Tutorial Hot Chips 01

Silicon Architectures for Wireless Systems – Part 1

Jan M. Rabaey
BWRC
University of California @ Berkeley
http://www.eecs.berkeley.edu/~jan
The Fibonacci Law on Wireless Growth

“The number of worldwide wireless subscribers (in tens of millions) grows as a Fibonacci series”

Source: Goldman-Sachs
From Handsets to Mobile Devices

Internet access the most important driver (text, graphics, multimedia)

Berkeley Infopad, One of the first wireless internet appliance 1990-1996
A Smorgasboard of Choices

2nd generation
- IS136
- GSM
- IS95

Generation 2.5
- IS136+
- GPRS

3rd generation
- Edge
- WCDMA
- IS95-B

Packets
- NARROWBAND
- WIDEBAND

9.6Kbps
- 2nd generation

64-384Kbps
- Generation 2.5

384-2000Kbps
- 3rd generation

Networks:
- NARROWBAND CIRCUIT VOICE
- WIDEBAND PACKET DATA

Technologies:
- GSM
- IS136
- IS95
- IS95-B
- GPRS
- WCDMA
- Edge
- IS136+
- cdma2000

Generations:
- 2nd generation
- Generation 2.5
- 3rd generation
And the Alternatives

**Metropolitan Wireless Networks**
- Various proprietary solutions
  - Ricochet (up to 138 Kb/sec)
  - Flarion Flash-OFDM (> 384 kBits/sec)

**Wireless LANs**
- 802.11 (b,a); Hyperlan
- from 1 to 56 Mbit/sec
- Restricted to the 50 meter range (at present)
(Projected) Growth in 802.11 WLAN

Units, k

Source: Cahners In-Stat 2001
The New Internet

The 1990s: Conquering the World
The Network revolution

The 2000s: Extending toward the Small
Enabled by integration and wireless connectivity

Pre-1990: Client-Server Systems
The Post-PC Era: The Distributed Approach to Information Processing
The emergence of ad-hoc wireless networks – the wire replacement

- Bluetooth, HomeRF
  ✓ up to 800 kBit/sec

- Sensor networks
  ✓ Low data-rates
The Evolving Wireless Scene

<table>
<thead>
<tr>
<th>Range</th>
<th>Data Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1m</td>
<td>1Kb</td>
</tr>
<tr>
<td>10m</td>
<td>10Kb</td>
</tr>
<tr>
<td>100m</td>
<td>100Kb</td>
</tr>
<tr>
<td>1km</td>
<td>1Mb</td>
</tr>
<tr>
<td>10km</td>
<td>10Mb</td>
</tr>
<tr>
<td>100Mb</td>
<td>100Mb</td>
</tr>
</tbody>
</table>

- **Cellular (WAN)**
  - 3G Cellular
  - 2.5 G Cellular
  - Cellular (WAN)

- **802.11 (LAN)**
  - 802.1a

- **802.15 (PAN)**
  - Bluetooth
  - Sensor networks

- **More bit/sec**
Compelling Issues in Wireless (1)

Ubiquitous services put wireless spectrum at a premium

- Effective use of aether hampered by standardization and fragmentation
- Current spectral efficiency far below theoretical limits
- Emerging Solutions
  - Adoption of better spectrum utilization techniques (interference cancellation, multi-path fading mitigation and exploitation)
  - Multi-functional, adaptive systems
- But … huge appetite for computations
Evolution of MOPS Requirements in Cellular

Single 384 kbps UTRA W-CDMA Channel

<table>
<thead>
<tr>
<th>Function</th>
<th>MOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital RRC Channel</td>
<td>3600</td>
</tr>
<tr>
<td>Searcher</td>
<td>2100</td>
</tr>
<tr>
<td>RAKE</td>
<td>2050</td>
</tr>
<tr>
<td>Maximal Ratio Combiner</td>
<td>24</td>
</tr>
<tr>
<td>Channel Estimator</td>
<td>12</td>
</tr>
<tr>
<td>AGC, AFC</td>
<td>10</td>
</tr>
<tr>
<td>Deinterleaver</td>
<td>15</td>
</tr>
<tr>
<td>Turbo Coder</td>
<td>90</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>7901</td>
</tr>
</tbody>
</table>

Source: IEEE Comm Theory Workshop, May 1999
The Cost of Approaching Shannon’s Bound

The Bliss and Challenge of Error Coding

Relative Complexity vs. SNR (db)

1/2 LDPC, N=10^7, 1100 iterations

8/9 Capacity Bound

2/3 Capacity Bound

1/2 Capacity Bound

8/9 LDPC, N=4k
1, 3, and 5 iterations

1/2 Turbo, v=4, N=64k 1, 2, and 3 iterations

2/3 Turbo, v=4, N=64k 1, 2, and 3 iterations

8/9 Turbo, v=4, N=4k

1/2 Conv. Code, v=4, N=64k

2/3 Conv. Code, v=4, N=64k

8/9 Conv. Code, v=3, N=4k

for BER of 10^-5

Courtesy Engling Yeo, UCB

Berkeley Wireless Research Center
Dealing with Non-ideal Channels (e.g., fading)

- Multi-antenna approach exploits multi-path fading by sending data along good channels
- Results in large theoretical improvements in bandwidth efficiency for fading channels
- But…computationally hungry
The Cost of Dealing with Non-ideal Channels

<table>
<thead>
<tr>
<th>Data rate per user</th>
<th>Spectral efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8 Mb/s</td>
<td>0.9 b/s/Hz</td>
</tr>
<tr>
<td>1.6 Mb/s</td>
<td>1.8 b/s/Hz</td>
</tr>
<tr>
<td>1.9 Mb/s</td>
<td>2.1 b/s/Hz</td>
</tr>
<tr>
<td>3.8 Mb/s</td>
<td>4.2 b/s/Hz</td>
</tr>
<tr>
<td>5.6 Mb/s</td>
<td>6.3 b/s/Hz</td>
</tr>
</tbody>
</table>

* Assume 25 MHz bandwidth and 28 users

Source: Ning Zhang, UCB
Shannon beats Moore’s law

Algorithmic Complexity (Shannon’s Law)

Processor Performance (~Moore’s Law)

1G
2G
3G


Courtesy: Ravi Subramanian (Morphics)
Single-Chip DSPs are Lagging...

While algorithms are beating Moore’s law!

Moore’s law: x 1.58/year
DSP Trend: x 1.4/year

Source: TI
Digital Processor Performance

Courtesy of Ravi Subramanian (Morphics)
The Law of Diminishing Returns

• More transistors are being thrown at improving general-purpose CPU and DSP performance

• Fundamental bounds are being pushed
  – limits on instruction-level parallelism
  – limits on memory system performance

• Returns per transistor are diminishing
  – new architectures realizing only 2-3 instructions/clock
  – increasingly large caches to hide DRAM latency
Some observations

• Von-Neuman style instruction set architectures were perceived when switching devices and interconnections were extraordinarily expensive, and multiplexing-in-time provided the most economical solution
  – Intel 4004: 2000 transistors, 1 MHz clock frequency, 1 metal layer

• This led to the “clock-speed” affixation, which in fact is only a secondary measure of performance

• Power is rapidly becoming a limiting factor
  – Newest processors are including thermal sensors and automatic slow-down (throttling) using pipeline bubbles and nop’s to combat overheating and meltdown
“The Last Meter Problem”
Ubiquitous wireless networking requires steep reduction in cost and energy dissipation

• To be acceptable, radio cost has to be below 1$
• Frequent battery replacement on 100’s of devices unacceptable
• Technology not likely to be of major help
Energy to Play a Major Role

Algorithmic Complexity (Shannon’s Law)

Processor Performance (~Moore’s Law)

Battery Capacity

Courtesy: Ravi Subramanian (Morphics)
Energy Trends in DSPs

Source: TI
Energy Trends in DSPs

Gene (Frantz)’s Law

Source: Gene Frantz (TI)
Energy to Play a Major Role
A holistic perspective

Energy = upper bound on the amount of available computation

- Total Energy of Milky Way Galaxy: $10^{59}$ J
- Minimum switching energy for digital gate (1 electron@100 mV): $1.6 \times 10^{-20}$ J (limited by thermal noise)
- Upper bound on number of digital operations: $6 \times 10^{78}$
- Operations/year performed by 1 billion 100 MOPS computers: $3 \times 10^{24}$
- Energy consumed in 180 years assuming a doubling of computational requirements every year.
Putting energy in perspective

• **Energy cost of digital computation**
  - 1999 (0.25µm): 1pJ/op (custom) … 1nJ/op (μproc)
  - 2004 (0.1µm): 0.1pJ/op (custom) … 100pJ/op (μproc)
    • Factor 1.6 per year; Factor 10 over 5 years
    • Assuming reconfigurable implementation: 1 pJ/op

• **Energy cost of communication**
  - 1999 Bluetooth
    • 1 nJ/bit transmission energy (thermal limit 30 pJ/bit)
      • 2.4 GHz band, 10 m distance
    • Overall energy: 170 nJ/bit reception / 150 nJ/bit transmission (!)
    • Standby power: 300 µW
  - 2004 Radio (10 m)
    • Only minor reduction in transmission energy
    • Reduce transceiver energy with at least a factor 10-50
The Changing Metrics

• **Power and/or Energy** have become dominant drivers
  – Limiting factor for performance and reliability in wall-plugged applications
  – Enabler for wide-spread use of distributed computing and data access

• Energy reduction requires joint optimization process between application and implementation
The Changing Metrics

- Design complexity, and “context complexity” is sufficiently high that design verification is a major limitation on time-to-market
- Cost of fabrication facilities and mask making has increased significantly
  - NRE cost of new design has increased significantly
- Physical effects (parasitics, reliability issues, power management) are increasingly significant in the design process
  - These must now be considered explicitly at the circuit level

Towards Fewer, but more Flexible and Reusable Silicon Platforms
The Changing Metrics

Performance as a Functionality Constraint
(“Just-in-Time Computing”)
Challenges in Single-Chip Radio Design

Application | Network | Mac/Data Link | Physical + RF

source data sec | streams msec | packets µsec | bits nsec

Data and Time Granularity

UI | Call Setup | Slot Allocation | Synchronization

Data

Control

Radio

Berkeley Wireless Research Center
The Software Radio

- Idea: Digitize (wideband) signal at antenna and use signal processing to extract desired signal
- Leverages of advances in technology, circuit design, and signal processing
- Software solution enables flexibility and adaptivity, but at huge price in power and cost
- 16 bit A/D converter at 2.2 GHz dissipates 1 to 10 W
The Mostly Digital Radio

RF input ($f_c = 2\text{GHz}$)

RF filter

LNA

chip boundary

RF input

$\cos[2\pi(2\text{GHz})t]$ and $\sin[2\pi(2\text{GHz})t]$

A/D

A/D

Digital Baseband Receiver

I (50MS/s)

Q (50MS/s)

Analog

Digital
The Opportunities …

- **Scaling of technology, of course**
  - Performance doubles every 18 months
  - Energy reduced by 10 every 5 years

- **The system-on-a-chip**
  - Integration of heterogeneous functions on a single die leads to new solutions

- **Novel architectures**
  - Processor architectures exploiting concurrency (e.g. VLIW)
  - Reconfigurable hardware
  - Network-on-a-chip

- **Novel circuit solutions**
  - Voltage as a design variable

- **Novel design methodologies**
  - Communication-based Design
  - Platform-based Design
The Ideal “Radio-on-a-Chip” Platform

Combines performance, flexibility and energy-efficiency

- Heterogeneous
- Supports massive concurrency
- Matches the computational model
- Operates at minimum supply voltage and clock frequency
- Provides flexibility only where needed and desirable and at the right granularity

Reconfigurable DataPath

Reconfigurable State Machines

Embedded \( uP + \) DSPs

FPGA

Dedicated DSP

Reconfigurable DataPath
The System-on-a-Chip Nightmare

“Femme se coiffant”
Pablo Ruiz Picasso
1940
The "Board-on-a-Chip" Approach
Current Integrated Wireless Transceivers

A "Board-on-a-Chip" Approach
Example: Single-Chip Bluetooth

Source: Alcatel, ISSCC 2001

40 mm²
CMOS 0.25µm
5 metals
MiM-capacitors
200Ω/□ Resistors
Flash
System-on-a-Chip
A Renaissance in Design

Applications
Multimedia
Consumer
Communications

Design
Methodology
Hard+Soft

Convergence

Implementation
Fabrics
Silicon reuse
Silicon networks

Aart De Geus
DAC’99
A Central Theme: Raising the Reuse Factor

Reuse comes in generations

<table>
<thead>
<tr>
<th>Generation</th>
<th>Reuse element</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1\textsuperscript{st}</td>
<td>Standard cells</td>
<td>Well established</td>
</tr>
<tr>
<td>2\textsuperscript{nd}</td>
<td>IP blocks</td>
<td>Being introduced</td>
</tr>
<tr>
<td>3\textsuperscript{rd}</td>
<td>Architecture</td>
<td>Emerging</td>
</tr>
<tr>
<td>4\textsuperscript{th}</td>
<td>IC</td>
<td>Early research</td>
</tr>
</tbody>
</table>

Source: Theo Claasen (Philips) – DAC 00
Platform-Based Design

“Only the consumer gets freedom of choice; designers need freedom from choice”

- A platform is a restriction on the space of possible implementation choices, providing a well-defined abstraction of the underlying technology for the application developer.
- New platforms will be defined at the architecture-micro-architecture boundary.
- They will be component-based, and will provide a range of choices from structured-custom to fully programmable implementations.
- Key to such approaches is the representation of communication in the platform model.

Source: R.Newton
Hardware Platforms

Hardware Platform: not only a fully specified SoC but also a **family of architectures** that share some **common feature**:

A Hardware Platform is a family of architectures that satisfy a set of architectural constraints imposed to allow the re-use of hardware and software components.

The stronger the constraints the more component re-use but stronger constraints imply fewer architectures to choose from!
Hardware Platforms Not Enough!

- Hardware platform has to be abstracted
- Interface to the application software is API
- Software layer performs abstraction:
  - Programmable cores and memory subsystem with RTOS
  - I/O subsystem via Device Drivers
Software Platforms

- Application Software
- Platform API
- System API
- RTOS
- BIOS
- Device Drivers
- Network Communication
- Hardware Platform
- Software Platform
- Input devices
- Output Devices
- I/O
- Network
The Platform Tension

Application Space

Architectural Space
The Platform Approach

Source:
Alberto Sangiovanni-Vincentelli
Example: Philips Nexperia™ DVP

MIPS™ TriMedia™

MIPS CPU
D$
I$
PRxxxx

DEVICE I/P BLOCK

DEVICE I/P BLOCK

DEVICE I/P BLOCK

DEVICE I/P BLOCK

SDRAM

MMI

TriMedia CPU
D$
I$
TM-xxxx

DEVICE I/P BLOCK

DEVICE I/P BLOCK

DEVICE I/P BLOCK

DEVICE I/P BLOCK

PI BUS

DVP MEMORY BUS

PI BUS

DVP System Silicon

General Purpose RISC Processor
• 50 to 300+ MHz
• 32-bit or 64-bit

Library of Device Blocks
• Image coprocessors
• DSPs
• UART
• 1394
• USB

• ...and more

VLIW Media Processor:
• 100 to 300+ MHz
• 32-bit or 64-bit

Nexperia System Busses
• PI bus
• Memory bus
• 32-128 bit

Flexible architecture for digital video applications

Source: Theo Claasen (Philips) – DAC 00
Nexperia™ Scalability

- Single architecture, multiple product configurations
  - Processor core options - TM32, TM64, MIPS32, MIPS64 ...
  - Device block options
- Highly programmable to weakly programmable

Source: Theo Claasen (Philips) – DAC 00
An Example of an Instantiation

Combines Trimedia VLIW with Configurable media co-processors

Philips Nexperia NX-2700
A programmable HDTV media processor
Pleiades: Digital Wireless Platform

Source: Berkeley Wireless Research Center
The Architectural Choices

Flexibility

1/Efficiency

- Dedicated Logic
  - Direct Mapped Hardware
  - Hardware Reconfigurable Processor
  - Satellite Processor
  - Satellite Processor

- Prog Mem
  - μP
  - MAC Unit
  - Addr Gen
  - Software Programmable DSP

- General Purpose μP
  - Prog Mem

Berkeley Wireless Research Center
An Attractive Option: Multi-Processor System-on-a-Chip

"The New NAND gate"

Number of Processor per SOC (Xtensa + Other Processors)

1 Core 2 Cores 3-10 Cores 10 to 100 100+ Processors

Number of Design Starts

Copyright Tensilica, Inc 2001

Courtesy: Chris Rowen, Tensilica
The Energy-Flexibility Gap

- **Embedded Processors**
  - SA110: 0.4 MIPS/mW
- **ASIPs**
- **DSPs**
  - 2 V DSP: 3 MOPS/mW
- **Reconfigurable Processor.Logic**
  - Pleiades: 10-80 MOPS/mW
- **Dedicated HW**

Energy Efficiency (MOPS/mW or MIPS/mW) vs. Flexibility (Coverage)
Orthogonalizing Communication from Behavior

- Historically lots of work on Behavior
  - hierarchy well established
  - several descriptions available (with variable levels of precision)
  - synthesis available
- Communication less well investigated
  - hard to separate from behavior, usually intertwined
  - telecomm protocols are the best existing example
- Need to understand Formalism, Abstraction, and Decomposition for communication

DEFINED BY MODELS OF COMPUTATION!
Communication-based Design
An Integrated Radio Processor (TCI)

- Memory Sub-system
- Baseband Processing
- Fixed Protocol Stack
- Programmable Protocol Stack
The “Network-on-a-Chip”

<table>
<thead>
<tr>
<th></th>
<th>dot product</th>
<th>vector sum w/ scalar mult.</th>
<th>IIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-bus</td>
<td>50</td>
<td>50</td>
<td>138</td>
</tr>
<tr>
<td>Mesh</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best</td>
<td>8.7</td>
<td>5.2</td>
<td>24.6</td>
</tr>
<tr>
<td>Worst</td>
<td>17.7</td>
<td>14.7</td>
<td>43.4</td>
</tr>
<tr>
<td>H. Mesh</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best</td>
<td>4.7</td>
<td>3.8</td>
<td>18.8</td>
</tr>
<tr>
<td>Worst</td>
<td>11.1</td>
<td>10.2</td>
<td>31.3</td>
</tr>
</tbody>
</table>
Summary and Perspective

• Technology scaling is redefining the term “complexity”
• System-on-a-Chip fosters renaissance in processor architecture, opening the door for new models and combinations thereof:
  Platform and Communication Based Design
• SOC for wireless driven by new set of metrics: how to simultaneously optimize flexibility, cost, energy, and performance?
• Application-Architecture Exploration is Focal Part of Implementation Methodology