Configurable Processors for SOC Design

Contents crafted by
Technology Evangelist Steve Leibson
Tensilica, Inc.
Why Listen to This Presentation?

- Understand how SOC design techniques, now nearly 20 years old, are breaking down under the continued application of Moore’s Law
- Learn about the significant differences between fixed-ISA microprocessor cores for SOC design and configurable cores
- Unlearn some outdated design habits
Today’s Standard SOC Design Techniques are Failing

- HDLs and logic synthesis developed at a time when ASICs were ~200,000 gates
  - 90nm silicon puts 200K gates in 1 mm², easily accommodates 500K-gate blocks, and 100M max gate counts per chip
  - Hand-coded HDL productivity has not and cannot keep pace

- Verification now >70% of the total SOC design effort
  - Not enough verification engineers in the world
  - The only solution is massive adoption of pre-verified IP blocks

- Cost of fixing SOC bugs is rising
  - A silicon respin costs more than $1M US (1.5M Euro)

- Late hardware/software integration
  - Has not changed in 35 years

- SOC designs much more complex
  - Example: Multiple audio, video compression standards
Increasing SOC Design Complexity Spurs Thirst for Processing Speed

- New multimedia standards set the pace for a rapid rise in computational power
- The tried-and-true way to get more processor performance has been clock rate increases achieved through Denard (Classical) Scaling
  - Often misidentified as Moore’s Law
  - Moore’s Law (2x more transistors every ~18 months) – Lives!
  - Denard Scaling died with the 90nm node
    - Transistors no longer getting faster as quickly
    - Leakage effects from low-Vt transistors make low-power operation tough
Data Speed Increases Every Year. Why?

IEEE 802.11

W-CDMA

HSDPA

Super3G

LTE

4G

Data speed [bps]

Year

Fixed network services

Mobile services

Copyright (C) 2007 NTT DoCoMo, Inc. All rights reserved
Evolving Audio Capability

**Technologies**

- Monaural
- Stereo
- Simple reproduction of 3D sound environment
- Spatial reproduction of 3D sound environment
- Bidirectional data transmission / Real Time
- High-quality sound

**Reality**

- Low coding/decoding time lags for 3D signaling
- 3D acoustic sound recording
- Broadband Wireless Access
- IP Packet Transfer (VoIP)
- High-end Broadband Coding
- Real World Compatible Sound Source (When turning from back to front sound moves to reflect user’s position)
- User has a larger range of movement
- User’s fixed location (Speakers)
- User’s position (Headphones) (Sound is heard from the correct side: left/right)
Evolving 3D Image Transmission

- Low coding and decoding time lags / 3D signaling
- Real time 3D picture recording

- User has a larger range of movement
- Multiple perspectives / parallax movements (user will be able to see the direction he is facing)

- User’s fixed location (can only see from user’s fixed location)
- One viewpoint (can only see from the viewer’s perspective)

Bidirectional data transmission / Real time

Multi-perspective stereo picture display

Display diversification

3D picture coding

Stereo picture display

High quality 2D picture (TV quality)

Video Picture

- Broadband Wireless Access
- IP Packet Transfer / High-end Broadband coding

Technologies
Rise in Packaged Microprocessor Clock Rate Over Time

Microprocessor Clock Rate over Time

- Intel 8008
- Intel 8080
- Intel 8086
- Intel 80186/80286
- Motorola 68000
- Motorola 68020
- Intel i386DX
- Intel i486DX
- Intel Pentium
- Intel Pentium Pro
- AMD K6
- DEC Alpha 21064
- DEC Alpha 21164
- AMD Athlon
- Intel Pentium III
- Intel Pentium 4
- Intel Pentium 4

Introduction Date

- 1971
- 1972
- 1973
- 1974
- 1975
- 1976
- 1977
- 1978
- 1979
- 1980
- 1981
- 1982
- 1983
- 1984
- 1985
- 1986
- 1987
- 1988
- 1989
- 1990
- 1991
- 1992
- 1993
- 1994
- 1995
- 1996
- 1997
- 1998
- 1999
- 2000
- 2001
- 2002
- 2003
- 2004
- 2005
Clock Frequency Inflection Point
End of the Road for Clock-Rate Ramp

Based on International Technology Roadmap for Semiconductors 12/03
What Replaces Clock-Rate Increases?

- Processor configuration
- New registers and register files sized to the data
  - 24 bits for audio
  - 56 bits for encryption
- New function units to match application
  - 24-bit multipliers for audio
  - Time-to-live field extraction/increment for networking
  - Viterbi butterfly for mobile phones
  - SAD (sum of absolute differences) pixel processing for video
ByteSwap: A Simple Example
ByteSwap in C and ASM

unsigned ss = (s<<24) | ((s<<8)&0xff0000) | ((s>>8)&0xff00) | (s>>24);

slli a9, a14, 24
slli a8, a14, 8
srli a10, a14, 8
and a10, a10, a11
and a8, a8, a13
or a8, a8, a9
extui a9, a14, 24, 8
or a10, a10, a9
or a10, a10, a8

Form intermediate result bits 24-31 in register a9
Shift 32-bit word left by 8 bits, save in register a8
Shift 32-bit word right by 8 bits, save in register a10
Form intermediate result bits 9-15 in register a10
Form intermediate result bits 16-23 in register a10
Form intermediate result bits 16-31, save in register a8
Extract result bits 0-7, save in register a9
Form result bits 0-15, save in register a10
Form final result in register a10
Define ByteSwap Instruction in TIE, Use BYTESWAP as C Intrinsic

One instruction versus nine

operation BYTESWAP {out AR outR, in AR inpR} { }
{
    wire [31:0] reg_swapped = {inpR[7:0],inpR[15:8],inpR[23:16],inpR[31:24]};
    assign outR = reg_swapped;
}
Define ByteSwap Instruction in TIE, Use BYTESWAP as C Intrinsic

One instruction versus nine

operationBYTESWAP {out AR outR, in AR inpR} { }
{
    wire [31:0] reg_swapped = {inpR[7:0],inpR[15:8],inpR[23:16],inpR[31:24]};
    assign outR = reg_swapped;
}

- Writing this one line of TIE code and submitting it to the processor generator...
- Automatically adds the appropriate hardware to the pipeline and instruction decoder
- Automatically adds the instruction to the compiler, assembler, debugger, linker, and ISS
Viterbi Butterfly Function Unit

- Unaugmented processor needs 42 cycles/butterfly
- Viterbi function unit ~11K gates
- Augmented processor needs 0.16 cycles/butterfly
- 250x speedup
SIMD SAD Function Unit for Video

- Unaugmented processor needs 641M cycles to compute SAD for QCIF image @ 15fps
- Augmented processor computes SAD for 16 pixels/cycle and needs 14M cycles to compute SAD for QCIF image @ 15fps
- 46x speedup
I/O Just as Important as Computation: Think Outside the Bus
Where did the Bus Originate?
World’s First Commercial Microprocessor
*The Intel 4004, Circa 1971*

16 pins – necessitating a multiplexed bus
And processors have been pin-limited ever since
Wide Interconnect and Wire Density: What’s Practical?

- Do the Math:

<table>
<thead>
<tr>
<th>Layer</th>
<th>Wire Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 mm</td>
<td>100,000+ wires/square mm</td>
</tr>
<tr>
<td>1 mm</td>
<td>200,000+ wires/square mm</td>
</tr>
</tbody>
</table>

8-10 Metal Layers, ITRS 2006 wire spacing

- 90nm: 100,000+ wires/square mm
- 65nm: Almost 200,000 wires/square mm
Keys to Efficient MP
Flexible range of system-interconnect topologies

Shared Bus
- Processor Master
- Processor Master
- Processor Master
- Processor Master
- Input Device Slave
- Output Device Slave
- Memory Slave
- Memory Slave

On-chip Routing Network
- Processor Master
- Processor Master
- Processor Master
- Routing Node
- Input Device Slave
- Output Device Slave
- Memory Slave
- Global Memory Slave
- Global I/O Slave
- Global I/O Slave

Cross-Bar
- Processor Master
- Processor Master
- Processor Master
- Input Device Slave
- Output Device Slave
- Memory Slave
- Memory Slave

Application-specific
- Processor Master
- Dual-Port Memory
- Queue
- Global Memory Slave
- Input Device Slave
Getting Your Data Off of the Bus: Dual-Ported Memory
“Instantaneous” Data Teleportation: Direct-Port Connection
Fast, Wide Data Transfers: FIFO-Queue Connection

Diagram showing the connection between Data Memory, Processors, and Instruction Memory, with arrows indicating 'Full,' 'Push,' 'Pop,' and 'Empty' states.
Multiple Queue Destinations

Producer Processor

Address Decode

Consumer Processor 1

Consumer Processor 2

Consumer Processor 3

Consumer Processor 4
True “Many-Core” System-on-Chip
192 Xtensas Per Chip in Cisco CRS-1 Terabit Router

- 192 Xtensa network-processor cores per Silicon Packet Processor
- Up to 400,000 processors per system

Complete 32-bit processor: 25K gates
12 processors per cluster
16 processor clusters per chip
EPSON’s printer controller chip with six heterogeneous, asymmetric, configurable VLIW cores + legacy controller + I/O
Data-plane is More than DSP
A Recent Example: Server Engines BladeEngine

From EETimes: July 31, 2007

• Next generation high-volume server chipset
• Processor-based acceleration of TCP/IP and network/storage protocols
• 8 configurable data processors + ARM management processor
1. SOC complexity will continue to increase
2. Programmability essential to flexible system design
3. Many board-level system-design techniques are not appropriate for SOC design
4. Hand-coded HDL with proper verification cannot produce 100M-gate SOCs in any realistic schedule
5. The years of ever-increasing clock rate are behind us
6. Use of multiple processors and high-interconnectivity system topologies is essential to fully exploiting nanometer silicon