Mass Market Applications for MPSOC and Design Space Exploration Challenge

1. Baseband
2. Multimedia
3. Spinoffs

Ulrich Ramacher
COM IN
<table>
<thead>
<tr>
<th>App.</th>
<th>Range</th>
<th>Rate</th>
<th>Freq.</th>
<th>Mod.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAN</td>
<td>WiFi Ultra wideband</td>
<td>10 m</td>
<td>480 Mbps</td>
<td>2.45 – 5.8 GHz</td>
</tr>
<tr>
<td>Wireless LAN</td>
<td>802.11 a/b/g</td>
<td>80 m</td>
<td>11 Mbps (b)</td>
<td>2.45 &amp; 5.8 GHz</td>
</tr>
<tr>
<td>LAN</td>
<td>Wireless LAN</td>
<td>50 - 150 m</td>
<td>100 Mbps ... 600 Mbps (DL)</td>
<td>2.45 &amp; 5.8 GHz</td>
</tr>
<tr>
<td>MAN</td>
<td>Wireless MAN</td>
<td>1-5 km</td>
<td>3...50 Mbps (DL)</td>
<td>2.45 GHz</td>
</tr>
<tr>
<td>MAN</td>
<td>Wireless MAN</td>
<td>1-5 km</td>
<td>63 Mbps (DL)</td>
<td>2.3 GHz</td>
</tr>
<tr>
<td>MAN</td>
<td>Wireless MAN</td>
<td>1+ km</td>
<td>100 Mbps (DL)</td>
<td>2.3 GHz</td>
</tr>
<tr>
<td>WAN</td>
<td>Cellular</td>
<td>-</td>
<td>384 kbps</td>
<td>1.25, 2.2, 5, 10, 20 GHz</td>
</tr>
<tr>
<td>WAN</td>
<td>3G LTE</td>
<td>-</td>
<td>7 Mbps</td>
<td>0.8 MHz, 1.6 GHz</td>
</tr>
</tbody>
</table>

Source: Philips
Classical Solution
ASIC-centric and DSP-assisted

Nokia 6630: WDCMA/EDGE, camera, video, internet access.

TI 3G BB (7.07 x 7.26 mm²)
TI 90nm low-voltage technology – TI 90 nm baseband
Gauge Points in L90

1. GSM/GPRS/EDGE  2. Bluetooth  3. HSDPA  4. 802.11a/b/g
5. DVB-H  6. WiMAX, WiFi 802.11n

No re-use between standards with multiple radios

Multiple radio
How close or far are these 3 solutions
In terms of area, power, flexibility ???

Flexibility

Multi-DSP-centered and Accel.-assisted

- SIMD DSP 1
- SIMD DSP n
- ARM
- Shared Mem
- Acc 1
- Acc n
- Mem

Reconfigurable Architectures

- CU 1
- CU n
- DSP
- rDPU 1
- rDPU n
- Mem
- Shared Mem

ASIC-centered and DSP-assisted

- DSP
- Mem
- Macro 1
- Macro 2
- Macro n

Software Parts

- PHY
- L1

Design Goal

- Highest degree of flexibility
  @acceptable area & power
  Simpler Programming Model

- Reduced area @ acceptable power
  Limited Flexibility
  Complex Programming Model

- Lowest power consumption
  No flexibility (e.g. GSM only)

Icera, Infineon, Philips, Sandbridge...

Cogent, Mercury, Morphics, Morpho,
Quicksilver, Televersal, 3P+1...

Acc: Accelerator
CU: Control Unit

L1: Layer-1 Control
Mem: Memory
PHY: Physical Layer

rDPU: Reconf. Data Processing Unit
SIMD: Single Instruction Multiple Data
Next generation communication cell phones
-- multi-band multi-standard, infinitely many applications

Man-Machine Interface

- Touch screen
  - display
  - virtual keyboard
  - handwriting recogn.
  - graphical user interf.

- Speaker & Microphone
  - hands-free
  - speech recognition

- Camera
  - video
  - still pictures

Communication
- multi bands
- multi standards
- video phone
- fax
- e-mail
- m-commerce

Applications
- Digital Assistant
  - organizer
  - info manager
  - text processing
  - spread sheets & presentations
  - www browser
  - voice recorder
  - scanner
  - virtual file system

Travel & Fun
- localisation, navigation
- travel assistant
- electronic key
- video camera
- digital camera
- picture editing
- MP3 player
- e-book
- games

SW / HW

System Architecture

?
Application roadmap (advanced)

1. Text to Speech (100 MIPS)
2. 3D-Animation (800 MIPS)
3. Keyword Spotting (100 MIPS)
4. Lip Reading (140 MIPS)
5. Face/Object Recognition (1400 MIPS)
6. Gaze Recognition (140 MIPS)
7. Mimics Recognition (200 MIPS)
8. Scene Analysis/AI (500 MIPS)
9. Dialog Modem Overall: ~3400 MIPS
Results of quantitative evaluation

- Performance requirements increase
- Control part grows in coming codecs
- Data part reducible by assembler coding and small ISA extensions
- Cache capacity mainly needed for control parts

- Fine granular interaction between control and data processing
- Code complexity increases (47,000 lines of code for H.264)
Multi Processor Architecture

- SDRAM
- Memory Interface
- L1 unified, pipelined cache, 128 KB (2.8 mm², 18 mW)
- GPP + Media Ext. (0.3 mm², 15 mW)
- ARM Instruction Set + Media Extensions
- L1 caches optimized for power
- L2 cache optimized for hit rate
- MIMD Coprocessor (total: 4.8 mm², 120 mW, 200 MHz) 90 mW

- Competitive Area/MIPS and Power/MIPS through sophisticated Media Extensions and memory system
- Differentiation hardly possible
- Costly ARM architecture license necessary
Infotainment Applications

Inter vehicle communication
- 802.11a, UTRA TDD

Telematic
- navigation
- traffic services
- toll collect / traffic monitoring
- services

Satellites
- Galileo, GPS

Wireless communication
- GSM, GPRS, EDGE, UMTS

Broadcast
- digital radio (DAB, DRM)
- digital TV (DVB-T, DVB-H, DMB, MediaFLO)

Communication
- voice, videophone, text
- internet

Entertainment
- media playback
- games
- internet

HMI
- display(s), speakers
- switches, control elements, key pad
- microphone, camera

Media drives
- DVD, CD, memory cards

Vehicle
- sensors
Customers’ Application Roadmap Requires Programmable and High-Performance Automotive Processors

- Higher complexity of traffic
- More processing tasks
- Higher reliability
- Larger observation area

MIPS

- 2003
- 2005
- 2007
- 2009

Companies:
- Aisin Seiko
- Toyota
- Ficosa
- Volvo
- Iteris
- Mitsubishi
- Nissan
- Audi

Features:
- Lane departure warning
- Blind spot checker
  - (low)
  - (high)
- Blind spot checker
- Parking assistant
- ACC
  - I
  - II
  - III
- Airbag control motion
- Night view
- Door assistant
- Car driver dialog
- Autopilots

Technologies:
- Total circumspection
- Inside
- Rear
- In front
Overtake-Checker and Door-Opener Assistant

Performance on VIP
- Image Size 120x80
- Gauss-Filter (7x7)
- Gradient-Filter
- Edge Detection
- Contour Linking

- 600 fps (total)*
- 1100 fps (edge detection)*
- 1600 fps (linking)*
- 4,3 GMACS (gauss filter)
- 67% of peak performance (gauss filter)

*VIP (80MHz), OAK (53MHz), SDRAM (30MHz)
Face Detection & Recognition

- Leading edge approach of face detection (University of Bochum)
- Detection of face regions (a)
- Pre-selecting of frontal faces (b)
- Face recognition (c,d)
- Elastic graph matching
- Gabor Wavelet Transform

Performance on VIP

- Image size 128x128
- Gabor wavelet transform
  (17x17 kernel, 3 scales, 8 orientations)
- Graph matching
  (16 nodes, 9 reference images)

- measured 11,3 GMAC/s
- 88% of peak performance
MPEG2 Encoder MP@ML (720x576)

Performance on VIP
- 26.7 frames/s
- fully standard compatible
  VIP (80MHz), OAK (53MHz), SDRAM (30MHz)

File with YUV Signal

Motion Estimation → DCT → Quantization → Zig-zag scan VLC/RLC

.mpg file on host
Mass Markets

- scenarios
  - cell phones, PDAs
  - set-top-box, home router, CPE
  - infotainment, driver ass. systems
  - wireless sensor networks (homes, cars, plants)
- multitude of physical media
  - cable, telefone line, radio waves
  - many standards
- modern markets
  - new applications, services, features show up every year
  - time from customer's final spec to ramp-up ~ 12 mths

➔ terminals must support ever more standards
architectures should be “pro-active”
convergence of applications calls for re-use of (modules of) architectures
Architect’s Challenge: Design Space Exploration

Lack of
-- systematic approach
-- benchmark

Where to enter design space?
MuSIC Baseband Processor

- **4 SIMD Execution Units**
- **Long Instruction Word**
  - Computation Slot
  - Memory Slot
  - Communication Slot
- **4 Interleaved Threads**
- **Multi-tasked GP core RC1632 to control SIMD**
- **Local Data Memories instead of L1 Cache**

- **3-Level Memory Hierarchy**
  - external DRAM/Flash
  - shared memory
  - local memories
- **GP Core for L1 Ctrl & MAC**
- reconf. accelerators for FIR and Channel Encoding/Decoding
Technology: 90 nm CMOS
Area: 57 mm²
Power: 330 mW (est.)
2.4M Gates, 768KB Memory

MT SIMD Core
Decoder
rec. Accelerator

ARM Core

FIR
rec. Accelerator

WCDMA 384kb/s
802.11b 11Mb/s

Tape-out in June 06

NEVER stop thinking
SW Architecture

- **Applications**
  - Hand-over
  - Service Selection

- **Radio Layer**
  - Multithreaded C-programs for phy layer ctrl
    - WCDMA
    - 2G- GSM/GPRS/EDGE
    - WLAN

- **C and assembly programs for signal processing**
  - Common modules

- **Operating System**
  - ILTIS (Infineon Lightweight Real-Time Operating System)
    - multithreading, synch., periph. and accel. drivers, IPC

- **Protocol stacks**
  - L2/L3 control
    - 3G-WCDMA (HSDPA)
    - 2G-GSM/GPRS/EDGE
    - WLAN

- **Hardware**
  - Programmable SIMD Cores and accelerators
  - ARM 926

- **Not considered**
  - partly implemented

**Alpha release**
## Tools

### System Modellierung
- **Simulink** (Mathworks)
- **MLDesigner** (ML Design Technologies)
- **CoCentric (COSSAP)** (Synopsys)
- **Metropolis/Ptolemy** (UC Berkeley)
- **UML Tool** (multiple suppliers)
- **SPW** (CoWare)

### Manual Code Generation

- **ILTOS** (IFX)
- **RC1632 Compiler (U-PDB)**
- **SIMD Compiler** (IFX)
- **ARMCC** (ARM)
- **Nucleus** (Mentor)
- **Planned tool support:** Automatic control code generation from models of radio standards

- **Partitioning (Multithreading)**
- **Scheduling (Synchronization)**
- **SIMD & LIW (assembly implementation)**

### SystemC Virtual Prototype (CoCentric Systemstudio)
- **RC1632 Simulator** (IFX w. LISATek/CoWare)
- **SIMD Simulator** (IFX)
- **ARMulator** (ARM)
- **SIMD Cores and TV / EQ - accelerators**
- **ARM 926**

**Shared Memory / System-Bus**
## New Class of Computing Architecture

### Server, Laptops

<table>
<thead>
<tr>
<th>Example</th>
<th>XEON</th>
<th>MuSIC-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power (typ.)</td>
<td>110 W</td>
<td>0.35 W</td>
</tr>
<tr>
<td>Clock</td>
<td>3 GHz</td>
<td>300 MHz</td>
</tr>
<tr>
<td>Cores</td>
<td>2</td>
<td>16 RISC+32 SIMD proc’s</td>
</tr>
<tr>
<td>Threads/core</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>gen.purp. GIPS</td>
<td>12</td>
<td>4.8</td>
</tr>
<tr>
<td>spec. purp. GIPS</td>
<td>24 (mmx)</td>
<td>9.6 (SIMD)</td>
</tr>
<tr>
<td>MB</td>
<td>17</td>
<td>0.9</td>
</tr>
<tr>
<td>Area (65nm)</td>
<td>435 mm²</td>
<td>43 mm²</td>
</tr>
</tbody>
</table>

### #all/y

<table>
<thead>
<tr>
<th></th>
<th>100 mill.</th>
<th>1 bill.</th>
</tr>
</thead>
</table>

---

*Example XEON
Power (typ.) 110 W
Clock 3 GHz
Cores 2
Threads/core 2
gen.purp. GIPS 12
spec. purp. GIPS 24 (mmx)
MB 17
Area (65nm) 435 mm²

MuSIC-2
Power (typ.) 0.35 W
Clock 300 MHz
Cores 16 RISC+32 SIMD proc’s
Threads/core 1
gen.purp. GIPS 4.8
spec. purp. GIPS 9.6 (SIMD)
MB 0.9
Area (65nm) 43 mm²

#all/y 100 mill. 1 bill.*
General Purpose-VLIW architecture for MM

4-channel architecture

Fetch
- L1 ICache

Decode
- Instruction Decode

Read
- Register Read

Execute
- ALU & Mult. & MME
- ALU & Ld/St & MME
- ALU & Ld/St & MME
- ALU & MME

Memory
- L1 DCache

Writeback
- Register Write
GP-VLIW Core is coprocessor

- OS support functions remain at ARM
- ARM/Java code compatible

IP increased largely

ARM+Co-Processor for Video and Graphics

QCIF videophone,
H.264 CIF play,
MPEG4 CIF play,
MPEG4 QCIF play/record,
MP3/AAC play/record,
low end games, office

L90:
4.6 mm²,
200 – 400/100 MHz, 45 - 140 mW,
380 kG, 96 KB
(w/o EBU, Peripherals)
Processor roadmap

- 4-ch**, 400 MHz, 1.2 V, 230 mW
- 4-ch**, 200 MHz, 0.9 V, 65 mW
- 2-ch**, 200 MHz, 0.9 V, 35 mW
- 2-ch**, 67 MHz, 0.9 V, 10 mW

** GP-VLIW excl. ARM926 with MMU, System Control, 100 MHz, 0.85 V (~ 15 mW)

- ARM11***, 350 MHz, 1V, ~120 mW
- ARM9*, 200 MHz, 1.2V, ~50 mW

* ARM926 incl. Move, MMU, System Control
*** ARM1136 incl. Media Ext., MMU, System Control
Binary Compile Technology is Key to CP Evolution

- **Examples**
  - FX32! by DEC: X86 to Alpha
  - Code Morphing by Transmeta: X86 to Crusoe or Efficeon
  - Intel: IA32 to IA64
  - Elbrus: IA32 and IA64 to E2k
  - ...

**VLIW Core is the only CPU**
- OS support functions integrated into GP-VLIW subsystem
- off-line/on-line binary compilation of Java/Arm executables

⇒ alternative to ARM, Xscale, …
## Multi-Standard CP Platform - NGA

<table>
<thead>
<tr>
<th>Management SW</th>
<th>Protocol Stacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSM GPRS</td>
<td>Edge UMTS</td>
</tr>
<tr>
<td>Edge UMTS</td>
<td>WCDMA/HSDPA</td>
</tr>
<tr>
<td>WCDMA/HSDPA</td>
<td>Bluetooth</td>
</tr>
<tr>
<td>Bluetooth</td>
<td>UWB</td>
</tr>
</tbody>
</table>

### Mapper, Compiler, Scheduler

- **Run-Time System**
  - Array of SIMD DSPs
  - ARM
  - ASIPs Filter, Turbo-Viterbi, LDPC, ...

### Operating System

- **GP VLIW+ext**
- **ARM**
- **High-end Graphics**

### Protocol Stacks

- Office
- Games
- Com Apps
- Speech
- Graphics
- MM Player
- MP3
- AAC
- MPEG4
- ... DVB-H
- H.264
Multi-Standard CP Platform - Tomorrow

Hand-over Service Selection
- 3G
- 802.11 a, b, g, n

Cognitive Radio
- LTE
- WiBro, WiMAX
- BT, UWB
- CR L1-SP

Protocol Stacks
- Office
- Speech
- MP3
- H.264
- Games
- Graphics
- AAC
- MPEG4
- Com Apps
- MM Player

Mapper, Assembler, Compiler, Scheduler

Multi-Band RF
- GP VLIW +BBext
- GP VLIW +BBext
- GP VLIW +MMU

Operating System, Dynamic Binary Compile

Shared Memory Hierarchy

High-end Graphics

Video Phone; Video-conference

Surveillance

Stop thinking
Design Criteria

area
power
scalability
flexibility = reusability in future designs, fixing errors
simplicity of programming model

• weights
• priorities

\[\Rightarrow\] architecture
Case 1: search for lowest power architecture provided area < bound

- Start from entry point

  1-1 ASIC  \rightarrow  area too high, power consumption minimal

- multiplex bas. func. comp.

  \rightarrow  area reduced, power consumption increased

- if area > bound, introduce new func. comp. more fine-grain than bas. func. comp.

  \rightarrow  area reduced, power consumption increased, some flexibility
Case 2a: search for most flexible architecture provided area, power < bounds

- Start from entry point

  ![fpga diagram](image)

  high flexibility, area too high, power consumption too high

- replace several clusters of fpga cells by reduced fpga cells

  flexibility reduced, area reduced, power consumption reduced

- if area > bound, replace by new fpga cells more coarse grain than previous ones

  flexibility further reduced, area further reduced, power consumption further reduced
Case 2b: search for most flexible architecture provided area, power < bounds

- Start from entry point

- choose smallest array of GP processing units with highest clock that fits power budget

Task partitioning

area, latency too high?
Design Space Exploration

• add instructions to GP ISA

result depends on initial task-partitioning

→ cost function for throughput and communication
System Design Framework

function modeling in C

- WCDMA, DVB-H, UWB,
- WLAN, WIMAX, …
- CR

architecture modeling in SystemC

- multiple program-mable processors
- reconfigurable DP+IC architecture
- fpga-inspired architectures

Common Research Agenda:
- benchmarks
- real-time in f&a space
- systematic design space exploration
- code generators
- „Future-proof“ framework for Function&Architecture Modeling (Matlab, Simulink, BeeCube, Metropolis, …?)