Transistor Revolution

- Transistor – Bardeen (Bell Labs) in 1947
- Bipolar transistor – Schockley in 1949
- First bipolar digital logic gate – Harris in 1956
- First monolithic IC – Jack Kilby in 1959
- First commercial IC logic gates – Fairchild 1960
- TTL – 1962 into the 1990’s
- ECL – 1974 into the 1980’s
MOSFET Technology

- MOSFET transistor - Lilienfeld (Canada) in 1925 and Heil (England) in 1935
- CMOS – 1960’s, but plagued with manufacturing problems
- PMOS in 1960’s (calculators)
- NMOS in 1970’s (4004, 8080) – for speed
- CMOS in 1980’s – preferred MOSFET technology because of power benefits
- BiCMOS, Gallium-Arsenide, Silicon-Germanium
- SOI, Copper-Low K, …
Introduction

• Why is designing digital ICs different today than it was before?
• Will it change in future?
The First Computer

The Babbage Difference Engine
(1832)
25,000 parts
cost: £17,470
ENIAC - The first electronic computer (1946)
The Transistor Revolution

First transistor
Bell Labs, 1948
The First Integrated Circuits

Bipolar logic
1960’s

ECL 3-input Gate
Motorola 1966
Intel 4004 Micro-Processor

1971
1000 transistors
1 MHz operation
Intel Pentium (IV) microprocessor

2001
42 million transistors
2 GHz operation

0.13\(\mu\) process
55 million transistors
2.4GHz clock
145mm\(^2\)
Moore’s Law

• In 1965, Gordon Moore predicted that the number of transistors that can be integrated on a die would double every 18 to 14 months (i.e., grow exponentially with time).

• Amazingly visionary – million transistor/chip barrier was crossed in the 1980’s.
  – 2300 transistors, 1 MHz clock (Intel 4004) - 1971
  – 16 Million transistors (Ultra Sparc III)
  – 42 Million, 2 GHz clock (Intel P4) - 2001
  – 140 Million transistor (HP PA-8500)
# State-of-the Art: Lead Microprocessors

<table>
<thead>
<tr>
<th>Processor</th>
<th>Alpha 21264B</th>
<th>AMD Athlon</th>
<th>HP PA-8600</th>
<th>IBM Power3-II</th>
<th>Intel Pentium III</th>
<th>Intel Pentium 4</th>
<th>MIPS R12000</th>
<th>Sun Ultra-II</th>
<th>Sun Ultra-III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock Rate</td>
<td>833MHz</td>
<td>1.33GHz</td>
<td>552MHz</td>
<td>450MHz</td>
<td>1.0GHz</td>
<td>1.7GHz</td>
<td>400MHz</td>
<td>480MHz</td>
<td>900MHz</td>
</tr>
<tr>
<td>Cache (I/D/L2)</td>
<td>64K/64K</td>
<td>64K/64K/256K</td>
<td>512K/1M</td>
<td>32K/64K</td>
<td>16K/16K/256K</td>
<td>12K/8K/256K</td>
<td>32K/32K</td>
<td>16K/16K</td>
<td>32K/64K</td>
</tr>
<tr>
<td>Issue Rate</td>
<td>4 issue</td>
<td>3 x86 instr</td>
<td>4 issue</td>
<td>4 issue</td>
<td>3 x86 instr</td>
<td>3 ROPs</td>
<td>4 issue</td>
<td>4 issue</td>
<td>4 issue</td>
</tr>
<tr>
<td>Pipeline Stages</td>
<td>7/9 stages</td>
<td>9/11 stages</td>
<td>7/9 stages</td>
<td>7/8 stages</td>
<td>12/14 stages</td>
<td>22/24 stages</td>
<td>6 stages</td>
<td>6/9 stages</td>
<td>14/15 stages</td>
</tr>
<tr>
<td>Out of Order</td>
<td>80 instr</td>
<td>72ROPs</td>
<td>56 instr</td>
<td>32 instr</td>
<td>40 ROPs</td>
<td>126 ROPs</td>
<td>48 instr</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Rename regs</td>
<td>48/41</td>
<td>36/36</td>
<td>56 total</td>
<td>16 int/24 fp</td>
<td>40 total</td>
<td>128 total</td>
<td>32/32</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>BHT Entries</td>
<td>4K x 9 bits</td>
<td>4K x 2 bits</td>
<td>2K x 2 bits</td>
<td>2K x 2 bits</td>
<td>≥ 512</td>
<td>4K x 2 bits</td>
<td>2K x 2 bits</td>
<td>512 x 2 bits</td>
<td>16K x 2 bits</td>
</tr>
<tr>
<td>TLB Entries</td>
<td>128/128</td>
<td>280/288</td>
<td>120 unified</td>
<td>128/128</td>
<td>321/64D</td>
<td>128I/64D</td>
<td>64 unified</td>
<td>64I/64D</td>
<td>128I/512D</td>
</tr>
<tr>
<td>Memory B/W</td>
<td>2.66GB/s</td>
<td>2.1GB/s</td>
<td>1.54GB/s</td>
<td>1.6GB/s</td>
<td>1.06GB/s</td>
<td>3.2GB/s</td>
<td>539 MB/s</td>
<td>1.9GB/s</td>
<td>4.8GB/s</td>
</tr>
<tr>
<td>Package</td>
<td>CPGA-588</td>
<td>PGA-462</td>
<td>LGA-544</td>
<td>SCC-1088</td>
<td>PGA-370</td>
<td>PGA-423</td>
<td>CPGA-527</td>
<td>CLGA-787</td>
<td>1368 FC-LGA</td>
</tr>
<tr>
<td>IC Process</td>
<td>0.18μ 6M</td>
<td>0.18μ 6M</td>
<td>0.25μ 2M</td>
<td>0.22μ 6M</td>
<td>0.18μ 6M</td>
<td>0.18μ 6M</td>
<td>0.25μ 4M</td>
<td>0.29μ 6M</td>
<td>0.18μ 7M</td>
</tr>
<tr>
<td>Die Size</td>
<td>115mm²</td>
<td>117mm²</td>
<td>477mm²</td>
<td>163mm²</td>
<td>106mm²</td>
<td>217mm²</td>
<td>204mm²</td>
<td>126mm²</td>
<td>210mm²</td>
</tr>
<tr>
<td>Transistors</td>
<td>15.4 million</td>
<td>37 million</td>
<td>130 million</td>
<td>23 million</td>
<td>24 million</td>
<td>42 million</td>
<td>7.2 million</td>
<td>3.8 million</td>
<td>29 million</td>
</tr>
<tr>
<td>Est mfg cost</td>
<td>$160</td>
<td>$62</td>
<td>$330</td>
<td>$110</td>
<td>$39</td>
<td>$100</td>
<td>$125</td>
<td>$70</td>
<td>$145</td>
</tr>
<tr>
<td>Power (max)</td>
<td>75W*</td>
<td>76W</td>
<td>60W*</td>
<td>36W*</td>
<td>30W</td>
<td>64W(TDP)</td>
<td>25W*</td>
<td>20W*</td>
<td>65W</td>
</tr>
<tr>
<td>Availability</td>
<td>1Q01</td>
<td>1Q01</td>
<td>3Q00</td>
<td>4Q00</td>
<td>2Q00</td>
<td>2Q01</td>
<td>2Q00</td>
<td>3Q0</td>
<td>4Q00</td>
</tr>
</tbody>
</table>
Moore’s Law in Microprocessors

Transistors on lead microprocessors double every 2 years

2X growth in 1.96 years!

Courtesy, Intel
Evolution in DRAM Chip Capacity

4X growth every 3 years!

- Human memory
- Human DNA

- Encyclopedia
- 2 hrs CD audio
- 30 sec HDTV

- 4X growth every 3 years!
Die Size Growth

Die size grows by 14% to satisfy Moore’s Law

~7% growth per year
~2X growth in 10 years

Courtesy, Intel
Clock Frequency

Lead Microprocessors frequency doubles every 2 years

Doubles every 2 years

Courtesy, Intel
Lead Microprocessors power continues to increase

Courtesy, Intel
Power will be a major problem

Power delivery and dissipation will be prohibitive

Courtesy, Intel
Power Density

Power density too high to keep junctions at low temp

Courtesy, Intel
Not Only Microprocessors

Cell Phone

Digital Cellular Market (Phones Shipped)

<table>
<thead>
<tr>
<th>Year</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>48M</td>
</tr>
<tr>
<td>1997</td>
<td>86M</td>
</tr>
<tr>
<td>1998</td>
<td>162M</td>
</tr>
<tr>
<td>1999</td>
<td>260M</td>
</tr>
<tr>
<td>2000</td>
<td>435M</td>
</tr>
</tbody>
</table>

(data from Texas Instruments)
As a result of the continuously increasing integration density and decreasing unit costs, the semiconductor industry has been one of the fastest growing sectors in the worldwide economy.
Industry Trends

- Video-on-demand
- Speech processing/ recognition
- Wireless/cellular data communication
- Data communication
- Multi-media applications
- Consumer electronics
- Portable computers
- Mainframe computers
- Personal computers
- Network computers

Large
Centralized
Expensive

Small / Portable
Distributed
Inexpensive
Industry Trends

High performance
Low power dissipation
Wireless capability
etc...

More portable, wearable, and more powerful devices for ubiquitous and pervasive computing...
Design Productivity Trends

Complexity outpaces design productivity

Courtesy, ITRS Roadmap
## Technology Directions: SIA Roadmap

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature size (nm)</td>
<td>180</td>
<td>130</td>
<td>100</td>
<td>70</td>
<td>50</td>
<td>35</td>
</tr>
<tr>
<td>Mtrans/cm²</td>
<td>7</td>
<td>14-26</td>
<td>47</td>
<td>115</td>
<td>284</td>
<td>701</td>
</tr>
<tr>
<td>Chip size (mm²)</td>
<td>170</td>
<td>170-214</td>
<td>235</td>
<td>269</td>
<td>308</td>
<td>354</td>
</tr>
<tr>
<td>Signal pins/chip</td>
<td>768</td>
<td>1024</td>
<td>1024</td>
<td>1280</td>
<td>1408</td>
<td>1472</td>
</tr>
<tr>
<td>Clock rate (MHz)</td>
<td>600</td>
<td>800</td>
<td>1100</td>
<td>1400</td>
<td>1800</td>
<td>2200</td>
</tr>
<tr>
<td>Wiring levels</td>
<td>6-7</td>
<td>7-8</td>
<td>8-9</td>
<td>9</td>
<td>9-10</td>
<td>10</td>
</tr>
<tr>
<td>Power supply (V)</td>
<td>1.8</td>
<td>1.5</td>
<td>1.2</td>
<td>0.9</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>High-perf power (W)</td>
<td>90</td>
<td>130</td>
<td>160</td>
<td>170</td>
<td>174</td>
<td>183</td>
</tr>
<tr>
<td>Battery power (W)</td>
<td>1.4</td>
<td>2.0</td>
<td>2.4</td>
<td>2.0</td>
<td>2.2</td>
<td>2.4</td>
</tr>
</tbody>
</table>

For Cost-Performance MPU (L1 on-chip SRAM cache; 32KB/1999 doubling every two years)

http://www.itrs.net/ntrs/publntrs.nsf
## Increasing Function Density

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TECHNOLOGY</td>
<td>130 nm</td>
<td>100 nm</td>
<td>70 nm</td>
<td>50 nm</td>
<td>35 nm</td>
</tr>
<tr>
<td>CHIP SIZE</td>
<td>400 mm²</td>
<td>600 mm²</td>
<td>750 mm²</td>
<td>800 mm²</td>
<td>900 mm²</td>
</tr>
<tr>
<td>NUMBER OF TRANSISTORS (LOGIC)</td>
<td>400 M</td>
<td>1 Billion</td>
<td>3 Billion</td>
<td>6 Billion</td>
<td>16 Billion</td>
</tr>
<tr>
<td>DRAM CAPACITY</td>
<td>2 Gbits</td>
<td>10 Gbits</td>
<td>25 Gbits</td>
<td>70 Gbits</td>
<td>200 Gbits</td>
</tr>
<tr>
<td>MAXIMUM CLOCK FREQUENCY</td>
<td>1.6 GHz</td>
<td>2.0 GHz</td>
<td>2.5 GHz</td>
<td>3.0 GHz</td>
<td>3.5 GHz</td>
</tr>
<tr>
<td>MINIMUM SUPPLY VOLTAGE</td>
<td>1.5 V</td>
<td>1.2 V</td>
<td>0.9 V</td>
<td>0.6 V</td>
<td>0.6 V</td>
</tr>
<tr>
<td>MAXIMUM POWER DISSIPATION</td>
<td>130 W</td>
<td>160 W</td>
<td>170 W</td>
<td>175 W</td>
<td>180 W</td>
</tr>
<tr>
<td>MAXIMUM NUMBER OF I/O PINS</td>
<td>2500</td>
<td>4000</td>
<td>4500</td>
<td>5500</td>
<td>6000</td>
</tr>
</tbody>
</table>
Why Scaling?

- Technology shrinks by ~0.7 per generation
- With every generation can integrate 2x more functions on a chip; chip cost does not increase significantly
- Cost of a function decreases by 2x
- But …
  - How to design chips with more and more functions?
  - Design engineering population does not double every two years…
- Hence, a need for more efficient design methods
  - Exploit different levels of abstraction
Notice: There is a need for structured design methodologies to handle the high level of complexity!
Major Design Challenges

- **Microscopic issues**
  - ultra-high speeds
  - power dissipation and supply rail drop
  - growing importance of interconnect
  - noise, crosstalk
  - reliability, manufacturability
  - clock distribution

- **Macroscopic issues**
  - time-to-market
  - design complexity (millions of gates)
  - high levels of abstractions
  - reuse and IP, portability
  - systems on a chip (SoC)
  - tool interoperability

<table>
<thead>
<tr>
<th>Year</th>
<th>Tech.</th>
<th>Complexity</th>
<th>Frequency</th>
<th>3 Yr. Design Staff Size</th>
<th>Staff Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>0.35</td>
<td>13 M Tr.</td>
<td>400 MHz</td>
<td>210</td>
<td>$90 M</td>
</tr>
<tr>
<td>1998</td>
<td>0.25</td>
<td>20 M Tr.</td>
<td>500 MHz</td>
<td>270</td>
<td>$120 M</td>
</tr>
<tr>
<td>1999</td>
<td>0.18</td>
<td>32 M Tr.</td>
<td>600 MHz</td>
<td>360</td>
<td>$160 M</td>
</tr>
<tr>
<td>2002</td>
<td>0.13</td>
<td>130 M Tr.</td>
<td>800 MHz</td>
<td>800</td>
<td>$360 M</td>
</tr>
</tbody>
</table>
TASARIMDA YARATICILIK
TASARIMDA YARATICILIK