLP
	– Embedded components in WLP
	– CTE mismatch compensation for large die and fan-out die
System in Package	– Cross talk due to increased circuit density
	–Thermal density and hot spots
	– Heterogeneous integration of of different semiconductor materials
	– Heterogeneous integration of different circuit types (logic, memory, analog, RF, MEMS, power, passives etc.)
Coordinated design tools and simulators to address chip, package, and substrate co-design	– Mixed signal co-design and simulation at system level with heterogeneous integration
	– Integrated analysis tools for transient thermal analysis and integrated thermal mechanical analysis
	– Electrical (signal and power integrity with higher frequency/current and lower voltage operation)
	– Models for reliability prediction
Interposers and Embedded components	– CTE mismatch and warpage for large interposers
	– Embedded active devices
	– Electrical and optical interface integration, thermal interface
Thinned die packaging	– Handling technologies for thinned die and wafers (particularly for bumped wafers)
	– Testability

11.2.2 Table HI-1 Heterogeneous Integration Difficult Challenges <16nm

Table HI-1 Heterogeneous Integration Difficult Challenges

Difficult Challenges ≤16 nm	***Summary of Issues***
Close gap between chip and substrate, Improved Organic substrates	-Increased wireability at low cost
	-Improved impedance control and lower dielectric loss to support higher frequency applications
	-Silicon I/O density increasing faster than the package substrate technology
3D assembly and packaging	-Thermal management
	-Alignment/placement accuracy layer to layer
	- Test access for individual wafer/die
	-Cost of TSV and cost of Interposer
	-Bumpless interconnect architecture
Package cost does not follow the die cost reduction curve	– Wafer level packaging and 3D equipment cost is not scaling with product (transistor) cost
	– Increased device complexity requires higher cost packaging solutions
Small die with high pad count and/or high power density	– Electromigration at high current density for interconnect (die, package).
	– Thermal dissipation
	– Improved current density capabilities
High frequency die	– Lower loss dielectrics
	– "Hot spot" thermal management
Power Integrity	– Power delivery in stacked die
	– Reducing power supply voltage with high device switching currents

11.2.3 Heterogeneous Component Difficult Challenges

Difficult Challenges	***Summary of Issues***
Device Testing	Standard testing protocols are needed for defining performance parameters published in device datasheets. The current lack of standardization results in the inability of the customer to compare the cost and performance tradeoffs between the manufacturers.
	The cost of testing continues to rise yet system integrators expect prices to stay constant or lower even with increases in performance and function. Increasing performance and adding more functions requires higher accuracy tests and additional tests, which should normally increase testing costs.
	There is a continuing need to extend knowledge of the physics of failure of MEMS devices for accelerated reliability test methods.
Standard Process Modules	There has been a continuing discussion amongst MEMS designers and fabricators for decades about the pros and cons, and the possibility of moving towards standardization in fabrication process modules for MEMS. The discussion and interest on the need for this has been growing within the MEMS Industry Group but it has not yet gained enough momentum to produce specific actions.
Bandwidth for Connectivity	The new IPV6 has a total of 340 undecillion IP addresses, which is the equivalent of 3.4 with 38 zeros. This number of addresses will support the optimistic over 100 trillion connected devices in the world by 2030. However, if the envisioned trillion sensors are to communicate wirelessly there will be a shortage of wireless bandwidth unless new paradigms for managing and transmitting the data are developed.

11.3. OUTSIDE SYSTEM CONNECTIVITY

11.3.1 Table OSC1: Outside System Connectivity Difficult Challenges

Difficult Challenges 2015-2022	***Summary of Issues***
Achieving high performance energy efficient RF analog technology compatible with CMOS processing	Achieving high performance RF required reduced gate resistance which is difficult to achieve Integrating SiGe with CMOS is difficult to achieve high performance heterojunction bipolar transistors (HBT) Integrating III-Vs with CMOS it is difficult to achieve high performance III-V devices Increasing passive device functional density on chip; e.g., resistors, inductors, varactors, and capacitors
Deliver wireless capabilities to support a broad range of applications for IoT devices	Increasing antenna complexity to support multiple applications Security: While solutions to avoid tampering / intercept of RF communications will probably mainly rely on software solutions, and hardware technologies beyond the scope of the Outside System Connectivity Focus Team (e.g., cryptography), we cannot exclude the possibility that security concerns will have an impact on RF technology requirements
Reducing cost of Optical Interconnects	Reducing the large number of components (hundreds) that are expensive Reducing the cost of single mode connected optical devices Increasing optical interconnect density while reducing cost and power Reducing optical device power Increasing Density of MUX/DeMUX and reducing the size of this functionality. Developing low cost joining methods to provide the sub-micron location tolerances needed over the life cycle of single mode structures.
Increase the bandwidth density (bits/mm^2) of Optical Interconnects	Increasing the density of DeMUX with multiple wavelengths, polarizations, etc. Increasing the operating frequency of modulators while reducing power consumption Increasing the operating frequency of detectors while reducing power Simultaneously fabricating and joining multiple parallel fibers or waveguides
Improve the reliability of optical Interconnects	Determine factors that limit the reliability of lasers for single mode OI Improve the reliability of connector technology Determine factors that limit the reliability of modulators, MUXes, DeMuxes
Difficult Challenges 2023-2030	***Summary of Issues***
Agreeing on Optical technology standards	By 2023, multiple optical implementations are likely to be in use implying that standardization could result in cost saving. Will standards emerge or will multiple proprietary solutions prevail?
Processing information in the optical domain	Regenerate digital signals in the optical domain without returning to the electronic domain, a capability that is likely to require non-linear optical materials. Perform logic operations in the optical domain Switch optical streams (circuit switching) in the optical domain.
Establish a System Level Reliability Strategy	Identify a strategy for reliable O/I with point to point connectivity Identify a strategy for reliable O/I with LAN connectivity
Utilize the "Z" dimension more effectively	Optical devices are often linear or planar yet much could be done, especially to reduce size, utilizing the 3 rd. dimension. Design, but specially fabrication in "Z" is "hard". Some type of 3D printing might enable this technical solution.

11.3.2 Table OSC13: RF & AMS Difficult Challenges

RF and Analog/Mixed-Signal (RF and AMS) Technologies Difficult Challenges	Summary of Issues
CMOS Technologies	Many of the materials-oriented and structural changes being invoked in the digital roadmap degrade or alter RF and analog device behavior. Complex tradeoffs in optimization for RF, HF, and AMS performance occur as different mechanisms emerge as limiting factors. Examples include series resistances at gate, source and drain, that greatly affect parasitic impedances and the impact of such local interconnect parasitics on f_{MAX}. Fundamental changes of device structures, e.g., multiple-gates and silicon-on-insulator (SOI), to sustain continued digital performance and density improvements greatly alter RF and AMS characteristics. Such differences, along with the steady reduction in supply voltages, pose significant circuit design challenges and may drive the need to make dramatic changes to existing design libraries.
Group IV Bipolar Technologies	Even though it is a challenge for the HS-NPN to increase the unity current gain cut-off frequency f_T by more aggressive vertical profiles, it is less of a challenge to achieve $f_{MAX} > f_T$,. What is unclear today is, how large the ratio f_{MAX}/f_T needs to be for future circuit applications. That is, the challenge is to determine what this ratio should be by using the "plateau" technologies" for the next roadmap and appropriate benchmark circuits. Since lateral scaling requirements for HBTs are significantly relaxed compared with those for MOSFETs, vertical profile fabrication under the constraints of overall process integration appears to be the bigger challenge. The reduction of imperfections and the increase of current carrying capability of the emitter and collector contact metallization are further challenges that need to be met by process engineers on the way to achieving the physical limits of this and any other technology.
III-V Compound Semiconductor Technologies	The unique challenges are yield (manufacturability), substrate size, thermal management, integration density, dielectric loading, and reliability under high fields. Challenges common with Si-based circuits include improving efficiency and linearity/dynamic range, particularly for power amplifiers. A major challenge is increasing the functionality of power amplifiers in terms of operating frequency and modulation schemes while simultaneously meeting increasingly stringent linearity requirements at the same or lower cost.

11.4. MORE MOORE

11.4.1 Table MM2: Process integration difficult challenges.

Near-Term 2015-2022	*Summary of Issues*
1. Scaling Si CMOS	• Scaling of fully depleted SOI and multi-gate (MG) structures • Implementation of gate-all-around (nanowire) structures • Controlling source/drain series resistance within tolerable limits • Further scaling of EOT with higher K materials (K > 30) • Threshold voltage tuning and control with metal gate and high-K stack • Inducing adequate strain in advanced structures
2. Implementation of high-mobility CMOS channel materials	• Basic issues same as Si devices listed above • High-K gate dielectrics and interface state (D_{it}) control • CMOS (*n*- and *p*-channel) solution with monolithic material integration • Epitaxy of lattice-mismatched materials on Si substrate • Process complexity and compatibility with significant thermal budget limitations
3. Scaling of DRAM and SRAM	• DRAM— • Adequate storage capacitance with reduced feature size; implementing high-κ dielectrics • Low leakage in access transistor and storage capacitor; implementing buried gate type/saddle fin type FET • Low resistance for bit- and word-lines to ensure desired speed • Improve bit density and lower production cost in driving toward $4F^2$ cell size • SRAM— • Maintain adequate noise margin and control key instabilities and soft-error rate • Difficult lithography and etch issues
4. Scaling high-density non-volatile memory	• Endurance, noise margin, and reliability requirements • Multi-level at < 20 nm nodes and 4-bit/cell MLC • Non-scalability of tunnel dielectric and interpoly dielectric in flash memory – difficulty of maintaining high gate coupling ratio for floating-gate flash • Few electron storage and word line breakdown voltage limitations • Cost of multi-patterning lithography • Implement 3-D NAND flash cost effectively • Solve memory latency gap in systems
5. Reliability due to material, process, and structural changes, and novel applications.	• TDDB, NBTI, PBTI, HCI, RTN in scaled and non-planar devices • Gate to contact breakdown • Increasing statistical variation of intrinsic failure mechanisms in scaled and non-planar devices • 3D interconnect reliability challenges • Reduced reliability margins drive need for improved understanding of reliability at circuit level • Reliability of embedded electronics in extreme or critical environments (medical, automotive, grid...)
Long-Term 2023-2030	• *Summary of Issues*
1. Implementation of advanced multi-gate structures	• Fabrication of advanced non-planar multi-gate and nanowire MOSFETs to below 10 nm gate length • Control of short-channel effects • Source/drain engineering to control parasitic resistance • Strain enhanced thermal velocity and quasi-ballistic transport
2. Identification and implementation of new memory structures	• Scaling storage capacitor for DRAM • DRAM and SRAM replacement solutions • Cost effective installation of high density 3-D NAND (512 Gb – 4 Tb) with high layer numbers or tight cell pitch

	• Implementing non-charge-storage type of NVM cost effectively • Low-cost, high-density, low-power, fast-latency memory for large systems
3. Reliability of novel devices, structures, and materials.	• Understand and control the failure mechanisms associated with new materials and structures for both transistor and interconnect • Shift to system level reliability perspective with unreliable devices • Muon-induced soft error rate
4. Power scaling	• V_{dd} scaling while supplying sufficient current drive • Controlling subthreshold current or/and subthreshold slope • Margin issues for low V_{dd}
5. Integration for functional diversification	• Integration of multiple functions onto Si CMOS platform • 3-D integration

11.4.2. Table MM1: Interconnect difficult challenges.

Critical Challenges	**Summary of Issues**
Materials Mitigate impact of size effects in interconnect structures	Line and via sidewall roughness, intersection of porous low-κ voids with sidewall, barrier roughness, and copper surface roughness will all adversely affect electron scattering in copper lines and cause increases in resistivity.
Metrology Three-dimensional control of interconnect features (with its associated metrology) will be required	Line edge roughness, trench depth and profile, via shape, etch bias, thinning due to cleaning, CMP effects. The multiplicity of levels, combined with new materials, reduced feature size and pattern dependent processes, use of alternative memories, optical and RF interconnect, continues to challenge.
Process Patterning, cleaning, and filling at nano dimensions	As features shrink, etching, cleaning, and filling high aspect ratio structures will be challenging, especially for low-κ dual damascene metal structures and DRAM at nano-dimensions.
Complexity in Integration Integration of new processes and structures, including interconnects for emerging devices	Combinations of materials and processes used to fabricate new structures create integration complexity. The increased number of levels exacerbates thermomechanical effects. Novel/active devices may be incorporated into the interconnect.
Practical Approach for 3D Identify solutions which address 3D interconnect structures and other packaging issues	Three-dimensional chip stacking circumvents the deficiencies of traditional interconnect scaling by providing enhanced functional diversity. Engineering manufacturable solutions that meet cost targets for this technology is a key interconnect challenge.

11.4.3. Table MM2: GAA manufacturing difficult challenges.

Difficult challenges	Opportunities and issues
Transition from fin to GAA	• Extension of fin processes • Need to deal with higher aspect ratio starting topography
Strain engineering for GAA devices	• Continued use of embedded epitaxy for channel mobility boost • GAA mobility enhancements • Integration of dual channel materials • Usage of SOI substrates and stress conversion through condensation techniques • In vertical GAA architectures new techniques are needed to induce channel stress (e.g. reintroduction of stressed liners)
Junction engineering	• reducing junction concentration and achieving dopant redistribution • Conformal doping solutions are needed, increased importance for GAA or NW architectures
High mobility material integration	• different materials needed for NFET and PFET which leads to significant challenges to co integrate the materials • process solutions need to be compatible with material requirements such as low temperatures needed for post processing steps and new requirements on limiting the material losses in the subsequent processing steps • the materials used need to have low defectivity requirements (no killer defects in the channel)
Starting substrates	• GAA architectures can potentially be easier integrated and with less parasitics on SOI, sSOI, thin SOI substrates • Substrates for high mobility solutions – cost and defectivity are issues that will need to be addressed
Etch	• high aspect ratio – deep trenches, high pillars • high selectivity requirements compatible with the aggressive ground rules • GAA architectures and the need to eliminate parasitics drive the requirements for directional etching • Improved LER, etch bias and loading • Gate recess control and uniformity – driven by the aggressive ground rules and yield requirements
Material deposition	• Very high conformality processes are needed to wrap-around gate materials around the wires • High aspect ratio fill capability is needed (taller and thinner structures) • Gate fill solutions for sub 20 nm gates with acceptable gate resistance • Contact metal deposition solutions for increased aspect ratio contacts (vertical devices) • TDDB requirements (smaller distances needed between gate and contacts drive the requirements for the insulator materials used)
Cleans	• Good clean or removal without any residual defects or material removal
CMP	• Control and uniformity for gate and contact • Higher starting topography • Both non selective and highly selective slurries

11.4.4. Table MM3: Metrology difficult challenges.

Difficult Challenges ≥ 10 nm	*Summary of Issues*
Factory level and companywide metrology integration for real-time *in situ*, integrated, and inline metrology tools.	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. Need for continued development of robust sensors and process controllers; Data management that allows integration of add-on sensors. Reduction of scrap, increased product quality and cycle time.
Starting materials metrology and associated manufacturing metrology	Existing capabilities will not meet Roadmap specifications. Very small particles must be detected and properly sized. Capability for SOI, III-V, GeOI wafers needs enhancement. Challenges come from the extra optical reflection in SOI and the surface quality. CD, film thickness, and defect detection are impacted by thin SOI optical properties and charging by electron and ion beams. Impurity detection (especially particles) at levels of interest for starting materials and reduced edge exclusion for metrology tools.
Control of new process technology such as Directed Self Assembly (DSA) Lithography, multi-patterning, complicated 3D structures such as FinFET & MuGFET transistors, capacitors and contacts for memory, and 3D Interconnect are not ready for their rapid introduction.	Increased adoption of FinFET transistor technology has placed renewed emphasis on the near term need for in-line metrology for dimensional, compositional, and doping measurements. The materials properties of block co-polymers for DSA result in new challenges for lithography metrology. The increased use of multi-patterning techniques introduces the need to independently solve a large set of metrics to fully characterize a multi-patterning process. 3D Interconnect comprises a number of different approaches. New process control needs are not yet established. For example, 3D (critical dimension (CD) and depth) measurements will be required for trench structures including capacitors, devices, and contacts. Traditional metrology instruments do not have the range and resolution required for accurate TSV measurement.
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 engineered thin films and interface layers as well 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. Carrier mobility characterization will be needed for stacks with strained silicon and SOI, III-V, GeOI, and other substrates, or for measurement of barrier layers. Metal gate work function characterization is another pressing need.
Measurement test structures and reference materials.	The area available for test structures is being reduced, especially in the scribe lines. Measurements on test structures located in scribe lines may not correlate with in-die performance. Overlay and other test structures are sensitive to process variation, and test structure design must be improved to ensure correlation between measurements in the scribe line and on chip properties. Standards institutions need rapid access to state of the art development and manufacturing capability to fabricate relevant reference materials.
Difficult Challenges < 10 nm	
Nondestructive, production worthy wafer and mask-level metrology for CD measurement for 3D structures, overlay, defect detection, and analysis	Surface charging and contamination interfere with electron beam imaging. CD measurements must account for overall feature profile. Metrology tool imaging resolution must improve to be able to discern 3D information. It is important to have both imaging and scattering techniques available for any given process control situation. Focus, exposure, and etch bias control will require better precision and 3D capability.
New strategy for in-die metrology must reflect across chip and across wafer variation.	Correlation of test structure variations with in-die properties is becoming more difficult as devices shrink. Sampling plan optimization is key to solving these issues.
Statistical limits of sub-12 nm process control	Controlling processes where the natural stochastic variation limits metrology will be difficult. Examples are low-dose implant, thin-gate dielectrics, surface, and sidewall and edge roughness of very small structures. Complementary, and hybrid metrology combined with state of the art statistical analyses would be required to reduce the measurement uncertainty.
Structural and elemental analysis at device dimensions and measurements for *beyond CMOS, and emerging materials and devices.*	Materials characterization and metrology methods are needed for control of interfacial layers, dopant positions, defects, and atomic concentrations relative to device dimensions. One example is 3D dopant profiling. Measurements for self-assembling processes are also required.
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.
Mask Defects	Mask defects, especially for EUV will continue to be a challenge. These include non-visible defects, film thickness non-uniformity, phase separation, and reflectivity.
DSA	Key measurands such as size, location, and alignment need to be better defined. Some of the measurands are material and system dependent. Many of the materials are similar enough that identifying a property with the required contrast may be difficult. A key question seems to be if we can detect low densities of surface and buried defects.

11.4.5. Table MM4: Metrology potential solutions.

Potential Solutions	Examples
Improved Resolution	Better resolution for current technologies such as CD-SEM, CD-AFM, and optical CD among others. Increased range for high resolution instruments and vice –versa (This will greatly increase the ability to measure features such as TSVs) • *Introduction of aberration-corrected low energy SEM column* • *Utilization of high energy SEM* • *Enhanced CD-AFM tip technology* • *Reduced spot size and uniform intensity for optical instruments* • *Use of multi-column electron beam instrument for defect inspection.* • *Increased use of data fusion or image stitching to increase range.*
Improved X-ray metrology for CD and films characterization	Higher brightness sources. • *An X-ray source with >100x brightness of conventional rotating anode sources can enable new techniques such as CD-SAXS and X-ray tomography* • *Improved throughput and increased utilization of already-mainstream X-ray metrology solutions such as HR-XRD, XRF, TXRF, XRR, and XPS among others.*
3D Metrology	Non raster capabilities for scanning instruments. • *This would allow extraction of information from non-orthogonal axes.* Multi head/column for scanning instruments: • *Each head could extract different type of information (dimensional, material ...etc.),* • *Faster measurements of large areas of the wafer.* Hybrid Metrology: • *Increased use of a combination of instruments to achieve the desired resolution, speed, or low levels of uncertainty needed to characterize different aspects of a feature.*

11.4.6. Table MM5: Reliability difficult challenges.

Near-Term 2015-2022	***Summary of issues***
Reliability due to material, process, and structural changes, and novel applications.	TDDB, NBTI, PBTI, HCI, RTN in scaled and non-planar devices. Gate to contact breakdown. Increasing statistical variation of intrinsic failure mechanisms in scaled and non-planar devices. 3D interconnects reliability challenges. Reduced reliability margins drive need for improved understanding of reliability at circuit level. Reliability of embedded electronics in extreme or critical environments (medical, automotive, grid...).
Long-Term 2023-2030	***Summary of issues***
Reliability of novel devices, structures, and materials.	Understand and control the failure mechanisms associated with new materials and structures for both transistor and interconnect. Shift to system level reliability perspective with unreliable devices. Muon induced soft error rate.

11.5. BEYOND CMOS

11.5.1 Emerging Research Devices Difficult Challenges

Difficult Challenges – 2020– 2030	*Summary of Issues and opportunities*
Scale high-speed, dense, embeddable, volatile/nonvolatile memory technologies to replace SRAM and FLASH in appropriate applications.	SRAM and FLASH scaling in 2D will reach definite limits within the next several years (see PIDS). These limits are driving the need for new memory technologies to replace SRAM and FLASH memories. Identify the most promising technical approaches to obtain electrically accessible, high-speed, high-density, low-power, (preferably) embeddable volatile and nonvolatile memories. The desired material/device properties must be maintained through and after high temperature and corrosive chemical processing. Reliability issues should be identified & addressed early in the technology development.
Extend CMOS scaling	Develop 2nd generation new materials to replace silicon (or InGaAs, Ge) as an alternate channel and source/drain to increase the saturation velocity and to further reduce V_{dd} and power dissipation in MOSFETs while minimizing leakage currents for technology scaled to 2020 and beyond. Develop means to control the variability of critical dimensions and statistical distributions (e.g., gate length, channel thickness, S/D doping concentrations, etc.) Accommodate the heterogeneous integration of dissimilar materials. The desired material/device properties must be maintained through and after high temperature and corrosive chemical processing. Reliability issues should be identified & addressed early in this development.
Extend ultimately scaled CMOS as a platform technology into new domains of application.	Discover and reduce to practice new device technologies and primitive-level architecture to provide special purpose optimized functional cores (e.g., accelerator functions) heterogeneously integrable with CMOS.
Continue functional scaling of information processing technology substantially beyond that attainable by ultimately scaled CMOS.	Invent and reduce to practice a new information processing technology eventually to replace CMOS as the performance driver. Ensure that a new information processing technology has compatible memory technologies and interconnect solutions. A new information processing technology must be compatible with a system architecture that can fully utilize the new device. Non-binary data representations or non-Boolean logic may be required to employ a new device for information processing, which will drive the need for new system architectures. Bridge the gap that exists between materials behaviors and device functions. Accommodate the heterogeneous integration of dissimilar materials. Reliability issues should be identified & addressed early in the technology development.
Invent and reduce to practice long term alternative solutions to technologies that address existing MtM ITRS topical entries.	The industry is now faced with the increasing importance of a new trend, "More than Moore" (MtM), where added value to devices is provided by incorporating functionalities that do not necessarily scale according to "Moore's Law". Heterogeneous integration of digital *and* non-digital functionalities into compact systems that will be the key driver for a wide variety of application fields, such as communication, automotive, environmental control, healthcare, security, and entertainment.

11.6. FACTORY INTEGRATION

11.6.1. Table FI-3 Factory Integration Difficult Challenges

Difficult Challenges through 2023	*Summary Of Issues*
1. Responding to rapidly changing, complex business requirements	• Increased expectations by customers for faster delivery of new and volume products (design → prototype and pilot → volume production) • Rapid and frequent factory plan changes driven by changing business needs • Ability to load the fab within manageable range under changeable market demand, e.g., predicting planning and scheduling in real-time • Enhancement in customer visibility for quality assurance of high reliability products; tie-in of supply chain and customer to Factory Information and Control Systems (FICS) operations • Addressing the Big Data issues, thereby creating an opportunity to uncover patterns and situations that can help prevent or predict unforeseeable problems difficult to identify such as current equipment processing / health tracking and analytical tools • To strengthen information security: Maintaining data confidentiality (the restriction of access to data and services to specific machines/human users) and integrity (accuracy/completeness of data and correct operation of services), while improving availability (a means of measuring a system's ability to perform a function in a particular time) contradictive to needs of data availability.
2. Managing ever increasing factory complexity	• Quickly and effectively integrating rapid changes in process technologies • Complexity of integrating next generation equipment into the factory • Increased requirements for high mix factories. Examples are (1) significantly short life of products that calls frequent product changes, (2) the complex process control as frequent recipe creations and changes for process tools and frequent quality control criteria due to small lot sizes, (3) managing load on tools • Manufacturing knowledge and control information needs to be shared as required among factory operation steps and disparate factories in a secure fashion • Need to concurrently manage new and legacy FICS software and systems with increasingly high interdependencies • Ability to model factory performance to optimize output and improve cycle time for high mix factories • Need to manage clean room environment for more environment susceptible processes, materials, and, process and metrology tools • Addressing need to understand and minimize energy resource usage and waste; determining what the energy usage profile actually is; e.g., need to integrate fab management and control with facilities management and control • Comprehending increased purity requirements for process and materials • Providing a capability for more rapid adaptation, re-use and reconfiguration of the factory to support capabilities such as rapid new process introduction and ramp-up. This includes a challenge of supporting evolution of a FI communication infrastructure to support emerging capabilities beyond interface A. • Communication protocols developed for semiconductor manufacturing are not aligned with trends in information technology communication such as web services. • Meeting challenges in maintaining yield and improving maintenance practices resulting from movement to new process materials that may be corrosive, caustic, environmentally impacting, molecularly incompatible etc.

	• Addressing factory integration challenges to assess and integrate EUV systems into the factory infrastructure • Address process hazard management issues • Addressing Airborn Molecular Contamination (AMC) challenges through possibly changing factory operation approach (e.g., maintaining vacuum in specific areas), as well as providing necessary interfaces, information and technologies (e.g., virtual metrology and APC). • Maintaining equipment availability and productivity, and minimizing equipment variability, while managing increase in sensors and systems, and associated data volume increases within and outside the equipment, coordinated to support new paradigms (e.g., management of energy expended by the equipment and the fab in general, augmenting reactive capabilities with predictive) • Linking yield and throughput prediction into factory operation optimization; incorporating incoming disturbances such as power and materials purity variability into the prediction equation. • Achieving real-time simulation of all fab operations as an extension of existing system with dynamic updating of simulation models • Understanding and managing queue times (time between operations/segments) and production of product within those times to achieve acceptable product quality • Managing and protecting IP, avoiding security issues such as malware attacks, and protection of the facility's instrumentation and control systems from attack • Addressing FI issues associated with implementing emerging technology revolutions (rather than evolutions) in achieving production targets, including rapid integration of new tools, components and materials, leveraging existing infrastructure, ramping up on new technology ramp-up • Achieving compatibility of existing systems that are largely reactive with emerging predictive paradigms of operation, such as Predict Maintenance and yield prediction • Addressing shifting focus from line width pitch shrinks, to 3D and emerging disruptive technologies
3. Achieving financial growth targets while margins are declining	• Developing metrics on performance of factory integration systems and understanding how these metrics translate to factory financial information • Achieve waste reduction continuous improvement targets, e.g., through equipment cycle time reduction, and reduction in power consumption • Improving efficiency of factory operations through tighter integration with supply chain, e.g., to achieve lean manufacturing targets • Incorporating product priority into factory integration planning and operations to achieve financial objectives
4. Meeting factory and equipment reliability, capability, productivity and cost requirements per the Roadmap	• Increased impacts that single points of failure have on a highly integrated and complex factory • Achieving better communication between equipment suppliers and users with respect to equipment requirements and capabilities • Improved bi-direction information exchange between equipment and factory systems to achieve equipment and factory reliability, capability and productivity objectives • Design-in of equipment capability visualization in production equipment; design-in of APC (R2R control, FD , FC and SPC) to meet quality requirements • Equipment data, analytics and visualization to support equipment health monitoring (EHM) • Developing and implementing methods that reduce the use of NPW (non-product wafers) and the associated lost production time

	• Reducing undesired wait-time waste; developing wait-time waste reporting for tools; providing standardized equipment wait-time waste metrics reporting to support fab-wide equipment wait-time waste management • Augmenting reactive with a predictive paradigm for scheduling, maintenance and yield management • Meeting tighter and more granular control requirements such as wafer-to-wafer (e.g., single-wafer oriented) and within wafer utilizing technologies such as virtual metrology • Yield mining techniques that support root cause analysis for determination of contributions to yield loss in the process stream. • Addressing the move towards "lights out" human-less operation in the fab to meet goals such as contamination levels. • More comprehensive traceability of individual wafers to identify problems to specific process areas • Standards for supply chain traceability of spares, e.g., for better understanding of lifetime of spares • Standards and best practices to support providing degradation characteristics of components from suppliers for improved tracking and predicting of failures • Comprehensive management that allows for automated sharing and re-usages of complex engineering knowledge and contents such as process recipes, APC algorithms, FD and C criteria, equipment engineering best known methods
5. Cross leveraging factory integration technologies across boundaries such as 300mm and 450mm to achieve economy of scale	• Addressing the potential data explosion and other big data issues associated with crossing a technology boundary • Ensuring that we take advantage of the technology change to implement the appropriate factory integration enhancements such as control system paradigm shift. • Understanding the software roadmap for moving across these technology boundaries. 450mm era: Effecting architectural and other changes as necessary at an affordable cost to maintain or improve wafer-throughput-to-footprint levels in migration to 450mm
6. Addressing unique challenges in the move to 450mm (where 300mm technologies cannot always be leveraged)	• Understanding and addressing 450mm implications on issues such as product logistics (e.g., AMHS), utility usage and factory design and layout • Determining and specifying what technologies (such as prediction) become required (as opposed to just desired) as they must be leveraged to achieve 450mm quality and production goals.

Difficult Challenges Beyond 2023	*Summary of Issues*
1. Meeting the flexibility, extendibility, and scalability needs of a cost-effective, leading-edge factory	• Evaluating and implementing revolutionary disruptive technologies such as distributed autonomous control at the appropriate time to maximize cost competitiveness • Determining the appropriate time to move to 450mm for all high volume commodity production • Consider the possibility of self-evolving and self-configuring FI technologies such as data analysis and prediction where software (re-)configuration tasks are greatly reduced • Adoption of augmented reality capabilities for enhanced human machine interaction • Cost and task sharing scheme on industry standardization activity for industry infrastructure development • Achieving the "prediction vision" of a state of fab operations where (1) yield and throughput prediction is an integral part of factory operation optimization, and (2) real-time simulation of all fab operations occurs as an extension of existing system with dynamic updating of simulation models.
2. Increasing global restrictions on environmental issues	• Addressing the move towards global regulations • Developing methods for increasing material reclamation • Proactively addressing future material shortages, such as non-renewable chemicals
3. Post-conventional Semiconductor manufacturing uncertainty	• Uncertainty of novel device types replacing conventional CMOS and the impact of their manufacturing requirements on factory design • Timing uncertainty to identify new devices, create process technologies, and design factories in time for a low risk industry transition • Potential difficulty in maintaining an equivalent 0.7× transistor shrink per year for given die size and cost efficiency

11.6.2. Table FI-4 Key Focus Areas and Issues for FI Functional Areas Beyond 2015

Functional Area	***Key technology focus and issues***
Factory Operations (FO)	1. Systematic productivity improvement methodology of the current "lot-based" manufacturing method prior to 450mm insertion 2. Challenges in moving to smaller lot and single wafer aspects of factory operations 3. Interdisciplinary factory productivity improvement method such as systematic factory waste visualization of manufacturing cycle times and factory output opportunity losses 4. Extendable and reconfigurable factory service structure
Production Equipment (PE)	1. 450mm production tool development 2. for integration into the factory information system; supporting bridge capabilities to 450mm 3. Determining context data set for equipment visibility 4. Equipment health monitoring (EHM) and fingerprinting to support improved uptime. 5. Run rate (throughput) improvement and reduction of equipment output waste that comes from NPW and other operations 6. Improving equipment data quality and data accessibility to support capabilities such as APC and e-Diagnostics 7. 7) Develop equipment capabilities to support the move to a predictive mode of operation (including virtual metrology, predictive maintenance, predictive scheduling and yield prediction and feedback); examples include reporting equipment state information, time synchronization, and equipment health monitoring (EHM) and reporting. 8. Migrate to a mode of operation where APC is mandatory for proper execution of process critical steps 9. Design, Develop and implement (standardized where appropriate) capabilities for utility (e.g., electricity) reduction such as support for idle mode, improved scheduling, and communication between host and equipment for energy savings
Automated Material Handling Systems (AMHS)	1. Reduction in average delivery times, 2. Avoid tool starvation 3. More interactive control with FICS and PE for accurate scheduled delivery, including (predictive) scheduling/dispatch, maintenance management, and APC 4. Aim for continuous improvement in reliability and corresponding minimization of downtime 5. 450mm specific AMHS issues 6. AMHS interaction with other wafer transport and storage systems such as sorter and load port
Factory Information and Control Systems (FICS)	1. Increased reliability of FICS systems such as maintenance management 2. Increased FICS performance for more complex factory operation, such as decision speed and accommodating larger data sets 3. Enhanced system extendibility including extensibility across Fabs 4. Utilize FICS information to achieve waste-reduction (e.g., wait-time waste, unscheduled downtime, and wafer scrap) and sustainability (e.g., resource conservation) 5. Facilitate enhancement of reactive with predictive approach to operations (e.g., planning and scheduling, maintenance, virtual metrology and yield prediction and feedback) 6. Determining approaches to control (e.g., distributed versus centralized) and when to institute disruptive control systems changes (e.g., at 450mm introduction) 7. Achieving minimum downtime, seamless transition, and uninterrupted operations in production throughout the software upgrade process
Facilities	1. Continuous improvement to maintain facility systems viability 2. Minimization of facilities induced production impacts 3. Facility cost reduction 4. Determining and addressing emerging technology requirements such as AMC (Airborne Molecular Contamination) control, 450mm, 3D, etc. 5. Maintaining safety in facilities operations (e.g., in response to a seismic event) 6. Even more aggressive focus on environmental issues and optimization to environmental targets. Facility utility reduction
Augmenting Reactive with Predictive (ARP)	1. Improved data quality to support effective prediction 2. Prediction solutions tied to application financials for optimized benefit 3. Integration of predictive functions (data, algorithm, user interface, and cross-leveraging capabilities) as an augmentation of existing systems 4. Move to real-time simulation of all fab operations occurring as an extension of existing system with dynamic updating of simulation models
Big Data (BD)	1. Optimization of data storage volumes and data access to achieve FI objectives and enable applications to plug and play

	2. Speed improvement in collecting, transferring, storing and analyzing data 3. Software optimization to gather data from multiple systems and sources for analysis resulting on actionable decisions 4. Data quality improvements to address issues of time synchronization, accurate compression / uncompression, and merging of data form multiple sources collected at potentially varying data rates 5. Migrating from relational data storage infrastructure to largely big data friendly infrastructure such as Hadoop, along with small relational component. 6. Algorithm development and implementation to support emerging capabilities such as predictive and machine learning
Control Systems Architecture (CSA)	1. Addressing evolutionary aspects of control system and control system architectures such as granularity, speed, quality, and capability. 2. Addressing potentially revolutionary aspects of control systems and control systems architectures such possible moves to cloud computing, distributed/autonomous control, and artificial intelligence enhanced control 3. Addressing the framework to integrate monitoring and closed loop control tied to all semiconductor manufacturing key performance indices – including engineering and manufacturing control levels. 4. Addressing the capability to integrate with supply-chain framework for value chain control.
Environmental Safety and Health (ESH)	1. The roadmapping process will continue to quantify factory environmental factors 2. Roadmapping from 2015 will include, new materials, sustainability and green chemistry 3. Provide proactive engagement with stakeholder partners and reset strategic focus on the roadmap goals. 4. Continue focus on factory, and supply chain safety for employees and the environment
Yield Enhancement (YE)	1. The road mapping focus will move from a technology orientation to a product/application orientation. 2. Airborne molecular contamination (AMC), packaging, liquid chemicals and ultra-pure water were identified as main focus topics for the next period. 3. Electrical characterization methods, Big Data and modeling will become more and more important for yield learning and yield prediction.

11.6.3. Table FI-13 Crosscut Issues Relating to Factory Integration

Crosscut Area	***Factory integration related key challenges***
Front end Process (FEP)	Factory and FEP teams will continue to work on AMC requirements. FEP and Factory Integration will work on 450 mm challenges. Energy conservation effort: such as equipment sleep mode for energy conservation and the 1.5 mm wafer edge exclusion for long term challenge to starting material and SOI. ARP will impact all FE process in some way; coordinate roadmaps to make sure FEPs are moving toward "prediction ready", e.g., by providing necessary data
Lithography	Continuing to understand EUVL (power, consumables) requirements from FI perspective; completely different factory design is expected. Fast reticle change; reticle storage issues and reticle buffering due to small lots. AMC relative to the reticle and tighter process control needs. Lithography DFM needs. EFM may be added as it is confirmed as mask quality detractor. Predictive scheduling is important to lithography as it is often the critical process to maintaining throughput.
ESH	See new ESH subsection in FI chapter
Metrology	Comprehensive metrology roadmap to be jointly defined. AMC, temperature, and humidity control remain crosscut issues Virtual Metrology is an emerging cross-cut issue. The role of VM in metrology will be increasing; VM may become an integral part of some metrology offerings. Metrology capabilities will become part of the prediction engine input (e.g., for throughput projections) and output (e.g., for VM tuning). Metrology requirements on ESD and EMI could impact FI targets
Yield Enhancement	See new YE subsection in FI chapter
Test	Big data and prediction requirements and solutions will impact and provide solutions for Test.