A Generic Profiling Infrastructure for the Hyperbolic PDE Solver Engine ExaHyPE
Part II

Fabian Güra

www.exahype.eu

November 2016

The project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 671698 (ExaHyPE).

Introduction

What this presentation is not about:

- **Numerics** deep dive into ADER-DG, Limiting, etc. (first $\sim$ 30 pages of the thesis)
- No beautiful **pictures**, convergence **plots**, application **examples**, ...
- **Demo session** ;-(

Introduction

What this presentation is not about:

- **Numerics** deep dive into ADER-DG, Limiting, etc. (first $\sim$ 30 pages of the thesis)
- No beautiful **pictures**, convergence **plots**, application **examples**, ...
- **Demo session** ;-(

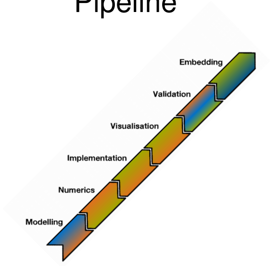
What this presentation is about:

- Review: **Context** and **motivation** behind the project
- **Hardware performance monitoring** (HPM) on x86
- A **generic profiling infrastructure** for ExaHyPE
- Preliminary **profiling results** and two **case studies** focusing on metrics-driven **performance engineering**

Context & Motivation

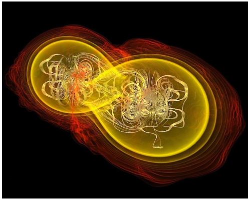
Important aspects in the context of Scientific Computing:

Simulation Pipeline



Exascale Computing

Hyperbolic Balance Laws



ExaHyPE: The Project

Vision:

- Three pillars of scientific progress:
  Theory, Experiment and **Simulation**
- Programming (exascale) supercomputers is a key challenge
- The ExaHyPE project seeks to address the **software aspect** of supercomputer development
  - Development of new mathematical and algorithmic approaches
  - Initial focus on applications in geo- and astrophysics
  - In correspondence with Europe's **2020 exascale strategy**
- Goal: Become **the engine** for large HCL simulations!

ExaHyPE: The Project

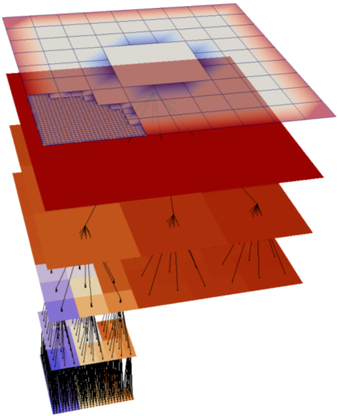
Objectives:

- **Energy efficiency** on tomorrow's supercomputing hardware
- Scalable algorithms through well-balanced **dynamic adaptivity**
- **Compute-bound** simulations in spite of slow memory and networks
- Extreme parallelism on **unreliable hardware**

ExaHyPE: The Project

Approach:

- High-order space-time Discontinuous Galerkin method (**ADER-DG**) with a-posteriori FVM based subcell limiting
- Dynamically adaptive Cartesian grids (**AMR**), space filling curves and dynamic load balancing (**Peano**)
- Hardware specific optimization of dominant compute kernels (**libxsmm**)



Adaptive grids in Peano
(via Tobias Weinzierl, 2014)

Profiling for ExaHyPE: Motivation

You can't optimize for what you don't measure!

Profiling provides metrics to

- obtain a **baseline**,
- **guide** and **track progress** of optimization efforts,
- **compare** current status to other state of the art solutions.

We don't need a third one of these on campus!

Profiling and Energy-aware Computing in Modern x86 Systems

The Current Prevalence of x86 in High Performance Computing

Architecture	abs.	rel.
x86	468	93.6%
Power	23	4.6%
SPARC	7	1.4%
Sunway	2	0.4%

(a) CPU Architecture

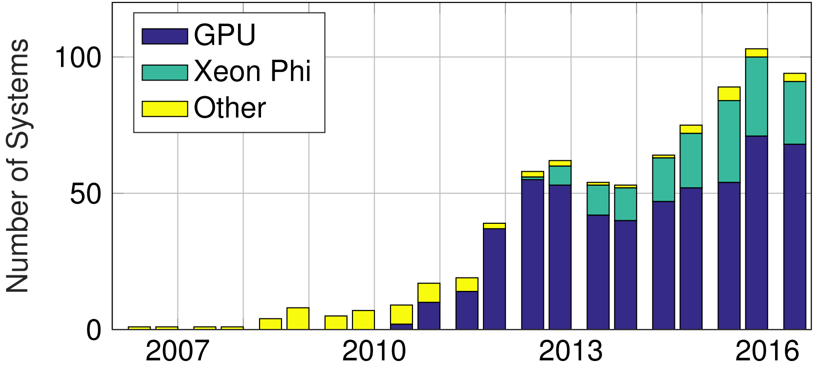
Accelerator	abs.	rel.
None	406	81.2%
GPU	66	13.2%
Xeon Phi	23	4.6%
Other	5	1.0%

(b) Accelerator Cards

Table: Distribution of CPU architectures and accelerator cards of the supercomputers listed in the June 2016 Top500 list. Systems that have both GPUs and Xeon Phi accelerator cards are listed as “Other.”

Profiling and Energy-aware Computing in Modern x86 Systems

The Current Prevalence of x86 in High Performance Computing?



Hardware Performance Monitoring on x86 Platforms

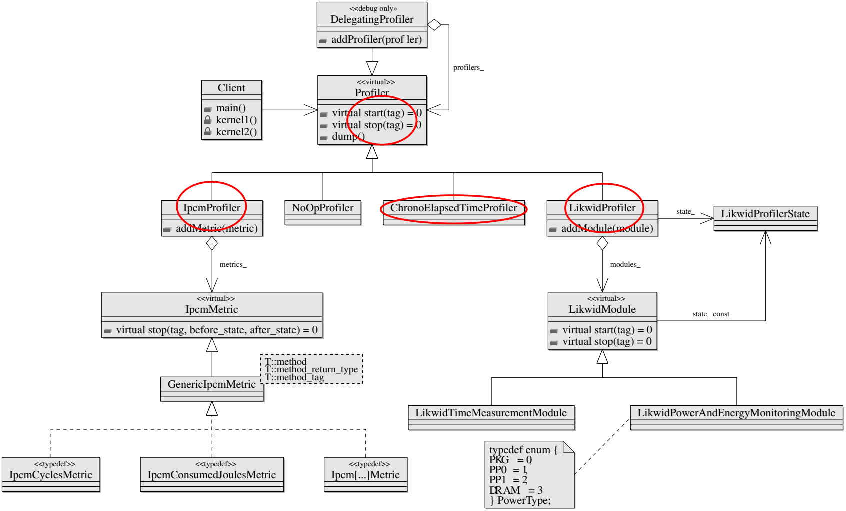
- History: 1st gen. Pentium (1993) collects metrics on interaction between hardware and code.
- Access undocumented/proprietary $\rightarrow$ reverse engineering
- General Principle: Performance counters programmable via module specific registers (MSRs)
- Today: Separate performance monitoring units (PMUs) for cores and shared resources
- Advantage: Low overhead and unintrusive
- Disadvantage: Not part of ISA, i.e. unstable and undocumented
- Use libraries (PAPI, LIKWID, . . .)

Energy Monitoring in x86

- Context: Underprovisioning of power supply in datacenters
- History: Sandy Bridge (2011) and Bulldozer (2013) support prescription of dynamic power limits (RAPL, APM)
- Control task gives rise to on-chip power estimation
- Exposed via MSRs, separate for PKG, DRAM, ?
- Good: Precise ($\sim \mu J$) and valid ($\rightarrow$ literature)
- Bad: Update rate ($\leq$1 kHz)

Energy Monitoring in x86

Profiling for ExaHyPE: Architecture



Implementations

Profiling for ExaHyPE: Metrics

What we can measure at the moment:

- Kernel call counts
- Cycles and wall clock time
- RAM reads/writes
- **Flop/s** (estimate)
- **Energy consumption** (Core, RAM, package; estimate)
- Cache hits/misses/ratio (I, L2, L3)
- **Cycles lost** due to cache misses (L2, L3)
- Instructions retired
- Branch prediction ratio
- ...

Overheads

Complete Simulation

The seventh order itch

Finding the juiciest fruits

reliably.

Conclusion

Key aspects:

1. Exascale Computing is about both hardware and **software**
2. ExaHyPE as an **answer to future challenges** in Scientific Computing
3. **Frameworks** are necessary to manage the increasing complexity in HPC
4. Profiling is important: **Measure, then optimize!**

A Generic Profiling Infrastructure for the Hyperbolic PDE Solver Engine ExaHyPE
Part II

Fabian Güra

www.exahype.eu

November 2016

The project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 671698 (ExaHyPE).