Roadmap
Past, Present and Future

Paolo Gargini
Chairman ITRS2.0
Fellow IEEE, Fellow I-JSAP
Intel Fellow (1995-2012)
Multiple Stories

1. **Introduction**
2. 1998. ITRS 1.0
   1. Equivalent Scaling
3. 2000. NNI.
5. 2005-6. MPU power limits
6. 2006. More than Moore
   1. Heterogeneous Integration
7. 2010. First Selection of post CMOS devices
9. 2014. Scaling acceleration
10. 2014-15. ITRS 2.0
11. 2015. Post CMOS map of devices
12. 2016. IRDS
13. 2017-2021. 3D POWER Scaling

April 2016  SPCC
Second Update of Moore’s Law

International Electron Device Meeting, December 1975

April 2016

SPCC

P. Gargini
Dennard Scaling

<table>
<thead>
<tr>
<th>Device or Circuit Parameter</th>
<th>Scaling Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device dimension tox, L, W</td>
<td>1/K</td>
</tr>
<tr>
<td>Doping concentration Na</td>
<td>K</td>
</tr>
<tr>
<td>Voltage V</td>
<td>1/K</td>
</tr>
<tr>
<td>Current I</td>
<td>1/K</td>
</tr>
<tr>
<td>Capacitance eA/t</td>
<td>1/K</td>
</tr>
<tr>
<td>Delay time per circuit VC/I</td>
<td>1/K</td>
</tr>
<tr>
<td>Power dissipation per circuit VI</td>
<td>1/K²</td>
</tr>
<tr>
<td>Power density VI/A</td>
<td>1</td>
</tr>
</tbody>
</table>

Dennard’s 1974 paper summarizes transistor or circuit parameter changes under ideal MOSFET device scaling conditions, where K is the unitless scaling constant.

The benefits of scaling: as transistors get smaller, they can switch faster and use less power. Each new generation of process technology was expected to reduce minimum feature size by approximately 0.7x (K ~1.4). A 0.7x reduction in linear features size provided roughly a 2x increase in transistor density.

Dennard scaling broke down around 2004 with unscaled interconnect delays and our inability to scale the voltage and the current due to reliability concerns.

But our the ability to etch smaller transistors has continued spawning multicore designs.
Moore’s Law and Dennard’s Scaling Laws Convergence

=> 30% LINEAR FEATURE REDUCTION

S=0.7

50%

50% AREA REDUCTION
GENERATION TO GENERATION
Phase 1
First Age of Scaling
(Self-aligned Silicon Gate)
IC Industry at a Glance
(1975-2003)

Driver  Cost/transistor  ->  50% Reduction

How  2x Density/2 years (Moore)

Method  Geometrical Scaling (Dennard)
The Incredible Shrinking Silicon Technology of the 90’s

- **1995**: 0.35 μm
- **1997**: 0.25 μm
- **1999**: 0.18 μm
Gate Dielectric Scaling

Monolayers

1997 NTRS
SPCC

From My Files

April 2016

P.Gargini
The Focus Center Research Program

- A set of national multi-university teams
- Long-range research vision & agendas
Multiple Stories

1. Introduction

2. **1998. ITRS 1.0**
   1. Equivalent Scaling

3. 2000. NNI.


5. **2005-6. MPU power limits**

6. 2006. More than Moore
   1. Heterogeneous Integration

7. 2010. First Selection of post CMOS devices


9. 2014. Scaling acceleration

10. 2014-15. ITRS 2.0

11. 2015. Post CMOS map of devices

12. 2016. IRDS

13. 2017-2021. 3D POWER Scaling

April 2016

SPCC

P.Gargini
1998 ITRS Update

- Participation extended to: EECA, EIAJ, KSIA, TSIA at WSC on April 23, 1998
- 1st Meeting held on July 10/11, 1998 in San Francisco
- 2nd meeting held on December 10/11, 1998 at SFO
- 50% of tables in 1997 NTRS required some changes
- 1998 ITRS Update posted on web in April 1999
Phase 2
Second Age of Scaling
(Equivalent Scaling)
The Ideal MOS Transistor

- Fully Surrounding Metal Electrode
- Fully Enclosed, Depleted Semiconductor
- High-K Gate Insulator
- Low Resistance Source/Drain
- Band Engineered Semiconductor

April 2016

SPCC

P.Gargini
From Strategy to Implementation

- ITRS
- Technology Needs
- Possible Solutions
  - Consortia
  - Researchers
  - Suppliers
  - Detailed Solutions
  - Suppliers
  - IC Makers
  - OEM

Implementation

ITRS 7/11/1998
IC Industry at a Glance (2003->2021)

Driver: Cost/transistor-> 50% Reduction

How: 2x Density/2 years (Moore)

Method: Equivalent Scaling (ITRS1.0)
The Start of the ITRS

http://www.itrs2.net

1992 NTRS
1994 NTRS
1997 NTRS

Europe
Japan
Korea
Taiwan
USA

1998 ITRS Update
1999 ITRS
2000 ITRS Update
2001 ITRS
2002 ITRS Update

April 2016
SPCC
P.Gargini
High-k/Metal-Gate
(year 2000)

SRC/ISMT Front End Processes Research Center
Transistor Innovations Enable Technology Cadence

Four year pace of introduction of Equivalent Scaling into production

Source: Intel

April 2016

SPCC
## Incubation Time

<table>
<thead>
<tr>
<th></th>
<th>Early Invention</th>
<th>Focused Research</th>
<th>Introduction Manufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>HKMG</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<11 years

April 2016

SPCC
## Sortable table

<table>
<thead>
<tr>
<th>Model</th>
<th>Core</th>
<th>Clock Speed</th>
<th>Thermal Design Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pentium 4 1.3</td>
<td>Willamette (180 nm)</td>
<td>1.3 GHz</td>
<td>51.6 W</td>
</tr>
<tr>
<td>Pentium 4 1.4 (Socket 423)</td>
<td>Willamette (180 nm)</td>
<td>1.4 GHz</td>
<td>54.7 W</td>
</tr>
<tr>
<td>Pentium 4 1.4 (Socket 478)</td>
<td>Willamette (180 nm)</td>
<td>1.4 GHz</td>
<td>55.3 W</td>
</tr>
<tr>
<td>Pentium 4 1.5 (Socket 423)</td>
<td>Willamette (180 nm)</td>
<td>1.5 GHz</td>
<td>57.8 W</td>
</tr>
<tr>
<td>Pentium 4 1.5 (Socket 478)</td>
<td>Willamette (180 nm)</td>
<td>1.5 GHz</td>
<td>57.9 W</td>
</tr>
<tr>
<td>Pentium 4 1.6 (Socket 423)</td>
<td>Willamette (180 nm)</td>
<td>1.6 GHz</td>
<td>61 W</td>
</tr>
<tr>
<td>Pentium 4 1.6 (Socket 478)</td>
<td>Willamette (180 nm)</td>
<td>1.6 GHz</td>
<td>60.8 W</td>
</tr>
<tr>
<td>Pentium 4 1.7 (Socket 423)</td>
<td>Willamette (180 nm)</td>
<td>1.7 GHz</td>
<td>64 W</td>
</tr>
<tr>
<td>Pentium 4 1.7 (Socket 478)</td>
<td>Willamette (180 nm)</td>
<td>1.7 GHz</td>
<td>63.5 W</td>
</tr>
<tr>
<td>Pentium 4 1.8 (Socket 423)</td>
<td>Willamette (180 nm)</td>
<td>1.8 GHz</td>
<td>66.7 W</td>
</tr>
<tr>
<td>Pentium 4 1.8 (Socket 478)</td>
<td>Willamette (180 nm)</td>
<td>1.8 GHz</td>
<td>66.1 W</td>
</tr>
<tr>
<td>Pentium 4 1.9 (Socket 423)</td>
<td>Willamette (180 nm)</td>
<td>1.9 GHz</td>
<td>69.2 W</td>
</tr>
<tr>
<td>Pentium 4 1.9 (Socket 478)</td>
<td>Willamette (180 nm)</td>
<td>1.9 GHz</td>
<td>72.8 W</td>
</tr>
<tr>
<td>Pentium 4 HT 560</td>
<td>Prescott (90 nm)</td>
<td>3.6 GHz</td>
<td>115 W</td>
</tr>
<tr>
<td>Pentium 4 HT 560J</td>
<td>Prescott (90 nm)</td>
<td>3.6 GHz</td>
<td>115 W</td>
</tr>
<tr>
<td>Pentium 4 HT 561</td>
<td>Prescott (90 nm)</td>
<td>3.6 GHz</td>
<td>115 W</td>
</tr>
<tr>
<td>Model</td>
<td>Core</td>
<td>Clock Speed</td>
<td>Thermal Design Power</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Core 2 Duo E4200</td>
<td>Conroe (65 nm)</td>
<td>1.6 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E4300</td>
<td>Conroe (65 nm)</td>
<td>1.8 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E4400</td>
<td>Conroe (65 nm)</td>
<td>2.0 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E4500</td>
<td>Conroe (65 nm)</td>
<td>2.2 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E4600</td>
<td>Conroe (65 nm)</td>
<td>2.4 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E4700</td>
<td>Conroe (65 nm)</td>
<td>2.6 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E8190</td>
<td>Wolfdale (45 nm)</td>
<td>2.66 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E8200</td>
<td>Wolfdale (45 nm)</td>
<td>2.66 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E8300</td>
<td>Wolfdale (45 nm)</td>
<td>2.83 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E8400</td>
<td>Wolfdale (45 nm)</td>
<td>3.0 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E8500</td>
<td>Wolfdale (45 nm)</td>
<td>3.16 GHz</td>
<td>65 W</td>
</tr>
<tr>
<td>Core 2 Duo E8600</td>
<td>Wolfdale (45 nm)</td>
<td>3.33 GHz</td>
<td>65 W</td>
</tr>
</tbody>
</table>
IEEE, ISSCC: Transistor’s 60th year commemorative supplement
Multiple Stories

1. Introduction
2. 1998. ITRS 1.0
   1. Equivalent Scaling
3. 2000. NNI.
5. 2005-6. MPU power limits
6. 2006. More than Moore
   1. Heterogeneous Integration
7. 2010. First Selection of post CMOS devices
9. 2014. Scaling acceleration
10. 2014-15. ITRS 2.0
11. 2015. Post CMOS map of devices
12. 2016. IRDS
13. 2017-2021. 3D POWER Scaling
National Nanotechnology Initiative (NNI)

“My budget supports a major new National Nanotechnology Initiative, worth $500 million.... the ability to manipulate matter at the atomic and molecular level. Imagine the possibilities: materials with ten times the strength of steel and only a small fraction of the weight -- shrinking all the information housed at the Library of Congress into a device the size of a sugar cube -- detecting cancerous tumors when they are only a few cells in size. Some of our research goals may take 20 or more years to achieve, but that is precisely why there is an important role for the federal government.”

--President William J. Clinton
January 21, 2000
California Institute Of Technology
Nanoelectronics: An International Perspective

By Mihail C. Roco
NSF and NNI

Paolo Gargini
ITRS Chairman
Intel

Discoveries and innovations in nanotechnology are flourishing worldwide. Centers of excellence and research networks with long-term programs supporting nanoelectronics, nanomagnetics, and nanophotonics have been created in the United States (U.S.), Europe, Japan, and other parts of the world. The National Nanotechnology Initiative (NNI) has provided a long-term scientific focus, a partnership approach, and a means of environmentally responsible funding the field in the United States since 2000. The program also inspired and partially motivated nanotechnology R&D activities in about 60 other countries. Creative programs of similar investment scale are underway in Europe and Asia. In Europe, the FP7 program organizes all the EU research initiatives including in nanoelectronics into four categories: Cooperation, Ideas, People, and Capacities. For each objective, there is a specific program corresponding to the main areas of EU research policy. In addition, a European Commission program in Future Emerging Technologies (FET) addresses the disruptive approaches to nanoelectronics. In Japan, the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) promotes nanotechnology research and development and supports a network among researchers to provide cross-sectional, comprehensive support across research institutions and research fields. For example, MEXT also provides opportunities for outside researchers to use large and special facilities and equipment. Japan also has the New Energy and Industrial Technology Development Organization (NEDO), which contributes research and development activities in a variety of nanoelectronic programs, e.g., their Next-Generation Semiconductor Materials and Process Technology (MIRAI) Project.

In the first ten years, the research focus in the U.S. has been on uncovering nanoscale phenomena and on synthesizing nanostructured components to improve existing products. For illustration, researchers and manufacturers have placed functional
Nanoelectronics Research Initiative

Benchmarking of Beyond-CMOS Exploratory Devices for Logic Integrated Circuits

Dmitri E. Nikonov and Ian A. Young, 2015, IEEE J. on Exploratory Solid-State Computational Devices and Circuits

Switching energy vs. delay of a 32-bit adder
Moore’s Law: A Path Forward

Bill Holt
Executive Vice President, Intel
General Manager, Technology and Manufacturing Group

© 2016 IEEE
International Solid-State Circuits Conference
The Changing Measure of Improvement

Source: ITRS 2011

1.1: Moore's Law: A Path Forward
The Changing Measure of Improvement

Sources: ITRS 2011, Nikonov and Young, IEEE JxCDC, 1, 3-11 (2015); Manipatruni, Nikonov and Young, Arxiv cond-mat 1512.05428 (2015)
The Changing Measure of Improvement

Sources: ITRS 2011, Nikonov and Young, IEEE JxCDC, 1, 3-11 (2015);
Manipatruni, Nikonov and Young, Arxiv cond-mat 1512.05428 (2015)
Multiple Stories

1. Introduction
2. 1998. ITRS 1.0
   1. Equivalent Scaling
3. 2000. NNI.
5. 2005-6. MPU power limits
6. 2006. More than Moore
   1. Heterogeneous Integration
7. 2010. First Selection of post CMOS devices
9. 2014. Scaling acceleration
10. 2014-15. ITRS 2.0
11. 2015. Post CMOS map of devices
12. 2016. IRDS
13. 2017-2021. 3D POWER Scaling

April 2016  SPCC  P.Gargini  36
MM+MtM=Heterogeneous Integration

More than Moore: Diversification

Interacting with people and environment

Non-digital content System-in-package (SiP)

Combining SoC and SiP: Heterogeneous Integration

Information Processing

Baseline CMOS: CPU, Memory, Logic

More Moore: Miniaturization

Beyond CMOS

2006
On January 9, 2007 Steve Jobs announced the iPhone at the Macworld convention, receiving substantial media attention,[16] and that it would be released later that year. On June 29, 2007 the first iPhone was released.
A WiFi-only model of the tablet was released in April 2010, and a WiFi+3G model was introduced about a month later.
Smartphone Marketshare Trends

- Smartphone Unit Sales (M)
- Smartphone Share of Total Cellphone Shipments

- 485M; +67%
- 712M; +47%
- 975M; +37%
- 1,260M; +29%

Source: Nokia, IC Insights
Multiple Stories

1. Introduction
2. 1998. ITRS 1.0
   1. Equivalent Scaling
3. 2000. NNI.
5. 2005-6. MPU power limits
6. 2006. More than Moore
   1. Heterogeneous Integration
7. 2010. First Selection of post CMOS devices
9. **2014. Scaling acceleration**
10. 2014-15. ITRS 2.0
11. 2015. Post CMOS map of devices
12. 2016. IRDS
13. 2017-2021. 3D POWER Scaling

April 2016        SPCC        P.Gargini 41
2D → 3D
22 nm Tri-Gate Transistor

Mark Bohr, Kaizad Mistry, May 2011

SPCC
Question

How many more technology generations can Equivalent Scaling be extended for?
Multigate FET Offers a Simple Way for Scaling and Improving Performance
Transistor Fin Improvement

Taller and Thinner Fins for Increased Drive Current and Performance

Mark Bohr, August 11, 2014

April 2016

SPCC

P. Gargini 46
Transistor Fin Improvement

Reduced Number of Fins for Improved Density and Lower Capacitance

Mark Bohr, August 11, 2014
Fin FET
Moore’s Law
Acceleration
Interconnects

22 nm Process

14 nm Process

80 nm minimum pitch

52 nm (0.65x) minimum pitch

52 nm Interconnect Pitch Provides Better-than-normal Interconnect Scaling
Offsetting Wafer Cost with Density

Source: Intel
© 2016 IEEE
International Solid-State Circuits Conference
Technology Node Scaling

Today’s Challenge

Technology Node (nm)

2013 ITRS
SPCC
Micron/ Intel 20-nm 64G MLC NAND Flash
NAND Relative Wafer Cost
Electron number of FG

- GCR=0.7, Tono=15 nm, Tox=9nm
- GCR=0.65, Tono=13 nm, Tox=8nm

Ne: Electron Number

Technology Node: F (nm)
Samsung's 32-Layer 3D V-NAND Memory Chip

Source: PC Perspective, Samsung
Vertical Logic Architecture
3D Moore’s Law Acceleration
3D Architecture
Multiple Stories

1. Introduction
2. 1998. ITRS 1.0
   1. Equivalent Scaling
3. 2000. NNI.
5. 2005-6. MPU power limits
6. 2006. More than Moore
   1. Heterogeneous Integration
7. 2010. First Selection of post CMOS devices
9. 2014. Scaling acceleration
10. 2014-15. ITRS 2.0
11. 2015. Post CMOS map of devices
12. 2016. IRDS
13. 2017-2021. 3D POWER Scaling

April 2016  SPCC  P.Gargini
Phase 3
Third Age of Scaling
(3D Power Scaling)
IC Industry at a Glance (2021->203X)

Driver  ➔ Cost/transistor & power reduction

How  ➔ 2x Density/2 years (Moore)

Method  ➔ 3D Power Scaling (ITRS2.0)
The Different Ages of Scaling
(Different methods for different times)

① Geometrical Scaling (1975-2003)
- Reduction of horizontal and vertical physical dimensions in conjunction with improved performance of planar transistors

② Equivalent Scaling (2003~2021)
- Reduction of only horizontal dimensions in conjunction with introduction of new materials and new physical effects. New vertical structures replace the planar transistor

③ 3D Power Scaling (2021~203X)
- Transition to complete vertical device structures. Heterogeneous integration in conjunction with reduced power consumption become the technology drivers
Beyond 2020

- Customized Functionality
- Outside System Connectivity
- System Integration
- Heterogeneous Integration
- More than Moore
- More Moore
- Beyond Moore

ITRS 2012

April 2016
21th Anniversary of TRS

http://www.itrs2.net

Europe
- 1998 ITRS Update
- 2003 ITRS
- 2008 ITRS Update
- 2013 ITRS

Japan
- 1999 ITRS
- 2004 ITRS Update
- 2009 ITRS
- 2010 ITRS Update

Korea
- 2000 ITRS Update
- 2005 ITRS
- 2010 ITRS Update

Taiwan
- 2001 ITRS
- 2006 ITRS Update
- 2011 ITRS

USA
- 2002 ITRS Update
- 2007 ITRS
- 2012 ITRS Update
Beyond 2020

- Customized Functionality
- Outside System Connectivity
- System Integration
- Heterogeneous Integration
- More than Moore
- More Moore
- Beyond Moore
- Manufacturing

Beyond 2020

- System Integration
- Outside System Connectivity
- Heterogeneous Integration
- More than Moore
- Beyond Moore
- More Moore
- Manufacturing

From ITRS to ITRS 2.0

April 2014

SPCC
ITRS 2.0

http://www.itrs2.net

April 2016

SPCC

P.Gargini 69
Q: How do we get back to exponential performance scaling?

*IEEE Rebooting Computing Initiative*
<table>
<thead>
<tr>
<th>Categories</th>
<th>Year</th>
<th>2015</th>
<th>2017</th>
<th>2019</th>
<th>2021</th>
<th>2023</th>
<th>2025</th>
<th>2027</th>
<th>2029</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Cores (K)</td>
<td></td>
<td>360</td>
<td>1044</td>
<td>3008</td>
<td>4935</td>
<td>5825</td>
<td>7578</td>
<td>8967</td>
<td>10602</td>
</tr>
<tr>
<td>Total Memory Storage (PB)</td>
<td></td>
<td>300</td>
<td>1559</td>
<td>4676</td>
<td>14029</td>
<td>42088</td>
<td>126264</td>
<td>378792</td>
<td>1136377</td>
</tr>
<tr>
<td>Area of One Server Building (MSF)</td>
<td>0.5</td>
<td>0.9</td>
<td>1.6</td>
<td>2.2</td>
<td>2.2</td>
<td>2.42</td>
<td>2.42</td>
<td>2.42</td>
<td></td>
</tr>
<tr>
<td>Global Total Power Consumed by Datacenter (MkW)</td>
<td></td>
<td>779.8</td>
<td>839</td>
<td>1004.7</td>
<td>1137.2</td>
<td>1380.7</td>
<td>1635.6</td>
<td>2044.3</td>
<td></td>
</tr>
<tr>
<td>Total Switching Capability of Datacenter BW (Tb/s)</td>
<td></td>
<td>1000</td>
<td>2512</td>
<td>6309</td>
<td>10500</td>
<td>15849</td>
<td>51199</td>
<td>39811</td>
<td>63096</td>
</tr>
<tr>
<td>Data Center Efficiency (GFLOPS/W)</td>
<td></td>
<td>2.4</td>
<td>4.9</td>
<td>10</td>
<td>17</td>
<td>24</td>
<td>33.9</td>
<td>47.9</td>
<td>67.8</td>
</tr>
<tr>
<td>Server Units/Rack</td>
<td></td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Number of Cores/socket</td>
<td></td>
<td>18</td>
<td>29</td>
<td>44</td>
<td>59</td>
<td>74</td>
<td>93</td>
<td>117</td>
<td>147</td>
</tr>
<tr>
<td>Power /Single Server Unit (W)</td>
<td>700</td>
<td>700</td>
<td>700</td>
<td>700</td>
<td>700</td>
<td>700</td>
<td>700</td>
<td>700</td>
<td>700</td>
</tr>
<tr>
<td>Main Memory / Single Server Unit (GB)</td>
<td></td>
<td>32</td>
<td>45</td>
<td>64</td>
<td>76</td>
<td>91</td>
<td>108</td>
<td>129</td>
<td>154</td>
</tr>
<tr>
<td>Power Consumed by Cores/single socket (W)</td>
<td></td>
<td>165</td>
<td>194</td>
<td>149</td>
<td>145</td>
<td>141</td>
<td>137</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>Network Bandwidth/Unit (Gb/s)</td>
<td>40</td>
<td>40</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>Network Bandwidth/Rack (Gb/s)</td>
<td>1600</td>
<td>1600</td>
<td>1800</td>
<td>4000</td>
<td>4000</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
</tr>
<tr>
<td>Rack Switch Capacity BW (Gb/s)</td>
<td>1200</td>
<td>1200</td>
<td>3000</td>
<td>3000</td>
<td>3000</td>
<td>12000</td>
<td>12000</td>
<td>12000</td>
<td>12000</td>
</tr>
<tr>
<td>Power Efficiency (Grid delivery/Data Center Use)</td>
<td></td>
<td>0.55</td>
<td>0.57</td>
<td>0.59</td>
<td>0.61</td>
<td>0.63</td>
<td>0.65</td>
<td>0.67</td>
<td>0.69</td>
</tr>
<tr>
<td>Power Consumed by Networking and Switching (MkWh)</td>
<td></td>
<td>21.91</td>
<td>51.05</td>
<td>138.26</td>
<td>87.66</td>
<td>138.93</td>
<td>220.19</td>
<td>348.97</td>
<td>553.08</td>
</tr>
<tr>
<td>Power Consumed by Storage (MkWh)</td>
<td></td>
<td>0.0657</td>
<td>0.259</td>
<td>0.449</td>
<td>0.778</td>
<td>1.347</td>
<td>2.334</td>
<td>4.043</td>
<td>7.004</td>
</tr>
<tr>
<td>Power Consumed for facility cooling (MkWh)</td>
<td></td>
<td>38.99</td>
<td>55.02</td>
<td>86.13</td>
<td>237.31</td>
<td>264.26</td>
<td>307.03</td>
<td>374.89</td>
<td>482.56</td>
</tr>
</tbody>
</table>
## Mobile

<table>
<thead>
<tr>
<th>Input Metrics</th>
<th>2015</th>
<th>2017</th>
<th>2019</th>
<th>2021</th>
<th>2023</th>
<th>2025</th>
<th>2027</th>
<th>2029</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of AP cores</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Number of GPU cores</td>
<td>14</td>
<td>22</td>
<td>34</td>
<td>54</td>
<td>86</td>
<td>138</td>
<td>220</td>
<td>352</td>
</tr>
<tr>
<td>Max frequency of any Component in System (GHz)</td>
<td>2.7</td>
<td>2.9</td>
<td>3.2</td>
<td>3.4</td>
<td>3.7</td>
<td>4</td>
<td>4.3</td>
<td>4.7</td>
</tr>
<tr>
<td>Number of Mega pixels in Display</td>
<td>2.1</td>
<td>2.1</td>
<td>3.7</td>
<td>8.8</td>
<td>8.8</td>
<td>33.2</td>
<td>33.2</td>
<td>33.2</td>
</tr>
<tr>
<td>Band Width between AP and Main memory (Gb/s)</td>
<td>45.3</td>
<td>45.3</td>
<td>89.6</td>
<td>89.6</td>
<td>148.2</td>
<td>148.2</td>
<td>148.2</td>
<td>148.2</td>
</tr>
<tr>
<td>Number of Sensors</td>
<td>14</td>
<td>16</td>
<td>20</td>
<td>21</td>
<td>21</td>
<td>22</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Number of Antennas</td>
<td>11</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Number of ICs</td>
<td>7-10</td>
<td>7-10</td>
<td>7-10</td>
<td>7-10</td>
<td>7-10</td>
<td>7-10</td>
<td>7-10</td>
<td>7-10</td>
</tr>
<tr>
<td>Cellular data rate growing ~1.3x/year (MB/s)</td>
<td>17.6</td>
<td>12</td>
<td>21.63</td>
<td>40.75</td>
<td>40.75</td>
<td>40.75</td>
<td>40.75</td>
<td>40.75</td>
</tr>
<tr>
<td>Wi-Fi data rate evolving with standards (Mb/s)</td>
<td>867</td>
<td>8</td>
<td>867</td>
<td>7000</td>
<td>7000</td>
<td>28000</td>
<td>28000</td>
<td>28000</td>
</tr>
<tr>
<td>PCB area of main Components (cm²)</td>
<td>62</td>
<td>69</td>
<td>76</td>
<td>84</td>
<td>93</td>
<td>103</td>
<td>103</td>
<td>103</td>
</tr>
<tr>
<td>Board power averaged at ~7%/year (mW)</td>
<td>4274</td>
<td>4706</td>
<td>5183</td>
<td>5708</td>
<td>6287</td>
<td>6926</td>
<td>7630</td>
<td>8406</td>
</tr>
</tbody>
</table>
# NAND Flash

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

## Year of Production

<table>
<thead>
<tr>
<th>Year</th>
<th>2015</th>
<th>2017</th>
<th>2019</th>
<th>2021</th>
<th>2024</th>
<th>2027</th>
<th>2030</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Half Pitch (Calculated Half pitch) (nm)</strong></td>
<td>24</td>
<td>20</td>
<td>17</td>
<td>14</td>
<td>8.4</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td><strong>DRAM cell size (μm²)</strong></td>
<td>0.00346</td>
<td>0.00240</td>
<td>0.00116</td>
<td>0.0005</td>
<td>0.00048</td>
<td>0.00028</td>
<td>0.00024</td>
</tr>
<tr>
<td><strong>DRAM cell FET structure</strong></td>
<td>RCAT+Fin</td>
<td>RCAT+Fin</td>
<td>VCT</td>
<td>VCT</td>
<td>VCT</td>
<td>VCT</td>
<td>VCT</td>
</tr>
<tr>
<td><strong>Cell Size Factor: a</strong></td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td><strong>Array Area Efficiency</strong></td>
<td>0.55</td>
<td>0.55</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td><strong>V_{int} (support FET voltage) [V]</strong></td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>Support min. $V_{in}$ (25°C, $G_{m,max}$, $V_d=55mV$)</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
</tr>
<tr>
<td><strong>Minimum DRAM retention time (ms)</strong></td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td><strong>DRAM soft error rate (fits)</strong></td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td><strong>Gb/1chip target</strong></td>
<td>8G</td>
<td>8G</td>
<td>16G</td>
<td>16G</td>
<td>32G</td>
<td>32G</td>
<td>32G</td>
</tr>
<tr>
<td>YEAR OF PRODUCTION</td>
<td>2015</td>
<td>2017</td>
<td>2019</td>
<td>2021</td>
<td>2024</td>
<td>2027</td>
<td>2030</td>
</tr>
<tr>
<td>--------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Logic device technology naming</td>
<td>P70M56</td>
<td>P48M36</td>
<td>P42M24</td>
<td>P32M20</td>
<td>P24M12G1</td>
<td>P24M12G2</td>
<td>P24M12G3</td>
</tr>
<tr>
<td>Logic industry &quot;Node Range&quot; Labeling (nm)</td>
<td>&quot;16/14&quot;</td>
<td>&quot;11/10&quot;</td>
<td>&quot;8/7&quot;</td>
<td>&quot;6/5&quot;</td>
<td>&quot;4/3&quot;</td>
<td>&quot;3/2.5&quot;</td>
<td>&quot;2/1.5&quot;</td>
</tr>
<tr>
<td>Logic device structure options</td>
<td>finFET</td>
<td>finFET</td>
<td>finFET</td>
<td>finFET</td>
<td>VGAA, M3D</td>
<td>VGAA, M3D</td>
<td>VGAA, M3D</td>
</tr>
<tr>
<td></td>
<td>FDSOI</td>
<td>FDSOI</td>
<td>LGAA</td>
<td>VGAA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**LOGIC DEVICE GROUND RULES**

<table>
<thead>
<tr>
<th></th>
<th>2015</th>
<th>2017</th>
<th>2019</th>
<th>2021</th>
<th>2024</th>
<th>2027</th>
<th>2030</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPU/SoC Metalx ½ Pitch (nm) [1,2]</td>
<td>28.0</td>
<td>18.0</td>
<td>12.0</td>
<td>10.0</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td>MPU/SoC Metal0/1 ½ Pitch (nm)</td>
<td>28.0</td>
<td>18.0</td>
<td>12.0</td>
<td>10.0</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Contacted poly half pitch (nm)</td>
<td>35.0</td>
<td>24.0</td>
<td>20.0</td>
<td>16.0</td>
<td>12.0</td>
<td>12.0</td>
<td>12.0</td>
</tr>
<tr>
<td>L₀: Physical Gate Length for HP Logic (nm) [3]</td>
<td>24</td>
<td>18</td>
<td>14</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>L₀: Physical Gate Length for LP Logic (nm)</td>
<td>26</td>
<td>20</td>
<td>16</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>FinFET Fin Half-pitch (new) =0.75 or 1.0 M0/M1 (nm)</td>
<td>21.0</td>
<td>18.0</td>
<td>12.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FinFET Fin Width (nm)</td>
<td>8.0</td>
<td>8.0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FinFET Fin Height (nm)</td>
<td>42.0</td>
<td>42.0</td>
<td>42.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Footprint drive efficiency - finFET</td>
<td>2.1</td>
<td>2.1</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral GAA Lateral Half-pitch (nm)</td>
<td></td>
<td>12.0</td>
<td>10.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral GAA Vertical Half-pitch (nm)</td>
<td></td>
<td>12.0</td>
<td>9.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral GAA Diameter (nm)</td>
<td></td>
<td>6.0</td>
<td>6.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Footprint drive efficiency - lateral GAA, 3x NWs stacked</td>
<td></td>
<td>2.4</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical GAA Lateral Half-pitch (nm)</td>
<td></td>
<td>10.0</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Vertical GAA Diameter (nm)</td>
<td></td>
<td>6.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>Footprint drive efficiency - vertical GAA, 3x NWs stacked</td>
<td></td>
<td>2.8</td>
<td>3.9</td>
<td>3.9</td>
<td>3.9</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td>Deface effective width - [nm]</td>
<td></td>
<td>92.0</td>
<td>90.0</td>
<td>56.5</td>
<td>56.5</td>
<td>56.5</td>
<td>56.5</td>
</tr>
<tr>
<td>Device lateral half pitch (nm)</td>
<td></td>
<td>21.0</td>
<td>18.0</td>
<td>12.0</td>
<td>10.0</td>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Device width or diameter (nm)</td>
<td></td>
<td>8.0</td>
<td>6.0</td>
<td>6.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>
Conclusions

- **“Geometrical Scaling”** led the IC Industry for 3 decades
  
  **ITRS 1.0**
  
  Cooperative and distributed research and manufacturing methods highlighted by **ITRS** emerged as cost effective means of reducing costs since the mid-90s
  
  FCRP, NRI, Sematech, IMEC and Government organizations actively cooperated in **advanced research**
  
- **“Equivalent Scaling”** saved the Semiconductor Industry since the beginning of the previous decade
  
  Preliminary evaluation of post-CMOS candidates published in 2010
  
  **ITRS 2.0**
  
  - **“3D Power Scaling”** is the next phase of (accelerated) scaling
  
  Post CMOS devices and emerging architectures are being jointly evaluated—>ITRS/IEEE RC—>IRDS