Systems-on-a-Chip (SoCs)

An introduction to SoCs using the Cell® processor as a case-study

*Presented by Michael Fuller*
What is a system-on-a-chip?

“An SoC is an ASIC that integrates, on a single silicon die, processors, memories, logic, analog, and I/O functions previously implemented as multiple discrete chips.”

RF Micro Devices Bluetooth SoC

Requires only 8 external components (6 caps, 1 inductor, 1 BPF)
Cypress PSoC™ Mixed-Signal Array

- Single chip containing programmable digital and analog function blocks
- Programmable interconnect
- Low Cost: $0.05 - $10
- 24 MHz, 4 MIPS, 8-bit μCPU
- Flash program memory
- SRAM data memory
- Configurable I/O

Cypress PSoC™ Mixed-Signal Array

• Digital System:
  – PWMs (8-32 bit)
  – Counters (8-32 bit)
  – Timers (8-32 bit)
  – UART w/ selectable parity
  – SPI master and slave (4 each)
  – Pseudo Random Seq. Generator
  – More...

Cypress PSoC™ Mixed-Signal Array

- Analog System:
  - ADCs (6- to 14-bit resolution)
  - Filters (2, 4, 6, 8 pole BP, LP, notch)
  - Amplifiers (selectable gain to 48x)
  - Instr. Amplifiers (sel. gain to 93x)
  - Comparators (16 sel. Thresholds)
  - DACs (6- to 9-bit resolution)
  - High current output drivers (40 mA)
  - 1.3 V reference
  - DTMF dialer
  - Modulators
  - Correlators
  - Peak Detectors
  - Many other topologies possible...

Issues for SoC Designs

- Semiconductor technology
  - Transistor type (CMOS, SOI, etc.)
  - Min. feature size (0.5 micron or 65 nm?)
  - Gate-oxide thickness
  - Leakage current
  - $V_t$ (high, low, multiple)
  - Analog transistors (long-channels, thicker oxide)?
  - Analog process elements (polycaps, precision resistors)
  - Single or multiple well
  - Metal stack (# layers, thin/thick, ↑layers ➔ ↑cost)
  - Memory process features (e.g., trench caps)
Issues for SoC Designs

• Packaging
  – Flip-chip?
  – Wirebond?
  – Number of I/O pins?
  – Number and distribution of power pins?
  – Thermal management
Issues for SoC Designs

• Power
  – Digital Power
  – Analog Power
    • Must be isolated from noisy digital circuitry and supplies
  – Multiple supplies (multiple VDDs, bulk biases, etc.)
  – Routing issues
  – Power management!
    • Sleep mode, deep-sleep mode, standby mode, idle mode, low-power mode, high-performance mode, etc., etc.
    • Throttling, voltage scaling, frequency scaling
Issues for SoC Designs

• Memory
  – DRAM
  – SRAM
  – Flash
  – ROM, PROM, EEPROM
  – L1/L2 cache
  – data memory, instruction memory
  – Memory interface (bus architecture, control, DMA)
  – Analog memory
Issues for SoC Designs

• Clocking
  – On-chip or off-chip?
  – Multiple clocks?
  – Asynchronous function blocks?
  – Routing and distribution
  – Skew management
  – PLLs
  – Frequency dividers/synthesizers
Issues for SoC Designs

- Testing and Verification
  - Pre-Tapeout
    - Simulate huge design (100s of millions of transistors)
    - Simulate both analog and digital components (some interaction)
    - Noise sims, power sims, thermal sims
  - Post-tapeout
    - Test/Debug modes
    - Test/Debug hardware (extra buffers, config. registers)
    - JTAG
    - Calibration
      - Laser trimming
      - Configuration registers
      - Redundant memory blocks
The CELL processor
Joint effort by IBM Corp., Sony, and Toshiba

- **CELL** = Cell Broadband Engine Architecture
- Intended for concurrent real-time multimedia and conventional computing applications
- $400 million budget
- 400 engineers
- 90-nm SOI process
- 10 simultaneous threads on 9 processors
- 256 GigaFLOPS @ 4 GHz
- 235 million transistors
- 221 mm² chip area
- 60-80 watts @ 4 GHz

Implemented on a 90-nm silicon-on-insulator process with eight levels of copper interconnect, the CELL processor can deliver a single-precision compute throughput of over 256 GFLOPS when running at 4 GHz.

[http://www.elecdesign.com/Files/29/9748/Figure_02.jpg](http://www.elecdesign.com/Files/29/9748/Figure_02.jpg)
First intended application:

Sony PlayStation 3

Release Date:
➢ early November 2006
Toshiba demonstrates the Cell processor simultaneously decoding 48 SDTV format MPEG-2 streams in 2005.
Main Components of the CELL

- Power Processing Element (PPE)
- Synergistic Processing Elements (SPEs)
- Element Interconnect Bus (EIB)
- Memory Interface Controller (MIC)
- Bus Interface Controller (BIC)
System Concepts

- **SPEs are the orchestra**
  - Each SPE has greater bandwidth & performance than the PPE
  - Use SPE for frequently repeated tasks with good data locality
  - Use data flow paradigms to orchestrate SPE task execution

- **PPE is the conductor**
  - The PPE controls and sets up the global synchronization
  - The OS runs on PPE allocating resources, controlling devices and providing system services
  - Use the PPE for less frequently repeated tasks with more random access data access

- **The application developer is the composer**
  - You choose how to assign tasks between SPE and PPE
  - Data flow and synchronization are critical for maximum performance

Cell Processor Architecture

Power Processor Element (PPE)
(64 bit PowerPC with VMX)

I/O Controller

Memory Controller

RAM

I/O Controller

DRAM

SPE 1

SPE 2

SPE 3

SPE 4

SPE 5

SPE 6

SPE 7

SPE 8

Dual "configurable" High speed I/O channels
(76.8 GBytes per second in total)

Dual 12.8 GByte per second memory busses give Cell huge memory bandwidth. (25.6 GBytes per second in total)

EIB (Element Interconnect Bus) is the internal communication system.

© Nicholas Blachford 2005

http://www.blachford.info/computer/Cell/Cell_Arch.gif
Processor Block Diagram

XDR Mem
25.6 GB/s

XIO

MIC
Memory Interface Controller

SPE<sub>1</sub>
LS (256KB)
DMA

SPE<sub>3</sub>
LS (256KB)
DMA

SPE<sub>5</sub>
LS (256KB)
DMA

SPE<sub>7</sub>
LS (256KB)
DMA

I/O

Flex-IO<sub>1</sub>

Flex-IO<sub>0</sub>

PPE

L1 (32 KB I/D)

L2 (512 KB)

SPE<sub>0</sub>
LS (256KB)
DMA

SPE<sub>2</sub>
LS (256KB)
DMA

SPE<sub>4</sub>
LS (256KB)
DMA

SPE<sub>6</sub>
LS (256KB)
DMA

EIB is 4 ring buses up to 96B per clock, 2 in each direction.

Total 76.8 GB/s

I/O

http://research.scea.com/research/html/CellGDC05/08.html
Cell Chip Layout
PPE

*Power Processing Element*

- Dual-threaded, 64-bit core
- 32-kB instruction/data L1 caches
- 512-kB L2 cache
- Power Architecture family compliance:
  - Integer unit
  - Floating point unit
  - VMX unit
  - MMU unit
- Handles DMA requests for SPEs
  - via memory-mapped IO controllers
SPE
Synergistic Processing Element

• Data and instructions stored in private 256kB local storage (LS)
• SPEs share system memory (L2 cache) with PPE
• Asynchronous data transfers between LS and main storage (L2 cache)
  – Communication overlaps with computation ➔ real-time operation

http://www.realworldtech.com/includes/images/articles/cell-3.gif
EIB

Element Interconnect Bus

- Four 16-byte-wide data rings
  - Each support 3 simultaneous data transfers
- 96 bytes per processor cycle

Physical interleaving of bus to minimize noise coupling

http://www.realworldtech.com/includes/images/articles/cell-9.gif
MIC
Memory Interface Controller

- Supports 2 Rambus XDR™ memory banks/channels
- Each bank 36 bits wide
- 2 XDR channels, 4 DRAM chips
  - 25.2 GB/s maximum bandwidth!
- MIC is asynchronous to processor and I/O interfaces
  - Speed-matching SRAM buffers/logic
  - Two clock domains
  - Transceiver training sequence req’d
  - Provides greater flexibility

http://www.realworldtech.com/includes/images/articles/cell-10.gif
BIC
Bus Interface Controller

- Rambus FlexIO™ interface
- 44.8 GB/s outbound
- 32 GB/s inbound
- TX/RX asynchronous to proc./memory
- Supports multiple configurations w/o chip redesign or repackaging
- BIC provides asynch. interface b/n EIB and 2 I/Os
  - SPEs can snoop I/O activity via EIB
  - speed-matching SRAM buffers/logic
  - Three clock domains
Pervasive Unit

- Contains all global logic for:
  - Basic functional operation
  - Lab debug
  - Manufacturing test
- Control for clock generation and distribution
- Power-on-reset (POR) engine
  - Contains debug mode allowing single-step, skipped, or out-of-order instruction sequences
- Performs performance analysis of all functional units on chip via trace/debug bus
- Provides logic used for programming and eFuses
  - Array repair and chip customization during manufacturing test
PMU
Power Management Unit

• Software controls to reduce chip power when full processing not needed

• OS has the power to:
  – Throttle...
  – Pause...
  – Stop...
  – ...single or multiple units or entire chip
Two types of thermal sensors:
- External linear diode
  - For controlling external cooling mechanisms
- 10 digital thermal sensors (DTSs)
  - Distributed throughout chip
  - Temp of each element can be individually controlled by throttling

Software controls TMU settings for each element
- Four temperature settings for each element:
  - Throttling stops
  - Throttling starts
  - Element shut down
  - Chip clocks shut down
Digital Thermal Sensor

Programmable temp ref.

2°C relative steps

Element temp. sensing diode
Technology Issues

• Additional HVT device
• Analog designs:
  – ↑ channel length
  – ↑ gate oxide thickness
  – Body-contacted devices
    • Vt control
  – Chip decoupling caps
• Interconnect
  – Too few layers → ↑ area
  – Too many layers → ↑ $$$

<table>
<thead>
<tr>
<th></th>
<th>FEOL</th>
<th>90 nm PD-SOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voltage</td>
<td>1.0 V</td>
<td></td>
</tr>
<tr>
<td>Thin oxide device</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxide thickness</td>
<td>10.5 A</td>
<td></td>
</tr>
<tr>
<td>Channel length</td>
<td>46 nm</td>
<td></td>
</tr>
<tr>
<td>Threshold voltages</td>
<td>Reg. $V_T$, High $V_T$, Super High $V_T$</td>
<td></td>
</tr>
<tr>
<td>Thick oxide device</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxide thickness</td>
<td>22.5 A</td>
<td>analog</td>
</tr>
<tr>
<td>Channel length</td>
<td>140 nm</td>
<td></td>
</tr>
<tr>
<td>Threshold voltages</td>
<td>Reg. $V_T$</td>
<td></td>
</tr>
<tr>
<td>Decoupling cap oxide thickness</td>
<td>15A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BEOL</td>
<td>Local Interconnect plus 8LM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5-1X layers, 1-2X layer, 2-6X layers</td>
</tr>
</tbody>
</table>
Packing Issues

- Flip-chip PBGA
- Decoupling caps mounted on back against chip (close as possible to power pins)
  - Facilitated by fine-pitch mechanical drilling
  - Ensures robust mid-freq response
- Lid in thermal contact with die
- 3,349 C4s (solder bumps)
Clock Distribution

- Three distinct clock distribution systems
  - Processors
  - BIC
  - MIC
- Main high-freq grid
  - Covers 85% of chip
  - Buffered tunable trees
  - Final levels drive common grid
- Clock grids interleaved
  - Logic circuits can use all grids
- 850 individually tunable buffers (!)

- 3 clock grids
- 1,100 clock buffers
- 19.4 meters of metal!
Clock jitter and skew @ 16 GHz

• Avg.(?) Jitter
  – 12.7 ps peak-to-peak
  – 1.57 ps rms

• Min. Jitter
  – 9.48 ps peak-to-peak
  – 1.09 ps rms

• Skew
  – < 12 ps

Fig. 13. Main clock distribution normalized arrival time.
Buffers/latches/FFs

- The point:
  - Minimize power consumption of clocked elements as a whole
  - Provide special high-speed constructs for critical paths
  - Single global clock, local clock splitters
Buffers/latches/FFs

- Support multiple operating modes
  - Local clock activate mode
  - Scan mode (scan)
  - Test mode (testhold_b)
  - State preservation

Fig. 14. Basic local clock components.

Fig. 15. Standard master–slave flip-flop.
Buffers/latches/FFs

- Support multiple operating modes
  - Local clock activate mode
  - Scan mode (scan)
  - Test mode (testhold_b)
  - Pulse mode (Fig. 17) ➔ “chicken-switch” used if race conditions occur
SRAM

• L1 cache:
  – 32-kB
  – 3-cycle latency
  – Parity checking
  – 64-byte write
  – 16-byte read
  – 16 cells / local BL
  – 8LBLs / GBL

• L2 cache:
  – 512-kB
  – 1/8 active at a time
  – 280-bit reads supported
  – 2-cyc write / 3-cyc read

• SPE Local Stores
  – Four 64-kB arrays
  – 4-cyc write / 6-cyc read
Noise Analysis

• Macro-level
  – Transistor-level simulator
  – Netlist w/ parasitics extr. from macro layout
  – Static noise analysis
  – Macro noise abstract formed
    • Input noise tolerance
    • Input capacitance
    • Output resistance

• Unit/chip-level
  – Input:
    • Macro abstract
    • Timing analysis
  – Global netlist w/ parasitics extr. from global layout
  – Dynamic noise analysis using:
    • Resistors, caps, voltage sources
    • Timing information
  – Failure reported if noise level exceeds input tolerance
**Power Analysis**

Fig. 25. (a) Macro and net power for a typical work load. (b) Total power.

**Total Average Power: 60-80 Watts**

Fig. 26. Percent of local clock buffers active for idle, typical, and high-power workload.
Thermal Analysis

- Analytic modeling techniques based on
  - Classical steady-state thermal diffusion theory
  - Transient analysis under various workloads
- Considered ext. lateral thermal spreading through the package
- Analysis carried out early in design cycle
  - Affected floorplan design and thermal sensor design/placement
Conclusion

(It works.)

Fig. 28. First pass hardware results in the laboratory.
Except where stated, all figures and information on the CELL processor is from: