GEMSCLAIM
Greener Mobile Systems By Cross Layer Integrated Energy Management

Thomas Fahringer, University of Innsbruck, Austria

Consortium:
University of Innsbruck
Queen’s University Belfast
RWTH Aachen University
Politenica University of Timisoara

Start: September 1st, 2012
Duration: 3 years
Scientific background

- Personal computing: desktop computers → mobile systems

- Mobile systems are becoming more energy demanding
  - Radio interfaces
  - High computational load (games, multimedia apps.)

- “Major challenge for the mobile society in the next decade is that battery capacity is not keeping pace with Moore’s Law.“ by M. Muller, CTO of ARM

- Trade-off between energy and performance for mobile devices and many other computing systems:
  - desktop computing, cloud computing, high performance computing
Key challenges and potential impact

Project duration: Sept. 2012 – Aug. 2015

Cross layer energy management and optimization for mobile devices: HW/simulator, OS, compiler

- Energy-aware optimizing and parallelizing compiler
- Energy-proportional operating system
- Customizable HW modeling with energy monitoring facilities
  - simulator
  - FPGA

Potential impact

- control trade-off between energy optimization and performance
- additional 30% energy saving for mobile terminals
Consortium and synergies

Cross-layer power optimization

Project Coordination

Compiler/RTL/OS interface, Multi-objective optimisation
Energy accounting, abstractions, SLAs

Physical prototyping
Continuous energy monitors & sensors

Virtual prototyping & hardware design
Instruction-level power modelling

OS-FPGA, OS-VP interface, lightweight RTL

Energy-efficient programming and energy optimisation of mobile workloads at/across levels
## Work plan and milestones

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Delivery month</th>
<th>WP involved</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>6</td>
<td>WP1, WP2</td>
<td>Requirements specification and design of the GEMSCLAIM HW/SW environment</td>
</tr>
<tr>
<td>M2</td>
<td>15</td>
<td>WP1, WP3-WP7</td>
<td>Early prototype of the GEMSCLAIM HW/SW environment (able to simulate/execute benchmarks applications)</td>
</tr>
<tr>
<td>M3</td>
<td>30</td>
<td>WP1, WP3-WP7</td>
<td>Final prototype of the GEMSCLAIM HW/SW environment (able to perform energy savings benchmarking)</td>
</tr>
<tr>
<td>M4</td>
<td>36</td>
<td>WP1, WP7</td>
<td>Benchmarks fully optimized for energy (targeted savings: 30%) on FPGA with GEMSCLAIM SW; project completion</td>
</tr>
</tbody>
</table>
Project Management

- Internal project meetings: every 6 - 8 months
  - Timisoara, Feb. 2014, Innsbruck, Sept 2014; Belfast, April 2015
- Skype meetings
  - once a month
- Scientists exchange:
  - bilateral meetings (1-2 weeks)
  - extend project meetings by a few days for technical work
  - 2 months visit by Phd student
- Deliverables and documents
  - internal review
  - Annual reports for chist-era offices and national funding agencies
- Financial reporting
  - Once a year
- document repository
  - dropbox
- software repository
  - gitlab server in Belfast
- webpage
  - www.gemsclaim.eu
Project Management

- Quality Assurance
  - roles and responsibilities for each WP
  - time plan controlled by project coordinator
    - redmine
    - interaction with WP leaders
  - risk management
    - regular updates with input from WP leaders
- SW Engineering
  - Agile software development
  - Test-driven development with focus on requirements and goals
Resources and Funding

- For projects 1 and 2

<table>
<thead>
<tr>
<th>N°</th>
<th>Partner</th>
<th>Person. months</th>
<th>Total costs in €</th>
<th>Percentage of requested budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UIBK</td>
<td>40,375</td>
<td>143,598,24</td>
<td>43,13</td>
</tr>
<tr>
<td>2</td>
<td>QUB</td>
<td>26,4</td>
<td>262,944</td>
<td>58,3</td>
</tr>
<tr>
<td>3</td>
<td>UPT</td>
<td>23,314</td>
<td>71,226,00</td>
<td>28,49</td>
</tr>
</tbody>
</table>
Simulation based Power Models

- Power Models for Buses and Memories estimated from hardware counters
- Utilization counters ($u_{active}$, $u_{read}$, $u_{write}$) updated live during simulation
- Power coefficients $c_i$ calibrated offline with real hardware

---

CHIST-ERA Project Follow-up Seminar, Madrid, Spain, March 18-20, 2015
Parallel SystemC Simulation

- Contemporary heterogeneous multi-core system
  - System parameters (number/type of cores, mem size...) statically configurable
- GEMSCLAIM simulator is based on SystemC
  - Exchanging reference SystemC with ICE’s SCOPE SystemC kernel enables parallel simulation

Instruction set simulators are automatically mapped to simulation threads
- Each thread simulates on its own and maintains its own simulation time
- Maximum allowed time difference ($t_{la}$) is annotated by the system designer
- Time-Decoupling may induce a timing error during cross thread communication
Simulation speed vs. accuracy

- Test setup (parallel simulator uses 4 threads):
  - 8 RISC + 8 VLIW system
  - Application: “ocean” from SPLASH-2 benchmark suite

- 3x speedup possible with ~1% error in timing
- Timing error increases linearly with time decoupling
Validation of power models by physical measurements

- **Aim**
  - Validate and calibrate the power models implemented by GEMSCLAIM simulator by physical measurements using FPGA prototyping

- **Measurement methodology**
  - FPGA components power profiling: LUT, PLL, DSP, BRAM
  - PD_RISC core instruction set power profiling
  - Power profiling of Microblaze cores, DDR3 memory, and AXI interconnects
    - Instruction level power profiling of MB core
    - Transaction level power profiling of DDR3 memory and interconnects

- **Measuring samples**

![Graph of Pd-LDR vs. LDR rate](image)

\[ y = 0.0101x + 0.0864 \]
\[ R^2 = 0.996 \]

\[ y = 0.0106x - 0.0261 \]
\[ R^2 = 0.9881 \]
Validation of power models by physical measurements

Results

- Power consumption correlation matrix between simulation values and physical measurements

<table>
<thead>
<tr>
<th></th>
<th>130nm (faraday)</th>
<th>90nm (faraday)</th>
<th>65nm (faraday)</th>
<th>45nm (Nangate)</th>
<th>45nm (Atlys)</th>
</tr>
</thead>
<tbody>
<tr>
<td>130nm (faraday)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90nm (faraday)</td>
<td>92.11%</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65nm (faraday)</td>
<td>98.95%</td>
<td>92.53%</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45nm (Nangate)</td>
<td>98.67%</td>
<td>93.16%</td>
<td>98.37%</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>45nm (Atlys)</td>
<td>85.78%</td>
<td>95.47%</td>
<td>84.34%</td>
<td>86.39%</td>
<td>1</td>
</tr>
</tbody>
</table>

- Power and energy models for DDR3
  - \( P_{\text{DRAM}} = C_1u_{\text{RD}} + C_2u_{\text{WR}} + P_s \)
  - \( E_{\text{DRAM}} = N_{\text{RD}}E_{\text{RD}} + N_{\text{WR}}E_{\text{WR}} + P_sxt \)

Publications

- 2 joint papers
  - 1 accepted, 1 submitted
- 1 single partner paper

CHIST-ERA Project Follow-up Seminar, Madrid, Spain, March 18-20, 2015
Hardware prototyping and energy monitor

- Load and Energy Monitor (LEM) architecture
  - Shared bus based infrastructure (adapted from Wishbone bus)
  - Components
    - Central LEM
    - LEM Bus
    - LEM Sensors
Rethinking the OS

Energy slices (first-class resource)

Processor time slices (within energy budget)

Application A

Application B

CHIST-ERA Project Follow-up Seminar, Madrid, Spain, March 18-20, 2015
GEMSCALEM OS services

- POSIX threading interface
  - Shared Memory abstraction
- Manage energy consumption of HW components
  - LEM
- Export PM interfaces to the application layer
  - RAPMI (resource allocation and power management interface)
- Implement energy-aware scheduling policies
  - Pre-emptive multithreading
  - Energy Interrupt
  - Synergy with Application Runtimes
- Ports
  - Virtual Platform (simulator)
  - FPGA Microblaze MPSoC
  - Legacy platforms (Linux-based)
RAPMII: Component-level power management

- **RAPMII**
- **Sensor control**
  - Start, Stop
  - Read
  - Set sampling rate
- **HW control**
  - DVFS
  - Halt cores

- **rapmi_ctrl**(mode, rate, power_level)
  - sensor mode
  - rate
  - power_level
- **rapmi_start()**
  - Start sensor measuring
- **rapmi_stop()**
  - Stop sensor measuring
- **rapmi_read**(rapmi_t *pm)
  - Read sensor values
- **rapmi_set_freq**(freq)
  - Set core frequency
- **rapmi_get_freq()**
  - Returns core frequency
Rethinking OS scheduling

- Distribute *energy* slices instead of time slices
  - Energy interrupts (LEM)
- Priority-based policies
  - Interactive applications (User assisted)
- Multi-objective optimization
  - Energy and Performance
  - Multi-program workloads
- HW and SW re-configuration
  - HW: DVFS, Halt cores
  - SW: DCT (Dynamic Concurrency Throttling) – Synergy between OS and Application Runtimes

CHIST-ERA Project Follow-up Seminar, Madrid, Spain, March 18-20, 2015
Some results

- ODROID-XUE platform
  - Cores, Memory power sensors
  - Heterogeneous cores
- Multi-program workloads
  - Per-program characterization
  - Synergistic scheduling (OS runtime and Application Runtime)
- Different optimizing schedulers
  - Energy (e)
  - Delay (d)
  - Energy-delay products (edp, e^2dp, ed^2p)
The GEMSCLAIM Compiler is a source-to-source C compiler
- supports C programs annotated with OpenMP+
- energy-aware extension to OpenMP

Programs are analyzed and divided into meaningful code regions
- individually tunable and annotated with additional metadata

Execution orchestrated by the runtime system
- handles dynamic decision making and scheduling
- communicates with the GEMSCLAIM OS.
Compiler runtime system

- Compiler runtime fully ported to relevant platforms
  - GEMSCLAIM virtual platform
  - ARM boards (e.g. ODROID XU+E)
  - X86/Linux baseline version

- Supports *upcalls* from OS/low-level layer
  - E.g. `irt_set_dop(uint32)` allows OS to control per-program degree of parallelism
  - ➔ more effective parallelism management than simple thread packing!
OpenMP extensions (OpenMP+)

- Now fully supported in compiler and runtime system:
  - Region construct (handling code outside OpenMP regions)
  - Objective clause (defining multiple optimisation objectives and constraints)
  - Param clause (define tunable parameters for compiler optimisation)

- New: *quality of service* constraints
  - Can be mapped to user-defined parameters
  - Example usage in video decoder:

```c
#pragma omp parallel for schedule(dynamic)
    objective(E : T<1/f_rate; Q<3) param(scaling, range(1:8:1))
for (int y=0; y<rows; y+=2*scaling)
  for (int x=0; x<cols-2; x+=scaling) {
    ...
    if(scaling > 1) { ... }
  }
```
e-optimizer

- Multi-objective optimization for OpenMP+ programs
  - Dynamically searches best configuration for given goals and constraints
  - Combines random sampling (to prevent local minima) with multi-dimensional hill climbing (to quickly converge)
  - Up to 77% energy savings on mobile and 31% on desktop

Mobile platform, 704x576 resolution

Desktop platform, 1408x1152 resolution
GEMSCLAIM Publications 2012 and 2013

- A Multi-Objective Auto-Tuning Framework for Parallel Codes Herbert Jordan, Peter Thoman, Juan J. Durillo, Simone Pellegrini, Philipp Gschwandtner, Thomas Fahringer, and Hans Moritsch. SC '12 November 11 - 15, 2012


GEMSCLAIM Publications 2014 and 2015

- Low-Cost Hardware Infrastructure for Runtime Thread Level Energy Accounting, Marius Marcu, Oana Boncalo, Madalin Ghenea, Alexandru Amaricai, Cosmin Cernazanu, Energy Efficiency with Heterogeneous Computing Workshop (EEHCO), Amsterdam, Holland, Jan. 2015
- Hardware support for performance measurements and energy estimation of OpenRISC processor, Submitted to 10th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI), Timisoara, Romania, May 2015
- Low-Cost Hardware Infrastructure for Runtime Thread Level Energy Accounting, UPT, RWTH and QUB, Invited to IET CDT Journal, Deadline Mar. 2015
- Low-Cost Hardware Infrastructure for Runtime Shared Memory Thread Level Energy Accounting, Marius Marcu, Oana Boncalo, Madalin Ghenea, Jan Henrik Weinstock; Rainer Leupers, Submitted to 8th ACM International Systems and Storage Conference (SYSTOR), Haifa, Israel, May 2015
- OpenMPE: A Language Extension for Application-level Energy Awareness. Ferdinando Alessi, Peter Thoman, Giorgis Georgakoudis, Thomas Fahringer, and Dimitrios S. Nikolopoulos, (submitted)
Important GEMSCLAIM project objectives

- Cross layer energy management and optimization for mobile devices: HW/simulator, OS, compiler
- Control trade-off between energy optimization and performance
- Additional 30% energy saving for mobile terminals
- GEMSCLAIM achieved all of that.
Conclusions and way forward

- Cross-layer energy management and optimization requires breaking barriers between layers
  - Common abstractions, metrics
  - Synergetic optimization approaches
- Measuring energy consumption remains challenging
  - Machine-specific, intrusive, coarse-grain
  - Hybrid modelling & measurement approaches are the only viable
- Software needs to evolve to break the energy wall
  - Energy optimization should become explicit
  - Sensitivity of energy to software structures is not well understood
- **GEMSCLAIM provides promising solutions for these challenges in a holistic approach, demonstrated on a physical HW/SW substrate**
Sustainability

- Joint FETHPC proposal
- Part of the SW under Apache 2.0 license (open source)
- Parts of the GEMSCLAIM OS Runtime are used in current research projects (FP7 NanoStreams) and are a foundation for future research proposals under the Horizon 2020 research programme
- Parallel simulation technology developed in GEMSCLAIM will be used for upcoming industry collaboration projects.
- The GEMSCLAIM VP will be used as a basis for future research projects proposed in the context of Horizon-2020 Research Programme of the EU.

www.gemsclaim.eu